

COORDENAÇÃO

Carlos Guardado da Silva
Jorge Revez > Luís Corujo



ORGANIZAÇÃO DO CONHECIMENTO NO HORIZONTE 2030

DESENVOLVIMENTO SUSTENTÁVEL E SAÚDE



Atas do V Congresso ISKO Espanha-Portugal

> Universidade de Lisboa > Faculdade de Letras
> 25 e 26 de novembro de 2021



ORGANIZAÇÃO DO CONHECIMENTO NO HORIZONTE 2030



DESENVOLVIMENTO SUSTENTÁVEL E SAÚDE

Atas do V Congresso ISKO Espanha-Portugal

> Universidade de Lisboa > Faculdade de Letras
> 25 e 26 de novembro de 2021

Colecção CA – Ciência Aberta

Direcção: Jorge Revez

Títulos publicados:

- 1 **Organização do Conhecimento no Horizonte 2030:**
Desenvolvimento Sustentável e Saúde: Atas do V Congresso ISKO Espanha-Portugal

Carlos Guardado da Silva > Jorge Revez > Luís Corujo

COORDENAÇÃO

ORGANIZAÇÃO DO CONHECIMENTO NO HORIZONTE 2030

DESENVOLVIMENTO SUSTENTÁVEL E SAÚDE

Atas do V Congresso ISKO Espanha-Portugal

> Universidade de Lisboa > Faculdade de Letras
> 25 e 26 de novembro de 2021



Edições Colibri

Biblioteca Nacional de Portugal
– *Catálogo na Publicação*

Silva, Carlos Guardado da, Revez, Jorge & Corujo, Luís (Eds.). (2021). *Organização do Conhecimento no Horizonte 2030: Desenvolvimento Sustentável e Saúde: Atas do V Congresso ISKO Espanha-Portugal*, Universidade de Lisboa. Faculdade de Letras, 25 e 26 de novembro de 2021. Lisboa: Centro de Estudos Clássicos, Colibri.
(CA – Ciência Aberta ; 1)
ISBN 978-989-566-137-4
CDU 025.4(063)

Título: Organização do Conhecimento no Horizonte 2030:
Desenvolvimento Sustentável e Saúde:
Atas do V Congresso ISKO Espanha-Portugal

Coordenação: Carlos Guardado da Silva, Jorge Revez, Luís Corujo

Edição: Centro de Estudos Clássicos, Faculdade de Letras, Universidade de Lisboa

Capa: Raquel Ferreira

DOI: <https://doi.org/10.51427/10451/50067>

Esta publicação é financiada por Fundos Nacionais através da FCT – Fundação para a Ciência e a Tecnologia, I.P., no âmbito do projecto UIDB/00019/2020

Lisboa, novembro de 2021

In memoriam

Emilia Currás
Maria da Graça Simões

Num deserto sem água
Numa noite sem lua
Num país sem nome
ou numa terra nua

Por maior que seja o desespero
Nenhuma ausência é mais funda do que a tua.

(Sophia de Mello Breyner Andresen, *No mar novo*, 1958)

Organização

ISKO – The International Society for Knowledge Organization – Capítulo Ibérico
Universidade de Lisboa, Faculdade de Letras, Centro de Estudos Clássicos

Coordenação Geral

Carlos Guardado da Silva (Universidade de Lisboa)

Jorge Revez (Universidade de Lisboa)

Luís Corujo (Universidade de Lisboa)

Comissão Organizadora

Ana Célia Rodrigues, Universidade Federal Fluminense, Brasil

Ana Lúcia Terra, Universidade de Coimbra, Portugal

Blanca Rodríguez Bravo, Universidad de León, Espanha

José António Frías Montoya, Universidad de Salamanca, Espanha

Luís Corujo, Universidade de Lisboa, Portugal

Maria Cristina V. Freitas, Universidade de Coimbra, Portugal

Maria Manuel Borges, Universidade de Coimbra, Portugal

Maria Teresa Costa, Universidade de Lisboa, Portugal

Pedro Estácio dos Santos, Universidade de Lisboa, Portugal

Paula Ochôa, Universidade Nova de Lisboa, Portugal

Rodrigo Furtado, Universidade de Lisboa, Portugal

Rosa San Segundo Manuel, Universidad Carlos III de Madrid, Espanha

Comissão Científica

Alexandre Faben, Universidade Estadual do Espírito Santo, Brasil

Ana Célia Rodrigues, Universidade Federal Fluminense, Brasil

Ana Lúcia Terra, Universidade de Coimbra, Portugal

Blanca Rodríguez Bravo, Universidad de León, Espanha

Carlos Guardado da Silva, Universidade de Lisboa, Portugal

Críspulo Travieso Rodríguez, Universidad de Salamanca, Espanha

Daniel Martínez-Ávila, Universidad Carlos III de Madrid, Espanha

Fernanda Ribeiro, Universidade do Porto, Portugal

Fidélia Ibekwe, Aix-Marseille Université, França

Francisco Javier García Marco, Universidad de Zaragoza, Espanha

Gercina Lima, Universidade Federal de Minas Gerais, Brasil

Isidoro Gil Leiva, Universidad de Murcia, Espanha

Jesús Gascón García, Universitat de Barcelona, Espanha

Jesús Tramullas, Universidad de Zaragoza, Espanha,

Jorge Revez, Universidade de Lisboa, Portugal

José Antonio Frías Montoya, Universidad de Salamanca, Espanha

José Antonio Moreiro González, Universidad Carlos III de Madrid, Espanha

José Augusto Chaves Guimarães, Universidade Estadual Paulista, Brasil

Luís Corujo, Universidade de Lisboa, Portugal

Maria Cristina V. Freitas, Universidade de Coimbra, Portugal

María José López-Huertas, Universidad de Granada, Espanha

María Luísa Alvite Díez, Universidad de León, Espanha

Maria Manuel Borges, Universidade de Coimbra, Portugal

Maria Teresa Costa, Universidade de Lisboa, Portugal

Mario Barité, Universidad de la República, Uruguai

Moises Rockembach, Universidade Federal do Rio Grande do Sul, Brasil

Natália Bolfarini Tognoli, Universidade Federal Fluminense, Brasil

Olívia Pestana, Universidade do Porto, Portugal

Paula Ochôa, Universidade Nova de Lisboa, Portugal

Renato Rocha Souza, Fundação Getúlio Vargas, Brasil

Rosa San Segundo Manuel, Universidad Carlos III de Madrid, Espanha

Stephanie Colombo, Universidad de la República, Uruguai

Thiago Henrique Bragato Barros, Universidade Federal do Rio Grande do Sul, Brasil

SUMÁRIO

Apresentação	15
Presentación.....	19
Health KOSs in a Data-driven Age Marcia L. Zeng	23
Os desafios para a formação e o desenvolvimento profissional na organização do conhecimento à luz da agenda 2030 Natália Bolfarini Tognoli.....	39
A produção científica sobre a agenda 2030 indexada na <i>web of science</i> e <i>scopus</i> : ciências sociais e ciência da informação Teresa Costa, Luísa Alvim.....	53
Identificação como base para a organização do conhecimento arquivístico: contribuições para o debate sobre acesso à informação nos arquivos municipais do Brasil na agenda 2030 Alexandre Faben, Ana Célia Rodrigues, Carlos Guardado da Silva	67
Modelo-base de vocabulário controlado sobre violência doméstica contra crianças e adolescentes Rita Isabel Figueira Costa, Maria Cristina Vieira de Freitas, Daniel Martínez-Ávila	79
Sistema de organização do conhecimento e pós-disciplinaridade: mudança climática no contexto da agenda 2030 Marcos Gonçalves Ramos, Priscila Ramos Carvalho, Rosali Fernandez de Souza.....	99
Organização do conhecimento na Câmara dos Deputados Brasileira: estudo teórico-metodológico para categorização das proposições sobre direitos em construção das mulheres Carla Maria Martellote Viola, Luana Farias Sales	113
O objetivo justiça e paz da agenda 2030/ONU por meio de dados do PORDATA Camila Mattos da Costa.....	125
Organização da informação e do conhecimento em ambientes digitais: uma apresentação do Herbário virtual IAN da Embrapa Cristiane Pantoja de Moraes	135
The cultural aspects of knowledge organization toward a sustainable development Asmaa Bouaamri, Ágnes Hajdu Barát.....	145
Aplicação de metadados na padronização de registros de ocorrência de espécies no contexto da ciência cidadã para a biodiversidade: um estudo de caso Filipi Miranda Soares, Raíssa Yuri Hamanaka	179

Un mapeo terminológico del dominio covid-19 con base en bibliometría y garantía académica Mario Barité, Exequiel Fontans	191
A nova plataforma PubMed: análise dos peritos Maria Luz Antunes, Carlos Lopes, Licínio Roque	205
Indexação e repositórios institucionais: enquadramentos, definições e traços discursivos Marta Filipa Pinheiro dos Reis Luro, Maria Cristina Vieira de Freitas	217
Ciência da informação e inteligência artificial: um caminho para arquivos e bibliotecas inteligentes Moisés Rockembach.....	235
Modelos de requisitos para sistemas de organização do conhecimento Luís Corujo, Jorge Revez	243
Organización y descripción de información en wikis semánticos: wikibase, semantic mediawiki y cargo Jesús Tramullas, Ana I. Sánchez-Casabón, Piedad Garrido-Picazo.....	263
O uso da inteligência artificial na organização do conhecimento: uma revisão sistemática Roberta Jerônimo da Silva.....	273
O gênero cinematográfico enquanto ferramenta de auxílio a categorização de filmes de terror em serviços de <i>video on demand</i> Rafael Rodrigo do Carmo Batista, Thiago Henrique Bragatto Barros, Rita do Carmo Ferreira Laipelt.....	285
Uma proposta de modelo conceitual para estruturar uma base de dados de lições aprendidas em projetos Mauricio Augusto Cabral Ramos Junior.....	295
Identificación y validación de formatos de ficheros orientada a la preservación digital de documentos. De Tika a Droid María-José Baños-Moreno, María González Balanza, Javier Lurquí López, Francisco-José Valentín-Ruiz	307
The brazilian current research information system: brcris Luiz Pinto, Washington Luís Ribeiro de Segundo, Luc Quoniam, Thiago Magela Rodrigues Dias.....	319
Organização do conhecimento no Brasil: uma análise do domínio pelas teses de doutorado em ciência da informação Adilson Luiz Pinto, José Augusto Chaves Guimarães.....	331
O <i>software</i> livre e a descrição arquivística no meio digital: o uso do AtoM em Portugal e no Brasil Susana Sofia Cunha, Maria Cristina Vieira de Freitas.....	343

Um modelo semântico baseado em ontologia para o cris brasileiro Vivian S. Silva, Tales Moreira, Thiago M. Dias, Josir Gomes, Washington C. Segundo	361
As redes sociais de colaboração internacional de autores impacto nas revistas e na difusão do conhecimento em ciência da informação Ana Gouveia Coelho	371
Divulgação <i>online</i> de coleções de fotografia: a utilização do AtoM em Portugal Susana Sofia Cunha.....	383
O impacto da inteligência artificial nos serviços de informação: inovação e perspectivas para as bibliotecas Liliana Isabel Esteves Gomes, Viviana Fernández Marcial Miguel Nuno Marques dos Santos	393
Tras la iconografía e iconología de las representaciones del libro en las artes: diseño de una herramienta para su descripción y catalogación Carlos Díaz-Redondo, José Antonio Frías, Elvira-Julieta Miguélez-González.....	407
Museus brasileiros com coleções de arte: análise dos instrumentos de registro para bens culturais Camila Aparecida da Silva	425
Glossário terminológico definitivo para sistemas de organização do conhecimento Luander Falcão, Maristela Sanches Lima Mesquita, Benildes Coura M. S. Maculan	435
Importância do controle de vocabulário em um repositório institucional: estudo de caso Isidoro Gil-Leiva, Isaque Katahira, Mariângela Spotti Lopes Fujita.....	447
La ordenación por relevancia de los resultados de búsqueda de información científica: posición y frecuencia. Estudios de caso Isidoro Gil-Leiva.....	461
Modelo semântico de informações corporativas para tratamento de objetos multimídia Sergio de Castro Martins	473
A inclusão das mulheres na Brigada Militar no RS: a análise do domínio como subsídio para construção de uma taxonomia Carine Melo Cogo Bastos; Thiago Henrique Bragato Barros.....	491
Os nativos digitais e modernização das bibliotecas universitárias: tendência ou modismo? Thais Batista Zaninelli Sandra Gomes de Oliveira Reis, Ana Luise Moura	497
Os benefícios da folksonomia com curadoria para a recuperação da informação em bibliotecas universitárias Aline Araújo, Camilla Oliveira	503

Caracterização conceitual do tratamento temático da informação na ciência da informação no Brasil: análise da produção científica no período de 2001 a 2020 Lais Pereira de Oliveira, Maria Cláudia Cabrini Grácio e Daniel Martínez-Ávila.....	509
Estudo da paisagem rural e sua rede de conceitos: análise de domínio a partir das perspectivas da comunidade discursiva Vera Lucia Punzi Barcelos Capone, Rosa Inês de Novais Cordeiro.....	521
Mapa teórico conceitual para a construção de categorias em projetos transdisciplinares Andréa Doyle, Vera Dodebei.....	535
Epistemología feminista y organización del conocimiento en el contexto de isko ibérico Carlos Cândido de Almeida, Rosa San Segundo Manuel, Daniel Martínez-Ávila.....	543
Sistemas de organização do conhecimento e o conceito de regime de informação: relações possíveis Ana Flávia Dias Zammataro, Ana Cristina de Albuquerque.....	559
Semiótica documental: retomando um diálogo Alexandre Robson Martines, Carlos Cândido de Almeida	569
Pragmatismo e categorias fenomenológicas de Peirce: um debate sobre a noção de conceito Alexandre Robson Martines, Valdirene Aparecida Pascoal, Carlos Cândido de Almeida	583
Cunho evidencial e indicial da fotografia na organização do conhecimento arquivístico Bruno Henrique Machado, Rafael Semidão, Telma Campanha de Carvalho Madio.....	595
Afirmando Pereira: a <i>integrative levels classification</i> na representação do assunto ficcional Patrícia de Almeida, Claudio Gnoli.....	605
As taxonomias navegacionais facetadas e a produção científica da ciência da informação: tendências temática e diacrónica (2011-2020) Ana Gouveia Coelho, Gercina Ângela de Lima Maria Manuel Borges.....	617
Representação da informação em arquivos deslocados: uma metassíntese de literatura qualitativa L. S. Ascensão de Macedo, Carlos Guardado da Silva e Maria Cristina Vieira de Freitas ...	635
Epistemologia da organização do conhecimento: um novo velho paradigma Carlos Cândido de Almeida	663
Disciplinas e teorias da ciência da informação: a organização do conhecimento em foco Richele Grengue Vignoli, Carlos Cândido de Almeida	681

Reflexões dos conceitos de organização do conhecimento e organização da informação pelo prisma de Bräscher e Café: a emergência de um paradigma no GT2 do Enancib? Marcos Oliveira da Costa, Thiago Henrique Bragato Barros.....	693
A organização do conhecimento em repositórios institucionais: uma análise da literatura recente publicada em periódicos de biblioteconomia e ciência da informação Mariângela Spotti Lopes Fujita, Carmen Agustín-Lacruz, Jéssica Beatriz Tolare, Ana Lúcia Terra, Gema Bueno-de-La-Fuente.....	703
O conceito de informação peirceana e sua atualidade nos estudos da informação Valdirene Aparecida Pascoal, Maria Eunice Quilici Gonzalez, Carlos Cândido de Almeida ...	717
Sistemas de organização do conhecimento e humanidades digitais: possíveis interlocuções a partir da abordagem da análise do domínio Ana Cristina de Albuquerque, Ania Rosa Hernadez Quintana.....	727
Aproximações entre organização do conhecimento e humanidades digitais Laura Mariane de Andrade, Paula Regina Dal'Evedove	739
Epistemologia social e organização do conhecimento: contribuições as abordagens culturais Wilson Roberto Veronez Júnior, Carlos Cândido de Almeida, Daniel Martínez-Ávila, Sonia Maria Troitiño Rodriguez	751
A teoria do conceito em Hegel aplicada à organização do conhecimento: um estudo dialógico e dialético Marco André Feldman Schneider, Marco Antonio de Carvalho Bonetti, Gustavo Silva Saldanha, Fernanda Valle, Diogo Xavier	763
Arquivologia e organização do conhecimento: uma análise nos anais da isko Brasil, isko internacional e isko ibérico Wilson Roberto Veronez Júnior, Daniel Martínez-Ávila, Sonia Maria Troitiño Rodriguez	775
Produção do conhecimento científico em organização do conhecimento da arquivologia no Brasil Rosale de Mattos Souza, Natália Araujo Lima	787
A política de indexação para a organização do conhecimento em museus: aplicação do protocolo verbal individual Raul de Azevedo Carvalho, Luciana di Paula Andrade da Fonseca, Franciele Marques Redigolo, Mariângela Spotti Lopes Fujita	793
Análise da produção e da colaboração científica na revista <i>Knowledge Organization</i> (2016-2020) Ana Beatriz Silva, Maria Cristina Vieira de Freitas	805

Elementos de interdisciplina en los cursos de posgrado iberoamericanos de organización del conocimiento: protocolo de investigación Mario Barité, Varenka Parentelli	825
Inovações dos recursos educacionais abertos: o caso da Universidade de São Paulo, Brasil Célia Regina de Oliveira Rosa	839
Organización del conocimiento en el diseño curricular académico. Perspectivas desde los nuevos grados en documentación implantados en España María Luisa Alvite Díez	851
Personal knowledge organizing through online collaborative writing tools Dalbert Marques Oliveira, Ana Lúcia Terra	863
Indexação de imagens fotográficas raras: um olhar sobre o álbum da construção do Canal do Mangue na cidade do Rio de Janeiro, Brasil Melina de Brito dos Santos, Rosali Fernandez de Souza, Jeorgina Gentil Rodrigues	875
Organização do conhecimento em arquivos: o uso da identificação arquivística na classificação de documentos Natália Bruno Rabelo, Clarissa Moreira dos Santos Schmidt	883

APRESENTAÇÃO

Quando o Outono de 2021 avança, sobre nós, “irreversivelmente o tempo foge”¹, como o poeta romano Vergílio bem notara, registando-o em *Geórgicas* (Liv. III). É o fim do Ano II da Pandemia COVID-19, pautado por uma crise sanitária, económica e social à escala global, que nos revela, no século XXI, o quanto é frágil a vida coletiva, assim como insuficientes os esforços e os projetos humanos para uma vida equilibrada e sustentável. Pela primeira vez na história, é uma crise vivida de forma universal, em direto na TV e na Internet, com uma profusão de informação – uma Infodemia – nunca antes vista, com impactos que serão certamente duradouros e persistentes. Uma crise que a todos, de alguma forma, afetou e permanece afetando, e na qual a informação e o conhecimento tiveram, como continuam a ter, um papel preponderante. Como em outras crises globais, o desenvolvimento técnico e científico caminhou de mãos dadas com a tragédia de milhões de vidas humanas perdidas. Este cenário pandémico soma-se a uma crise climática anterior, provocada pelas alterações do clima terrestre, pelo aquecimento global, pela desestabilização dos ecossistemas e pela mitigação da biodiversidade. No entanto, saúde e clima são duas faces de um mesmo problema, que coloca desafios acrescidos no âmbito do Desenvolvimento Sustentável.

O Desenvolvimento Sustentável, que se desenvolve a partir de uma perspetiva eminentemente ambiental, que emerge na década de 60 do século XX, é um conceito cunhado em 1987, através da Comissão Mundial do Meio Ambiente e Desenvolvimento. Neste ano, na *Declaração Brundtland*, a Organização das Nações Unidas (ONU) definia o Desenvolvimento Sustentável como o desenvolvimento que satisfaz as necessidades do presente sem comprometer a capacidade das gerações futuras para satisfazer as suas, visando a qualidade de vida e respeitando os limites da capacidade dos ecossistemas. Com a evolução social e política, o conceito amplia-se e consolida-se, alcançando uma definição próxima do seu conteúdo semântico atual em 1994, quando, o *International Council for Local Environmental Initiatives* define Desenvolvimento Sustentável como o desenvolvimento que proporciona serviços económicos, sociais e ambientais básicos a todos os membros de uma comunidade, sem comprometer a viabilidade dos sistemas sociais, naturais e manufaturados de que aqueles dependem. O conceito afirmar-se-ia, sobretudo, a partir do início do século XXI com a 2.^a Cimeira da Terra, que teve lugar entre 24 de agosto e 4 de setembro de 2001, em Joanesburgo (África do Sul), sob o título *Cimeira Mundial para o Desenvolvimento Sustentável*, integrando os pilares económico, social e ambiental, bem como com as *Prepcom*, conferências preparatórias da Cimeira de Joanesburgo. Em 2015, um conjunto alargado de metas foi consolidado nos 17

¹ “fugit irreparabile tempus”.

Objetivos do Desenvolvimento Sustentável, que formam a Agenda 2030 da ONU. Esta engloba problemas universais, como a fome ou a pobreza, dois *topoi* indissociáveis e aqueles que naturalmente assumem maior relevância, e almeja desenvolver aspetos essenciais como a saúde, o acesso a água potável e saneamento, a educação, a igualdade de género, as energias renováveis, o trabalho digno, a inovação, a redução das desigualdades, as cidades sustentáveis, que incluem a produção e o consumo sustentáveis. Procura, ainda, combater as alterações climáticas, proteger os oceanos e os ecossistemas terrestres, bem como alcançar a paz e a justiça, num esforço global de parcerias para o desenvolvimento e a melhoria da qualidade de vida das populações.

A Agenda 2030 da Organização das Nações Unidas é uma agenda global, para enfrentar problemas globais, que coloca o desenvolvimento sustentável no topo das prioridades internacionais e constitui-se como uma visão comum para a Humanidade. O Conhecimento é uma dimensão central nos 17 Objetivos de Desenvolvimento Sustentável. A informação e as tecnologias digitais são entendidas como essenciais para a comunicação global e determinantes para a aceleração do progresso humano com vista a um desenvolvimento pleno da Sociedade do Conhecimento. Na linha desta agenda, o Horizonte Europa, que pretende desenvolver a investigação e a inovação no espaço europeu entre 2021 e 2027, tem como um dos componentes centrais da sua visão ajudar a alcançar os Objetivos de Desenvolvimento Sustentável.

A obra, que agora se apresenta, reúne um importante conjunto de contributos científicos trazidos até Lisboa no âmbito do *V Congresso Espanha-Portugal da ISKO*, o evento que reúne bianualmente o Capítulo Ibérico da ISKO, uma comunidade de investigadores dos países ibéricos, que abraça e dialoga com os seus colegas da América Latina. Não são apenas as línguas em comum que nos unem, mas laços de amizade já antigos, formando, assim, uma comunidade que é já ibero-americana. Esta obra recebe o mesmo título do evento que lhe deu origem: *Organização do Conhecimento no Horizonte 2030: Desenvolvimento Sustentável e Saúde: Atas do V Congresso ISKO Espanha-Portugal*. Neste título e neste tema, bem como nos eixos de investigação propostos aos autores, quisemos deixar claro que a Organização do Conhecimento deve fazer um esforço de investigação que se alinhe não apenas com a problemática da pandemia COVID-19, mas sobretudo com a questão global do Desenvolvimento Sustentável, dando também o contributo, que lhe cabe dar.

Perante estes desafios, pretendia-se que os participantes apresentassem investigações em curso no domínio da Organização do Conhecimento e refletissem em conjunto sobre diversas temáticas, que incluem o desenvolvimento sustentável, a saúde, os sistemas de organização e conhecimento, aspetos teóricos e epistemológicos, bem como a formação e o desenvolvimento profissional. Apoiar a implementação da Agenda 2030 através do conhecimento científico e facilitar a sua transferência para a sociedade é, na sua essência, o contributo do presente encontro científico. Na senda dos congressos anteriores, este pretende mostrar a estreita relação entre a investigação e os problemas do mundo atual, revelando a natureza transversal da Organização do Conhecimento, enquanto área de saber basilar para todo o campo da informação e do conhecimento.

Quando a Faculdade de Letras da Universidade de Lisboa recebeu o testemunho para a organização deste evento, estávamos ainda longe de adivinhar o que seria o

mundo dali a alguns meses. Mesmo assim, com todas as incertezas no horizonte, lançámos o sítio oficial (<https://isko2021.letras.ulisboa.pt/>), em novembro de 2020, bem como a chamada de comunicações (*Call for Papers*). Recebemos cerca de 100 propostas de comunicações e posters, o que nos deixou confiantes para a realização do evento, mesmo que ainda não tivéssemos decidido se seria em modo presencial ou virtual. O trabalho da Comissão Científica na avaliação por pares foi de excelência e rigor, aportando inúmeras sugestões de melhoria para os trabalhos apresentados. Cerca de 60% das propostas foram admitidas, algumas das quais indicadas para poster, tendo permitido reunir neste livro 65 textos de elevada qualidade. A este *corpus* somam-se duas das quatro conferências (*keynotes*) apresentadas no evento, aqui registadas sob a forma de texto, de Marcia Zeng e Natália Bolfarini Tognoli, respetivamente intituladas *Health Knowledge Organization Systems in a Data-driven Age (Sistemas de Organização do Conhecimento da Saúde na Era dos Dados)* e *Os desafios para a formação e o desenvolvimento profissional na Organização do Conhecimento à luz da Agenda 2030*, que muito enriquecem este trabalho.

A decisão de realizar um evento presencial foi o resultado de uma vontade de voltar a reunir a ‘nossa’ comunidade de investigadores, possibilitando, por esta forma, o regresso da riqueza da discussão e do debate, que sabemos não serem iguais quando nos encontramos em formato virtual. Tão somente porque o virtual quando não divide, como defendera Manuel Castell em a *Sociedade em Rede* (vol. 1), separa e afasta.

Este volume, que inaugura uma nova coleção nascida no seio do Grupo de Ciência da Informação do Centro de Estudos Clássicos da Faculdade de Letras da Universidade de Lisboa, a Coleção CA – *Ciência Aberta*, é uma coedição do Centro de Estudos Clássicos e das Edições Colibri. *Ciência Aberta* integrará apenas obras em acesso aberto, em formato eletrónico e cumprindo os princípios fundamentais da identificação de objetos digitais, como a inclusão do número DOI (*Digital Object Identifier*). A publicação *Organização do Conhecimento no Horizonte 2030* fica depositada no Repositório da Universidade de Lisboa, um repositório institucional que integra a rede alargada de repositórios digitais que formam uma parte importante dos novos ecossistemas abertos de informação científica.

Para que fosse possível a realização do *V Congresso ISKO Espanha-Portugal*, bem como a publicação *Organização do Conhecimento no Horizonte 2030*, que reúne os resultados da investigação apresentados em Lisboa, é devido um penhorado agradecimento a Rodrigo Furtado, diretor do Centro de Estudos Clássicos e, simultaneamente, diretor do Programa em Ciências da Documentação e Informação da Faculdade de Letras da Universidade de Lisboa. Um agradecimento que se estende a Ana Matafome, no secretariado do referido Centro de Estudos Clássicos, e a Pedro Estácio, Chefe da Divisão da Biblioteca.

Por último, porque a ISKO – The International Society for Knowledge Organization é efetivamente um espaço de encontros (científicos e humanos) e de partilha (de conhecimento, informação e afetos), que o Capítulo Ibérico sempre soube valorizar, o espírito e o coração dos membros da ISKO – Capítulo Ibérico une os presentes e os ausentes. Assim se compreendem os dois espaços de homenagem das duas figuras ibéricas ímpares da Organização do Conhecimento: Emilia Currás e Maria da Graça Simões. O sentir dos membros do Capítulo foi trazido pela voz e o gesto generoso

dos colegas e Amigos: José Antonio Frías e Rosa San Segundo, para aquela; Maria Manuel Borges, para esta. Querendo interpretar este mesmo sentir, *Organização do Conhecimento no Horizonte 2030* é dedicado à sua memória, tão somente porque, recorda-nos Sophia de Mello Breyner Andresen, “numa noite sem lua”, Emilia Curras / Maria da Graça Simões, “nenhuma ausência é mais funda do que a tua” (1958).

Lisboa, 26 de novembro de 2021.

Carlos Guardado da Silva, Jorge Revez, Luís Corujo

PRESENTACIÓN

Avanza el otoño de 2021 y, sobre nosotros, "el tiempo se escapa irremediablemente", como señaló el poeta Vergilio en las *Geórgicas* (Libro III). Es el final del Año II de la Pandemia COVID-19, marcado por una crisis sanitaria, económica y social a escala mundial que nos revela, en el siglo XXI, lo frágil que es nuestra vida colectiva y lo insuficientes que son los esfuerzos y proyectos humanos para una vida equilibrada y sostenible. Por primera vez en la historia, vivimos una crisis de forma universal, seguida en directo por televisión y por Internet, con una profusión de información -una Infodemia- nunca antes vista, cuyo impacto será, sin lugar a duda, duradero y persistente. Se trata de una crisis que, de una u otra manera, nos afectó a todos, nos sigue afectando, y en la que la información y el conocimiento han jugado un papel protagonista. Al igual que en otras crisis mundiales, el desarrollo técnico y científico fue de la mano de la tragedia de millones de vidas humanas perdidas. Este escenario pandémico se suma a una crisis anterior provocada por los cambios en el clima terrestre, el calentamiento global, la desestabilización de los ecosistemas y la disminución de la biodiversidad. Sin embargo, la salud y el clima son dos caras de un mismo problema, lo que plantea mayores desafíos en el contexto del Desarrollo Sostenible.

El Desarrollo Sostenible, que se enuncia desde una perspectiva eminentemente medioambiental, surgida en los años 60 del siglo XX, es un concepto acuñado en 1987, a través de la Comisión Mundial de Medio Ambiente y Desarrollo. En ese año, en la *Declaración de Bruntland*, la Organización de las Naciones Unidas (ONU) definió el Desarrollo Sostenible como un desarrollo que satisface las necesidades del presente sin comprometer la capacidad de las generaciones futuras para satisfacer las suyas, teniendo como objetivo la calidad de vida y respetando los límites de la capacidad de los ecosistemas. Con la evolución social y política, el concepto se amplía y consolida, alcanzando una definición cercana a la actual en 1994, cuando el Consejo Internacional de Iniciativas Ambientales Locales define Desarrollo Sostenible como el desarrollo que proporciona servicios económicos, sociales y ambientales básicos a todos los miembros de una comunidad, sin comprometer la viabilidad de los sistemas sociales, naturales y manufacturados de que aquellos dependen. El concepto se afirmaría, sobre todo, desde principios del siglo XXI con la II Cumbre de la Tierra, celebrada entre el 24 de agosto y el 4 de septiembre de 2001, en Johannesburgo (Sudáfrica), bajo el título de Cumbre Mundial para el Desarrollo Sostenible, integrando los pilares económico, social y medioambiental, así como con las Prepcom, conferencias preparatorias de la Cumbre de Johannesburgo. En 2015, un amplio conjunto de objetivos se consolidó en los 17 Objetivos de Desarrollo Sostenible que forman la Agenda 2030 de la ONU. Esto abarca problemas universales como el hambre y la pobreza, dos *topoi* inseparables

que naturalmente adquieren mayor relevancia, y pretende desarrollar aspectos esenciales como la salud, el acceso al agua potable y al saneamiento, la educación, la igualdad de género, las energías renovables, el trabajo decente, la innovación, la reducción de las desigualdades, las ciudades sostenibles, que incluyen la producción y el consumo sostenibles. También busca combatir el cambio climático, proteger los océanos y los ecosistemas terrestres, y lograr la paz y la justicia, en un esfuerzo global de alianzas para el desarrollo y la mejora de la calidad de vida de las poblaciones.

La Agenda 2030 de las Naciones Unidas es una agenda global para hacer frente a los problemas mundiales, que sitúa el desarrollo sostenible a la cabeza de las prioridades internacionales y constituye una visión común para la Humanidad. El conocimiento es una dimensión central en los 17 Objetivos de Desarrollo Sostenible. La información y las tecnologías digitales se entienden como esenciales para la comunicación global y determinantes para acelerar el progreso humano hacia el pleno desarrollo de la Sociedad del Conocimiento. En consonancia con esta Agenda, Horizonte Europa, que pretende desarrollar la investigación y la innovación en Europa entre 2021 y 2027, tiene como componente central de su visión ayudar a alcanzar los Objetivos de Desarrollo Sostenible.

La obra que ahora se presenta recoge un importante conjunto de aportes científicos traídos a Lisboa en el marco del *V Congreso España-Portugal de la ISKO*, el evento que bianualmente reúne al Capítulo Ibérico de la ISKO, una comunidad de investigadores de los países ibéricos, que acoge y dialoga con sus colegas de América Latina. No sólo nos unen las lenguas comunes, sino viejos lazos de amistad, formando así una comunidad que ya es iberoamericana. Esta obra recibe el mismo título del evento que le dio origen: *Organización del Conocimiento en Horizonte 2030: Desarrollo Sostenible y Salud: Actas del V Congreso ISKO España-Portugal*. Con este título y tema, así como con los ejes de investigación propuestos a los autores, hemos querido dejar claro que la Organización del Conocimiento debe hacer un esfuerzo de investigación que esté alineado no sólo con la problemática de la pandemia del COVID-19, sino, sobre todo, con la cuestión global del Desarrollo Sostenible, dando también la contribución que le corresponde.

Ante estos retos, se pretendía que los participantes presentaran las investigaciones en curso en el ámbito de la Organización del Conocimiento y reflexionaran juntos sobre diversos temas, entre los que se encuentran el desarrollo sostenible, la salud, la organización y los sistemas de conocimiento, los aspectos teóricos y epistemológicos, así como la formación y el desarrollo profesional. Apoyar la implementación de la Agenda 2030 a través del conocimiento científico y facilitar su transferencia a la sociedad es, en esencia, la contribución de esta reunión científica. Siguiendo la senda de los congresos anteriores, éste pretende mostrar la estrecha relación entre la investigación y los problemas del mundo actual, revelando el carácter transversal de la Organización del Conocimiento, como área de conocimiento fundamental para todo el ámbito de la información y el conocimiento.

Cuando la Facultad de Letras de la Universidad de Lisboa recibió el testigo de la organización de este evento, aún estábamos lejos de adivinar cómo sería el mundo dentro de unos meses. Aun así, con todas las incertidumbres en el horizonte, lanzamos la página *web* oficial (<https://isko2021.letras.ulisboa.pt/>) en noviembre de 2020,

así como la convocatoria de ponencias. Recibimos unas 100 propuestas de ponencias y pósters, lo que nos hizo confiar en la celebración del evento, aunque aún no habíamos decidido si la modalidad sería presencial o virtual. El trabajo del Comité Científico en la revisión por pares fue excelente y riguroso, aportando numerosas sugerencias de mejora para los trabajos presentados. Alrededor del 60% de las propuestas fueron aceptadas, algunas de las cuales fueron indicadas para póster, lo que nos ha permitido reunir en este libro 65 textos de gran calidad. Además de este *corpus*, dos de las cuatro conferencias magistrales presentadas en el evento, registradas aquí en forma de texto, por Marcia Zeng y Natália Bolfarini Tognoli, tituladas respectivamente *Sistemas de organización del conocimiento en salud en una Era impulsada por los datos* y *Los desafíos para la formación y el desarrollo profesional en la organización del conocimiento a la luz de la Agenda 2030*, enriquecen enormemente este trabajo.

La decisión de celebrar un evento presencial fue el resultado de un deseo de reunir de nuevo a "nuestra" comunidad de investigadores, permitiendo, de esta manera, el retorno de la riqueza de la discusión y el debate, que sabemos que no son los mismos cuando nos reunimos en un formato virtual. Tanto es así que lo virtual, cuando no divide, como sostiene Manuel Castell en *La sociedad red* (vol. 1), separa y distancia.

Este volumen, que inaugura una nueva colección nacida en el seno del Grupo de Ciencias de la Información del Centro de Estudios Clásicos de la Facultad de Letras de la Universidad de Lisboa, la Colección CA - *Ciência Aberta*, es una coedición del Centro de Estudios Clásicos y de Edições Colibri. *Ciência Aberta* sólo incluirá obras de acceso abierto, en formato electrónico y que cumplan con los principios fundamentales de identificación de objetos digitales, como la inclusión del número DOI (*Digital Object Identifier*). La publicación *Organização do Conhecimento no Horizonte 2030 = Organización del Conocimiento en Horizonte 2030* está depositada en el Repositorio de la Universidad de Lisboa, un repositorio institucional que integra la extensa red de repositorios digitales que forman parte importante de los nuevos ecosistemas de información científica abierta.

Para hacer posible la celebración del *V Congreso ISKO España-Portugal*, así como la publicación *Organização do Conhecimento no Horizonte 2030 = Organización del Conocimiento en el Horizonte 2030*, que recoge los resultados de las investigaciones presentadas en Lisboa, hay que agradecer a Rodrigo Furtado, director del Centro de Estudios Clásicos y, simultáneamente, director del Programa de Ciências de Documentação e Informação de la Facultad de Letras de la Universidad de Lisboa. Gracias también a Ana Matafome, en la secretaría del Centro de Estudios Clásicos, y a Pedro Estácio, Jefe de la División de Bibliotecas.

Por último, optamos por la presencialidad porque ISKO - The International Society for Knowledge Organization - es efectivamente un espacio de encuentro (científico y humano) y de intercambio (de conocimientos, información y afectos), que el Capítulo Ibérico siempre ha sabido valorar. Además, el espíritu y el corazón de los miembros del ISKO - Capítulo Ibérico une a los presentes y a los ausentes. En ese sentido se pueden entender los homenajes a dos figuras ibéricas únicas en la Organización del Conocimiento: Emilia Currás y Maria da Graça Simões. Los sentimientos de los miembros del Capítulo hacia estas dos personalidades fueron

representados por la voz y el gesto generoso de colegas y Amigos: José Antonio Frías y Rosa San Segundo, en homenaje a la primera; y María Manuel Borges, en homenaje a la segunda. Queriendo interpretar este mismo sentimiento, la obra *Organização do Conhecimento no Horizonte 2030* está dedicada a su memoria simplemente porque, como recuerda Sophia de Mello Breyner Andresen, “en una noche sin luna”, Emilia Curras / Maria da Graça Simões, “ninguna ausencia es más profunda que la tuya” (1958).

Lisboa, el 26 de noviembre de 2021.

Carlos Guardado da Silva, Jorge Revez, Luís Corujo

HEALTH KOSs IN A DATA-DRIVEN AGE

Marcia L. Zeng

Kent State University, USA.

Invited Keynote, ISKO 2021 LISBOA

1. Background [Health KOSs]	24
1a. The prerequisite for data normalization in health data exchange	24
1b. Semantical enrichments of sharable data using standard health KOSs	25
1c. The necessities of various KOS standards for different processes	26
1d. Purpose and functions of health KOSs – using ICD as a model	27
2. Challenges & Actions in the Data-driven Age	28
2a. Dealing with the problem of information overload	28
2b. Eliminating semantic conflicts	28
2c. Implications of the major health KOSs (and others) during the COVID-19 pandemic.....	29
3. Transforming Big Data into Smart Data Using Semantics	32
3a. About Smart Data	32
3b. Semantic technologies involved in Smart Data approaches	33
3c. Ensuring the semantic interoperability of health KOSs.....	34
Conclusion	37
Sources	37
References	37

Standardized health KOSs [including taxonomies, classifications, thesauri, code lists, and ontologies, known as “knowledge organization systems (KOS)”] have played critical roles in communication and information exchange in the data-driven environment as they have been used by all healthcare systems all over the world. During pandemics like COVID-19, it is more challenging to deal with information overload and semantic conflict as it directly impacts the whole domain of healthcare. After introducing the challenges, this essay will highlight the efforts of major health KOSs, especially the *International Statistical Classification of Diseases and Related Health Problems* (usually referred as *International Classification of Diseases [ICD]*), revealing their essential roles in health information exchange and communication during the COVID-19 pandemic. It is clear that the standardized

health KOSs, which are the foundations of successful semantic interoperability, allow the world to compare and share data in a consistent and standard way between institutions, across regions and countries, and over periods of time. This essay is based on my related papers recently published and in-progress (Zeng, 2017, 2018; Zeng et al., 2020; Hong & Zeng [2022]).

1. Background [Health KOSs]

1a. The prerequisite for data normalization in health data exchange

In healthcare workflows, data and information have to be exchanged among many parties, including medical professionals, medical technicians and paramedics, clinical laboratory technologists and technicians, government agencies, international organizations, academic researchers, medical center administrators, patients and families, and insurance providers.

For example, when a pandemic like COVID-19 starts, public health laboratories need to respond quickly to test the new virus. It is clear that, in healthcare information systems, data normalization is fundamental for any subsequent data analysis, information management, and decision-making. The Association of Public

Figure 1. COVID-19 Data Exchange on the AIMS Platform

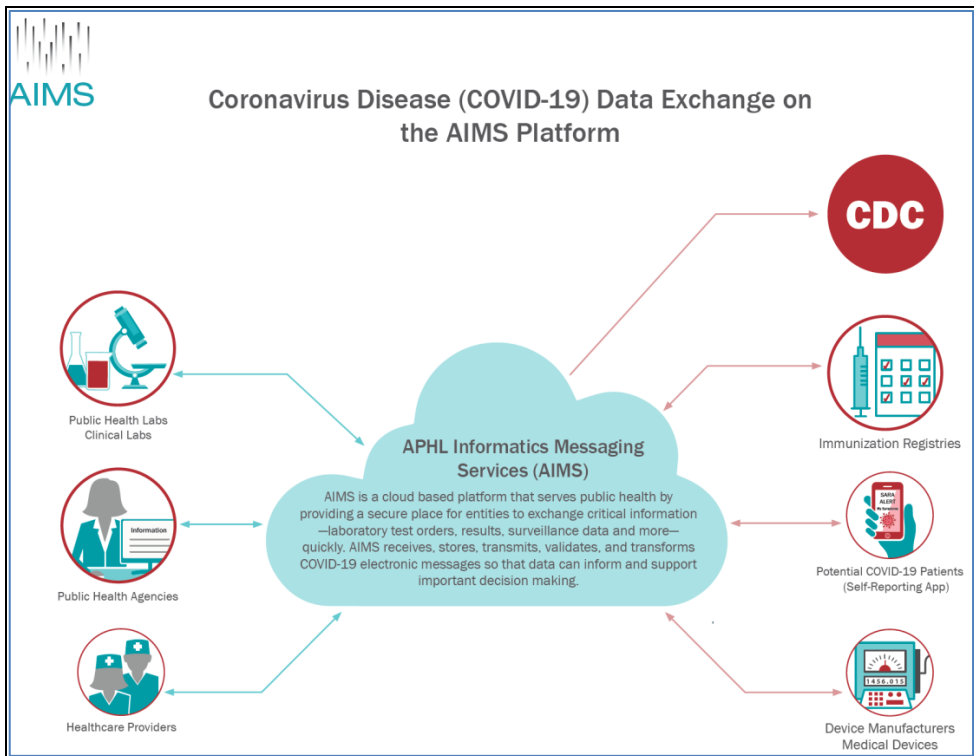


Image Source: Association of Public Health Laboratories. 2020a.

Health Laboratories (APHL) in the USA provides an electronic messaging services platform called AIMS for labs to send and receive Health Level Seven (HL7) messages, and exchange data between providers and Centers for Disease Control and Prevention (CDC) as well as other public health organizations. APHL provides a coding guideline for COVID-19 to help labs generate HL7 Messages with standard terminologies. The diagram (Figure 1) shows a dataflow of the public health network in the USA.

1b. Semantical enrichments of sharable data using standard health KOSs

An adequate public health response is predicated on the testing data, which are enabled to communicate this information by using the standard health KOSs, as demonstrated in the following image of the HL7 message (Figure 2). Health Level Seven (HL7) refers to the seventh level (the application level) of the ISO seven-layer communications model for Open Systems Interconnection (OSI). HL7 is the workhorse of electronic data exchange in the clinical domain and arguably the most widely implemented standard for healthcare in the world. This messaging standard allows the exchange of clinical data between systems. It is designed to support a central patient care system as well as a more distributed environment where data resides in departmental systems (HL7 International, n.d.).

Figure 2: A sample HL7 message carrying COVID-19 related data coded with LOINC and SNOMED CT.

Example: COVID-19 HL7 data messaging - Sample HL7 messages for lab data exchange
 Provided by the ASSOCIATION OF PUBLIC HEALTH LABORATORIES, March 2020.
<https://www.aphl.org/programs/preparedness/Crisis-Management/Documents/Forms/AllItems.aspx>
2019nCoV_PHLIPsample_2.3.1_Detected

```

MSH|^~&|STARLIMS.AR.STAG^2.16.840.1.114222.4.3.3.2.5.2^ISO|AR.LittleRock.SPHL^2.16.840.1.114222.4.1.20
083^ISO|US.WHO.Collab.LabSys^2.16.840.1.114222.4.3.3.7^ISO|CDC-EPI.Surv
Branch^2.16.840.1.114222.4.1.10416^ISO|20191203100718-
0600|ORU^R01^ORU_R01||170703|T|2.3.1|||||PHLIP_ORU_v1.0.2^PHIN_Profile_ID^2.16.840.1.114222.4.10.3^ISO
PID||PID13295037^^^STARLIMS.AR.STAG&2.16.840.1.114222.4.3.3.2.5.2&ISO^PI||~^S||20000101|F||^A
R^72016^USA
ORC|RE|1905700000256-12^PHLIP-Test-EHR^2.16.840.1.113883.3.72.5.24^ISO|1905700000256-
176^STARLIMS.AR.STAG^2.16.840.1.114222.4.3.3.2.5.2^ISO|CM|||||Little Rock General Hospital Lab|2217
Trancas^Little Rock^AR^72205
OBR|1|1905700000256-12^PHLIP-Test-EHR^2.16.840.1.113883.3.72.5.24^ISO|1905700000256-
176^STARLIMS.AR.STAG^2.16.840.1.114222.4.3.3.2.5.2^ISO|68991-9^Epidemiologically important info
pn^LN||20191125201900-
0600||||20191126|ORH&Other&HL70070|1412941681^Smith^John^C^DR^^NPI&2.16.840.1.113883.4.6&ISO^L|^
^PH^^952^4863332||||F
OBX|1|CX|LAB202^Unique Specimen
ID^PHINQUESTION||1905700000256^^STARLIMS.AR.STAG&2.16.840.1.114222.4.3.3.2.5.2&ISO||||F||20191203
100718-0600
ORC|RE|1905700000256-13^PHLIP-Test-EHR^2.16.840.1.113883.3.72.5.24^ISO|1905700000256-
177^STARLIMS.AR.STAG^2.16.840.1.114222.4.3.3.2.5.2^ISO|CM|||||Little Rock General Hospital Lab|2217
Trancas^Little Rock^AR^72205
OBR|2|1905700000256-13^PHLIP-Test-EHR^2.16.840.1.113883.3.72.5.24^ISO|1905700000256-
177^STARLIMS.AR.STAG^2.16.840.1.114222.4.3.3.2.5.2^ISO|94306-8^SARS-CoV-2 RNA Pnl XXX
NAA+probe^LN||20191125201900-
0600||||20191126|ORH&Other&HL70070&Nasopharyngeal|1412941681^Smith^John^C^DR^^NPI&2.16.840.1.1
13883.4.6&ISO^L|^PH^^952^4863332||||20191203081920-0600||F
OBX|1|CE|94307-6^SARS-CoV-2 N gene XXX Q1 NAA N1^LN|260373001^Detected^SCT||||F||20191203081920-
0600
OBX|2|CE|94308-4^SARS-CoV-2 N gene XXX Q1 NAA N2^LN|260373001^Detected^SCT||||F||20191203081920-
0600
OBX|3|CE|68993-5^Human RNase P RNA XXX Q1
NAA+probe^LN||260373001^Detected^SCT||||F||20191203081920-0600
    
```

Annotations in the image:

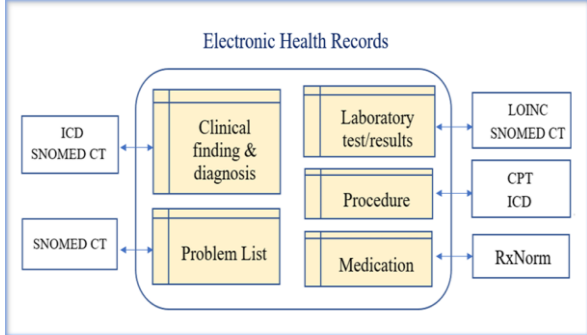
- Blue callout: LOINC code and name (points to 260373001^Detected^SCT)
- Red callout: SNOMED CT code and name (points to 94306-8^SARS-CoV-2 RNA Pnl XXX)
- Blue callout: LOINC code and name (points to 260373001^Detected^SCT)

Source: Annotated based on the sample file provided by the Association of Public Health Laboratories, 2020b.

HL7 requires using standardized terminologies to represent health data, as illustrated in Figure 2. Besides developing its own standardized code sets to identify administrative data such as gender code, data type, and status codes, HL7 has employed existing standardized health KOSs to support the disambiguation of health information exchanges. This process begins with *Logical Observation Identifiers Names and Codes* (LOINC) codes used in the laboratory for the questions being asked about viral specimens. That information must then be passed over to *Systematized Nomenclature of Medicine–Clinical Terms* (SNOMED-CT) which is used to code the results of those tests, e.g., positive or negative. The above HL7 case record example, based on the norms in the USA, demonstrates the usage of LOINC and SNOMED CT to semantically enrich the shareable and reusable data. Through HL7 messaging, information is then passed to other systems, with clear semantics for the codes.

1c. The necessities of various KOS standards for different processes

With the realities of everyday healthcare, various KOS standards are needed for different processes. This illustration (Figure 3) shows the usage of them in Electronic Health Records (EHR). Table 1 provides a list of the commonly used KOS standards in health information exchanges.

Figure 3. Usage of different standardized health KOSs in Electronic Health Records (EHR).	Table 1. Commonly used health KOSs
 <p style="text-align: center;">Source: Zeng et al., 2020.</p>	<p><i>International Classification of Diseases (ICD) WHO </i></p> <p><i>Systematized Nomenclature of Medicine–Clinical Terms (SNOMED-CT) International Health Terminology Standards Development Organization </i></p> <p><i>Logical Observation Identifiers Names and Codes (LOINC) LOINC.org </i></p> <p><i>Current Procedural Terminology (CPT) American Medical Association </i></p> <p><i>RxNorm, a normalized naming system for generic and branded drugs US National Library of Medicine </i></p>

As shown in the figure, ICD and SNOMED CT are mainly used to normalize the diagnostic process and health problem list. There are others that have been developed to encode medical lab tests and clinical observations (LOINC), medical

procedures (CPT), and names for clinical drugs (RxNorm). They have a crucial role in healthcare information systems to support data normalization and semantic linking and sharing.

The KOSs listed in the above figures are only based on practices in the USA. In fact, nearly 200 biomedical KOS vocabularies from around the world are included in the Unified Medical Language System (UMLS)'s Metathesaurus, a large, multi-purpose, and multi-lingual biomedical vocabulary database. In another terminology repository, BioPortal, the world's most comprehensive repository of biomedical ontologies, the number of ontologies is near 900 as of Oct. 2021. These numbers indicate the need for standardized KOS vocabularies in diverse circumjacent communities. Both Metathesaurus and BioPortal enable cross-vocabulary mapping that facilitates knowledge and semantically interoperable data throughout knowledge sources developed independently to meet diverse needs.

1d. Purpose and functions of health KOSs – using ICD as a model

ICD, with its full title being *International Statistical Classification of Diseases and Related Health Problems*, is an international standard for classifying diseases and reporting health conditions, as well as a foundation for health trends and statistical analysis. It is steeped in history, international cooperation, and improvement over the last 150 years (Salcido, 2015). ICD has been published and maintained by the World Health Organization (WHO) since ICD-6. The current version is ICD-10, with ICD-11 to be used starting in 2022.

The purpose of the ICD is to allow for the systematic recording, analysis, interpretation, and comparison of mortality and morbidity data collected in different countries or areas and at different times. “The ICD has been designed to address the needs of a broad range of use cases: Mortality, morbidity, epidemiology, casemix, quality and safety, primary care” (World Health Organization, 2019, 1.1.1 “Intended use”). More importantly, as a KOS, it ensures **semantic interoperability** and **reusability** of recorded data for the different use cases beyond mere health statistics, including decision support, resource allocation, reimbursement, guidelines and more. Version 10 (ICD-10), released in 1994, has been widely used in healthcare systems worldwide. It defines diseases, disorders, injuries, and other related health conditions in a **hierarchical structure**, and has been gradually developed as a diagnostic classification standard for all epidemiology, health management, clinical, and research purposes. Furthermore, its development has augmented efforts to enhance its coverage and implement a unified **notation system** for enabling its usage across languages, cultures, and healthcare systems globally.

ICD and other KOS standards facilitate the collection and storage of data for analysis and evidence-based decision-making. They are indispensable for helping clinicians and public health officials in a number of ways, including understanding the cost of care delivery, improving outcomes, carrying out disease surveillance, identifying and assigning resources, conducting research, and performing predictive analytics, which can forecast future trends and needs.

2. Challenges & Actions in the Data-driven Age

2a. Dealing with the problem of information overload

“Information overload” is a term which refers to the difficulty a person can have in understanding an issue and making decisions, which can be caused by the presence of too much information (Toffler, 1970).

During times of crisis such as a global pandemic, users need immediate access to information on the status, impact, prevention, and control of the disease. They want to find accurate information while gaining a complete picture of its significance around the world. This is compounded by the fact that news reports from around the world are constantly being updated, providing users with yet more new information. This information may not be clearly delivered if the language or keywords referenced carry different meanings in different contexts. Moreover, one may also wish to trace the methods or criteria for collecting data used in data-driven messages and decision-making.

The challenge could grow dramatically when a report needs to be communicated across languages, regions, and cultures, especially if statistics were generated based on different contextual information and gathered from diverse resources.

2b. Eliminating semantic conflicts

Semantic conflict can occur within any data and information communication process. In the healthcare circumstance, semantic conflicts undoubtedly add more confusion to an environment already overwhelmingly characterized by heavy information overload. In terms of public health, the issue is much the same, particularly in a situation where public health is being addressed on a global scale.

A relevant example is the process of naming a new disease and the key variants of the virus that causes COVID-19. Several confounding factors must be considered, including the reuse of previous names that may share some similarities, or the adoption of a known accepted name which may carry different meanings at different times of history, and/or the inclusion of names of particular groups of people, places, or animals based on the cases reported earliest, and the competing interests assigning these terms. These conflicting and sometimes problematic names can be observed in use in major pandemics of the past, as well as in the most recent ones, such as the 2009 outbreak of H1N1 influenza in humans (“swine flu,” “pig flu,” “Mexican flu”) and the COVID-19 outbreak (“Wuhan Flu,” “China virus,” “Kong Flu,” “India variant,” “UK variant,” etc.) since early 2020. Such naming conventions persist even after the WHO assigned official terminology to the disease. This situation demonstrates that disease-naming-based semantic conflicts are not only an issue of the past, but also an issue of the present and the future, thus necessitating further investigation. Similar semantic conflicts may also exist within the KOSs, which are used in a variety of healthcare services, complicating the systems they support as well as the decision making that impacts patient health outcomes.

For the outbreak of a new viral disease, three most important names have to be decided: (1) the disease, (2) the virus, and (3) the species (International Committee on Taxonomy of Viruses, 2020). Establishing a name for a new disease provides a shared understanding for researchers and developers to discuss disease prevention, spread, transmissibility, severity, and treatment. Viruses are named based on their genetic structure to facilitate the development of diagnostic tests, vaccines, and medicines (World Health Organization, 2020). The WHO guidelines ensure that the name does not refer to a geographical location, an animal, an individual, or group of people, while still being pronounceable and related to the disease (World Health Organization 2015, 2020, 2021).

2c. Implications of the major health KOSs (and others) during the COVID-19 pandemic

Collaborating with health informatics researcher Dr. Hong and three others, we conducted a review of the prompt actions of the major health KOSs (refer to the full article Zeng et al., 2020). First, we created a timeline to expose the recent efforts to eliminate ambiguities and semantic conflicts through naming of the disease, led by the WHO and the actions of the ICD-10.

- Jan. 30. WHO declared the 2019 Novel Coronavirus (2019-nCoV) disease outbreak a public health emergency of international concern.
- Jan. 31. WHO Family of International Classifications (WHO-FIC) network's Classification and Statistics Advisory Committee (CSAC) convened an emergency meeting to discuss the creation of a specific code for this new type of coronavirus.
 - ICD-10 established a new emergency code “U07.1, 2019-nCoV, acute respiratory disease.”
- Feb. 11. The WHO officially announced the name of the disease, “COVID-19,” an acronym for “coronavirus disease 2019.”
 - A study group of the International Committee on Taxonomy of Viruses (ICTV) christened the novel virus as “severe acute respiratory syndrome coronavirus 2,” or “SARS-CoV-2” (International Committee on Taxonomy of Viruses, 2020).
 - The ICD-10 was updated with two emergency codes: “U07.1 COVID-19, virus identified” and “U07.2 COVID-19, virus not identified” (Figure 4).

Secondly, we collected the new codes and coding guidance released in March 2020 by major standardized health KOSs, which are used in healthcare workflows, in response to the WHO announcement of an official name for the disease. The health KOSs studied include ICD-10, CPT, SMOMED-CT, and LOINC (Refer to Table 1. “COVID-19 Coding Guidance” in Zeng et al., 2020).

Going beyond the healthcare workflow, KOSs are also essential in fighting misinformation and enabling meaningful analyses, reports, news, and decision making for everyone. As an example, the *Medical Subject Headings* (MeSH) of the

Figure 4. ICD-10 emergency codes for COVID-19

<ul style="list-style-type: none"> ▼ XXII Codes for special purposes <ul style="list-style-type: none"> ▼ U00-U49 Provisional assignment of new diseases of uncertain etiology or emergency use <ul style="list-style-type: none"> ▶ U04 Severe acute respiratory syndrome [SARS] <ul style="list-style-type: none"> ▼ U07 Emergency use of U07 <ul style="list-style-type: none"> U07.0 Vaping-related disorder U07.1 COVID-19, virus identified U07.2 COVID-19, virus not identified U07.3 Emergency use of U07.3 U07.4 Emergency use of U07.4 U07.5 Emergency use of U07.5 U07.6 Emergency use of U07.6 U07.7 Emergency use of U07.7 U07.8 Emergency use of U07.8 U07.9 Emergency use of U07.9 ▶ U08 Emergency use of U08 	<p>U07.1 COVID-19, virus identified COVID-19 NOS</p> <p>Use this code when COVID-19 has been confirmed by laboratory testing irrespective of severity of clinical signs or symptoms. Use additional code, if desired, to identify pneumonia or other manifestations.</p> <p>Excl.: Coronavirus infection, unspecified site (B34.2) Coronavirus as the cause of diseases classified to other chapters (B97.2) Severe acute respiratory syndrome [SARS], unspecified (U04.9)</p> <p>U07.2 COVID-19, virus not identified</p> <p>Use this code when COVID-19 is diagnosed clinically or epidemiologically but laboratory testing is inconclusive or not available. Use additional code, if desired, to identify pneumonia or other manifestations</p> <p>Excl.: Coronavirus infection, unspecified site (B34.2)</p> <p>COVID-19:</p> <ul style="list-style-type: none"> • confirmed by laboratory testing (U07.1) • special screening examination (Z11.5) • suspected but ruled out by negative laboratory results (Z03.8)
---	--

Source: ICD-10 browser. <https://icd.who.int/browse10/2019/en#/U07>. Image captured 2020-04-26.

U.S. National Library of Medicine added a Supplementary Concept Record (SCR) for COVID-19 to its MeSH Browser on February 13, right after the WHO and ICD-10's actions. In this new Supplementary Concept Record, in addition to the preferred lexical label "COVID-19" and a Unique ID, more than 10 entry terms are provided to facilitate subject access to the related information resources. It has been mapped to other related MeSH descriptors (Figure 5). Searching the MeSH ID on the Web brought 128 results of multilingual Wikipedia entries aligned with this MeSH identifier used in Wikipedia, based on the data we collected in May of 2020. Two other normally used KOSs are ICD and SNOMED CT in Wikipedia.

Figure 5. MeSH Supplementary Concept Record for COVID-19.

COVID-19 MeSH Supplementary Concept Data 2020	
Details Concepts	
MeSH Supplementary Unique ID	COVID-19 C000657245
RDF Unique Identifier	http://id.nlm.nih.gov/mesh/C000657245
Entry Term(s)	2019 novel coronavirus disease 2019 novel coronavirus infection 2019-nCoV disease 2019-nCoV infection COVID-19 pandemic COVID-19 virus disease COVID-19 virus infection COVID19 SARS-CoV-2 infection coronavirus disease 2019 coronavirus disease-19
Registry Number	0
Heading Mapped to	*Pneumonia, Viral *Coronavirus Infections *Pandemics
Note	A viral disorder characterized by high FEVER; COUGH; DYSPNEA; renal dysfunction and other symptoms of a VIRAL PNEUMONIA. A coronavirus SARS-CoV-2 in the genus BETACORONAVIRUS is the suspected agent. severe acute respiratory syndrome coronavirus 2
Indexing Information	
Date of Entry	2020/02/13
Revision Date	2020/04/07

Source: MeSH browser. <http://id.nlm.nih.gov/mesh/C000657245>. Image capture 2020-04-21. Replaced by <https://id.nlm.nih.gov/mesh/D000086382> in 2021.

Schema.org, an open, community-based, collaboratively developed vocabulary to promote the structuring of data on the Internet, originally founded by Google, Microsoft, Yahoo, and Yandex, quickly responded to the global pandemic by releasing versions 7.0 through 7.04 within one month in early 2020, in responding to the needs of the world. Such a shared vocabulary is crucial for webmasters and website developers to provide semantically enriched structured data regarding the contents of their web resources to eliminate semantic conflicts and handle information overload and misinformation. From the following examples, it is clear that those properties are very unique and most needed by the whole world.

- 2020-03-17, Schema.org Version 7.0
 - Motivated by the need to structure urgent, crisis-related information during the COVID-19 pandemic, a new type *SpecialAnnouncement* and its properties were created.
 - A subtype of *MedicalClinic*, labelled *CovidTestingFacility*, was implemented to describe locations where testing for COVID-19 is available. This also allowed reuse of existing Schema.org structured data, such as contact information, address, and operating hours for existing described instances.
 - A property, *hasDriveThroughService*, was created to indicate the presence of drive-through services, like drive-through virus testing.
 - Due to work at home policies and the migration to online work settings, a *VirtualLocation* type was developed, including a way to indicate that an event has moved from a physical location to an online or virtual location.

Other KOSs are also taking immediate action. The KOSs we investigated involving multiple ontologies took action in the pandemic, such as those based on the *Infectious Disease Ontology* (IDO) and the extended or revised ontologies that integrate new entities and relationships regarding COVID-19, including *Gene Ontology*, *Experimental Factor Ontology*, and the *Vaccine Ontology*. In addition, we paid attention to multilingual contexts, such as the *AGROVOC* of the Food and Agriculture Organization of the United Nations, Wikipedia, and Wikidata, as well as applications, services, and data-based analysis that are empowered by KOS (refer to the full paper of Zeng et al., 2020).

This study of the efforts, usages, and functions of the major health KOSs and other KOSs reveals their essential responsibilities in supporting health data exchange and information management, as well as ensuring consistency and interoperability of data collection and reuse among various providers and care settings.

These KOSs enable the world to compare and share data in a consistent and standard way— between institutions, across regions and countries, and over a period of time. Together they are contributing to the actions of eliminating semantic conflicts and avoiding information overload in real-world healthcare systems.

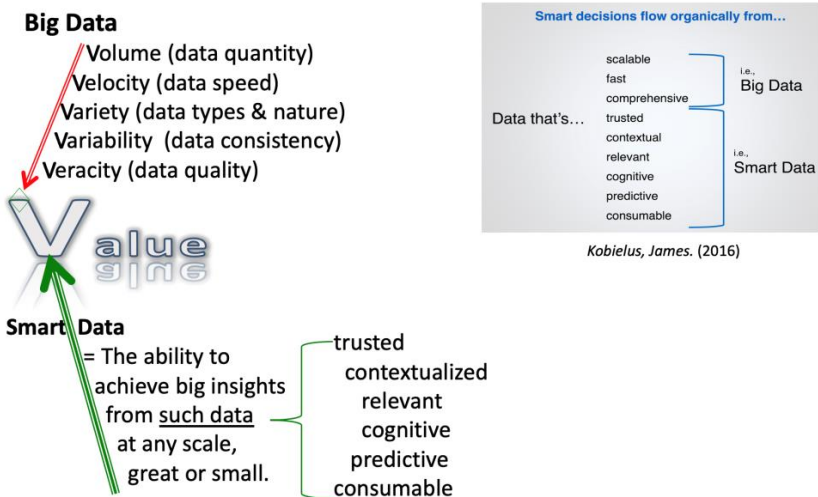
3. Transforming Big Data into Smart Data Using Semantics

3a. About Smart Data

Referring to the real cases, including the prediction of the spread of flu in real time during H1N1 in 2009, Parkinson’s disease, asthma, and others, the keynote by Dr. Amit Sheth at the 30th IEEE International Conference on Data Engineering 2014 was entitled “Transforming Big Data into Smart Data: Deriving Value via harnessing Volume, Variety and Velocity using semantics and Semantic Web.” He addressed that Smart Data is realized by extracting value from a variety of data and explained how Smart Data for the growing variety (e.g., social, sensor/IoT, health care) of Big Data enable a much larger class of applications that can benefit not just large companies but each individual. “In particular, we will need to utilize metadata, employ semantics and intelligent processing, and go beyond traditional reliance on ML and NLP” (Sheth, 2014).

Big data has been characterized by multiple “V”s: Volume, Velocity, Variety, Variability, and Veracity (Figure 6). Big Data can bring another big V: Value, if used appropriately, because it is now possible to find the hidden patterns, the unexpected correlations, and the surprising connections within large datasets through effective processing (Gardner, 2012). At the 2016 Smart Data Online conference, the keynote by James Kobielus, IBM's Big Data evangelist, emphasized that the realization of the last “V”, Value, is dependent on “Smart Data”, the “ability to achieve big insights from trusted, contextualized, relevant, cognitive, predictive, and consumable data at any scale, great or small” (Kobielus, 2016).

Figure 6. Smart Data in the context of Big Data



Source: Zeng, M.L. 2017 DOI: [10.1515/jdis-2017-0001](https://doi.org/10.1515/jdis-2017-0001)
 Compiled based on Kobielus, James. (2016, June). *The Evolution of Big Data to Smart Data*.
 Keynote at Smart Data Online 2016.

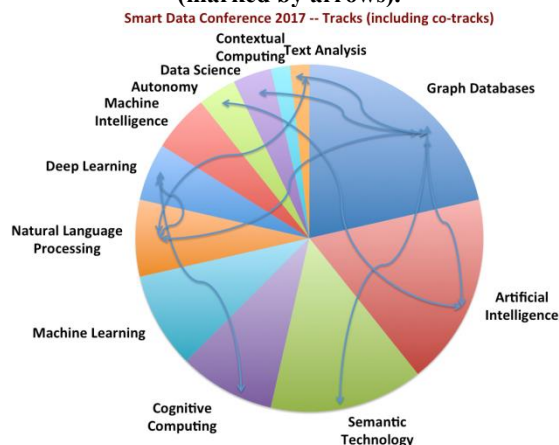
Simply speaking, Smart Data makes sense out of Big Data. It provides Value from harnessing the challenges posed by the volume, velocity, variety, and veracity of big data, in turn providing actionable information and improving decision making (Sheth, 2014). Smart Data is the way in which different data sources (including Big Data) are brought together, correlated, analyzed, etc., to be able to feed decision-making and action processes (Iafrate, 2015).

Extracting Value from Big Data characterized by the other “V”s presents both great challenges and inestimable opportunities. Only after it has been tamed through organization and integration processes is such data turned into Smart Data that reflects the research priorities of a particular discipline or field. These tamed results, as Smart Data inquiries, can then be used to provide comprehensive analyses and generate new products and services (Gardner, 2012; Mukerjee, 2014; Schöch, 2013; TiECON East, 2014.)

3b. Semantic technologies involved in Smart Data approaches

A look at the topics presented at Smart Data conferences since 2015 may provide a good overview of the technologies involved in Smart Data strategies for achieving big insights from trusted, contextualized, relevant, cognitive, predictive, and consumable data at any scale. These include cognitive computing, deep learning, machine learning, artificial intelligence, predictive analytics, graph databases, machine intelligence, voice processing, semantic technologies, autonomous vehicles, Big Data, data science, Internet of Things (IoT), text analysis, RDF, knowledge graphs, contextual computing, Linked Data, deep reasoning, ontologies, JSON-LD, common sense, natural language processing (NLP), and semantic search (DATAVERSITY, 2017). These topics are closely interrelated and overlapping. Meanwhile, the topics of the Smart Data conferences reflect the varied applications of the W3C standards for the Semantic Web, including, but not limited to, RDF, Linked Data, ontologies, graph databases, semantic search, and other semantic technologies. [Figure 7]

Figure 7. Smart Data Conference 2017 tracks, including combined co-tracks (marked by arrows).



Source: Zeng, 2017. Compiled according to the program of the 2017 Smart Data Conference. DATAVERSITY, 2017.

3c. Ensuring the semantic interoperability of health KOSs

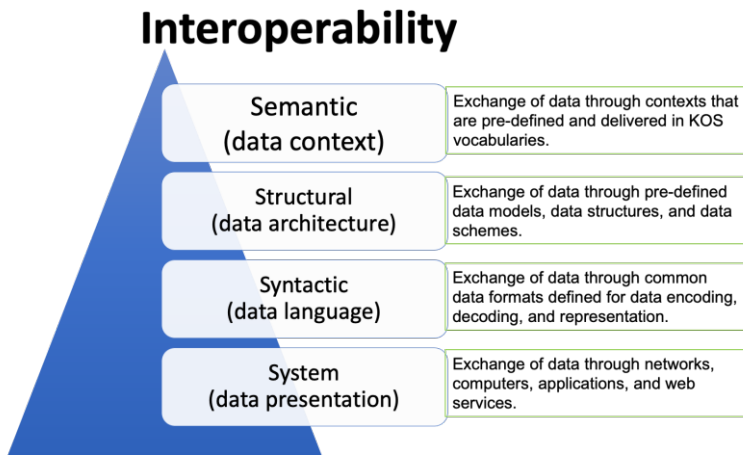
The HL7 message example (Figure 2) and HER (Figure 3) remind us of the key role of KOS in the interoperability layer for any data and information exchange and communication processes.

Witnessing the Web and distributed computing infrastructures gaining in popularity as a means of communication, Sheth (1999) brought attention to the changing focus on interoperability in information systems since the mid 1980s: from system, syntax, and structure to semantics. Ouksel and Sheth (1999) further laid out four types of heterogeneity corresponding to four types of potential interoperability issues:

- System: incompatibilities between hardware and operating systems.
- Syntactic: differences in encodings and representation.
- Structural: variances in data models, data structures, and schemas.
- Semantic: inconsistencies in terminology and meanings.

Semantic interoperability is the ability of different agents, services, and applications to communicate data, information, and knowledge, while ensuring accuracy and preserving the meaning of that same data, information, and knowledge (Zeng and Chan, 2009). Here, communicating could be in any of the forms listed: transfer, exchange, transformation, mediation, migration, integration, aggregation, etc. Figure 8 is an image I generated to interpret the four layers of interoperability in data exchanges. Semantic interoperability is situated in one of them. KOSs are situated in this layer. (Refer to the ISKO Encyclopedia article on “Interoperability.” Zeng, 2018.)

Figure 8. Interoperability layers

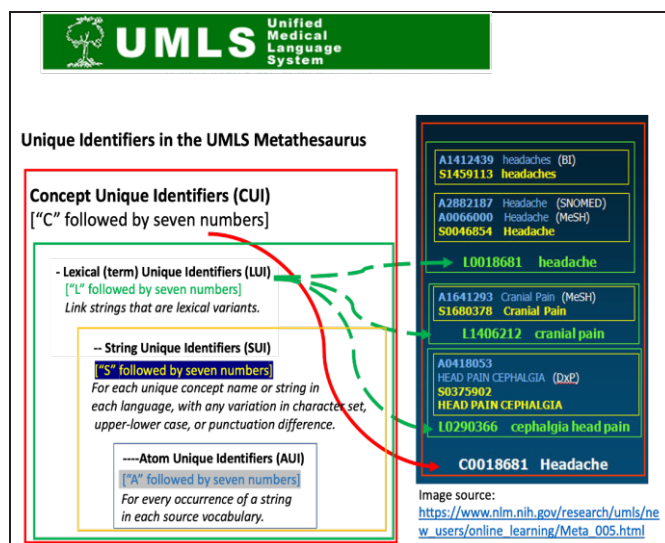


Source: Portion from Zeng, 2018. Figure 1.

One of the approaches to achieve semantic interoperability is demonstrated by the Unified Medical Language System (UMLS), the world's largest compendium of biomedical vocabularies, which "integrates and distributes key terminology, classification and coding standards, and associated resources to promote the creation of more effective and interoperable biomedical information systems and services, including electronic health records" (U.S. National Library of Medicine, 2021a). It is organized by concept, with links to similar names for the same concept from nearly 200 different KOS sources (distinct source families by language) from around the world (mainly in English, about 20 non-English, as of October 2021) (U.S. National Library of Medicine, 2021b). Major categories include diagnosis & clinical observations, procedures & supplies, diseases, comprehensive vocabularies/thesauri, and other categories. In addition, some vocabularies fall into more than one category (U.S. National Library of Medicine, 2016).

UMLS is a pioneer of using *identifiers* for the concepts and concept names it contains, in addition to retaining all identifiers that are present in the source vocabularies. As explained in the left side of the Figure 9, the concept unique identifier (CUI) is the one to unify the mapped entries from different KOSs. Other identifiers for lexical term (LUI), string (SUI), and atom (AUI) provide the contextual information for each entry from the original vocabularies. It also identifies useful relationships between concepts and preserves the meanings, concept names, and relationships from each vocabulary. Many relationships (primarily for synonyms), concept attributes, and some concept names are added by the NLM during Metathesaurus creation and maintenance, but essentially all the concepts themselves come from one or more of the source vocabularies (U.S. National Library of Medicine, 2009-).

Figure 9: Concept Unique Identifiers in the UMLS Metathesaurus



Source: Generated based on the guidelines and images captured from U.S. National Library of Medicine, 2016.

This methodology of using URIs to represent the concepts in a classification is officially implemented in the ICD-11, the newest version of ICD. ICD-11 is a vast expansion on ICD-10, updated to reflect vital advances in life science and medicine for the 21st century, and enhanced to integrate with electronic health information systems. Each of the ICD-11 entities has a URI which is independent of the classification notations. In addition to ensuring the ICD feature of language-independence and maintaining the classification structure, this URI practice will enhance its consistency and semantic interoperability in the implementations, while allowing binding of any desired language to the elements of its foundation component (World Health Organization, 2019, 1.2.6). It ensures consistency with traditional use cases of earlier ICD versions, because it has been built with the past revisions in mind. Past data analyses based on older versions of ICD can be linked to analyses of data based on ICD-11 (World Health Organization, 2019). In this example (Figure 10), each of the vaccine-related entries has its own “Foundation URI”.

Figure 10: Example of URIs used in ICD-11.

The image shows a screenshot of the ICD-11 browser interface. On the left, the 'search' tab is active, displaying a list of COVID-19 related entities. A red arrow points from the 'XM68M6 COVID-19 vaccine' entry in the search results to the 'browse' tab. The 'browse' tab shows a hierarchical list of entities, including 'XM68M6 COVID-19 vaccine', 'XM1NL1 COVID-19 vaccine, inactivated virus', 'XM5DF6 COVID-19 vaccine, live attenuated virus', 'XM9QW8 COVID-19 vaccine, non-replicating viral vector', 'XMOCX4 COVID-19 vaccine, replicating viral vector', 'XM5JCS COVID-19 vaccine, virus protein subunit', 'XM1J92 COVID-19 vaccine, virus like particle (VLP)', 'XM6AT1 COVID-19 vaccine, DNA based', and 'XM0GQ8 COVID-19 vaccine, RNA based'. On the right, the 'ICD-11's URIs' section displays the Foundation URIs for these entities: 'http://id.who.int/icd/entity/B94585096' for XM68M6, 'http://id.who.int/icd/entity/1104808441' for XM1NL1, and 'http://id.who.int/icd/entity/1316179031' for XM6AT1. Red arrows indicate the mapping from the entity names in the browse list to their respective URIs in the right-hand panel.

Source: Generated based on the images captured from ICD-11 for Mortality and Morbidity Statistics (Version 05/2021) browser <https://icd.who.int/browse11/l-m/en>.

Today, advanced technologies, under the umbrella of Big Data and Smart Data, allow for access and reuse of large volumes of diverse data; to unearth patterns and connections formerly hidden from view; to reconstruct the past; to discover the impact and value of qualitative and quantitative variables in both real and virtual environments; and to bring knowledge of the complex intricacies of human society to light. Challenges and opportunities co-exist, but it is certain that Smart Data, the ability to achieve big insights from trusted, contextualized, relevant, cognitive, predictive, and consumable data at any scale, will continue to have extraordinary value in the data-driven age.

Conclusion

Health KOSs have become even more critical in the data-driven age to aid in the frontline endeavors to overcome the obstacles of information overload and semantic conflicts that can occur during special historic and worldwide events like the COVID-19 pandemic. Health KOSs have played important roles in supporting health data exchange and information management, ensuring consistency and interoperability of data collection and reuse among various providers and healthcare settings, and facilitating data normalization, which is a fundamental requirement for any significant and needed subsequent data analysis and information management.

Sources

- Zeng, M. L. (2017). Smart data for Digital Humanities. *Journal of Data and Information Science*, 2(1), 1–12. <https://doi.org/10.1515/jdis-2017-0001>
- Zeng, M. L. (2018). Interoperability. *ISKO Encyclopedia of Knowledge Organization (IEKO)*, <https://www.isko.org/cyclo/interoperability.htm>
- Zeng, M. L., Hong, Y., Clunis, J., He, S., & Coladangelo, L. P. (2020). Implications of knowledge organization systems for health information exchange and communication during the COVID-19 pandemic. *Data and Information Management*, 4(3), 148-170.
- Hong, Y., & Zeng, M.L. [2022]. International Classification of Diseases (ICD). *ISKO Encyclopedia of Knowledge Organization (IEKO)* [in-review]

References

- Association of Public Health Laboratories. (2020a). *Coronavirus Disease (COVID-19) Data Exchange on the AIMS Platform*. [Infographic]. <https://www.aphl.org/aboutAPHL/publications/Documents/INFO-2020-AIMS-COVID-19-Data-Flow-Infographic.pdf>
- Association of Public Health Laboratories. (2020b). *2019nCoV_PHLIPsample_2.3.1_Detected_UPDATED3.3.20*. https://www.aphl.org/programs/preparedness/Crisis-Management/Documents/2019nCoV_PHLIPsample_2.3.1_Detected_UPDATED3.3.20.pdf
- DATAVERSITY Education, LLC. (2017). *2017 Smart Data Conference: San Francisco Bay, California*. Jan.30-Feb.1, 2017. <http://smartdata2017.dataversity.net/>
- Gardner, D. (2012). An ocean of data [Introduction]. In R. Smolan & J. Erwitte (Eds.), *The Human face of Big Data*. (pp. 14-17). Against All Odds Productions.
- HL7 International. (2007-2021). *HL7 Version 2 Product Suite*. Retrieved from https://www.hl7.org/implement/standards/product_brief.cfm?product_id=185
- Iafrate, F. (2015). *From Big Data to Smart Data*. ISTE Ltd and John Wiley and Sons, Inc.
- International Committee on Taxonomy of Viruses. (2020). *Naming the 2019 coronavirus*. ICTV. <https://talk.ictvonline.org/information/w/news/1300/page>
- Kobielus, J. (2016, July 13). *The Evolution of Big Data to Smart Data* [Keynote address]. Smart Data Online Conference 2016.
- Mukerjee, P. (2014). *Introduction to data science* [PowerPoint slides]. SlideShare. <http://www.slideshare.net/prithwis/01-intro2-datascienceyantrajaalblog>
- Ouksel, A. M., & Sheth, A. (1999). Semantic interoperability in global information systems. *ACM Sigmod Record*, 28(1), 5-12.

- Salcido, R. (2015). The ABCs of ICD. *Advances in Skin & Wound Care*, 28(2). <https://doi.org/10.1097/01.asw.0000459889.36550.d7>
- Schöch, C. (2013). Big? smart? clean? messy? Data in the humanities. *Journal of digital humanities*, 2(3), 2-13. <http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/>
- Sheth, A.P. (1999). Changing focus on interoperability in information systems: from system, syntax, structure to semantics. In *Interoperating geographic information systems* (pp. 5-29). Springer, Boston, MA. <http://personales.dcc.uchile.cl/~cguieterr/cursos/IC/S98-changing.pdf>
- Sheth, A. (2014, March). *Transforming Big Data into Smart Data: Deriving Value via harnessing Volume, Variety and Velocity using semantics and Semantic Web* [Keynote address]. 30th IEEE International Conference on Data Engineering, Chicago, IL, United States. DOI:10.1109/ICDE.2014.6816634
- TiECON East. (2014, May 29-30). Data is the new oil [Technology track]. 16th Annual TiECON East Conference, Cambridge, MA, United States.
- Toffler, A. (1970). *Future Shock*. New York, N.Y.: Bantam Books.
- U.S. National Library of Medicine. (2009-). Metathesaurus. In *UMLS® Reference Manual [Internet]*, (ch.2). Bethesda (MD): National Library of Medicine (US). <https://www.nlm.nih.gov/books/NBK9684/>
- U.S. National Library of Medicine. (2016). Unique Identifiers in the Metathesaurus. In *UMLS Basics – Metathesaurus*. Bethesda (MD): National Library of Medicine (US). https://www.nlm.nih.gov/research/umls/new_users/online_learning/Meta_005.html
- U.S. National Library of Medicine. (2021a). *Unified Medical Language System (UMLS)*. <https://www.nlm.nih.gov/research/umls/index.html>
- U.S. National Library of Medicine. (2021b). Statistics – 2021AA Release. In *UMLS Metathesaurus Release Documentation*. https://www.nlm.nih.gov/research/umls/knowledge_sources/meta_thesaurus/release/statistics.html
- World Health Organization. (2015). *World Health Organization Best Practices for naming of new human infectious diseases*. https://www.who.int/topics/infectious_diseases/naming-new-diseases/en/
- World Health Organization. (2019). *ICD-11 Reference Guide*. <https://icd.who.int/icd11refguide/en/index.html>
- World Health Organization. (2020). *Naming the coronavirus disease (COVID-19) and the virus that Causes it*. [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it)
- World Health Organization. (2021). *WHO announces simple, easy-to-say labels for SARS-CoV-2 Variants of Interest and Concern*. <https://www.who.int/news/item/31-05-2021-who-announces-simple-easy-to-say-labels-for-sars-cov-2-variants-of-interest-and-concern>
- Zeng, M. L., & Chan, L. M. (2009). Semantic interoperability. *Encyclopedia of Library and Information Sciences, Third Edition*, 4645–4662. Taylor and Francis Group.

OS DESAFIOS PARA A FORMAÇÃO E O DESENVOLVIMENTO PROFISSIONAL NA ORGANIZAÇÃO DO CONHECIMENTO À LUZ DA AGENDA 2030

Natália Bolfarini Tognoli¹,

¹Universidade Federal Fluminense (UFF), nataliatognoli@id.uff.br,
ORCID iD 0000-0002-0609-498X

Resumo

O presente trabalho aborda os desafios e as perspectivas para a formação e o desenvolvimento profissional na Organização do Conhecimento à luz da Agenda 2030, tendo como fios condutores da discussão o acesso à informação e a ética no contexto dos arquivos e bibliotecas. Pretende-se incitar uma reflexão sobre como a ética, enquanto um espaço para reflexão das normas e condutas do ser humano, pode conduzir a Organização do Conhecimento em direção à promoção de instituições eficazes, responsáveis e inclusivas, indo ao encontro do desenvolvimento sustentável promovido pela Agenda 2030. Para tanto, propõe-se uma estrutura ética amparada nos processos de Organização e Representação do Conhecimento contextualizados, na justiça social, na transparência e na privacidade.

Palavras-chave: Organização do Conhecimento, ética profissional, Agenda 2030, acesso à informação, arquivos, bibliotecas.

Introdução

Em setembro de 2015, buscando alcançar o desenvolvimento sustentável, a Organização para as Nações Unidas (ONU) desenvolveu a Agenda 2030 com 17 objetivos de desenvolvimento sustentável (ODS) com 169 metas para estimular um caminho mais sustentável e resiliente em áreas importantes para a humanidade e para o planeta nos próximos 15 anos.

Combinando os Objetivos do Milênio (ODM) – primeira iniciativa global de políticas para o desenvolvimento – com os processos resultantes da Rio+20, a Agenda 2030 tem como objetivo a integração de componentes necessários para a garantia de um desenvolvimento sustentável que pode ser levado a cabo por todos os países engajados em sua participação.

Esse desenvolvimento sustentável, que até a década de 1980 limitava-se aos debates ambientais, equilibra-se agora em três dimensões integradas: a econômica, a social e a ambiental, e pode ser definido como aquele que pode suprir as necessidades da geração atual sem comprometer os recursos das futuras gerações.

Especificamente no que tange à Organização do Conhecimento (OC), os desafios e as perspectivas que se apresentam para a formação e o desenvolvimento dos arquivistas e bibliotecários concentram-se, em sua maioria, no ODS 16 – Paz, Justiça e Instituições Eficazes – cujo desafio é a promoção de sociedades pacíficas e inclusivas por meio do acesso à justiça e à construção de instituições eficazes, responsáveis e inclusivas.

Pretendemos abordar aqui esses desafios e perspectivas tendo como fios condutores dessa discussão a ética e o acesso à informação no contexto das instituições coletoras de cultura como arquivos e bibliotecas que promovem a gestão institucional dos saberes (Homulos, 1990; Fernandes, 1996).

A escolha por esses fios condutores se justifica por dois motivos: *primeiro*, porque a ética, além de fundamentar as normas estabelecidas por um grupo social, permite problematizá-las, abrindo novas perspectivas de pensamento (Capurro, 2010). Ou seja, ao partirmos da perspectiva da ética para discutir os desafios enfrentados pelos arquivistas e bibliotecários no contexto da Agenda 2030, pretendemos lançar luz às reflexões emergentes no âmbito deontológico da OC; *segundo*, porque o acesso à informação além de figurar como um direito fundamental perpassando todos os ODS da Agenda 2030, só poderá ser efetivado a partir do momento em que as informações são disponibilizadas, não prescindindo, portanto, da organização.

A discussão que propomos perpassa necessariamente a abordagem sociocognitiva e epistemológica emergente na OC, que compreende o conhecimento do indivíduo para além de uma perspectiva cognitiva, compreendendo-o em uma perspectiva social, cultural e histórica – notadamente a partir dos estudos de análise de domínio impulsionados por Hjørland e Albrechtsen (1995), e está amparada em uma estrutura que apresenta os processos de Organização e Representação do Conhecimento contextualizados, a justiça social, a transparência e a privacidade como desafios éticos a serem enfrentados pelos arquivistas e bibliotecários, mais especificamente, por aqueles que executam os processos de organização do conhecimento.

A Organização do Conhecimento: um marco teórico-conceitual dinâmico

A Organização do Conhecimento, enquanto marco teórico-conceitual na Ciência da Informação (Tognoli, Milani e Guimarães, 2017), apresenta-se como um terreno fértil para o desenvolvimento de teorias e metodologias que visam organizar e representar o conhecimento com o objetivo de recupera-lo, atuando como mediadora entre a produção e o uso da informação, seja no âmbito do contexto – a exemplo da Arquivologia – seja no âmbito do conteúdo – a exemplo da Biblioteconomia.

A busca por um aprofundamento teórico e pelo desenvolvimento de metodologias é fruto de um terceiro momento que se seguiu à concepção da OC como arte, de forma intuitiva, e como técnica, fruto de uma concepção positivista e utilitarista que prevaleceu durante o século XIX (Guimarães, 2003) e que encontra na *International Society for Knowledge Organization* (ISKO), a partir da década de 1980, um espaço integrador para as discussões interdisciplinares acerca das dimensões conceituais e metodológicas da OC.

Essa interdisciplinaridade que se efetiva na OC a partir das discussões advindas dos campos da Ciência da Informação, Lógica, Psicologia, Linguística, entre outros, para o desenvolvimento de produtos, processos e sistemas de OC, nos permite caracterizá-la como um campo dinâmico (Guimarães, 2017) indo de uma perspectiva Aristotélica e ontológica de organização, baseada na Teoria da Classificação de Ranganathan e na Teoria do Conceito de Dahlberg, para uma abordagem socio-cognitiva em uma perspectiva epistemológica (Hjørland, 2002) que incorpora o contexto de produção, organização e uso do conhecimento, a partir das relações estabelecidas entre o sujeito e o objeto do conhecimento.

O contexto é, como se pode observar, um elemento determinante das circunstâncias de produção e uso dos documentos e, por esse motivo, um pressuposto para os processos, instrumentos e produtos da OC que, em sua essência, são os responsáveis pelo estabelecimento de “pontes” entre essas duas circunstâncias – produção e uso. Assim, no âmbito da OC, conhecer e controlar o contexto é fundamental tanto para compreendê-la enquanto um marco teórico-conceitual, tanto para compreender o lugar e o sentido dos documentos nos sistemas de organização do conhecimento.

Essa perspectiva contextual, que é muito mais antiga e discutida no domínio dos arquivos, está ligada à incorporação de um olhar sociocognitivo e cultural, ampliado a partir da perspectiva da análise de domínio (Hjørland e Albrechtsen, 1995; Beghtol, 1995; Hjørland, 2002, 2004; Tennis, 2003, 2012; Mai, 2005; Smiraglia, 2012, 2015; Albrechtsen, 2015; Guimarães e Tognoli, 2015; López-Huertas, 2015), e às comunidades discursivas nas quais a Organização do Conhecimento atua.

Portanto, ao considerar a análise de domínio a partir de um viés sociológico-epistemológico na CI, as ferramentas, conceitos e estruturas informacionais são moldados segundo comunidades discursivas que fazem parte da divisão social do trabalho. Consequentemente, o foco desloca-se dos paradigmas físico – com ênfase no objeto – e cognitivo – com ênfase no indivíduo – para o paradigma social, que enxerga o sujeito, contextualizado, em sua relação com o objeto-informação (Smit, 2012; Hjørland, 2008; 2002). Logo, como nos mostra Hjørland (2008), diferentes pontos de vista precisam de diferentes formas de organização, ou seja, cada domínio possui necessidades específicas.

Assim, o conhecimento deve ser compreendido como um produto social, uma necessidade social e um dínamo social (Baritè, 2001) e os sistemas que o representam como construções sociais, moldadas por um contexto proposital e cultural, assim como pelo tempo e espaço (Doyle et al., 2015).

Essa perspectiva contextual ou relativista (Mai, 2013) da OC, abre espaço para as discussões sobre os desafios éticos enfrentados pelos arquivistas e bibliotecários em sua formação e atuação profissional, aqui em especial na construção de sistemas de organização e representação do conhecimento mais justos e transparentes, cujo principal objetivo é fornecer acesso às informações contextualizadas, o que vai ao encontro das seguintes metas do ODS 16: **16.10**. Assegurar o acesso público à informação e proteger as liberdades fundamentais, em conformidade com a legislação nacional e os acordos internacionais; **16.7** Garantir a tomada de decisão responsável, inclusiva, participativa e representativa em todos os níveis; **16.6** Desenvolver instituições eficazes, responsáveis e transparentes em todos os níveis (ONU, 2015).

A agenda 2030 e os desafios para a Organização do Conhecimento

O desafio ético 1: Sistemas de Organização do Conhecimento Contextualizados

A sociedade diversa e multicultural que temos hoje exige das instituições instrumentos de acesso à informação que desempenhem um papel político, ético e social para além de uma ferramenta neutra para a recuperação da informação (Mai, 2013) e que evidenciem um conhecimento contextualizado no tempo e no espaço. Como consequência, deve-se pensar a OC a partir de uma das dez premissas de Barité (2001), quando o autor afirma que toda OC é artificial, provisória e determinista.

O debate sobre a não-neutralidade e a universalidade nos sistemas de organização do conhecimento não é recente, encontrando um marco teórico na obra de Berman (1971), *Prejudices and Antipathies*, e ecoando nos estudos que se seguiram sobre os problemas, preconceitos e desvios no tratamento temático e descritivo nas bibliotecas e nos arquivos (Olson, 2002, Milani, 2013, Mai, 2013, Szostak, 2014, Adler e Tennis, 2013, Adler, 2016, Duff; Harris, 2000). O tema, embora em debate há pelo menos 40 anos, está longe de se esgotar. Com o passar do tempo a confirmação de que os processos de organização (aqui em específico a classificação no contexto das bibliotecas e a descrição no contexto dos arquivos) precisam levar em conta o multiculturalismo e a diversidade da sociedade – para além do conteúdo dos documentos e de seu produtor – se faz cada vez mais presente nas discussões da área.

Especificamente no que toca aos produtos e instrumentos de OC, seja nos arquivos, seja nas bibliotecas, os desafios recaem sobre o reconhecimento dessas estruturas enquanto construções sociais que, como tal, são moldadas por um contexto de espaço e tempo e, portanto, toda representação acaba privilegiando e destacando alguns aspectos, conceitos, pontos de vista e experiências, enquanto marginaliza ou desloca seu olhar de outros (Mai, 2013).

Considerando que a escolha sobre como organizar e representar para recuperar e dar acesso às informações será sempre arbitrária, o tema da ética assume um papel fundamental, principalmente no que se refere à sua dimensão profissional (Guimarães, Milani e Evangelista, 2015), uma vez que o arquivista ou bibliotecário, conscientes de seu poder e dos dilemas que emergem dessa consciência, devem adotar uma postura ética em direção à construção de instrumentos e produtos de acesso à informação que possam minimizar os danos causados pelos processos de OC (Mai, 2003).

O caminho não é simples e, por isso, os códigos de ética devem funcionar como uma bússola para arquivistas e bibliotecários, subsidiando-os e amparando-os em suas decisões, enquanto um conjunto de valores aceitos pela comunidade profissional a serem seguidos. De acordo com Mai (2013), o código de ética não funciona como uma lista de respostas sobre como agir quando o profissional se depara com um dilema ético, mas como uma estrutura com a qual o profissional pode contar para se informar sobre suas decisões.

Além dos códigos de ética, Fox e Reece (2012) propõem seis modelos epistemológicos para subsidiar as discussões acerca das decisões éticas dos profissionais bibliotecários e que foram discutidos por Pinho e Milani (2020): Modelo utilitarista, modelo deontológico de Kant, o modelo de Rawls, a ética do cuidado, a ética de Derrida e a ética pragmática de Dewey. Ao final, compreendendo que cada modelo

apresenta prós e contras, os autores propõem uma estrutura ética para a OC: Dever do cuidado; Hospitalidade com suavizações; Direcionamento à consequência com ênfase na melhoria da prática; Tratamento das pessoas como fins com direitos e responsabilidades; e; Não prescrição de ações que sabemos que estão erradas.

Essa preocupação combinada entre as decisões baseadas nas epistemologias e nos códigos de ética vai ao encontro dos estudos de Mai (2013) e de Fallis (2007) para quem a construção de uma estrutura para sistemas de organização e representação do conhecimento éticos e responsáveis não deve prescindir da análise de domínio (Hjørland e Albrechtsen, 1995) e dos processos de tomada de decisão baseados no contexto plural e diverso.

Devemos observar que os valores éticos, enquanto preceitos eleitos pela sociedade ou por uma comunidade profissional a partir de suas raízes culturais e construídos ao longo da história (Pinho e Milani, 2020), além de não serem universais, também não são imutáveis. No entanto, cabe destacar que tanto no domínio da Arquivologia, quanto na Biblioteconomia, quando analisados os valores referentes à Organização do Conhecimento, o acesso à informação figura como um supervalor, a medida em que todos os demais valores concorrem para a sua consecução (Pinho; Milani, 2020, Guimarães, 2008, Guimarães; Milani; Evangelista, 2015, Silva; Tognoli; Guimarães, 2017), o que confirma nossa teoria inicial de tê-lo, ao lado da ética, como um dos fios condutores nessa discussão.

O desafio ético 2: sistemas de Organização do Conhecimento como promotores da justiça social nas instituições

Na esteira dos sistemas de organização do conhecimento mais inclusivos como parte de instituições eficazes na promoção do acesso à informação como um direito fundamental para subsidiar a tomada de decisão responsiva, participativa e representativa, emergem as discussões sobre a promoção da justiça social em arquivos e bibliotecas, compreendida aqui sob a égide do conceito de Young (1990) para quem “a justiça social significa a eliminação da dominação e opressão institucionalizadas.”

Para a autora, a justiça social não está apenas ligada à distribuição de bens materiais, mas de variáveis como a capacidade de tomada de decisão, a divisão de trabalho e a cultura. A justiça social, dessa forma, deve ser entendida para além da justiça distributiva, como um conjunto de relações sociais e não exclusivamente como distribuição de coisas.

Na tendência de dilatar o conceito de justiça social para além da esfera econômica, observa-se uma preocupação crescente também no âmbito das discussões sobre a formação e atuação profissional em OC em assegurar a justiça social por meio do reconhecimento do papel dos arquivos e bibliotecas no estabelecimento e legitimação de identidades e liberdades (Jimerson, 2007). Nestas instituições, a justiça social perpassa, também, valores éticos do profissional, seja no tocante ao processo de organização, seja no tocante à disseminação das informações preservadas no acervo.

A discussão da justiça social nas instituições perpassa os debates sobre a universalidade dos processos e sistemas de organização do conhecimento propostos pelo chamado WEBCHAM (*white, ethnically European, bourgeois, Christian/protestant,*

heterosexual, able-bodied, and male) *maistream*, e vai ao encontro do que discutimos sobre uma OC contextualizada.

Ao questionar a universalidade nos processos de organização e representação, a OC tem se pautado nas teorias críticas dentre as quais destacam-se as voltadas aos estudos de gênero, raça e decolonização. Assim, tem-se, por exemplo, tentativas de “desracializar” o esquema de classificação de Dewey, a partir da Teoria Crítica da Raça (Furner, 2007), a inclusão da ética do cuidado, da teoria feminista, como norteadora nos processos de organização do conhecimento arquivístico (Caswell e Cifor, 2016), os estudos pós-estruturalistas e a interseccionalidade nos estudos feministas (Olson, 1997; 2007; Fox, 2014; 2016), estudos sobre a teoria Queer e a organização e representação do conhecimento na comunidade LGBTQIA (Pinho; Guimarães 2012, Campbell, 2000, Brilmyer, 2018) e sobre as perspectivas indigenistas (Doyle et al., 2015).

Martínéz-Ávila et al. (2016) ao analisarem as correntes epistemológicas e implicações metodológicas de três teorias críticas na Organização do conhecimento – as epistemologias feministas, a teoria Queer e a Teoria Crítica da Raça – advogam que estas devem ser classificadas de acordo com o “outro” previamente marginalizado pelo sistema, para quem a voz foi dada. Nesse contexto, as teorias críticas buscam evidenciar o outro nos sistemas de classificação, indo ao encontro de uma perspectiva de organização voltada para a promoção da justiça social em arquivos e bibliotecas.

Concordamos com Caswell e Cifor (2016) na defesa de que os esforços por uma justiça social nessas instituições devem ser empreendidos no sentido de garantia da equidade e justa distribuição dos recursos, desde a avaliação, aquisição, classificação, descrição e indexação até preservação e difusão para o acesso, estando em consonância com o conceito de justiça distributiva, com vistas a promover a democratização das instituições custodiadoras, como os arquivos e bibliotecas, e a inclusão de indivíduos e comunidades comumente marginalizadas e oprimidas.

Essa opressão, segundo Fox (2016) acontece na Organização do Conhecimento por meio de seus processos e sistemas, a exemplo da classificação, categorização e representações linguísticas, que tendem a considerar os usuários como grupos homogêneos que buscam as informações da mesma maneira, criando espaços hostis e terminologias ofensivas.

Em defesa da reparação como um elemento para garantia da justiça social nas instituições, autores como Adler (2016), Duff e Harris (2000) propõem o uso de taxonomias reparativas que respondam às injustiças propagadas durante anos nos sistemas de classificação e representação do conhecimento em arquivos e bibliotecas, cuja violência, segundo eles, tornou-se sistêmica. A proposta de padrões de descrições liberais (*Liberatory descriptive standards*), onde haja transparência e *accountability* em todos os processos, bem como a participação ativa dos usuários na representação do conhecimento, contribuiria para promover a inclusão de comunidades marginalizadas, democratizando o acesso e o uso às instituições e aos seus documentos, tornando-as, conseqüentemente, mais justas e inclusivas.

Nesse sentido, podemos dizer que a democratização do acesso à informação passa, necessariamente, pela democratização da construção de sistemas de organização do conhecimento, uma vez que é necessário compreender que os usuários vão até as instituições com diferentes propósitos, advindos de diferentes culturas e com

expectativas distintas. Incluí-los no processo de tomada de decisão é fundamental para o processo de acolhimento do usuário no arquivo ou na biblioteca, contribuindo para que ele possa se reconhecer enquanto parte de uma comunidade.

Para dar conta desse desafio, necessária se faz uma transformação para além da atuação de arquivistas e bibliotecários enquanto agentes ativos nessa mudança. A compreensão da justiça social como um imperativo ético a ser seguido por esses profissionais deve estar pautado em uma formação crítica e reflexiva, (Gilliland, 2011; Tognoli e Rocha, 2021, Adler e Harper, 2018), com abordagens pedagógicas que versem sobre as questões éticas e de diversidade dentro de uma estrutura de justiça social com o objetivo de promover uma formação acadêmica além das estruturas tradicionais da Organização do Conhecimento.

Ao abordar os aspectos políticos e sociais nos currículos dos cursos e disciplinas ligadas à Organização do Conhecimento, abre-se um espaço importante para que os estudantes compreendam as bases epistemológicas e ontológicas sobre as quais a OC é construída e como elas funcionam (Adler e Harper, 2018), o que pode ser uma oportunidade para discutir sobre como os sistemas de organização do conhecimento frequentemente reproduzem padrões dominantes e como isso afeta diretamente a circulação da informação, dificultando seu acesso.

Os arquivistas e bibliotecários, aqui em específico aqueles que executam os processos de organização do conhecimento devem, portanto, reconhecer que no âmbito dos arquivos e das bibliotecas, a justiça social pode funcionar como um instrumento para o reconhecimento das desigualdades sistêmicas e estruturais na produção e manutenção dos documentos, atuando diretamente nos sistemas de organização do conhecimento nessas instituições, uma vez que como defendem Adler e Harper (2018) as questões relacionadas à justiça social, diversidade e inclusão são inerentes à OC.

Reconhecendo a justiça social como um imperativo ético necessário, recentemente a *American Library Association* (ALA) adotou-o como nono princípio em seu código de ética e a *Society of American Archivists* (SAA) elaborou uma Declaração sobre Diversidade, Equidade e Inclusão, onde as questões voltadas à justiça social são contempladas.

O desafio ético 3: Sistemas de Organização do Conhecimento para a transparência e accountability

Os documentos, e aqui em específico os documentos de arquivo, são um produto necessário para o funcionamento de qualquer instituição organizada e desempenham um importante papel na transparência das ações, especialmente em sociedades democráticas comprometidas com a prestação de contas administrativa e histórica (Tognoli e Correa, 2020).

Nesse sentido, quando falamos de arquivo e de sociedades democráticas, é importante abordarmos dois conceitos fundamentais: a transparência pública e a *accountability*. A primeira é definida por Silva (2020) como o retorno informacional que o poder público presta à sociedade acerca da aplicação de seus recursos. É o ato de dar publicidade aos dados e às informações para os cidadãos de forma rápida e eficaz. Já a segunda é definida pelo *Glossary of archival and records terminology*, da

Associação de Arquivistas Americanos, como “a capacidade de responder, explicar ou justificar ações ou decisões pelas quais um indivíduo, organização ou sistema é responsável” (Pearce-Moses, 2019).

No contexto das instituições, a *accountability* significa o cumprimento de seus mandatos de forma eficiente e eficaz no âmbito jurídico, cultural e legal no qual atuam, ao passo que para o público significa uma maior transparência, abrindo caminho para a liberdade de informação e proteção da privacidade (Dirks, 2004).

Considerando que os documentos de arquivos são uma fonte importante para a memória nacional e cultural cujo valor probatório encontra-se em sua natureza orgânica, enquanto produto de uma atividade desempenhada por um indivíduo ou instituição, concordamos com Cunnigham e Phillip (2005) e Silva (2002) ao compreendermos os arquivos como um instrumento para ajudar na *accountability*, capacitando os cidadãos contra a má administração e a corrupção. Ainda, segundo Silva (2020), o estudo das organizações e da gestão de seus documentos fundamenta as discussões sobre o fluxo ideal de informações que possa contribuir para a efetividade da transparência da gestão.

Podemos vislumbrar, portanto, a *accountability* como um ponto de intersecção entre os arquivos e a democracia. Miguel (2005) discute o conceito a partir de um eixo horizontal, que está ligado ao controle que os poderes estabelecidos exercem uns sobre os outros, e vertical, ligado à necessidade que os representantes têm de prestar contas e submeter-se ao veredicto da população.

Nesse contexto, os arquivos desempenham um papel central nos mecanismos que controlam e asseguram a *accountability*, a representação e a responsividade dos representantes, essa última compreendida por Miguel (2005) como a disposição dos representantes de se sensibilizarem à vontade dos representados, adotando as políticas por eles preferidas.

No entanto, para que tenhamos uma efetividade na transparência do Estado, bem como arquivos que possam instrumentalizar a *accountability*, resultando em instituições transparentes, eficazes e responsáveis, como vislumbrado no ODS 16, é necessária a construção de sistemas de organização do conhecimento que contemplem as informações acerca dos contextos de produção, funcional e de uso desses documentos.

Especificamente no contexto dos arquivos, esse instrumento é o plano de classificação – compreendido aqui como um sistema de organização do conhecimento desenvolvido para organizar o conhecimento e a informação, com o objetivo de otimizar sua gestão e recuperação e que deve ser construído representando a rede de relações entre o documento, seu produtor, a atividade que lhe deu origem e os outros documentos com os quais guarda relações no mesmo conjunto documental (Mazzocchi, 2018, Tognoli e Correa, 2020). Essas relações, ao serem representadas no plano de classificação evidenciam não apenas o contexto de produção mas, também, o que Durante (2015) chamou de vínculo arquivístico.

No contexto das bibliotecas podemos relacionar a transparência e a *accountability* com a OC quando o bibliotecário expõe as decisões que o levaram à escolha de termos nas linguagens documentárias, bem como o registro das mudanças que eventualmente afetam os esquemas de classificação e as linguagens. (McDonald, 2020).

Consideramos assim que a transparência e a *accountability* nos arquivos e nas bibliotecas guardam uma relação estreita e necessária com os sistemas de organização do conhecimento e que os profissionais, conscientes dessa relação, têm em suas mãos uma das ferramentas necessárias para a construção de instrumentos que possibilitem o acesso às informações contextualizadas como um dos caminhos não apenas para a efetividade da transparência nas instituições, mas também para a inclusão.

É importante destacarmos, por fim, que na dimensão da ética profissional tanto a transparência quanto a *accountability* são consideradas valores fundamentais a serem seguidos pelos arquivistas de acordo com a Declaração de valores do profissional e com o Código de ética da Associação de Arquivistas Americanos (SAA *Core Values Statment and Code of Ethics*, 2020) que defendem que os profissionais ajudem a manter as provas documentais das ações de indivíduos, grupos e instituições preservando os documentos para que as futuras gerações possam buscar a *accountability*.

O desafio ético 4: Sistemas de Organização do Conhecimento e privacidade

No atual contexto, a privacidade, ao lado da proteção dos dados pessoais, apresenta um desafio ao profissional da Organização do Conhecimento que, ao garantir o acesso às informações nos arquivos e nas bibliotecas, deve respeitar a privacidade e os dados pessoais compartilhados entre os indivíduos que produzem e utilizam a informação e as instituições.

Esse desafio se faz ainda mais premente no âmbito do acesso às informações em ambientes digitais, notadamente a partir da emergência do Big data e da inteligência artificial que, nas últimas décadas, estabeleceu uma forte relação com a Ciência da Informação, especialmente no que tange aos aspectos como organização, gerenciamento e recuperação da informação.

Especificamente no contexto da Organização do Conhecimento, a inteligência artificial oferece um suporte técnico para a análise da informação, especialmente na classificação, descrição e indexação dos documentos em ambientes digitais, e sua instrumentalidade na recuperação resulta em uma maior exaustividade e especificidade, com resultados mais relevantes aos usuários (Guimarães et al., 2019).

No entanto, essa exaustividade e especificidade na recuperação vem acompanhada de uma personificação (Pariser, 2012), onde as informações pessoais são coletadas para oferecer um conteúdo personalizado aos usuários, de acordo com suas inferências e conhecimento, um processo que está carregado de interesses e manipulações. Nesse contexto, é necessário que sejam adotadas posturas éticas no que tange aos dados pessoais, de forma que os danos potenciais relacionados às violações de privacidade sejam minimizados.

Embora a discussão sobre a proteção dos dados pessoais tenha emergido com mais força a partir da promulgação do Regulamento Geral sobre a Proteção de Dados, em 2016, pelo Parlamento Europeu – influenciando um movimento similar em países como o Brasil, com a Lei Geral de Proteção de Dados (LGPD), sancionada em 2018 – as preocupações acerca da privacidade e do acesso aos documentos em arquivos e bibliotecas já estavam presentes nos códigos de ética dos profissionais

(International Federation of Library Associations and Institutions, 2011; Society of American Archivists, 1980) como valores fundamentais.

A noção da privacidade como um direito é adotada pela Assembleia Geral das Nações Unidas em 1948, no artigo 12 da Declaração Universal dos Direitos Humanos que a estabelece enquanto um direito humano, e ratificada no artigo 17 do Pacto Internacional dos Direitos Civis e Políticos (1966), adotado pela mesma instituição, o qual consagra que “Ninguém será objeto de ingerências arbitrárias ou ilegais em sua vida privada, sua família, seu domicílio ou sua correspondência, nem de ataques ilegais a sua honra e reputação [...] Toda pessoa tem direito à proteção da lei contra essas ingerências ou esses ataques” (Assembleia Geral das Nações Unidas, 1966).

A privacidade é um conceito fluido, variável com relação ao país e aos grupos culturais ou espaços de informação (Guimarães et al. 2019). Nas bibliotecas, as políticas de privacidade utilizadas para definir as práticas de coleção são estabelecidas levando em conta o usuário, e não a privacidade do autor (Tessler, 2004). Já nos arquivos, a responsabilidade vai além dos usuários, alcançando a proteção da privacidade dos dados dos produtores dos documentos.

No contexto da OC, especial atenção deve ser dada ao momento da descrição em arquivos, quando o arquivista deve equilibrar dois valores fundamentais na produção dos sistemas de organização do conhecimento arquivístico: o acesso e uso – promovendo-o de forma equitativa, minimizando as restrições – e a privacidade – levando em conta tanto a proteção dos dados dos usuários, quanto das pessoas envolvidas nos documentos representados.

Guimarães et al. (2019) ao analisarem dois sistemas de organização do conhecimento nos arquivos, a ISAD (G) e a ISAAR (CPF) à luz do Regulamento Geral sobre a Proteção de Dados, identificaram pontos conflitantes, especialmente nas áreas referentes ao contexto – como nome do produtor e história administrativa/biográfica – e ao conteúdo que apresentam informações sensíveis ligadas a relevância, precisão, *compliance*, consenso, categorias de dados pessoais, transparência, acesso e segurança. Assim, para evitar esses problemas, os autores sugerem que a descrição seja feita no nível da série, evitando a representação dos itens documentais que envolvam uma descrição mais detalhada, embora admitam que o nível mais alto no contexto da descrição também pode infringir nos direitos da proteção dos dados dos indivíduos.

Observamos, portanto, que os sistemas de organização do conhecimento, aqui em específico duas normas internacionais voltadas à representação do contexto e do conteúdo nos arquivos, ainda não levam em conta a proteção dos dados pessoais, o que impacta sobremaneira no direito à privacidade dos cidadãos, seja ele o usuário, seja o produtor dos documentos. Ainda, sua não observância nos sistemas de organização do conhecimento bem como em sua aplicação enquanto valor a ser seguido por arquivistas e bibliotecários, mais especificamente para aqueles voltados às práticas da OC, dificulta, conseqüentemente, a implementação efetiva da meta 16.10 que prevê a proteção das liberdades fundamentais em conformidade com a legislação nacional e os acordos internacionais.

Conclusões

Buscamos nesse trabalho abordar os desafios para a formação e o desenvolvimento profissional na Organização do Conhecimento à luz da Agenda 2030. Longe de esgotarmos o tema, essa apresentação se baseou em dois pilares para discussão: o acesso à informação como um direito básico e fundamental que perpassa todos os ODS da Agenda 2030, e em uma estrutura ética para arquivistas e bibliotecários voltada à Organização do Conhecimento, amparada em processos e sistemas de organização do conhecimento contextualizados, na justiça social, na transparência e na privacidade.

É importante dizer que o debate sobre o acesso às informações registradas e institucionalizadas, ou seja, documentadas e passíveis de serem socializadas a partir de sua inserção em uma instituição que lhe confere um selo de qualidade (Smit, 2012), se faz ainda mais necessário em um contexto pós-pandemia onde se observa os impactos das notícias falsas e projetos de desinformação que, acredita-se poderem ser minimizados a partir da confiança em instituições custodiadoras eficazes que garantem a autenticidade e a confiabilidade do contexto e do conteúdo dos documentos, a exemplo dos arquivos e das bibliotecas.

Nesse sentido, longe de apresentar respostas aos desafios e perspectivas que se apresentam à Organização do Conhecimento, em específico na dimensão da formação e desenvolvimento e atuação profissional, pretendemos incitar uma reflexão sobre como a ética, enquanto um espaço para reflexão das normas e condutas do ser humano – pensada de forma universal, sem desconsiderar os contextos sociais, históricos e geográficos (Capurro, 2010) – pode conduzir a Organização do Conhecimento em direção à promoção de instituições eficazes, responsáveis e inclusivas, indo ao encontro do desenvolvimento sustentável promovido pela Agenda 2030.

Referências

- Adler, M. & Tennis, J.T. (2013). Toward a Taxonomy of Harm in Knowledge Organization Systems. *Knowledge Organization* 40(4): 266-72. <http://dx.doi.org/10.7152/nasko.v4i1.14641>
- Adler, M. (2016). The Case for Taxonomic Reparations. *Knowledge Organization*, 43(8): 630-40. <https://doi.org/10.5771/0943-7444-2016-8-630>.
- Adler, M., & Harper, L.M. (2018). Knowledge Organization from a social justice perspective. *Library Trends*, 67(1): 52-73.
- Albrechtsen, H. (2015). This is not domain analysis. *Knowledge Organization*, 42(8): 557-61.
- Barité, M. (2001). Organización del conocimiento: um nuevo marco teórico-conceptual em Bibliotecología y Documentación. In Carrara, L. (Ed.), *Educação, universidade e pesquisa* (pp. 35-60) UNESP.
- Beghtol, C. (1995). Domain analysis, literary warrant, and consensus: the case of fiction studies. *Journal of the American Society for Information Science*, 46(1): 30-44.
- Berman, S. (1971). Prejudices and antipathies: a tract on the LC Subject Heads concerning people. MacFarland & Co.

- Brilmyer, G. (2018). Archival assemblages: applying disability studies' political/relational model to archival description. *Archival Science*, 18(2):95-118, <https://link.springer.com/article/10.1007/s10502-018-9287-6>.
- Campbell, G. (2000). Queer theory and the creation of contextual subject access tools for gay and lesbian communities. *Knowledge Organization* 27(3): 122-31.
- Capurro, R. (2010). Desafios teóricos y prácticos de la ética intercultural de la información. In G.H.A. Freire (Ed.), *Ética da Informação: conceitos, abordagens, aplicações* (pp.11-51). Ideia.
- Caswell, M., & Cifor, M. (2016). From human rights to feminist ethics: radical empathy in the archives. *Archivaria*, 81(spring): 23-43.
- Cunningham, A., & Phillips, M. (2005). Accountability and accessibility: ensuring the evidence of e-governance in Australia. *Aslib Proceedings*, 57(4): 301-17. <https://doi.org/10.1108/00012530510612059>
- Dirks, J. M. (2004). Accountability, History, and Archives: Conflicting Priorities or Synthesized Strands? *Archivaria* 57(spring): 29-49.
- Doyle, A.M., Lawson, K., & Dupont, S. (2015). Indigenization of Knowledge Organization at the Xwi7xwa Library. *Journal of Library and Information Studies*, 13(2): 107-134. [https://doi:10.6182/jlis.2015.13\(2\).107](https://doi:10.6182/jlis.2015.13(2).107).
- Duff, W. M. & Harris, V. (2002). Stories and Names: Archival Description as Narrating Records and Constructing Meanings. *Archival Science*, 2(3-4): 263-85.
- Duranti, L. (2015). Archival Bond. In L. Duranti & P. Franks (Eds.). *Encyclopedia of Archival Science*, Rowman & Littlefield (pp. 23-29).
- European Parliament. Council of the European Union. (2016). General Data Protection Regulation (EU) 2016-679. Brussels.
- Fallis, D. (2007). Information ethics for the twenty-first century library professionals. *Library Hi tech*, 25 (1): 23-36. <https://doi:10.1108/07378830710735830>.
- Fernandes, G. C. (1995). O objeto de estudo da Ciência da Informação. *Informare*,1(1):25-30.
- Fox, M. J. (2014). Enabling gender-inclusivity in essential questions." *Journal of Education for Library and Information Science* 55:241-9.
- Fox, M. J., & Reece, A. (2012). Which ethics? Whose morality?: An analysis of ethical standards for information organization. *Knowledge Organization*, 39 (5): 377-383.
- Fox, M.J. (2016). Priorities of Arrangement" or a "Hierarchy of Oppressions? *Knowledge Organization*, 43(5): 373-83.
- Furner, J. (2006). Dewey Deracialized: A critical race-theoretic perspective. *Knowledge Organization*, 34(3):144-168
- Gilliland, A. (2011). Neutrality, social justice and the obligations of archival education and educators in the twenty-first century. *Archival Science*, 11(3-4): 193-209. <http://dx.doi.org/10.1007/s10502-011-9147-0>.
- Guimarães, J. A. C. (2017). Organização do Conhecimento: passado, presente e futuro sob a perspectiva da ISKO. *Informação & Informação*, 22(2): 84 – 98.
- Guimarães, J. A. C., Fernandez-Molina, J.C., Pinho, F.A., & Milani, S.O. (2008). Ethics in the Knowledge Organization Environment: overview of values and problems in the LIS literature. In: C. Arsenault & J.T. Tennis, (Eds.). *Cultural and Identity in Knowledge Organization* (pp.340-346) ERGON.
- Guimarães, J. A. C., Milani, S. O., & Evangelista, I. V. (2015). Valores éticos na organização e representação do conhecimento. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, 20(1): 19-32.
- Guimarães, J.A.C and Tognoli, N.B. (2015). Provenance as a Domain Analysis Approach in Archival Knowledge Organization. *Knowledge Organization*, 42(8): 562-69.

- Guimarães, J.A.C., Tognoli, N. B., Bagatini, J. A., Martinez-Avila, D. (2019). Ethical challenges in Archival Knowledge Organization: the description of personal data for long-term preservation In D. Haynes & J. Vernau (Eds.), *The Human Position in an Artificial World: Creativity, Ethics and AI in Knowledge Organization*. Ergon-Verlog, (pp.113-126).
- Hjørland, B. & Albrechtsen, A. (1995). Toward a new horizon in Information Science: Domain-Analysis. *Journal of the American Society for Information Science*, 46(6), 400-25.
- Hjørland, B. (2002). Epistemology and the socio-cognitive perspective in information science. *Journal of the American Society for Information Science and Technology*, 53 (4): 257-270. <https://doi.org/10.1002/asi.10042>
- Hjørland, B. (2004). Domain analysis: a socio-cognitive orientation for Information Science research. *Bulletin of the American Society for Information Science and Technology*, 30(3):17-21.
- Hjørland, B. (2008). What is knowledge organization (KO)? *Knowledge Organization*, 35 (2-3): 86-101.
- Homulos, P. (1990). Museums to libraries: a Family of collectinmg institutions. *Art Libraries Journal*, 15(1): 11-13.
- Jimerson, R. C. (2007). Archives for All: professional responsibility and social justice. *The American Archivist*, 70(2): 252-281. [https:// doi/abs/10.17723/aarc.70.2.5n20760751v643m7](https://doi/abs/10.17723/aarc.70.2.5n20760751v643m7)
- Mai, J.E. (2013). Ethics, Values and Morality in Contemporary Library Classifications. *Knowledge Organization*, 40(4): 242-253.
- Martínez-Ávila, D., Semidão, R., & Ferreira, M. (2016) methodological aspects of critical theories in Knowledge Organization. *Knowledge Organization*, 43 (2): 118 – 125.
- McDonald, C. (2020). Call us by our name(s). Shifting representations of the transgender community in classificatory practices. In M. Lykke, T. Svarre, M. Skov & D. Martinez-Ávila (Eds), *Knowledge Organization at the Interface*. Ergon-Verlog, (pp. 290-294).
- Mazzocchi, F. (2018). Knowledge organization system (KOS). *Knowledge Organization* 45(1):54-78.
- Miguel, L. F. (2005). Impasses da Accountability: Dilemas e Alternativas da Representação Política. *Revista de Sociologia e Política* 25: 25-38. <https://doi.org/10.1590/S0104-44782005000200004>
- Milani, S.O. (2013). *Bias na Representação de Assunto: Uma Discussão de Oposições Binárias nos Functional Requirements for Subject Authority Data (FRSAD)*. (Tese de doutorado, Universidade Estadual Paulista “Júlio de Mesquita Filho, Marília, Brasil).
- Olson, H. A. (1997). The feminist and the Emperor’s new clothes: feminist deconstruction as a critical methodology for Library and Information Studies. *Library & Information Science Research*, 19(2):181-98. [https://doi.org/10.1016/S0740-8188\(97\)90042-6](https://doi.org/10.1016/S0740-8188(97)90042-6)
- Olson, H. A. (2002). The power to name: locating the limits of subject representation in libraries. Kluwer Academic.
- Olson, H. A. (2007). How we construct subjects: a feminist analysis. *Library Trends*, 56 (2):509-41. <http://hdl.handle.net/2142/4586>
- Organização das Nações Unidas (ONU). (2015) Transformando nosso mundo: a agenda 2030 para o desenvolvimento sustentável. ONU. Disponível em: <http://www.agenda2030.org.br/ods/16/>
- Pariser, E. (2011). *The Filter Bubble: what the internet is hiding from you*. Penguin Press.
- Pinho, F. A., & Guimarães, J.A.C. (2012). Male homosexuality in Brazilian indexing languages: some ethical questions. *Knowledge Organization* 39(5): 363-9.
- Pinho, F.A., & Milani, S.O. Ética em Organização do Conhecimento: categorização de termos fronteirços em relação a gênero e sexualidade. *Logeion: Filosofia da Informação*, 6(2): 84-103.

- Silva, A.P. da; Tognoli, N.B., & Guimarães, J.A.C. (2017). Os valores éticos na organização e representação do conhecimento arquivístico. *Brazilian Journal of Information Studies: Research Trends*, 11(1): 44-53.
- Silva, J.A. da. (2020). *A efetividade da transparência pública no Brasil à luz dos 'dispositivos de acesso à informação': entre possibilidades e limitações*. (Tese de doutorado, Universidade Federal Fluminense, Niterói, Brasil).
- Smiraglia, Richard P. (2012). Epistemology of Domain Analysis. In R. P. Smiraglia & H. Lee (Eds.), *Cultural Frames of Knowledge* (pp. 114-24). Ergon.
- Smit, J.W. (2012). A informação na Ciência da Informação. *InCID: Revista de Ciência da Informação e Documentação*, 3(2): 84-101.
- Society of American Archivists (SAA). (2012). Core Values Statement and Code of Ethics. <http://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics#core-values>.
- Szostak, R. (2014). Classifying for social diversity. *Knowledge Organization*, 41(2): 160-171.
- Tennis, J. T. (2003). Two axes of domains for domain analysis. *Knowledge Organization*, 30(3/4): 191-95.
- Tennis, J.T. (2012). What does a domain analysis look like in form, function, and genre? *Brazilian Journal of Information Science*, 6(1): 3-14.
- Tessler, C. Z. (2014). Privacy, Restriction, and Access: legal and ethical dilemmas". *SLIS Student Research Journal*, 4(1):1-9.
- Tognoli, N. B., & Correa, L.S. (2020). Knowledge Organization Systems as Accountability Tools in Archival Science. In M. Lykke, T. Svarre, M. Skov & D. Martinez-Ávila (Eds), *Knowledge Organization at the Interface*. Ergon-Verlog, (pp. 569-571).
- Tognoli, N. B., & Rocha, G.M.Q. (2021). A justiça social na Arquivologia: uma revisão de literatura. *ÁGORA: Arquivologia Em Debate*, 31(63), 1–23.
- Tognoli, N.B; Milani, S.O., & Guimarães, J.A.C. (2017, novembro). Elementos de indexação para arquivos permanentes: aproximações teóricas preliminares. Atas do III Congresso ISKO Espanha-Portugal e XIII Congresso ISKO Espanha, Coimbra, Portugal.
- Young, I. M. (1990). *Justice and the Politics of Difference*. Princeton University.

A PRODUÇÃO CIENTÍFICA SOBRE A AGENDA 2030 INDEXADA NA WEB OF SCIENCE E SCOPUS: CIÊNCIAS SOCIAIS E CIÊNCIA DA INFORMAÇÃO

Teresa Costa¹, Luísa Alvim²

¹Centro de Estudos Clássicos – Universidade de Lisboa;
CIDHEUS-UE, Portugal, teresa.costa.23@gmail.com, ORCID 0000-0003-4928-9047

²CIDHEUS-UE; CEIS20 – Universidade de Coimbra,
Portugal, mluisa.alvim@gmail.com, ORCID 0000-0001-9106-1658

Resumo

Este estudo descreve o comportamento bibliométrico da produção científica da área do conhecimento da Ciências Sociais, com destaque para a Ciência da Informação, sobre a temática Agenda 2030 e os Objetivos de Desenvolvimento Sustentável a nível mundial, indexada na Web of Science (WoS) e na Scopus entre 2015 e 2020. A Agenda 2030, fixada em 2015, apresenta os Objetivos de Desenvolvimento Sustentável abrangendo o desenvolvimento económico, cultural, ambiental e social. Este estudo dá a conhecer o cenário global sobre a temática enunciada e revela o estado de arte da área do conhecimento das Ciências Sociais em geral e da Ciência da Informação em particular. No total foram recuperados 981 artigos e com a presente investigação, pretende-se ver quais os Objetivos de Desenvolvimento Sustentável mais trabalhados em termos de investigação, as principais temáticas e as áreas geográficas. A investigação sobre a Agenda 2030 continua em expansão e desenvolvimento a nível internacional e admite-se que a produção científica reflita novas áreas de investigação.

Palavras-chave: Agenda 2030, Bibliometria, Objetivos de Desenvolvimento Sustentável, Ciência da Informação, Ciências Sociais.

Introdução

Este trabalho de investigação descreve o comportamento bibliométrico da produção científica da área do conhecimento da Ciências Sociais, com destaque para a Ciência da Informação, sobre a temática Agenda 2030 e os Objetivos de Desenvolvimento Sustentável a nível mundial, indexada na Web of Science (WoS) e na Scopus entre 2015 e 2020. Este estudo dá a conhecer o cenário global sobre a temática enunciada e revela o estado de arte da área do conhecimento das Ciências Sociais em geral e da Ciência da Informação em particular.

1. Agenda 2030 e os Objetivos de Desenvolvimento Sustentável

A Organização das Nações Unidas realizou a cimeira, em setembro de 2015, onde aprovou a Agenda 2030, com a concordância de 193 membros (Organização das Nações Unidas, 2015). Esta Agenda é ambiciosa e pretende que em todo o universo sejam alcançados os Objetivos de Desenvolvimento Sustentável (ODSs). São abordados 17 objetivos, com as respetivas metas, que dizem respeito às áreas económica, social e ambiental (Organização das Nações Unidas, 2018a).

De 2000 a 2015, os governos e as nações tinham aprovado os 8 Objetivos de Desenvolvimento do Milénio (Organização das Nações Unidas, 2000; 2010). Atualmente, e até 2030, são apresentados os 17 ODSs que dizem respeito ao mundo global. Os ODSs deverão ser implementados pelos estados, governos, instituições, setor privado e público, pelas pessoas individuais, etc. Todos têm um papel proactivo na sua implementação, do global ao local. O lema dos ODSs é transformar o nosso mundo através da integração nas tomadas de decisão, dos diversos setores da sociedade, dos ODSs para que contribuam com o seu poder de inovação para um futuro mais sustentável e inclusivo (Organização das Nações Unidas, 2015).

Os 17 ODSs e as suas 139 metas têm uma ambição universal, são integrados, interrelacionados e dizem respeito às áreas económicas, social, cultural e ambiental. A Agenda incrementa a necessidade de parcerias de todo o género para promover a prosperidade e o bem-estar de todos, a proteger o ambiente, acabar com a pobreza e a fome, combater as alterações climáticas, proteger a terra, a vida terrestre e a vida marinha, promover a igualdade de género, uma melhor educação, saúde, etc. Os países, a nível mundial, foram convidados a construir políticas, medidas e ações para o cumprimento dos objetivos (Organização das Nações Unidas, 2016; 2017; 2018; 2019; 2020).

Estes Objetivos estão interrelacionados pelo desenvolvimento sustentável que se aplica à economia, ao ambiente e à sociedade. A União Europeia sempre demonstrou sensibilidade para esta questão, alertando para uma política comum dos estados membros de coesão social, económica e ambiental (União Europeia. Conselho das Comunidades Europeias, 1992). Atualmente o Eurostat apresenta regularmente relatórios sobre a sustentabilidade e a monitorização do cumprimento dos ODSs (União Europeia. Eurostat, 2018; 2019; 2020).

Faltando uma década para atingir 2030, será necessário os estados e todos os setores da comunidade assumirem rapidamente soluções sustentáveis para os desafios dos ODSs – desde o erradicar a pobreza às mudanças climáticas, com ações globais e locais. As transformações sociais a nível mundial têm de acontecer respeitando as pessoas e o ambiente (Organização das Nações Unidas. Secretário-Geral, 2020).

O Relatório Anual de ODSs apresenta uma visão mundial de todas as etapas de implementação, privilegiando as áreas de desenvolvimento e as áreas em que é necessário ultrapassar dificuldades. Quatro anos após a adoção dos ODSs, este relatório observa avanços em algumas áreas, como a redução extrema da pobreza, a diminuição das taxas de mortalidade infantil e o aumento do acesso das pessoas à eletricidade, mas avisa que a resposta global ainda não é ambiciosa, porque existem muitas pessoas e países vulneráveis (Organização das Nações Unidas, 2020).

1.2. A Agenda 2030, as Ciências Sociais e a Ciência da Informação

Os 17 ODSs apontam para todos os setores da sociedade, visando a implementação de políticas sociais que favoreçam a promoção e o desenvolvimento da Agenda 2030. No presente estudo procurámos conhecer a produção científica, em particular os artigos científicos, publicados sobre esta temática no âmbito geral das Ciências Sociais, e em particular da Ciência da Informação, sobretudo por ser reconhecido e assumido o valor das bibliotecas, dos arquivos, dos centros de documentação, e outros na implementação das diretrizes da Agenda 2030 e dos 17 ODSs.

Mais do que uma Agenda ideológica, a Agenda 2030 é um compromisso de um ideal comum para a melhoria da condição humana em todos os aspetos, visando acima de tudo a justiça e igualdade social. E aqui as bibliotecas e centros de documentação assumem-se como parceiros fundamentais no e para o desenvolvimento sustentável e implementação da Agenda 2030.

Este trabalho de investigação descreve o comportamento bibliométrico da produção científica da área do conhecimento da Ciências Sociais, com destaque para a Ciência da Informação, sobre a temática Agenda 2030 e os Objetivos de Desenvolvimento Sustentável a nível mundial, indexada na Web of Science (WoS) e na Scopus entre 2015 e 2020. Este estudo dá a conhecer o cenário global sobre a temática enunciada e revela o estado de arte da área do conhecimento das Ciências Sociais em geral e da Ciência da Informação em particular.

Na bibliografia encontram-se duas situações sobre esta temática, uma investigação sobre os trabalhos que estas estruturas das bibliotecas efetuam e o papel que desempenham na sociedade; e uma outra sobre a aplicação do desenvolvimento sustentável às próprias estruturas, como por exemplo, as questões associadas à arquitetura, à sustentabilidade dos edifícios, à promoção e consumo sustentáveis, etc. Há outras linhas de investigação na área da Bibliometria e Agenda 2030, sendo relevantes os trabalhos de Nakamura, Pendlebury, Schnell, & Szomszor (2019).

Quanto à primeira situação, em que se dá conhecimento dos contributos destas estruturas para a implementação dos ODSs, convém referir o esforço da IFLA (International Federation of Library Associations and Institutions) para salientar que as bibliotecas e centros de informação têm um papel natural no acesso à informação e a serviços em rede que sustentam o desenvolvimento sustentável.

Anteriormente à aprovação desta Agenda, a IFLA através da Declaração de Lyon (IFLA, 2014) chamou a atenção aos estados membros das Nações Unidas para que assumissem um compromisso internacional de garantir que todos tenham acesso e possam compreender, usar e partilhar a informação necessária para promover o desenvolvimento e promover sociedades democráticas.

Em muitos lugares do mundo e em muitas comunidades locais, os centros de informação e documentação são o único espaço onde os cidadãos podem aceder a informações que as ajudam a melhorar a sua educação, a desenvolver novas competências, a encontrar empregos, a construir negócios, a tomar decisões informadas sobre a agricultura e a saúde ou obter informações sobre questões ambientais e para melhorarem as suas vidas (IFLA, 2017a). O acesso à informação não é um fim em si, mas é o impulsionador do progresso para capacitar as comunidades e sustentar a

igualdade entre todos os indivíduos como defendido na abordagem holística da Agenda 2030 (Garrido, M. & Wyber, 2017; 2019).

Para desenvolver os ODSs, as bibliotecas e centros de informação apoiam ações de literacia da informação, digital e outras, ajudam os governos locais a perceberem as necessidades em matéria de informação necessárias às comunidades, promovem serviços em rede, a inclusão digital e as Tecnologias da Informação e Comunicação (TIC), etc. (IFLA, 2017b). A IFLA difundiu o International Advocacy Programme (IAP) que tem como objetivo aumentar o nível de consciencialização sobre os ODSs dos profissionais de bibliotecas e promover o papel que as bibliotecas podem desempenhar no desenvolvimento a nível local e nacional (IFLA, 2017c).

2. Metodologia

Para a obtenção dos artigos a analisar foi realizada uma pesquisa nas bases de dados Scopus e Web of Science, do dia 13 de fevereiro de 2021, complementada com recursos à investigação documental (Silva, 2018). A expressão de pesquisa foi (“2030 Agenda”) OR ALL FIELDS: (“17 Sustainable Development Goals”) AND (Librar* OR Archiv*) realizada em “todos os campos” com o limite temporal 2015-2020. Numa primeira abordagem, limitaram-se os resultados da pesquisa à área das Ciências Sociais e posteriormente à tipologia documental artigo. Esta estratégia permitiu recuperar um total de 988 artigos, dos quais 835 eram da Scopus e apenas 153 da Web of Science. Destes verificou-se uma sobreposição de sete artigos comuns a ambas as bases de dados, tendo a lista final que serviu de base ao presente trabalho um conjunto de 981 artigos únicos.

A extração e análise dos dados de cada artigo foi feita com recurso ao programa Excel da Microsoft, tendo para tal sido criado um ficheiro com a seguinte informação: base de dados, ano, autor(es), título do artigo, revista de publicação, palavras-chave, resumo e idioma. A estes dados, obtidos a partir dos resultados, adicionou-se ainda o seguinte: ODSs, país/continente e Instituição (para obtenção de dados relativos às instituições da área da CI).

A análise e recolha da informação relativa a cada artigo foi uma tarefa morosa, que em muitos casos, para além da leitura do resumo de cada artigo, implicou o acesso e leitura do texto integral dos mesmos.

Após a recolha de toda a informação, os dados foram trabalhados com recurso a tabelas pivot para a elaboração das várias tabelas e figuras de suporte ao artigo e que ajudam não só na análise dos dados obtidos, mas também na sua visualização e interpretação.

Em termos metodológicos, recorreu-se à bibliometria que permitiu não só quantificar a informação recolhida, mas também descrevê-la e interpretá-la. Desta forma foi possível analisar a produção científica indexada na Scopus e Web of Science sobre a Agenda 2030 e os ODSs, na área das Ciências Sociais e da Ciência da Informação.

3. Resultados e discussão

Da pesquisa realizada nas duas bases de dados, Scopus e Web of Science, e no total dos cinco anos em análise, recuperaram-se 981 artigos sobre Ciências Sociais e os Objetivos para o Desenvolvimento Sustentável.

3.1. Número de artigos por ano

3.1.1. Ciências Sociais

Como se pode observar na figura 1, o número de artigos na área das Ciências Sociais por ano tem aumentado significativamente, passando de apenas dois artigos em 2015 para 353 em 2019 e 297 em 2020. No ano de 2016 foram 41, em 2017, 93 e em 2018 foram 195 artigos.

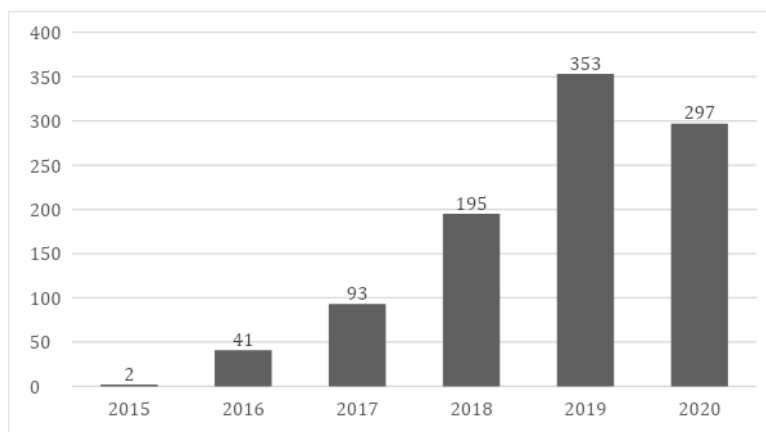


Figura 1: Número de artigos por ano

No final de 2019, como já se referiu, e no período temporal de seis anos, temos um total de 981 artigos sobre os ODSs na área das Ciências Sociais, sendo o número dos trabalhos científicos muito grande. Constata-se um cenário de investigação com uma rede de publicações a crescer, tanto na WoS como na Scopus, resultado do impacto mundial da implementação dos ODSs da Agenda 2030.

3.1.2. Ciência da Informação

Do total de artigos recuperados apenas 49 (5%) são produzidos no âmbito da Ciência da Informação (figura 2), em particular sobre o tema genérico *Bibliotecas*, pois como já anteriormente mencionado no âmbito dos ODSs, estas instituições apoiam ações diversas de literacia da informação, ajudam os governos locais a perceberem as necessidades em matéria de informação fundamentais às comunidades, promovem serviços em rede, a inclusão digital, o uso da Tecnologias da Informação e Comunicação, entre outros.

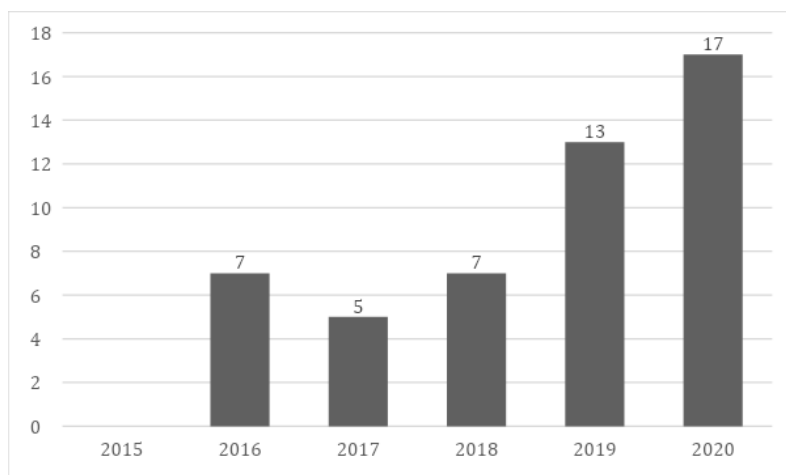


Figura 2: Número de artigos da área da Ciência da Informação por ano

Não se identificou qualquer artigo de 2015, no ano de 2016 há sete e em 2017 apenas cinco artigos. Porém, em 2019 e 2020 o número de artigos aumentou, podendo destacar-se o aumento de 2019 de sete artigos para os treze e de 2019 para 2020 de treze para dezassete, com um total de 49 artigos em seis anos. Acentua-se bem o interesse e o envolvimento das bibliotecas no apoio à divulgação e implementação dos ODSs junto das populações e, por consequência, o aumento de artigos científicos sobre este assunto. Anteriormente, Alvim (2019) afirmou, num trabalho incluindo os recursos exclusivos da WoS, que se considera um bom posicionamento deste domínio científico em termos de produção e revela-se um interesse muito positivo dos investigadores da Ciência da Informação nesta linha de investigação.

3.2. Número de artigos por Objetivo Desenvolvimento Sustentável

3.2.1. Ciências Sociais

Foi possível identificar e recuperar 981 artigos sobre os 17 ODSs, havendo mesmo alguns (117, correspondente a 12%) que referem os 17 Objetivos no seu conjunto.

Como se pode observar pela Tabela 1, e analisando os vários Objetivos individualmente, no âmbito das Ciências Sociais, é possível verificar que o objetivo com mais destaque e que mais artigos abordam é o Objetivo 4 – *Garantir uma educação inclusiva* (184), seguido do Objetivo 11 – *Tornar as cidades sustentáveis* (109), também o objetivo 3 – *Garantir uma vida saudável* merece destaque (74 artigos).

Tabela 1: Número de artigos por ODSs em Ciências Sociais

Objetivo		N.º de artigos
1	Acabar com a pobreza	20
2	Acabar com a fome	30
3	Garantir uma vida saudável	74
4	Garantir uma educação inclusiva	184
5	Alcançar a igualdade de género	48
6	Garantir água e saneamento	42
7	Garantir acesso à energia	17
8	Promover o crescimento económico e o trabalho	70
9	Promover a industrialização e fomentar a inovação	22
10	Reduzir a desigualdade	13
11	Tornar as cidades sustentáveis	109
12	Promoção de consumo sustentável	11
13	Combater alterações climáticas	28
14	Proteger a vida marinha	15
15	Proteger a vida terrestre	60
16	Promover a paz e a justiça	68
17	Fortalecer a parceria global	53
Todos os		
ODS	17 Objetivos de Desenvolvimento Sustentável	117
Total		981

Toda a Agenda 2030 reflete a visão da importância transversal de uma resposta educacional adequada. Também, as Ciências Sociais elegem o ODS 4 – *Garantir uma educação inclusiva* – como tema principal dos trabalhos de investigação. Neste objetivo estão incluídos subtemas, como garantir uma educação de qualidade, equitativa, promover oportunidades de aprendizagem permanente para todos. Sem dúvida, que os investigadores deste domínio científico consideram crucial a educação e a aprendizagem para a realização de muitos outros objetivos, sendo o fio condutor que atravessa a implementação de todos os 17 ODSs (Garrido & Wyber 2019).

Quanto ao ODS 11 – *Tornar as cidades sustentáveis* – é um tema muito focado nos artigos de Ciências Sociais. As questões associadas à urbanização excessiva, à expansão física das cidades, juntamente com os temas do crescimento da população, a falta de infraestruturas básicas nas cidades, as questões políticas relacionadas com um melhor planeamento urbano que apoiem a sustentabilidade (Organização das Nações Unidas, 2018b), são temas sociais que importam aos investigadores destas ciências.

O ODS 3 – *Garantir uma vida saudável e promover o bem-estar para todos* – em todas as idades, está em terceiro lugar a nível de números de artigos. As preocupações sobre a vida saudável, evitando doenças e mortes prematuras, como o financiamento dos sistemas de saúde e as questões das doenças transmissíveis na sociedade são assuntos muito investigados nas Ciências Sociais.

Seguem-se os objetivos abordados nas matérias de investigação, por ordem decrescente de número de artigos: ODS 8 – *Promover o crescimento económico e o trabalho* (70); ODS 16 – *Promover a paz e a justiça* (68); ODS 15- *Proteger a vida terrestre* (60 artigos); ODS 17 – *Fortalecer a parceria global* (53); ODS 5 – *Alcan-*

çar a igualdade de gênero (48); ODS 6 – Garantir água e saneamento (42); ODS 2 – Acabar com a fome (30); ODS 13 – Combater alterações climáticas (28); ODS 9 – Promover a industrialização e fomentar a inovação (22); ODS 1 – Acabar com a pobreza (20); ODS 7 – Fortalecer a parceria global (17); ODS 14 – Proteger a vida marinha (15); ODS 10 – Reduzir a desigualdade (13) e ODS 12 – Promoção de consumo sustentável (11).

3.2.2. Ciência da Informação

Relativamente aos artigos sobre Ciência da Informação (tabela 2), em que a temática relevante é sobre as bibliotecas em geral, a grande maioria aborda os 17 ODSs no seu conjunto (29), evidenciando assim o importante papel destas instituições como agentes fundamentais no acesso à informação, à literacia universal, ao acesso público à informação e às tecnologias da comunicação e património cultural no âmbito da Agenda 2030 (Bradley, 2016, p. 118).

Tabela 2: Número de artigos por ODSs em Ciência da Informação

	Objetivo	N.º de artigos
1	Acabar com a pobreza	1
4	Garantir uma educação inclusiva	9
5	Alcançar a igualdade de gênero	1
8	Promover o crescimento económico e o trabalho	1
10	Reduzir a desigualdade	2
11	Tornar as cidades sustentáveis	1
16	Promover a paz e a justiça	5
20	17 Objetivos de Desenvolvimento Sustentável	29
Total		49

Neste sentido, tem sido fundamental o papel da IFLA, como a principal entidade internacional que representa os interesses dos serviços de informação e dos seus utilizadores, que trabalha para promover a área das bibliotecas como força geradora de uma sociedade informada e letrada.

A IFLA, sendo uma confederação de associações de profissionais ligados às bibliotecas e arquivos, composta por cerca de 150 países, tem uma postura universal, e por outro lado, uma voz local através das associações de cada país, sobre a implementação dos ODSs no mundo tão diverso e desigual. Portanto, tem aconselhado, e feito um trabalho de *advocacy*, partilhando boas práticas e capacitando os profissionais para estas questões no seu todo, fazendo compreender a importância da Agenda 2030 e dos ODSs no seu todo, englobando todo o tipo de desenvolvimento económico, ambiental, cultural e social, em que nenhum ODS é relegado (IFLA, 2017a). Esta situação reflete-se na investigação que acaba, também, por focar-se nos trabalhos e nas práticas dos profissionais da informação que se empenham na implementação dos ODS nas suas bibliotecas.

Como se observa na tabela 2, a abordagem nesta área da Ciência da Informação, sobre os ODSs, é uma visão holística sobre os mesmos.

Dos restantes vinte artigos, nove destacam o papel das bibliotecas como agentes ativos na educação inclusiva (ODS 4), colaborando na aprendizagem ao longo da

vida e no apoio à educação desde o pré-escolar até ao ensino superior. Os outros objetivos abordados nos artigos não têm grande significado, o ODS 16 – *Promover a Paz e a Justiça* com cinco artigos, o ODS 10 – *Reduzir a desigualdade* com dois artigos, o ODS 1 – *Acabar com a pobreza*, o ODS 5 – *Alcançar a igualdade de género*, o ODS 8- *Promover o crescimento económico e o trabalho*, o ODS 11 – *Tornar as cidades sustentáveis*, com um artigo cada.

Observando os resultados, da tabela 3, por tipo de biblioteca e objetivo verifica-se que 18 destes artigos são sobre bibliotecas públicas. Segundo a IFLA, estas instituições encontram-se na vanguarda na implementação destes objetivos, com serviços de informação acessíveis em resposta a interesses e necessidade de diferentes grupos e instituições que compõem as comunidades em que se inserem. As suas ações favorecem o diálogo entre os cidadãos e o poder local, relação-chave que viabiliza a realização de projetos sociais em conformidade com os objetivos da Agenda 2030 (IFLA, 2017b).

Tabela 3: Número de artigos por ODSs e tipo de biblioteca

Objetivo	Tipo	N.º
1	Biblioteca pública	1
	Biblioteca ensino superior	3
	Biblioteca nacional	1
4	Biblioteca pública	2
	Bibliotecas	2
	Repositórios	1
5	Biblioteca ensino superior	1
8	Bibliotecas, Arquivos	1
10	Biblioteca pública	2
11	Biblioteca pública	1
	Arquivo	1
	Bibliotecas	1
16	Bibliotecas públicas	2
	Centros de informação	1
20	Biblioteca ensino superior	5
	Biblioteca itinerante	1
	Biblioteca nacional	1
	Biblioteca pública	10
	Bibliotecas	8
	Bibliotecas escolares	2
	Biblioteca, museus	1
	Bibliotecas sustentáveis	1
Total		49

Estas instituições podem contribuir de forma clara para os Objetivos da Agenda 2030, não como entidades isoladas, mas como centros aglutinadores e de participação social para, e com as comunidades em contacto direto com as administrações públicas e projetos de desenvolvimento local (Fernández-Hernández, S., Lobelle-Fernández, G., & Rivera, Z., 2018) (IFLA, 2018).

Também, as bibliotecas do ensino superior merecem destaque com nove artigos, sobressaindo o importante papel dos bibliotecários, no apoio ao ensino e à investiga-

ção, graças às suas competências ao nível da pesquisa e acesso à informação, bem como no uso das TIC.

Os bibliotecários são cada vez mais agentes fundamentais no processo educacional, graças às suas competências de pesquisa, mas também no uso e aplicação das TIC que a cada dia transformam a forma como a informação é criada e disseminada.

Estes resultados revelam o pensamento dos investigadores da Ciência da Informação sobre o enquadramento de realidades, sejam elas de bibliotecas públicas, escolares, nacionais, do ensino superior, itinerantes ou centros de informação.

3.3. Número de artigos por objetivos e por área geográfica

3.3.1. Ciências sociais

A Agenda 2030 visa uma cobertura global, universal, com aplicação em todo o Mundo. Deste modo, não é de estranhar que tenham sido recuperados artigos com uma visão global em termos geográficos, visando a aplicação dos ODSs nos vários continentes, países ou cidades.

Daí que do conjunto de artigos analisados, 35 são artigos que reportam a situações desenvolvidas em colaboração por vários continentes, países ou cidades.

Contudo, e excluindo os artigos que não fazem qualquer menção a países e/ou continentes (383 artigos), a maioria dos artigos relatam experiências na Europa (170) e em África (157) em termos gerais.

Os ODSs das Nações Unidas representam uma aspiração poderosa e partilhada por um futuro melhor, não sendo de estranhar que a maioria da investigação em Ciências Sociais, se concentre, por um lado, sobre experiências em África, e por outro, na Europa, onde se encontra uma confluência política e financeira com uma agenda crítica para investimentos coordenados e esforços conjuntos na implementação dos ODSs, através da União Europeia.

Relativamente a África, são os objetivos 4 – *Garantir uma educação inclusiva* (28 artigos) e 11 – *Tornar as cidades sustentáveis* (26 artigos) que mais se destacam com maior número de artigos. A mesma tendência verifica-se com os outros continentes, porém no caso da Europa, podem salientar-se igualmente os 26 artigos sobre os 17 ODSs em geral e os 15 sobre o ODS 15 – *Proteger a vida terrestre*.

Relativamente à Ásia recuperaram-se 129 artigos, destaque igualmente para os ODS 4 (22), mas também para o ODS 5 – *Alcançar a igualdade de género* (14), o ODS 3 – *Garantir uma vida saudável* (13), o ODS 8 – *Promover o crescimento económico* (11) e o trabalho e o 11 – *Tornar as cidades sustentáveis* (11).

De referir que no que concerne ao continente americano localizaram-se 97 artigos, a maioria dos artigos sobre a América do Sul ou América Latina (58 artigos). Por fim, e relativamente à Oceânia, há unicamente dez artigos, e apenas sobre os ODS: 3, 4, 8, 13 e 16.

Relativamente aos países, no caso do continente africano, a maioria dos artigos refere situações da África do Sul (22), Nigéria (14) e Gana (14).

No que respeita à Europa, é interessante verificar que a maioria dos artigos é sobre a aplicação dos ODSs na Europa em geral (26 artigos) ou na União Europeia (5). De referir que no que concerne ao continente americano e como já mencionado

a maioria dos artigos é sobre a América Latina, particularmente Brasil (15) e México (14). No caso da Ásia, destaque para a China (29) e Índia (15).

3.3.2. Ciência da Informação

No caso particular da Ciência da Informação, observa-se que a maioria dos artigos relata experiências ocorridas em África (14), com destaque para a Nigéria com quatro artigos. No caso da Europa, há 11 artigos com destaque para a Croácia com cinco.

Para o continente americano obtiveram-se apenas quatro artigos, todos sobre experiências na América do Sul (México, Caracas, Costa Rica e Chile). Recuperaram-se dois artigos sobre experiências no continente asiático (Jordânia e Cazaquistão) e apenas um para a Oceânia (Nova Zelândia). Do total dos 49 artigos que mencionam as bibliotecas e os centros de documentação nove têm uma abordagem focada em experiências comuns a vários países.

São também nove os artigos que não fazem nenhuma referência em termos de localização/aplicação geográfica, destacando-se de forma genérica o papel e contribuição das bibliotecas na aplicação dos ODS como um todo.

Os resultados demonstram que a investigação em Ciências da Informação, tal como aconteceu nas Ciências Sociais, trabalha e foca experiências cujos objetivos estão situados geograficamente sobretudo em África e na Europa. As dinâmicas indicam que o campo de intervenção da Ciência de Informação tem vindo a consolidar as perspetivas relativas à ligação da avaliação do desempenho com a sustentabilidade, constituindo a Agenda 2030 um desafio e, simultaneamente, uma oportunidade interdisciplinar para a convergência e o desenvolvimento de competências transversais (Ochôa & Pinto, 2019).

Conclusões

Desde 2015, que os líderes mundiais adotaram esta ambiciosa Agenda, com 17 Objetivos para o Desenvolvimento Sustentável. Este plano global visa transformar o mundo até 2030, para em última instância, construir vidas dignas para todos. É uma visão universal, integrada e transformadora para um mundo melhor. A publicação científica em Ciências Sociais, na Scopus e na WoS, entre 2015 e 2020, demonstrou uma análise sobre uma questão abrangente e multifacetada que a temática dos ODS e Agenda 2030 propõe.

Assim, este trabalho permitiu, entre outras matérias, fazer uma análise de número de artigos publicados, por anos, nas áreas das Ciências Sociais em geral e da Ciência da Informação em particular, que demonstrou uma evolução crescente no interesse dos investigadores sobre as questões da Agenda 2030 e ODSs, nos últimos cinco anos.

Outra particularidade a referir, é que dos 981 artigos recuperados, sob o ponto de vista da abordagem de conteúdos, maioritariamente, é a visão geral sobre os ODSs que prevalece. Esta visão holística confirma a filosofia dos mesmos que se compreendem de uma forma integral e indivisível.

De qualquer forma, os ODSs 4 – *Garantir uma educação inclusiva*, ODS 11 – *Tornar as cidades sustentáveis* e o ODS 3 – *Garantir uma vida saudável*, estão nas prioridades da investigação das Ciências Sociais. Provavelmente, porque estas áreas de trabalho são prioritárias em todo o mundo: a alfabetização universal; um mundo com o acesso equitativo e universal à educação de qualidade em todos os níveis; aos cuidados de saúde e proteção social, onde o bem-estar físico, mental e social estão assegurados; e onde as cidades e os habitats humanos são seguros e sustentáveis. E no caso particular da Ciência da Informação, o ODS 4 reforçando o papel das bibliotecas, em especial as públicas, enquanto agentes fundamentais no acesso à informação, à literacia universal, ao acesso público à informação.

De destacar ainda o carácter universal dos ODS com aplicação a nível global, e as Ciências Sociais não são exceção. Com base nos resultados obtidos, foi possível confirmar a universalidade de todos os ODS, tendo sido possível identificar artigos relativos aos vários continentes com destaque para África e Europa. A partir desta análise foi possível compreender a forma como cada país implementou os vários objetivos, dependendo das preferências políticas, dos desafios específicos do seu contexto socioeconómico e das suas próprias circunstâncias.

A análise da evolução da Agenda 2030 e da sua relevância académica, profissional e social entre investigadores das Ciências Sociais e da Ciência da Informação a partir das experiências/dinâmicas/relatos, em vários países de todo o mundo, permite salientar um início de uma base de construção de um campo de investigação, face ao tema do Desenvolvimento Sustentável e à Agenda 2030. A área de investigação da Ciência da Informação, apesar de estar a aumentar o número de artigos publicados, será necessário, no futuro, ter um posicionamento estratégico, face a esta temática, no debate científico internacional.

A consolidação interdisciplinar das perspetivas das Ciências Sociais, relativas ao tema do Desenvolvimento Sustentável, é um campo de investigação muito relevante nos próximos tempos, até 2030. Será necessário existir mais trabalhos de investigação para se avaliar o desempenho da implementação dos ODS, criar indicadores cientificamente comprovados e criar um pensamento teórico e conceptual sobre o enquadramento das realidades sociais, culturais, ambientais e económicas, de modo, que a investigação contribua para refletir, discutir e intervir nos desafios e compromissos da Agenda 2030.

Referências

- Alvim, L. (2019). Análise bibliométrica de publicações sobre o tema Agenda 2030 na área da Ciência da Informação indexadas na Web of Science. IX Encuentro Ibérico EDICIC 2019. Org. Facultat de Biblioteconomia i Documentació de la Universitat de Barcelona. Obtido de <http://hdl.handle.net/10760/34420>
- Bradley, Fiona. (2016). ‘A world with universal literacy’ The role of libraries and access to information in the UN 2030 Agenda.” *IFLA Journal*, 42.2: 118-125. Obtido de <https://journals.sagepub.com/doi/full/10.1177/0340035216647393>
- Fernández-Hernández, S., Lobelle-Fernández, G., & Rivera, Z. (2018). Las bibliotecas públicas por el desarrollo sostenible. *Revista Cubana de Información en Ciencias de la Salud (ACIMED)*, 29(2), 1-16. Obtido de http://scielo.sld.cu/pdf/ics/v29n2/a02_1247.pdf

- Garrido, M. & Wyber, S. (2017). *Development and Access to Information (DA2I)*. The Hague. Obtido de <https://da2i.ifla.org/sites/da2i.ifla.org/files/uploads/docs/da2i-2017-full-report.pdf>
- Garrido, M., & Wyber, S. (2019). *Development and Access to Information*. Obtido de <https://da2i.ifla.org/wp-content/uploads/da2i-2019-full-report.pdf>
- IFLA. (2014). *The Lyon Declaration: on access to information and development*. Obtido de <http://www.lyondeclaration.org/>
- IFLA. (2017a). *Access and opportunity for all: How libraries contribute to the United Nations 2030 Agenda*. The Hague: IFLA. Obtido de <https://www.ifla.org/files/assets/hq/topics/libraries-development/documents/access-and-opportunity-for-all.pdf>
- IFLA. (2017b). *IFLA Toolkit: Libraries, Development and the United Nations 2030 Agenda*. The Hague. Obtido de: <https://www.ifla.org/files/assets/hq/topics/libraries-development/documents/libraries-un-2030-agenda-toolkit-2017.pdf>
- IFLA. (2017c). *International Advocacy Programme (IAP) Update – October 2017*. Obtido de <https://www.ifla.org/files/assets/hq/topics/libraries-development/documents/iap-update-october2017.pdf>
- IFLA. (2018). *Libraries and the Sustainable Development Goals: a storytelling manual*. Den Haag: IFLA. Obtido de <https://www.ifla.org/files/assets/hq/topics/libraries-development/documents/sdg-storytelling-manual.pdf>
- Nakamura, M., Pendlebury, D., Schnell, J., & Szomszor, M. (2019). *Navigating the Structure of Research on Sustainable Development Goals*. Institute for Scientific Information. Obtido de <https://clarivate.com/g/sustainable-development-goals/>
- Ochôa, P., & Pinto, L. G. (2019). Agenda 2030 e o campo de intervenção da Ciência de Informação: dinâmicas de aprendizagem, envolvimento e desenvolvimento de competências. *Cadernos BAD*, (2). Obtido de <https://www.bad.pt/publicacoes/index.php/cadernos/article/view/2043/pdf>
- Organização das Nações Unidas. Secretário-Geral. (2020). *Remarks to the General Assembly on the Secretary-General's priorities for 2020*. New York. Obtido de <https://www.un.org/sg/en/content/sg/speeches/2020-01-22/remarks-general-assembly-priorities-for-2020>
- Organização das Nações Unidas. (2000). *United Nations Millennium Declaration*. New York. Obtido de https://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/55/2
- Organização das Nações Unidas. (2010). *Millennium Development Goals Report 2010*. New York. Obtido de http://mdgs.un.org/unsd/mdg/Resources/Static/Products/Progress2010/MDG_Report_2010_En.pdf
- Organização das Nações Unidas. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. New York: UN. Obtido de <https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf>
- Organização das Nações Unidas. (2016). *The Sustainable Development Goals report*. New York. Obtido de http://ggim.un.org/documents/The_Sustainable_Development_Goals_Report_2016.pdf
- Organização das Nações Unidas. (2017). *The Sustainable Development Goals report*. New York. Obtido de <https://unstats.un.org/sdgs/files/report/2017/thesustainabledevelopmentgoalsreport2017.pdf>
- Organização das Nações Unidas. (2018a). Sustainable Development Goals (SDGs). Obtido de <https://www.un.org/sustainabledevelopment/>
- Organização das Nações Unidas. (2018b). *The Sustainable Development Goals report*. New York. Obtido de <https://unstats.un.org/sdgs/files/report/2018/TheSustainableDevelopmentGoalsReport2018-EN.pdf>

- Organização das Nações Unidas. (2019). *The Sustainable Development Goals report*. New York. Obtido de <https://unstats.un.org/sdgs/report/2019/The-Sustainable-Development-Goals-Report-2019.pdf>
- Organização das Nações Unidas. (2020). *The Sustainable Development Goals report*. New York. Obtido de <https://unstats.un.org/sdgs/report/2020/>
- Silva, Carlos Guardado da. (2018). Investigação Documental. *Manual de Investigação Qualitativa: Conceção, Análise e Aplicações*. Lisboa: PACTOR, p. 103-123.
- União Europeia. Conselho das Comunidades Europeias. (1992). *Treaty on European Union*. Luxemburg. Obtido de https://europa.eu/european-union/sites/europaeu/files/docs/body/treaty_on_european_union_en.pdf
- União Europeia. Eurostat. (2018). *Sustainable Development in the European Union: Monitoring report on progress towards the SDGs in an EU context*. Luxembourg. Obtido de <https://ec.europa.eu/eurostat/documents/3217494/9237449/KS-01-18-656-EN-N.pdf/2b2a096b-3bd6-4939-8ef3-11cfc14b9329>
- União Europeia. Eurostat. (2019). *Sustainable Development in the European Union: Monitoring report on progress towards the SDGs in an EU context*. Luxemburgo. Obtido de <https://ec.europa.eu/eurostat/documents/3217494/9940483/KS-02-19-165-EN-N.pdf/1965d8f5-4532-49f9-98ca-5334b0652820>
- União Europeia. Eurostat. (2020). *Sustainable Development in the European Union: Monitoring report on progress towards the SDGs in an EU context*. Luxemburgo. Obtido de <https://ec.europa.eu/eurostat/web/products-statistical-books/-/KS-02-20-202>

IDENTIFICAÇÃO COMO BASE PARA A ORGANIZAÇÃO DO CONHECIMENTO ARQUIVÍSTICO: CONTRIBUIÇÕES PARA O DEBATE SOBRE ACESSO À INFORMAÇÃO NOS ARQUIVOS MUNICIPAIS DO BRASIL NA AGENDA 2030

Alexandre Faben¹, Ana Célia Rodrigues², Carlos Guardado da Silva³

¹Programa de Pós-Graduação em Ciência da Informação – PPGCI/UFF, Arquivo Geral da Cidade do Rio de Janeiro – AGCRJ, alexandrefaben@id.uff.br, ORCID iD 0000-0001-9011-632X

²Programa de Pós-Graduação em Ciência da Informação – PPGCI/UFF, anaceliarodrigues@id.uff.br, ORCID iD 0000-0003-0851-528X

³Universidade de Lisboa, Faculdade de Letras, Centro de Estudos Clássicos, carlosguardado@campus.ul.pt, ORCID iD 0000-0003-1490-8709

Resumo

Com a entrada em vigor da Lei de Acesso à Informação (LAI) no Brasil, em maio de 2012, cresce a demanda de acesso à informação e aos próprios documentos produzidos na administração pública. Nos municípios observa-se a maior dificuldade de sua implementação, pois há um inexpressivo número de arquivos municipais institucionalizados e uma ausência de políticas de gestão de documentos. Este estudo visa contribuir para os debates sobre o acesso à informação e a transparência administrativa na governança dos municípios. A pesquisa tem por objetivo investigar os instrumentos de gestão de documentos publicados e disponíveis no âmbito dos Arquivos Públicos Municipais brasileiros, como contribuição para o acesso à informação e a prestação de contas em contexto de transparência e para a qualidade da governança pública municipal. Espera-se que os resultados desta pesquisa possam trazer novos elementos para o debate sobre a Identificação como parâmetro para a Organização do Conhecimento Arquivístico. Conclui-se que a Organização do Conhecimento Arquivístico contribui de modo relevante para a qualidade da governança pública, particularmente a municipal, e desse modo para os objetivos da Agenda 2030 das Nações Unidas.

Palavras-chave: Identificação Arquivística, Organização do Conhecimento, Acesso à Informação, Arquivos Municipais, Agenda 2030.

Introdução

A complexidade das ações governamentais, a crescente produção de documentos, o avanço das tecnologias e a ampliação de demanda por informação, que caracterizam a sociedade hoje, apontam para a necessidade de políticas públicas voltadas

para a gestão de documentos e sistematização dos arquivos públicos, que asseguram racionalidade, eficiência, transparência e responsabilização administrativa.

No Brasil, os dispositivos expressos na Constituição Federal de 1988, onde se afirma que “compete à administração pública, na forma da lei, a gestão da documentação governamental e as providências para franquear sua consulta a quantos dela necessitem” (§ 2.º do art. 216), ofereceu os fundamentos para a aprovação da Lei Nacional de Arquivos e da Lei de Acesso à Informação. A Lei n.º 8.159, aprovada em 8 de janeiro de 1991, que dispõe sobre a política nacional de arquivos públicos e privados, conhecida como Lei Nacional de Arquivos, faz referência à gestão e ao acesso aos documentos, estabelecendo a rede de arquivos existentes nos níveis de governo. A Lei n. 12.527, de 18 de novembro de 2011, denominada Lei de Acesso à Informação (LAI), vem regulamentar o acesso à informação no âmbito da administração pública brasileira, fortalecendo os instrumentos de controle da gestão pública, o combate à corrupção e a transparência dos atos administrativos. Vigente desde 2012, a LAI tem como princípio basilar o direito amplo à informação, reconhecendo que documento e informação produzidos e conservados pela administração pública constituem bem público e, como tal, devem ser preservados e ter seu acesso garantido à sociedade.

Nas palavras de quem já foi a diretora-geral da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), Irina Bokova ressalta que “o acesso à informação é vital para a construção de sociedades mais inclusivas e sustentáveis, tendo papel fundamental no avanço da Agenda 2030 para o Desenvolvimento Sustentável. A garantia desse direito começa com o compromisso dos governos em desenvolver e reforçar políticas e leis sobre o direito à informação. Isso requer mecanismos eficientes de implementação e uma cultura de transparência nas instituições (UNESCO, 2017).

Os mecanismos de transparência das ações dos municípios previstos na legislação em vigor, possuem grande relevância neste contexto. O acesso da sociedade à informação sobre os gastos dos governo possibilita o controle social das ações governamentais. Além disso, contribui para inibir eventuais atos de corrupção e facilita a fiscalização e a responsabilização ao dar publicidade à informação (Pessanha; Faben, 2020).

A Lei n.º 12.527/2011, popularmente conhecida como Lei de Acesso à Informação (LAI), é o principal marco de regulamentação do direito fundamental de acesso à informação no Brasil. Esta lei vale para órgãos públicos integrantes da administração direta dos Poderes Executivo, Legislativo e Judiciário, incluindo as Cortes de Contas e o Ministério Público; as autarquias, as fundações públicas, as empresas públicas, as sociedades de economia mista e demais entidades controladas direta ou indiretamente pela União, Estados, Distrito Federal e Municípios. Entidades privadas sem fins lucrativos também são obrigadas a dar publicidade a informações referentes ao recebimento e à destinação dos recursos públicos por elas recebidos. Para os efeitos desta Lei, considera-se informação “dados, processados ou não, que podem ser utilizados para produção e transmissão de conhecimento, contidos em qualquer meio, suporte ou formato” (BRASIL, 2011).

A importância desta lei está associada à ampliação da “demanda social por documentos e informações que caracteriza o Estado brasileiro” e a sua função é a de

“fortalecer os instrumentos de controle da gestão pública, o combate à corrupção e a transparência dos atos administrativos” (Rodrigues, 2013, p. 65).

A Lei de Acesso à Informação traz em seu escopo que as informações públicas devem ser publicizadas como regra geral e o sigilo deve ser uma exceção. O direito à informação “vai além de um simples direito comum; trata-se de um dos direitos humanos fundamentais, que pertence a todos os indivíduos e lhes faculta exigir, receber e difundir informação”. Quando a informação está protegida, “fala-se então de um direito à informação, de um direito da informação e isto não importa apenas do ponto de vista legal, mas também do ponto de vista arquivístico” (Baudrit Barquero, 1999, p. 9, 12).

O direito à informação é um dos direitos humanos fundamentais, que pertence a todos os indivíduos, que lhes faculta a possibilidade de exigir, receber e difundir informação e, nesta perspectiva, direito à informação e acesso aos documentos e arquivos são conceitos muito relacionados para a garantia do exercício pleno da cidadania e para a transparência administrativa dos órgãos municipais (Rodrigues, 2019).

O arquivo público municipal tem por finalidade formular e implementar a política municipal de arquivos, por meio da gestão, do tratamento técnico, da transferência, do recolhimento, da preservação e da divulgação dos documentos arquivísticos, em qualquer suporte ou formato, garantindo, desta forma, pleno acesso à informação com vistas a subsidiar as decisões governamentais de caráter político-administrativo, apoiar o cidadão na defesa de seus direitos, além de fomentar o desenvolvimento científico e divulgar o patrimônio documental (CONARQ, 2014, p. 20).

Cabe ao gestor público promover a gestão dos documentos, preservar seus acervos, desenvolver e implantar políticas que deem acesso à informação aos cidadãos (Carli; Fachin, 2016).

Uma política pública com foco em soluções de gestão de documentos e acesso à informação se relaciona com o Objetivo de Desenvolvimento Sustentável (ODS) 16 – Paz, Justiça e Instituições Eficazes da Agenda 2030 da Organização das Nações Unidas (ONU), especificamente com as metas 16.6 – desenvolver instituições eficazes, responsáveis e transparentes em todos os níveis e com a 16.10 – assegurar o acesso público à informação e proteger as liberdades fundamentais, em conformidade com a legislação nacional e os acordos internacionais. A ONU Brasil recomenda que um plano de governo alinhado com o ODS 16 deve incluir “ações para garantir o acesso público à informação e proteger as liberdades fundamentais” e, além disso, “criar instituições eficazes, responsáveis e transparentes em todos os níveis”.

O Tribunal de Contas da União (TCU) indica que um bom sistema de governança pública deve ter como componente a *accountability*, que se refere ao “conjunto de mecanismos e procedimentos que levam os decisores governamentais a prestar contas dos resultados de suas ações, garantindo-se maior transparência e exposição das políticas públicas e na informação, especialmente nas de alta relevância, que impactem os negócios e que envolvam resultados, oportunidades e riscos. A transparência deve situar-se dentro dos limites de exposição que não sejam conflitantes com a salvaguarda de informações” (MATIAS-PEREIRA, 2010, p. 109).

A *accountability* envolve, “além do dever e da responsabilidade de prestar contas, o desejo de fazê-lo de forma voluntária” e sua efetividade só ocorre em contexto

de transparência, como salienta a *International Federation of Accountants* (IFAC) (BRASIL, TCU, 2014, p. 61-62).

As boas práticas de governança pública, com base na *accountability* e na transparência, estão fundamentadas no princípio de “dar transparência da organização às partes interessadas, admitindo-se o sigilo, como exceção, nos termos da lei”, identificando “as exigências normativas e jurisprudenciais de publicidade e as demandas por informação pelas partes interessadas.” (BRASIL, TCU, 2014, p. 61-62).

O atendimento aos dispositivos da LAI, que garantem o acesso à informação, depende da recuperação e do acesso aos documentos públicos, uma vez que, neles, é registrada a informação produzida, recebida e acumulada pelos órgãos e entidades da administração pública no exercício de suas funções e atividades.

Estas questões remetem à necessidade de reconhecimento da natureza do documento e da informação que circulam como ativo estratégico e produto da gestão administrativa e da necessidade de metodologia capaz de gerar tal conhecimento no âmbito da administração pública municipal brasileira como exigência para o efetivo cumprimento da legislação em vigor, perspectiva da Identificação como método de pesquisa que produz conhecimento arquivístico sobre os documentos e o órgão produtor.

Segundo Hjørland (2016) a Organização do Conhecimento (*Knowledge Organization*) trata de organizar e representar documentos, bem como os assuntos e conceitos tanto por humanos como por programas de computador. Para esses propósitos, são desenvolvidos regras e padrões, incluindo sistemas de classificação. O autor afirma que a organização do conhecimento nos arquivos deve, no entanto, ser considerada como parte da KO, salientando que o princípio específico mais importante da organização para este domínio é o princípio da proveniência.

A produção de conhecimento sobre o contexto de produção dos documentos e os vínculos que se revelam na sua forma e conteúdo compõem a base teórica da Identificação, estudos desenvolvidos por Rodrigues (2003, 2008).

Tognoli e Rodrigues (2018), ao estudarem as contribuições da Identificação para a Organização e Representação do Conhecimento, afirmam que a Identificação enquanto método para a arquivística, “permite a normalização das funções arquivísticas que apoiam os procedimentos da gestão documental, contribuindo para a organização e representação do conhecimento arquivístico” (TOGNOLI; RODRIGUES, 2018, p. 178).

Portanto, acessar documentos e informação classificados é base e condição para uma governança pública transparente e comprometida com uma cidadania responsável e inclusiva, perspectiva da gestão de documentos e dos arquivos que é garantida pelos seus instrumentos técnicos norteadores.

Neste contexto, cabe a seguinte questão para nortear a pesquisa: De que instrumentos de gestão dispõem os municípios que lhes permitam reconhecer e tratar documentos públicos para garantir a responsabilização e a prestação de contas em contexto de transparência?

A gestão de documentos é um processo arquivístico *continuum*, que se caracteriza como um conjunto de procedimentos aplicados no controle dos documentos durante todo o seu ciclo de vida, incidindo sobre o momento da produção e acumulação na primeira e segunda idades, ou seja, nos arquivos correntes e no intermediário, até a sua destinação final: eliminação ou preservação definitiva.

O Programa de Gestão de Documentos tem como elementos metodológicos a identificação, a classificação, a avaliação e o planejamento da produção documental (criação), aplicados aos documentos convencionais ou digitais. Tem como resultado a elaboração dos seguintes instrumentos, respectivamente: Sistemas e/ou Manual de Tipologia Documental, Plano de Classificação Funcional e Tabela de Temporalidade com classificação de acesso à informação. Recentemente, a Associação Brasileira de Normas Técnicas (ABNT) traduziu a ISO 15489-1: 2016, publicando a NBR ISO 15489-1: 2018, que define os conceitos e princípios a partir dos quais são desenvolvidas abordagens para produção, captura e gestão de documentos de arquivo.

Estes fundamentos da gestão de documentos devem ser observados para elaboração dos instrumentos legais e normativos, que estabelecem as diretrizes para as políticas de gestão de documentos e de arquivos do Brasil, contexto em que se inserem também os municípios.

Nesta perspectiva, a pesquisa tem por objetivo investigar os instrumentos de gestão de documentos publicados e disponíveis no âmbito dos Arquivos Públicos Municipais brasileiros, como contribuição para o acesso à informação e a prestação de contas em contexto de transparência e para a qualidade da governança pública municipal.

Trata-se de uma questão emergente no âmbito do debate científico e profissional, na busca de soluções para superar os desafios impostos ao direito de acesso à informação a partir dos requisitos da gestão de documentos, como contribuição para a *accountability* (responsabilização e prestação de contas) e transparência administrativa, em especial, no âmbito municipal com adesão à Agenda 2030 da ONU.

Metodologia

A pesquisa conta com uma abordagem metodológica qualitativa de modo descritivo, pois objetiva analisar e relacionar informações acerca da gestão de documentos, bem como o acesso à informação e transparência administrativa na governança dos municípios brasileiros para o cumprimento do ODS 16 da Agenda 2030 da ONU.

Assenta na investigação documental, conforme Silva (2021), quer para a revisão da literatura, não sistemática, quer para a pesquisa e análise dos sítios eletrônicos institucionais, para identificar os instrumentos de gestão de documentos (planos de classificação e tabelas de temporalidade) disponibilizados por Arquivos Municipais brasileiros. O levantamento de dados nos sítios eletrônicos dos Arquivos Municipais deu-se, primeiramente, a partir das informações disponíveis no Cadastro Nacional de Entidades Custodiadoras de Acervos Arquivísticos (CODEARQ) do Conselho Nacional de Arquivos (CONARQ) para identificar os Arquivos Municipais cadastrados. No segundo momento, foram analisados os sítios eletrônicos das instituições arquivísticas municipais identificadas no CODEARQ.

Fazem parte do universo desta pesquisa os Arquivos Públicos Municipais com instrumentos de gestão de documentos publicados e disponíveis em linha. A coleta dos dados foi realizada março de 2021, nos *sites* institucionais dos Arquivos, onde foi possível obter informações sobre a identificação dos instrumentos de gestão de documentos publicados pelos Arquivos. A sistematização dos dados foi realizada

por meio da elaboração de tabelas, sendo apresentados adiante. Assim que os *sites* institucionais dos arquivos foram localizados, foi realizado o levantamento de informações sobre a gestão de documentos: plano de classificação e tabela de temporalidade.

Resultados

Como resultados, observa-se que há uma disparidade no que se refere aos Arquivos Municipais, pois dos 5.570 municípios brasileiros registrados pelo Instituto Brasileiro de Geografia e Estatística (IBGE), apenas 88 possuem Arquivo Municipal institucionalizado, conforme os dados levantados no CODEARQ. Estes dados são ainda mais agravantes, quando apenas 3 destas instituições disponibilizam, nos *sites* institucionais dos Arquivos, os seus instrumentos de gestão de documentos. Nesse sentido, é urgente repensar qual é o papel que os Arquivos Públicos Municipais representam para a sociedade, tendo em vista a ausência das instituições na maior parte dos municípios brasileiros. Se ainda é urgente pensar o compromisso com a transparência administrativa dos atos do governo registrados nos documentos de arquivo, ao exercício pleno dos direitos de cidadania de acesso à informação, muito mais urgente se faz pensar o ODS 16 da Agenda 2030, para que os Arquivos sejam instituições eficazes e capazes de assegurar o acesso público à informação e proteger as liberdades fundamentais dos cidadãos.

Tabela 1: Arquivos Públicos Municipais cadastrados no CODEARQ:

Região	Estado	Arquivo
N O R T E	Acre	–
	Amapá	–
	Amazonas	Arquivo Público Municipal de Manaus
	Pará	Departamento de Arquivo Público Municipal de Curuçá
	Rondônia	–
	Roraima	–
	Tocantins	–
N O R D E S T E	Alagoas	Arquivo Público Municipal de Coruripe
	Bahia	Arquivo Histórico Municipal de Salvador; Arquivo Municipal de Vitória da Conquista; Arquivo Público Municipal Antonio Cerqueira Antunes (Conceição da Feira); Arquivo Público Municipal de Camaçari; Arquivo Público Municipal de Jacobina; Arquivo Público Municipal de Porto Seguro; Arquivo Público Municipal Hermenito Dourado (Irecê); Arquivo Público Municipal Maria Barreto Coutinho (Miguel Calmon); Arquivo Histórico Municipal Memória de Brumado; Arquivo Público Municipal de Caetité; Arquivo Público Municipal de Dom Basílio; Arquivo Público Municipal de Lauro de Freitas; Arquivo Público Municipal de Rio de Contas;

		Arquivo Público Municipal João Mangabeira (Ilhéus); Arquivo Público Municipal Sansão Geraldo Ribeiro (Ourolândia).
	Ceará	-
	Maranhão	-
	Paraíba	Arquivo Municipal da Prefeitura de Soledade
	Pernambuco	Arquivo Público Municipal de Águas Belas; Arquivo Geral de Olinda.
	Piauí	-
	Rio Grande do Norte	-
CENTRO OESTE	Distrito Federal	-
	Goiás	-
	Mato Grosso	-
	Mato Grosso do Sul	Arquivo Histórico de Campo Grande – ARCA
S U D E S T E	Espírito Santo	Arquivo Geral de Cariacica; Arquivo Geral Municipal da Serra; Arquivo Geral Municipal de Vitória; Arquivo Público Municipal de Domingos Martins.
	Minas Gerais	Arquivo da Câmara Municipal de Arinos; Arquivo da Cidade de Barbacena Visconde de Lima Duarte (Barbacena); Arquivo Geral da Prefeitura Municipal de Mariana; Arquivo Histórico Municipal de Muriaé; Arquivo Histórico-Geográfico Carangolense (Carangola); Arquivo Municipal de Uberaba; Arquivo Público da Cidade de Belo Horizonte; Arquivo Público de Itabira; Arquivo Público de Tiradentes; Arquivo Público de Uberlândia; Arquivo Público Municipal de Bonfim; Arquivo Público Municipal de Diamantina; Arquivo Público Municipal de Divinópolis; Arquivo Público Municipal de Dores de Guanhães; Arquivo Público Municipal de Ouro Preto; Arquivo Público Municipal de Varginha; Arquivo Público Municipal José Otaviano Pacheco (Formiga); Arquivo Público Municipal Olímpio Michael Gonzaga (Paracatu); Arquivo Central da Prefeitura Municipal de Paracatu.
	Rio de Janeiro	Arquivo Geral da Cidade do Rio de Janeiro; Arquivo Público e Histórico Municipal de Valença; Arquivo Público Municipal de Campos dos Goytacazes.

	São Paulo	Arquivo Municipal de Caraguatatuba Arino Sant'Ana de Barros (Caraguatatuba); Arquivo Histórico Municipal Felix Guisard Filho (Taubaté); Arquivo Municipal de Barueri; Arquivo Municipal de Campinas; Arquivo Municipal e Histórico Athayde Marcondes (Pindamonhangaba); Arquivo Público do Município de São José dos Campos; Arquivo Público e Histórico de Ribeirão Preto – Casa da Memória; Arquivo Histórico de São Paulo; Arquivo Público e Histórico do Município de Rio Claro – Oscar de Arruda Penteadado; Arquivo Público e Histórico Municipal Sérgio Buarque de Holanda (Porto Feliz); Arquivo Público Municipal de São Caetano do Sul; Arquivo Público Municipal de São José do Rio Preto; Arquivo Público Municipal Nilson Cardoso de Carvalho (Indaiatuba); Fundação Arquivo e Memória de Santos.
SUL	Paraná	Arquivo Público Municipal de Curitiba; Arquivo Público de Londrina
	Rio Grande do Sul	Arquivo Público e Histórico Municipal do Rio Grande; Arquivo Histórico de Porto Alegre – Moysés Vellinho; Arquivo Histórico do Município de Torres; Arquivo Histórico Municipal Antônio Soares da Fonseca (Gravataí); Arquivo Histórico Municipal de Alegrete – Miguel Jacques Trindade; Arquivo Histórico Municipal de Santa Maria; Arquivo Histórico Municipal João Spadari Adami (Caxias do Sul); Arquivo Público Municipal de Bagé; Arquivo Público Municipal de Novo Hamburgo; Arquivo Público Municipal de São Francisco de Assis.
	Santa Catarina	Arquivo Histórico de Balneário Camboriú; Arquivo Histórico de Joinville; Arquivo Histórico do Município de Campo Alegre – Eugênio João Herbst; Arquivo Histórico do Município de Florianópolis – Professor Oswaldo Rodrigues Cabral; Arquivo Histórico Eugênio Victor Schmöckel (Jaraguá do Sul); Arquivo Histórico José Ferreira da Silva (Blumenau); Arquivo Histórico Municipal Waldemar Rupp (Campo Novo); Arquivo Histórico Municipal de São José; Arquivo Histórico Pastor Wilhelm Lange (Guaramirim); Arquivo Público de Caçador; Arquivo Público do Município de Palhoça; Arquivo Público e Histórico do Município de Tubarão; Arquivo Público Municipal de Campo Belo do Sul; Arquivo Público Municipal de Chapecó.

Tendo em vista as cinco regiões brasileiras, a região nordeste conta com 19 Arquivos Municipais cadastradas no CODEARQ, onde nenhuma possui instrumentos de gestão de documentos disponíveis em linha, além de poucas informações sobre o funcionamento dos Arquivos. Seguindo o mesmo caminho, as regiões centro-oeste e norte possuem respectivamente, apenas 1 e 2 Arquivos Municipais cadastrados, também sem muitas informações facilmente acessíveis. A região sul possui 25 Arquivos Municipais cadastrados no total, onde apenas o Arquivo Público de Londrina possui instrumentos de gestão de documentos disponíveis em linha. Na região sudeste que contém 41 instituições arquivísticas cadastradas, somente a Fundação Arquivo e Memória de Santos conta com seus referidos instrumentos de gestão de documentos publicados e disponíveis em linha.

Tabela 2: Arquivos Públicos Municipais do Brasil que possuem instrumentos de gestão de documentos publicados

Região	Cidade/Estado	Arquivo	Instrumento de Gestão de Documentos	Sítio eletrônico
SUDESTE	Santos/SP	Fundação Arquivo e Memória de Santos	Tabela de Temporalidade	http://www.fundasantos.org.br/page.php?215
SUL	Londrina/PR	Arquivo Público de Londrina	Tabela de Temporalidade Plano de Classificação	https://portal.londrina.pr.gov.br/arquivo-publico-de-londrina/gestao-documental?start=2
	São Francisco de Assis/RS	Arquivo Público Municipal de São Francisco de Assis	Plano de Classificação Tabela de Temporalidade	https://www.saofranciscodeassis.rs.gov.br/pagina/1031/arquivo-publico-municipal

Sabemos que, embora, algumas instituições não disponibilizem os instrumentos de gestão de documentos, não quer dizer que eles não existam. Todavia, reforçamos a importância de publicizar o plano de classificação e tabela de temporalidade, como forma de contribuir com a transparência ativa e possibilitar o acesso à informação.

Neste sentido, é fundamental dotar a administração pública municipal de instrumentos de gestão de documentos, que permitam o acesso à informação contida nos documentos arquivísticos. A denominação correta dos documentos de arquivo, nos parâmetros da tipologia documental, é necessária para evidenciar a ligação dos documentos com as respectivas atividades e funções que lhes deram origem. Ressalta-se que a identificação dos documentos pode contribuir para a organização do conhecimento arquivístico, pois conhecer onde a informação está registrada possibilita o acesso e propicia a transparência administrativa.

Para garantir o acesso aos documentos e à informação, o Arquivo deve ser uma instituição responsável e eficiente. Para ser uma instituição responsável e inclusiva, o arquivo municipal deve ser transparente e garantir aos cidadãos o direito de acesso à informação e documentos públicos. Os prazos de guarda nas tabelas de temporalidade bem como a classificação de acesso contribuem para promover a *accountability*.

Conclusão

Em sequência dos resultados apresentados, fica claro que a administração pública municipal do Brasil tem um longo caminho a percorrer no que se refere à disponibilização de informação, que engloba desde a estruturação da instituição, passando pela metodologia de gestão adotada até sua referida legislação, que conta com as diretrizes das políticas municipais de gestão de documentos e de arquivos e leis de implementação.

Todavia, é, desde já, possível concluir que a não institucionalização do serviço de informação, Arquivo Municipal, na maioria dos municípios brasileiros é um sinal da ausência de organização da informação e do conhecimento arquivístico e, deste modo, um fator entrópico na gestão eficaz e eficiente da informação. Os próprios arquivos, quando existentes, também não são eles fomentadores da transparência como forma de promoção do acesso à informação.

Conseqüentemente, os municípios brasileiros, não reconhecem o papel dos arquivos na promoção da gestão pública com responsabilidade ética e na prestação de contas com transparência às entidades reguladoras e aos cidadãos. Ou seja, ainda não se reconhece na maioria dos municípios brasileiros o papel dos Arquivos na governança pública, no reforço e na consolidação de uma cidadania ativa e, em última instância, da qualidade de vida dos cidadãos.

A partir dessa pesquisa, é possível perceber que os instrumentos de gestão de documentos não são considerados pelos municípios como fator determinante para propiciar o acesso à informação, tendo em vista que para além da não publicização de planos de classificação e tabelas de temporalidade, os municípios consideram outros fatores para divulgar a informações nos portais de transparência.

Para ampliar o debate sobre a gestão de documentos em contexto de transparência e como desdobramento dessa pesquisa, pretende-se futuramente investigar a gestão de documentos no âmbito dos portais de transparência dos municípios brasileiros, com o objetivo de verificar a relação do acesso à informação com a identificação dos documentos públicos. Este debate é essencial para promover a organização do conhecimento arquivístico para o acesso à informação e a prestação de contas em contexto de transparência, de modo a contribuir para qualidade da governança pública municipal comprometida com a Agenda 2030 do Desenvolvimento Sustentável.

Referências

- Baudrit Barquero, Álvaro. (1999). El derecho a la información. In: Actas del Seminario Internacional de Normatividad Archivística, 1999. Lima: Archivo General de la Nación, 1999.
- Brasil. Constituição da República Federativa do Brasil. Brasília: Senado Federal, 1988. Recuperado de: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 08 ago. 2012.
- Brasil. Lei n. 12.527, de 18 de novembro de 2011. Lei de Acesso à Informação. Recuperado de: <http://www.acaoainformacao.gov.br/acaoainformacao.gov/acao-informacao-brasil/legislacao-integra-completa.asp>. Acesso em 08 ago. 2012.

- Brasil. Lei n. 8.159, de 8 de janeiro de 1991. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Brasília, DF, 9 jan. 1991. Seção 1, p. 457.
- Brasil. Tribunal de Contas da União. Referencial básico de governança aplicável a órgãos e entidades da administração pública / Tribunal de Contas da União. Versão 2 – Brasília: TCU, Secretaria de Planejamento, Governança e Gestão, 2014. Recuperado de: https://portal.tcu.gov.br/data/files/30/C3/5F/CA/EBDEF610F5680BF6F18818A8/Referencial_basico_governanca.PDF. Acesso em: 13 abr. 2020.
- Carli, Deneide Teresinha de; Fachin, Gleise Regina Bóries. (2016). Lei de acesso à informação nos municípios do extremo oeste de santa catarina. *Em Questão*, v. 22, n. 1, p. 8-36, 2016. DOI: 10.19132/1808-5245221.8-36 Acesso em: 29 ago. 2021
- CONARQ. (2014). Conselho Nacional de Arquivos. Criação e desenvolvimento de Arquivos Públicos Municipais: transparência e acesso à informação para o exercício da cidadania. Conselho Nacional de Arquivos – Rio de Janeiro: Arquivo Nacional, 2014. Recuperado de: http://conarq.arquivonacional.gov.br/images/publicacoes_textos/Cartilha_criacao_arquivos_municipais.pdf. Acesso em: março 2020.
- IBGE. Instituto Brasileiro De Geografia e Estatística. Áreas Territoriais. 2021. Recuperado de: <https://www.ibge.gov.br/geociencias/organizacao-do-territorio/15761-areas-dos-municipios.html?=&t=o-que-e>. Acesso em: 07 mar. 2021.
- IFAC. (2001). International Federation of Accountants. Governance in the public sector: a governing body perspective. In: *International public sector study n.º 13*.
- Hjørland, B. *Knowledge organization*, 43, n. 6: 475-484, 2016. Recuperado de: http://www.isko.org/cyclo/knowledge_organization. Acesso em: 17/05/2020.
- Matias-Pereira, J. (2010). A governança corporativa aplicada no setor público brasileiro. *Administração Pública e Gestão Social*, 2(1), 109-134.
- Nações Unidas no Brasil. Roteiro para a Localização dos Objetivos de Desenvolvimento Sustentável: Implementação e Acompanhamento no nível subnacional. Recuperado de: <https://nacoesunidas.org/wp-content/uploads/2017/06/Roteiro-para-a-Localizacao-dos-ODS.pdf>. Acesso em: 10 mar. 2021.
- Pessanha, Diogo Paiva; Faben, Alexandre. (2020). Transparência na administração pública e o direito fundamental de acesso à informação em tempos de pandemia. In: SEMINÁRIO INTERNACIONAL SOBRE DIREITOS HUMANOS FUNDAMENTAIS, 6, 2020, Niterói. *Anais...* Niterói: UFF. p. 93-118.
- Rodrigues, Ana Célia. (2008). *Diplomática contemporânea como fundamento metodológico da identificação de tipologia documental em arquivos*. São Paulo: Universidade de São Paulo. Tese (Doutorado em História Social). (Mestrado em História Social). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, SP. Recuperado de: www.teses.usp.br/teses/.../8/.../TESE_ANA_CELIA_RODRIGUES.pdf. Acesso em: 26 set. 2010.
- Rodrigues, Ana Célia. (2013). Os Arquivos Públicos e o Direito à Informação In: *História, Arquivos & Mídias Digitais*. 1 ed. Fortaleza, CE: Expressão Gráfica e Editora. p. 13-30.
- Rodrigues, Ana Célia. (2019). Gestão de documentos e o acesso à informação pública nos arquivos municipais. In: REIS, Philippe Arthur dos; ALMEIDA, Renata Geraissati Castro de. (org). *Arquivos e história: a cidade de Campinas e seus documentos*. 1. ed. – São Paulo: Alameda, 2019. 19 – 40 p. ISBN 978-85-7939-632-8.
- Silva, Carlos Guardado da. (2018). Investigação Documental. In: *Manual de Investigação Qualitativa: Conceção, Análise e Aplicações*. Lisboa: PACTOR. p. 103-123.
- Tognoli, Natalia Bolfarini; Rodrigues, Ana Célia. (2018). O aporte teórico-metodológico da diplomática para a organização e representação do conhecimento arquivístico. In: *Produção, tratamento, disseminação e uso recursos informacionais heterogêneos* [recurso

eletrônico]; diálogos interdisciplinares. Maria Luiza de Almeida Campos [et al.] (Org.). – Niterói: IACS/UFF. 241p.

UNESCO. (2017). Acesso à informação é vital para construção de sociedades mais inclusivas. Publicado em 26/09/2017. Recuperado de: <https://nacoesunidas.org/unesco-acesso-a-informacao-e-vital-para-construcao-de-sociedades-mais-inclusivas/>. Acesso em: 10 mar. 2021.

MODELO-BASE DE VOCABULÁRIO CONTROLADO SOBRE VIOLÊNCIA DOMÉSTICA CONTRA CRIANÇAS E ADOLESCENTES

Rita Isabel Figueira Costa¹, Maria Cristina Vieira de Freitas²,
Daniel Martínez-Ávila³

¹Universidade de Coimbra, rita_figueira93@hotmail.com, ORCID iD 0000-0003-0829-6144
Centro de Estudos Interdisciplinares do Século XX, Faculdade de Letras,

²Universidade de Coimbra, cristina.freitas@fl.uc.pt, ORCID iD 0000-0002-8849-8792

³Universidad Complutense de Madrid, danima27@ucm.es,
ORCID iD 0000-0003-2236-553X

Resumo

O objetivo geral do presente trabalho é apresentar um modelo de um vocabulário controlado sobre a violência doméstica contra crianças e adolescentes, tomando como base as relações hierárquicas, associativas e de equivalência entre termos, descritas na ISO 25964-1:2011, bem como as garantias literária, cultural e ética, que subsidiam a sua elaboração. O estudo segue uma abordagem qualitativa e apresenta um caráter exploratório e descritivo. A estratégia metodológica adotada é o estudo de caso. O público-alvo são os/as profissionais especializados/as e pessoas interessadas no tema. Consultou-se o quadro teórico de referência sobre a violência doméstica contra crianças e adolescentes e realizou-se um estudo prévio sobre os vocabulários controlados e as relações semânticas entre os respetivos termos. Construiu-se um modelo-base de vocabulário controlado, de caráter incremental, justificando-se a seleção dos termos com argumentos que vão ao encontro das garantias literária, cultural e ética. Trata-se de um primeiro intento de construção de um instrumento que permite organizar a informação relativa ao tema selecionado e que serve o propósito de facilitar a indexação e a procura de informação sobre o assunto. Constata-se que o vocabulário controlado elaborado reflete os referenciais socioculturais e éticos que lhe são subjacentes, circunscrevendo-se assim aos contextos que oferecem os subsídios para a sua construção. Conclui-se que as garantias constituem uma base sólida sobre a qual se erguem as justificações para as escolhas dos termos a incluir em vocabulários controlados.

Palavras-chave: violência doméstica contra crianças e adolescentes, vocabulário controlado, garantia literária, garantia cultural, garantia ética.

Introdução

A violência doméstica contra crianças e adolescentes é um crime público presente nas diversas sociedades e/ou grupos. Evidentemente, surge como um tema com uma mais expressiva cobertura e um mais adequado enquadramento nos países e regimes democráticos e com um quadro jurídico-legal e social avançado. Tratando-se a violência doméstica contra crianças e adolescentes de um tema controverso, que representa um flagelo social com um registo histórico considerável, tem todo o interesse desenvolverem-se Sistemas de Organização do Conhecimento (SOC) que permitam dar uma resposta eficaz às necessidades de utilizadores/as interessados/as nessas matérias. A denúncia, que se afigura como um importante instrumento de combate a esse tipo de crime, frequentemente, é realizada por profissionais da área da educação e da saúde, que devem atuar na linha de frente da prevenção e estar despertados/as para comportamentos suspeitos e para danos físicos e psicológicos resultantes da agressão, de ocorrência, não raro, reiterada.

No campo informacional, o enfoque normalmente é colocado em processos que levam ao tratamento, armazenamento e recuperação. É, sobretudo, na fase de recuperação que os vocabulários controlados devem desempenhar um papel crucial, potenciando o acesso eficiente e eficaz aos conteúdos informacionais contribuindo, assim, para a minimização da desinformação nessas matérias. A inclusão e a exclusão de termos nos vocabulários controlados não são aleatórias, devendo reger-se por regras e por um sistema de “garantias”, sendo estas as justificativas que apoiam e legitimam o processo de seleção dos termos.

O conceito de garantia foi introduzido, em 1911, por Edward Wyndham Hulme, para explicar especificamente a **garantia literária**. Em 1986, Clare Beghtol conferiu um sentido mais genérico ao conceito de garantia (Barité, 2011, p. 3). No âmbito da Organização do Conhecimento (OC), uma garantia apresenta um estatuto de autoridade, sendo o elemento de referência adotado pelo/a indexador/a para determinar os termos a incluir num vocabulário controlado e para selecionar os descritores eleitos como representativos de determinado tema pela sua frequência de ocorrência no léxico usado por uma área.

A garantia literária, embora útil, impõe limites às justificações a usar para termos situados em diferentes contextos. No palco de duas culturas distintas, por exemplo, o próprio conceito de violência ganha sentidos igualmente distintos embora as “personagens” (leia-se: crianças e adolescentes) sejam as mesmas. Aos olhos europeus, o casamento infantil praticado em diversos países em África é visto como um ato de violência e abuso de menores. Segundo os padrões sociais europeus este é considerado um ato deplorável, porque é culturalmente inaceitável que se infrinjam os direitos e a liberdade das crianças. No entanto, no caso concreto de África, em alguns contextos, este mesmo ato é social e culturalmente aceite, porque está enraizado nos costumes de diversos dos seus povos, onde é encarado como algo “natural”. Depreende-se, deste exemplo, que pessoas pertencentes a diferentes culturas necessitam de informação igualmente diferenciada. É precisamente neste sentido e para se proceder a uma validação contextual dos descritores a usar num vocabulário controlado, que se apoia na **garantia cultural**. Lee, em 1976, fundou este conceito e abriu um leque de considerações relativamente à importância do contexto cultural em que um SOC é

criado, remetendo-nos, inevitavelmente, para uma reflexão sobre a capacidade de nos colocarmos no lugar do outro/a e de repensarmos o significado do mesmo termo sob outro ponto de vista (Barité, 2011, p. 6). Quando há aquilo a que se pode chamar de uma descentralização do eu e de um esquecimento da bagagem cultural que carregamos, ficamos numa posição legítima para focar a atenção num termo visto por outra perspectiva. Este exercício de aceitação da diversidade e de tolerância relativamente aos valores, às crenças e ao conhecimento intrínseco de um povo é o que Beghtol (2002) apelida de **hospitalidade cultural**. Neste sentido, pode afirmar-se que a garantia cultural abarca posições epistemológicas pragmáticas (Barité & Rauch, 2020).

A par com os sistemas anteriores de garantia literária e cultural está a **garantia ética**. A ética, mormente associada às normas, aos padrões e aos valores que regem o comportamento humano (*ethos*), é o que nos permite distinguir, por exemplo, entre o bem e o mal. No campo da Ciência da Informação, a garantia ética lança mão de um conjunto de considerações que protegem a informação de processos de desrespeito pela dignidade humana, de desestruturação de valores e de discriminação cultural. A título de exemplo, fere mais a suscetibilidade humana utilizar o termo simples “violação” do que o termo composto “abuso sexual”, quando se pretende referir o ato de violar sexualmente um/a menor. Dado que os SOC adquirem cada vez mais importância nos tempos que correm (Golub et al., 2019, p. 205), na sua elaboração, terá toda a pertinência seguir as orientações éticas disponíveis. Smith (2001, pp. 520-524), por exemplo, apresenta-nos o conceito de *global information justice* (GIJ), que se baseia na Declaração Universal dos Direitos Humanos (DUDH) da Organização das Nações Unidas (ONU). Beghtol (2002, p. 518) acrescenta que a justiça global da informação é um ideal a alcançar, referindo a hospitalidade cultural como um excelente regulador ético a adotar pelos/as profissionais de informação, de forma a procurarem a imparcialidade cultural e a disponibilização de informação globalizada (Beghtol, 2002, p. 519). Criar um sistema universal será certamente uma utopia. Não obstante, será útil e necessário propor um SOC que seja mais permeável a estas variáveis culturais e aos preceitos éticos, preservando as identidades dos povos que se pretende representar, tendo como referente o seu próprio contexto.

Partindo destes aspetos teóricos fundamentais envoltos no manto dos SOC, o principal objetivo do presente trabalho é elaborar um modelo-base de vocabulário controlado sobre violência doméstica contra crianças e adolescentes, tomando como referência: i) as relações hierárquicas, associativas e de equivalência entre termos, tal como descritas na ISO 25964-1:2011 e ii) a problemática do sistema de garantias antes apresentado.

Metodologia

Dado o seu desenho microanalítico, procedimental e faseado, bem como a sua circunscrição a um contexto específico, sem a pretensão de atingir alguma generalização de resultados, mas com a intenção de responder a uma situação particular, este estudo segue os preceitos do método qualitativo e adota como estratégia o estudo de caso, que assume uma natureza exploratória e descritiva, no que tange ao processo, e instrumental e incremental, no que respeita aos resultados.

A fase inicial do processo de construção de um vocabulário controlado inclui a pesquisa detalhada de um quadro de referência teórico que possa ser o veículo para o levantamento de conceitos representativos do tema que se pretende modelar. A seleção dos termos, por sua vez, segue a garantia literária, especialmente nos seus desenvolvimentos mais recentes (Barité, 2018, 2019). No caso em apreço, partiremos de uma revisão aprofundada do conteúdo dos documentos que refletem com clareza o tema da violência doméstica contra crianças e adolescentes, com vista à identificação e à análise dos termos utilizados com maior frequência nessa literatura de referência (ANSI/NISO z39-19-2005, R2010, p. 16). Na escolha dos descritores, teremos em consideração os termos mais frequentemente utilizados no site da Associação Portuguesa de Apoio à Vítima (APAV), da Direção Geral da Saúde (DGS), do Centro de Estudos Judiciários (CEJ) e da *United Nations International Children's Emergency Fund* (UNICEF), por considerarmos que nos oferecem as garantias necessárias.

Tratando-se a violência doméstica contra crianças e adolescentes de um tema sensível, são diversos os conceitos com uma conotação negativa que surgem e que são expressos na literatura sobre o tema. Na construção deste modelo-base de vocabulário controlado, procuraremos não utilizar eufemismos e, por isso, a seleção dos termos recairá sobre aqueles que demonstram, sem filtros, a crueza do próprio tema e, desta forma, chamam a atenção e sensibilizam o público-alvo para este problema persistente nas sociedades atuais.

Assim, para a construção do vocabulário controlado realizamos primeiramente uma revisão de literatura, que nos auxiliou a integrar o tema no respetivo domínio de ocorrência. Seguidamente, analisamos e extraímos, de vários documentos inseridos no domínio da violência contra crianças e adolescentes selecionados e consultados para o efeito, os termos representativos, sendo estes posteriormente incluídos no vocabulário apresentado.

Durante o processo de escolha dos termos a incluir no vocabulário, tivemos em consideração a recomendação de disponibilizar a informação adequada às necessidades do público-alvo (ANSI/NISO z39-19-2005, R2010, p. 16). No caso em apreço, o público-alvo serão os/as profissionais especializados/as, sejam assistentes sociais, professores/as, bibliotecários/as, psicólogos/as, etc. que tenham interesse sobre o tema da violência doméstica contra crianças e adolescentes.

Como produtos do estudo, são apresentados: a proposta de um modelo-base de vocabulário controlado sobre o tema, no qual a seleção da terminologia a incluir foi legitimada e justificada pelas garantias literária, cultural e ética, apresentadas e discutidas, em breves linhas, na introdução deste artigo.

Resultados

Considerações gerais acerca do modelo-base de vocabulário controlado

As consultas realizadas aos documentos técnicos e científicos situados no domínio do vocabulário em causa revelaram que, na modelagem destas realidades, são mais facilmente encontráveis as listas de termos e os glossários do que os vocabulários controlados propriamente ditos, uma vez que a sua constituição exige o estabe-

lecimento de uma rede de relações que não apenas os diferenciam de outros vocabulários, como também funcionam como elementos de ligação entre os seus termos, de modo a proporcionar, efetivamente, a potenciação da recuperação da informação, sendo esta a pretensão subjacente à sua elaboração. O vocabulário controlado elaborado neste estudo e apresentado no Apêndice I deste artigo, parte assim da constatação de uma lacuna. Trata-se, pois, de um primeiro intento não exaustivo de constituição de um SOC utilizável no domínio da violência doméstica praticada contra crianças e adolescentes, com o propósito de colmatar esta mesma lacuna. É apresentado na forma e na estrutura recomendadas, dispondo-se os termos que o compõem em ordem alfabética e expondo-se os três tipos de relações necessárias ao controlo terminológico, como sejam: hierárquicas, associativas e de equivalência.

No que respeita à composição, foram incluídos termos descritores, termos não descritores, termos genéricos e termos específicos, seguidos das respetivas relações (Apêndice I). No processo de seleção dos termos descritores e não descritores foram dadas as devidas justificações, com base no sistema de garantias abordado no ponto introdutório deste estudo.

Seleção de termos descritores, não descritores e sistema de garantias

Para cumprir os objetivos deste artigo, selecionamos um conjunto de 16 termos (simples ou compostos) identificados como descritores e de 64 termos (simples ou compostos) identificados como não descritores, que apresentaremos, nas linhas a seguir, para tipificar a forma como foram encaminhadas as nossas reflexões sobre o uso do sistema de garantias na constituição do modelo-base pretendido.

Assim, começemos pelo termo **abuso**, que foi selecionado como descritor, porque cumpre a garantia literária (Costa 2021, p. 59), na medida em que possui uma forte expressão e, conseqüentemente, representatividade, na literatura da área, para além da sua evidente abrangência terminológica. De acordo com o Dicionário online de Língua Portuguesa da Porto Editora (2021), as definições de abuso são: “mau uso”; “uso excessivo ou injusto”; “ultraje ao pudor”; “insulto”. A estas definições, acresce o facto de que o abuso é um comportamento que ultrapassa, em muitos aspetos, o limite do que seria racionalmente, moralmente e culturalmente aceitável, sendo assim considerado porque, de facto, a ele subjaz uma moral que dita os seus contornos. Sob esta perspetiva, a seleção desse termo (abuso) incorpora justificações que também vão ao encontro das garantias cultural e ética. Quanto aos termos não descritores associados a este termo, verificamos que representam comportamentos associados a diferentes tipos de abuso: psicológico (injúria, insulto, provocação, depravação); físico (ataque); por omissão ou desleixo (negligência); sexual (violação, pedofilia).

Os descritores **adolescentes** e **crianças** cumprem, sobretudo, a garantia literária (Costa 2021, p. 60), sendo, por isto, os mais comumente utilizados, no quadro de referência consultado, para representar as faixas inferiores da pirâmide etária, sobre as quais se debruça o modelo aqui proposto. Os não descritores selecionados, tendo como base a mesma garantia, foram, neste caso: bebês e infantes (descritor: crianças); jovens e puberdade (descritor: adolescentes). É de notar, no entanto, que esses

conceitos podem ser diferentemente representados em diferentes culturas e/ou sociedades, sendo variáveis, inclusivamente, em função da época. Assim, a garantia cultural é também usada, neste caso, como justificção para o uso do termo adolescente, por exemplo, no atual contexto das sociedades capitalistas ocidentais consideradas economicamente desenvolvidas, nas quais essa proposta de modelo de vocabulário se insere.

O descritor **apoio** tem particular importância no contexto do presente estudo, na medida em que desempenha um papel fundamental na resposta social à violência doméstica. Não por acaso, uma das fontes de garantia literária (Costa 2021, p. 61) aqui usadas é a APAV. Este descritor também é orientado pelas garantias cultural e ética, na medida em que se verifica que as pessoas pertencentes a uma determinada cultura tendem a entreajudarem-se, por um lado, e a recriminarem aquelas que não o fazem, por outro, ou seja, por trás do apoio há um controle social exercido em nome supostamente do Bem de todo o grupo. Os não descritores selecionados para este descritor, neste caso, são: ajuda, amparo, assistência, proteção e suporte.

O descritor **crença**, conforme constatamos, pode gerar alguma ambiguidade. A crença é algo em que se acredita e está fortemente impregnada das vivências individuais. No enquadramento da violência doméstica, a crença é a convicção que leva à legitimação de atitudes violentas. Para além de este ser um conceito com uma ampla representação na literatura da área (Costa 2021, p. 62), é também dependente de um contexto, dado que é no seio de uma determinada cultura que se desenvolvem as perceções culturais e éticas do que se pode considerar como um ato violento. Há, portanto, uma predisposição cultural para a aceitação ou não da violência. Neste caso, o sistema de garantias (literária, cultural e ética) justifica, em conjunto, a escolha deste descritor. Os não descritores correspondentes são: conduta, confiança, dogma e fé.

O descritor **ambiente doméstico** remete-nos não apenas a um espaço, mas também a um conjunto de pessoas próximas à vítima, que gravitam no seu entorno, onde, normalmente, acontecem os casos de violência doméstica. Consideramos que este termo cumpre, sobretudo, a garantia literária (Costa 2021, p. 62). Os não descritores selecionados para este descritor são: agregado, familiar, intimidade e progenitores.

Os descritores **medo**, **vergonha**, **tristeza** e **isolamento** referem-se a sentimentos. A violência doméstica desperta na vítima uma série de sentimentos, sendo estes os mais frequentemente referidos na literatura de referência consultada (Costa 2021, p. 64), justificando-se, deste modo, a sua seleção a partir dos critérios inerentes à garantia literária. Os termos não descritores selecionados são: ameaça, intimidação, pânico, perigo, temor e terror (descritor: medo); constrangimento e descrédito (descritor: vergonha); desânimo e infelicidade (descritor: tristeza); abandono, desaconchego, desamparo e solidão (descritor: isolamento). No nosso entender, os sentimentos são pessoais e emocionais, mas seguem um código interpretacional incutido pela sociedade. Por exemplo, a vítima sente vergonha porque existe uma sociedade que a pode sancionar e criticar com base em valores morais. Os sentimentos formam-se no seio de um grupo social, por conseguinte, transferem-se entre elementos desse mesmo grupo, através de uma aprendizagem cultural. As pessoas tendem, assim, a perceberem e sentir a realidade, pela ótica do grupo a que pertencem. Neste

sentido, os termos descritores supracitados seguem também as garantias cultural e ética.

O descritor **perturbação** está vinculado a um significado médico. O mesmo Dicionário online de Língua Portuguesa da Porto Editora (2021) o define como: “desarranjo que ocorre numa função orgânica ou psíquica”. Qualquer ato de violência origina uma perturbação na vítima. Os tipos de perturbações que podem surgir estão relacionados com o ambiente sociocultural a que pertence a pessoa que sofre o ato de violência e também à forma particular como essa pessoa reage à perturbação que lhe é causada. Porquanto, há todo um processo de gestão das emoções que se molda ao referencial cultural que é inculcado na vítima do ato de violência. Para além disso, o termo perturbação é frequentemente utilizado na literatura da área (Costa 2021, p. 65), o que lhe confere uma garantia literária, para além da garantia cultural e ética que lhe é subjacente. Os termos não descritores selecionados para o descritor perturbação, neste caso, são: ansia e crise.

O descritor **violência**, fulcral para o vocabulário proposto, remete para um ato que fere a integridade física ou psicológica de alguém, desrespeitando-se os seus direitos enquanto pessoa. Relativamente aos termos que consideramos não descritores, verificamos que o seu uso é também frequente entre autores/as da área, corroborando-se, assim, o emprego da garantia literária (Costa 2021, p. 66). Como anteriormente já foi referido, a violência é um conceito que ganha diferentes interpretações, de acordo com o referencial cultural e ético que está enraizado nas pessoas. No entanto, considera-se que a violência acontece quando se ultrapassa a ténue linha que demarca os direitos de uma pessoa face aos direitos de outra pessoa, causando danos. Esta particularidade leva a que o descritor violência seja escolhido também em função das garantias cultural e ética. Neste caso, os termos não descritores correspondentes ao descritor violência são: agressão, controlo, crueldade, empurrões, força, hostilidade, maldade e ofensa.

O descritor **manipulação** tem também grande relevância no contexto em causa, porque representa um ato que pode encobrir a violência, através da mistificação da própria vítima ou de outras testemunhas do ato. Manipular é obrigar alguém, de forma subtil, sem que esta pessoa se aperceba, a fazer, dizer, ou pensar algo contrário ao que quer. Segundo o Dicionário online de Língua Portuguesa da Porto Editora (2021), a palavra manipulação significa, num sentido figurado: “influência nefasta e sub-reptícia exercida sobre um grupo ou um indivíduo, visando predispor-lo a determinado fim”. Este tipo de influência negativa é transversal às diferentes culturas e, obviamente, transforma a liberdade de ser e de estar. Assim, a seleção deste termo como descritor, no presente modelo, vai ao encontro das determinações das garantias literária (Costa 2021, p. 67), cultural e ética. Os termos não descritores selecionados, neste caso, são: chantagem, coação, inibição, opressão, persuasão, privação, proibição e tirania.

No seguimento do que foi dito relativamente ao descritor anterior (manipulação), selecionamos **dependência** como termo descritor, porque, para além da sua representatividade na literatura da área (garantia literária), é sabido que quando uma pessoa depende de outra torna-se mais fácil deixar-se manipular. É o que aponta a literatura que nos serve como quadro de referência. No caso concreto das crianças, elas dependem de uma pessoa adulta – o/a responsável – não possuindo autonomia.

Esta dependência pode ser propícia a silêncios que são convenientes para o/a agressor/a. No entanto, a questão da dependência e dos silêncios encontra-se sob o manto cultural e ético, porque há povos que, contrariamente a outros, não apenas fomentam, mas também reconhecem, desde cedo, o desenvolvimento da autonomia, incluindo nas crianças o valor da liberdade de expressão. Os termos escolhidos como não descritores, neste caso, são: obediência, subordinação e sujeição. Pelas justificações anteriores, somos a concluir que este descritor (dependência), tal como os anteriores, no vocabulário pretendido, deve seguir também as diretrizes das garantias cultural e ética.

O descritor **delinquência**, por sua vez, aponta para um comportamento que é comum na sociedade atual, que tanto pode ser um fator motivador, uma causa, como pode ser uma consequência da concretização de atos violentos. O Dicionário online de Língua Portuguesa da Porto Editora (2021) sugere as seguintes definições para este termo: “acto de cometer delitos, desobedecer a leis ou a padrões morais” e “conjunto de infrações penais cometidas numa sociedade, durante determinado período”. A delinquência, conforme sabemos, é influenciada por múltiplos fatores socioculturais e analisada à luz dos princípios morais e éticos. Neste sentido, trata-se de um descritor que cumpre as garantias literária (Costa 2021, p. 69), cultural e ética. Os termos não descritores, neste caso, são: criminalidade e marginalidade.

Finalmente, constata-se que a violência doméstica é um fenómeno transversal às diversas classes sociais, graus académicos, faixas etárias, etnias, localização geográfica, género, sexo, etc. Portanto, é legítimo ter uma preocupação acrescida no uso de termos que abarquem essas diferenças, o que justifica, por exemplo, a seleção do termo descritor **hematomas**, usado no jargão científico com frequência pelos/as profissionais de saúde, que são um dos públicos-alvo prováveis de um vocabulário controlado sobre violência doméstica. Este descritor cumpre, essencialmente, a garantia literária (Costa 2021, p. 69). Os não descritores selecionados, neste caso, a partir de expressões dominantes no senso comum, são: nódoas negras e pisaduras.

Justificações para as relações entre os termos

As relações hierárquicas são estabelecidas entre descritores. No modelo pretendido, consideramos perseguição, violência emocional, violência sexual, violência social e violência financeira como sendo descritores do tesouro construído.

Os termos específicos estão subordinados aos termos genéricos. No modelo pretendido, todos os termos genéricos selecionados são antecedidos pelo operador TG. De igual modo, todos os termos específicos são antecedidos pelo operador TE.

A violência, que é um termo genérico, no modelo proposto, pode ser subdividida nos seguintes termos específicos: violência emocional, violência financeira, violência física, violência sexual, violência social e perseguição (APAV, 2012).

A violência emocional é um termo genérico, ao qual correspondem os termos específicos abuso, manipulação, medo e perturbação (Costa 2021, p. 70). A violência física é um termo genérico, ao qual corresponde o termo específico abuso (Costa 2021, p. 71). A violência financeira é um termo genérico, ao qual corresponde o termo específico dependência (Costa 2021, p. 72). A violência sexual é um termo

genérico, ao qual corresponde o termo específico abuso (Costa 2021, p. 72). A perseguição é um termo genérico, ao qual corresponde o termo específico medo (Costa 2021, p. 73).

A violência social é “qualquer tipo de violência cometida por indivíduos ou pela comunidade, com uma finalidade social” (Tremblay, 2012, p. 1). Lima et al. (2006, p. 66) explicam a relação (TR) entre a violência emocional e a violência social: “Certamente que as formas de violência psicológica aqui apresentadas apenas tangenciam o universo relacional de muitas famílias e escolas que socializam crianças e adolescentes em ambientes marcados pelo desrespeito à individualidade de cada um. Pode-se perceber o abuso psicológico como importante potencializador da violência social. Pessoas que crescem em ambientes onde não há respeito à própria individualidade estão mais facilmente em situação de repetirem relações desrespeitosas na comunidade em que vivem”.

Consideramos também que o descritor apoio se relaciona com os descritores violência emocional e violência social na medida em que o apoio por parte das instituições sociais pode ajudar a combater este tipo de violência.

A delinquência relaciona-se com a violência física e o medo, na medida em que: “Vários autores vêm associando (obviamente de forma não-linear) punição física com delinquência, por causa de sentimentos ambíguos, confusos ou vingativos que ela gera, como angústia, raiva, ansiedade, medo, terror, ódio e hostilidade” (Lima et al., 2006, p. 34).

O descritor hematomas representa uma lesão consequente da violência física portanto estes dois descritores também estão ligados associativamente.

O estudo realizado por Garbin et. al (2012, p. 115) permitiu compreender as relações entre violência física, emocional e sexual: “No mesmo sentido, os achados deste estudo permitem afirmar que a violência física esteve significativamente associada à emocional, à sexual e à negligência emocional. Os casos de abuso sexual relacionaram-se também à violência emocional. Da mesma forma, violência e negligência emocionais.”

Uma criança que seja vítima de violência física, emocional ou sexual desenvolverá, certamente, um sentimento de tristeza e tenderá a isolar-se das outras pessoas. Concluímos que é legítimo relacionar associativamente os descritores violência física, violência emocional, violência sexual, tristeza e isolamento.

O termo casamento infantil relaciona-se com os termos abuso, adolescentes, ambiente doméstico, crianças, dependência, crença e tristeza. O casamento infantil é a união formal ou informal, na qual um dos elementos do casal é menor que 18 anos. Trata-se de um abuso dos Direitos Humanos, na medida em que a criança fica privada da infância e perde a sua liberdade (Nour, 2009, p. 51). Fala-se de violência estrutural, quando nos referimos às desigualdades, sobretudo econômicas e políticas, existentes na estrutura da sociedade e que têm impacto na vida das crianças e adolescentes em todo o mundo (Kent, 2016, p. 24).

Conclusões

Partindo dos conceitos de garantia literária, cultural e ética, presentes na literatura científica da área de OC, e das discussões que lhe são inerentes, e após a aplicação dos critérios de recolha, análise e seleção de termos a utilizar, tendo como base o quadro de referência do domínio a modelar (violência doméstica contra crianças e adolescentes), constata-se que o modelo-base de vocabulário controlado elaborado no âmbito deste estudo reflete também os referenciais socioculturais e éticos que lhe são subjacentes, circunscrevendo-se assim minimamente aos contextos que oferecem os subsídios para a sua construção. Conclui-se, igualmente, que as garantias constituem uma base sólida sobre a qual se erguem as justificações para as escolhas dos termos a incluir em vocabulários controlados.

Assim, espera-se que o modelo-base de vocabulário controlado construído no presente estudo possa contribuir, de algum modo, para organizar a informação relativa ao tema da violência doméstica contra crianças e adolescentes. Desta forma, surge como um primeiro intento que poderá ser útil para facilitar a indexação e a procura de informação sobre o tema, tratando-se de um esboço inicial do árduo trabalho de construção, por exemplo, de um tesouro, considerado um vocabulário controlado completo, por excelência.

A justificação para a sua elaboração prendeu-se com a necessidade de colmatar uma lacuna existente, uma vez que as buscas realizadas no âmbito desta investigação demonstraram haver uma ausência de vocabulários controlados especificamente virados para o tema da violência doméstica contra crianças e adolescentes.

Quanto aos resultados apresentados, como a própria definição de modelo indica, trata-se de um ensaio de um vocabulário controlado de pequena dimensão, que pretende servir como uma primeira referência para investigações futuras.

Assim sendo, e pela forma como foi construído, o presente vocabulário poderá ser complementado com novos estudos que alarguem o seu espetro inicial, propiciando a seleção e a inclusão de novos termos, bem como das relações semânticas correspondentes, aos já existentes.

Finalmente, reconhece-se que a maior limitação do presente estudo está no facto de o vocabulário apresentado estar composto, ainda, por um número reduzido de termos associados ao domínio que pretende modelar.

Referências

- ANSI/NISO Z 39.19-2005 (R2010). *Guidelines for the construction, format, and management of monolingual controlled vocabularies*. NISO Press
- Associação Portuguesa De Apoio À Vítima [APAV] (2012). *Violência Doméstica*. <https://apav.pt/vd/index.php/features2>
- Barité, M. (2011). La garantía cultural como justificación en sistemas de organización del conocimiento: aproximación crítica. *Palabra Clave (La Plata)*, 1(1), [2]-11
- Barité, M. (2018). Literary Warrant. *Knowledge Organization*, 45(6), 517-536
- Barité, M. (2019). Toward a General Conception of Warrants: First Notes. *Knowledge Organization*, 46(8), 647-55

- Barité, M., & Rauch, M. (2020). Cultural Warrant Old and New Sights from Knowledge Organization. In M. Lykke, T. Svarre, M. Skov, & D. Martínez-Ávila (Eds.), *Knowledge Organization at the Interface* (pp. 31-40). Ergon Verlag
- Beghtol, C. (2002). A proposed ethical warrant for global knowledge representation and organization systems. *Journal of Documentation*, 58(5), 507-532
- Costa, R.I.F. (2021). *Vocabulário controlado e relações semânticas sobre violência doméstica contra crianças e adolescentes: construção de um modelo* [Dissertação de mestrado, Universidade de Coimbra]
- Garbin, C., Queiroz, A., Rovida T., & Saliba, O. (2012) A violência familiar sofrida na infância: uma investigação com adolescentes. *Psicologia em Revista*, 18(1), 107-118
- Golub, K., Schmiede, R., Tudhope, D. (2019). Recent applications of Knowledge Organization Systems: introduction to a special issue. *International Journal on Digital Libraries*, 20, 205-207
- ISO 25964-1. (2011). *Information and documentation – thesauri and interoperability with other vocabularies – part 1: thesauri for information retrieval*. ISO
- Kent, G. (2016). *Ending Hunger Worldwide*. Routledge
- Lima, C., et al. (2006). *Violência faz mal à saúde*. Editora do Ministério da Saúde. http://bvsmis.saude.gov.br/bvs/publicacoes/violencia_faz_mal.pdf
- Nour, N. (2009). Child Marriage: A Silent Health and Human Rights Issue. *Reviews in Obstetrics & Gynecology*, 2(1), 51-56
- Porto Editora. (2021). *Infopedia: Dicionários Porto Editora*. <https://www.infopedia.pt/dicionarios/lingua-portuguesa/abuso>
- Smith, M. (2001) Global Information Justice: Rights, Responsibilities, and Caring Connections. *Library Trends*, 49(3), 519-537
- Tremblay, R. E. (2012) *Enciclopédia sobre o Desenvolvimento na Primeira Infância*. <http://www.encyclopedia-crianca.com/sites/default/files/syntheses/pt-pt/2676/violencia-social-sintese.pdf>

APÊNDICE I – VOCABULÁRIO CONTROLADO SOBRE VIOLÊNCIA DOMÉSTICA CONTRA CRIANÇAS E ADOLESCENTES

Abandono

USE: Isolamento

Abuso

TG: Violência emocional

TG: Violência sexual

TR: Adolescentes

TR: Ambiente Doméstico

TR: Casamento infantil

TR: Crença

TR: Crianças

TR: Dependência

TR: Tristeza

UP: Ataque

UP: Depravação

UP: Injúria

UP: Insulto

UP: Negligência

UP: Pedofilia

UP: Provocação

UP: Violação

Adolescentes

TR: Abuso

TR: Ambiente Doméstico

TR: Casamento infantil

TR: Crença

TR: Crianças

TR: Dependência

TR: Tristeza

TR: Violência estrutural

UP: Jovens

UP: Puberdade

Agregado

USE: Ambiente doméstico

Agressão

USE: Violência

Ajuda

USE: Apoio

Ambiente Doméstico

TR: Abuso

TR: Adolescentes

TR: Casamento infantil

TR: Crença

TR: Crianças

TR: Dependência

TR: Tristeza

UP: Agregado

UP: Familiar

UP: Intimidade

UP: Progenitores

Ameaça

USE: Medo

Amparo

USE: Apoio

Ânsia

USE: Perturbação

Apoio

TR: Violência emocional

TR: Violência social

UP: Ajuda

UP: Amparo

UP: Assistência

UP: Proteção

UP: Suporte

Assistência

USE: Apoio

Ataque

USE: Abuso

Bebés

USE: Crianças

Casamento Infantil

TR: Abuso

TR: Adolescentes

TR: Ambiente doméstico

TR: Crença

TR: Crianças

TR: Dependência

TR: Tristeza

Chantagem

USE: Manipulação

Coação

USE: Manipulação

Conduta

USE: Crença

Confiança

USE: Crença

Constrangimento

USE: Vergonha

Controlo

USE: Violência

Crença

TR: Abuso

TR: Adolescentes

TR: Ambiente Doméstico

TR: Casamento infantil

TR: Crianças

TR: Dependência

TR: Tristeza

UP: Conduta

UP: Confiança

UP: Dogma

UP: Fé

Crianças

TR: Abuso

TR: Adolescentes

TR: Ambiente Doméstico

TR: Casamento Infantil

TR: Crença

TR: Dependência

TR: Tristeza

TR: Violência estrutural

UP: Bebés

UP: Infantes

Criminalidade

USE: Delinquência

Crise

USE: Perturbação

Crueldade

USE: Violência

Delinquência

TR: Hematomas

TR: Medo

TR: Violência física

UP: Criminalidade

UP: Marginalidade

Dependência

TG: Violência financeira

TR: Abuso

TR: Adolescentes

TR: Ambiente Doméstico

TR: Casamento infantil

TR: Crença

TR: Crianças

TR: Tristeza

UP: Obediência

UP: Subordinação

UP: Sujeição

Depravação

USE: Abuso

Desaconchego

USE: Isolamento

Desânimo

USE: Tristeza

Desamparo

USE: Isolamento

Descrédito

USE: Vergonha

Dogma

USE: Crença

Empurrão

USE: Violência

Familiar

USE: Ambiente doméstico

Fé

USE: Crença

Força

USE: Violência

Hematomas

TR: Delinquência
TR: Medo
TR: Violência física
UP: Nódos negras
UP: Pisaduras

Hostilidade

USE: violência

Infantes

USE: Crianças

Infelicidade

USE: Tristeza

Inibição

USE: Manipulação

Injúria

USE: Abuso

Insulto

USE: Abuso

Intimidação

USE: Medo

Intimidade

USE: Ambiente doméstico

Isolamento

TR: Tristeza
TR: Violência emocional
TR: Violência física
TR: Violência sexual
UP: Abandono
UP: Desaconchego
UP: Desamparo
UP: Solidão

Jovens

USE: Adolescentes

Maldade

USE: Violência

Manipulação

TG: Violência emocional
UP: Chantagem

UP: Coação
 UP: Inibição
 UP: Opressão
 UP: Persuasão
 UP: Privação
 UP: Proibição
 UP: Tirania

Marginalidade

USE: Delinquência

Medo

TG: Perseguição
 TG: Violência emocional
 TR: Delinquência
 TR: Hematomas
 TR: Violência física
 UP: Ameaça
 UP: Intimidação
 UP: Pânico
 UP: Perigo
 UP: Temor
 UP: Terror

Negligência

USE: Abuso

Nódoas Negras

USE: Hematomas

Obediência

USE: Dependência

Ofensa

USE: Violência

Opressão

USE: Manipulação

Pânico

USE: Medo

Pedofilia

USE: Abuso

Perigo

USE: Medo

Perseguição

TE: Medo

Persuasão

USE: Manipulação

Perturbação

TG: Violência emocional

UP: Ânsia

UP: Crise

Pisaduras

USE: Hematomas

Privação

USE: Manipulação

Progenitores

USE: Ambiente doméstico

Proibição

USE: Manipulação

Proteção

USE: Apoio

Provocação

USE: Abuso

Puberdade

USE: Adolescente

Solidão

USE: Isolamento

Subordinação

USE: Dependência

Sujeição

USE: Dependência

Suporte

USE: Apoio

Temor

USE: Medo

Terror

USE: Medo

Tirania

USE: Manipulação

Tristeza

- TR: Abuso
- TR: Adolescentes
- TR: Ambiente doméstico
- TR: Casamento infantil
- TR: Crença
- TR: Crianças
- TR: Dependência
- TR: Isolamento
- TR: Violência emocional
- TR: Violência física
- TR: Violência sexual
- UP: Desânimo
- UP: Infelicidade

Vergonha

- TG: Violência social
- UP: Constrangimento
- UP: Descrédito

Violação

- USE: Abuso

Violência

- TE: Perseguição
- TE: Violência emocional
- TE: Violência financeira
- TE: Violência física
- TE: Violência sexual
- TE: Violência social
- UP: Agressão
- UP: Controle
- UP: Crueldade
- UP: Empurrões
- UP: Força
- UP: Hostilidade
- UP: Maldade
- UP: Ofensa

Violência emocional

- TG: Violência
- TR: Apoio
- TR: Isolamento
- TR: Tristeza
- TR: Violência física
- TR: Violência sexual
- TR: Violência social
- TE: Abuso
- TE: Manipulação

TE: Medo
TE: Perturbação

Violência estrutural
TR: Adolescentes
TR: Crianças

Violência financeira
TG: Violência
TE: Dependência

Violência física
TG: Violência
TE: Abuso
TR: Delinquência
TR: Hematomas
TR: Isolamento
TR: Medo
TR: Tristeza
TR: Violência emocional
TR: Violência sexual

Violência sexual
TG: Violência
TE: Abuso
TR: Isolamento
TR: Tristeza
TR: Violência emocional
TR: Violência física

Violência social
TG: Violência
TR: Apoio
TR: Violência emocional

SISTEMA DE ORGANIZAÇÃO DO CONHECIMENTO E PÓS-DISCIPLINARIDADE: MUDANÇA CLIMÁTICA NO CONTEXTO DA AGENDA 2030

Marcos Gonçalves Ramos¹, Priscila Ramos Carvalho²,
Rosali Fernandez de Souza³

¹Doutorando PPGCI IBICT/UFRJ, marcosgon54@hotmail.com,
ORCID iD 0000-0001-5807-0042

²Doutoranda PPGCI IBICT/UFRJ, carvalho.priscila@gmail.com,
ORCID iD 0000-0003-3042-1669

³Pesquisadora do IBICT e docente PPGCI IBICT/UFRJ,
rosali@ibict.br, ORCID iD 0000-0002-0890-7999

Resumo

O presente trabalho se inspirou nas abordagens de Beghtol, Buckland, Kleinberg e Hjørland, para analisar a representação do conhecimento sobre mudança climática e COVID-19, no contexto da Agenda 2030. A pesquisa buscou identificar o núcleo central do formato CERIF (Common European Research Information Format) alinhado a dois princípios FAIR (Encontrável e Acessível) na recuperação dos termos “*climate change* e COVID-19” nas bases de dados Scopus, Web of Science e Dimensions, no período 2010-2020. As abordagens 7 e 8 da análise de domínio de Hjørland juntamente com as ferramentas VOSviewer e CitNetexplorer foram utilizadas para analisar uma amostra de termos recuperados nas bases. Os dados provenientes do Conselho Nacional de Desenvolvimento Científico e Tecnológico do Brasil referentes à Plataforma Lattes e ao Diretório de Grupos de Pesquisa, coletados em novembro de 2020, foram acrescentados às análises. O resultado das bases de dados internacionais foi: 378 documentos na Scopus, 423 na Web of Science e 775 na Dimensions. O resultado da Plataforma Lattes recuperou o total de 14.448 currículos pelo termo “COVID-19”; 5.064 currículos pelo termo de busca “COVID-19 AND meio ambiente”; 679 currículos pelo termo de busca “COVID-19 AND mudança climática”. O resultado no Diretório de Grupos de Pesquisa pelo termo de busca “COVID-19” foi 233 grupos; para o termo “COVID-19 AND meio ambiente” foi apenas um grupo, e para o termo “COVID-19 AND mudança climática” não foi encontrado nenhum grupo. As conclusões sinalizam que mudança climática e COVID-19 podem ser considerados domínios emergentes que sugerem um possível modelo pós-disciplinar na relação entre projetos de pesquisa, pessoas e organizações.

Palavras-chave: Mudança climática, COVID-19, Análise de domínio, Pós-disciplinaridade, Complexidade.

Introdução

O encontro dos estudos sobre ecologia, climatologia e sustentabilidade indicam aspectos técnicos de mensuração e análise de dados provenientes de diversas fontes e reforçam tanto a importância dos sistemas de recuperação, análise e armazenamento de dados quanto às ações de comunicação e interação entre governos, pesquisadores e o público em geral. A Agenda 2030 é composta por 193 países signatários da declaração para o cumprimento de 17 Objetivos de Desenvolvimento Sustentável (ODS) compreendendo cerca de 169 metas e 231 indicadores macroeconômicos e sociais.

A proposta da Agenda 2030 demonstra o reconhecimento de novos arranjos na organização do conhecimento baseado na ação coletiva e integrada entre pessoas, instituições e órgãos governamentais. ODS 14 visa o gerenciamento holístico para tomada de decisões urgentes diante da situação de risco global da mudança climática. ODS 15 reforça a necessidade do compromisso entre povos, nações e organizações para proteger, recuperar e promover o uso sustentável dos ecossistemas terrestres de modo a reverter a degradação da perda de biodiversidade.

Os 17 ODS apontam para o conhecimento de base sustentável que admite uma arquitetura sócio-técnica complexa de comunicação e informação que viabilize a construção de conhecimentos coletivos através de regimes de trocas e colaboração. Nesse contexto, os Sistemas de Organização de Conhecimentos (SOC) devem buscar a sustentabilidade informacional (Barbalho, 2021) e formatos de comunicação que facilitem o diálogo entre agentes individuais, institucionais e governamentais.

A emergência sanitária mundial causada pela pandemia do vírus SARS COV-2, responsável pela doença COVID-19, e os efeitos da reversão do clima sobre o planeta suscitam encontros pós-disciplinares das áreas de conhecimento como Ciências Humanas, Ciências Sociais, Ciências da Saúde e Ciências Ambientais. Os sistemas de informação de pesquisa nessas áreas apresentam o duplo desafio seja no cruzamento de dados em termos da variabilidade em suas diferentes escalas, seja na velocidade da repercussão dos fenômenos do clima e seus efeitos no nível do regional ao global.

O último painel intergovernamental de mudança climática atesta a ação antropocêntrica da atividade humana pela alta taxa de concentração dos gases provenientes da atividade industrial na atmosfera e que incidem sobre o aumento da temperatura e no comportamento do clima (IPCC, 2021). Treze anos antes, Giddens (2008) levantou a importância da integração entre governo, entidades de pesquisa e indústria através de uma plataforma pública para discutir os aspectos políticos e econômicos, pois considerava o debate sobre mudança climática segmentado por setores especializados e sem um plano definido de governança.

Os resultados do IPCC e da proposta de Giddens (2008) se concretizaram na participação de pesquisadores de diversos países que colaboraram através das plataformas de dados de pesquisa para a modelagem específicas sobre o comportamento do clima, temperatura e atividade oceanográfica. Os resultados das modelagens do IPCC estão disponíveis em mapas interativos que facilitam a compreensão do públi-

co em geral, diante do risco crescente de desastres causados pela mudança climática e seus impactos econômicos, sociais e ambientais.

Bolzano et al (2020) comparam os cenários da mudança climática e da pandemia da COVID-19 como duas emergências globais que envolvem agenciamentos transnacionais na aplicação de recursos, necessidade de incentivos a projetos de pesquisa e de cooperação internacional, bem como a troca de dados e informações. Porém, os autores destacam que os cenários parecem condicionados a políticas de governabilidade divergentes, pois os principais agentes de tomada de decisão possuem diferentes níveis de percepção do risco e da abrangência destes fenômenos.

A partir da abordagem de Beghtol (1998) sobre os sistemas de classificação de disciplinas; Buckland (2012) o conceito de objeto digital; Kleinberg (2017) os níveis integrativos e complexidade do conhecimento; e Hjørland (1995, 2002) a análise de domínio, o presente trabalho visa analisar a representação do conhecimento no contexto da mudança climática e COVID-19, através do uso combinado da análise de domínio com as ferramentas cientométricas VOSviewer e CitNetexplorer, de maneira a perceber a emergência de domínios de conhecimento pós-disciplinares na organização do conhecimento.

O núcleo central do formato CERIF composto pelas entidades: projetos, pessoas e organizações, foi tomando como referência e alinhado aos dois princípios FAIR – Encontrável e Acessível (Wilkinson et. al, 2016) para fins de análise. A coleta de dados com base nos termos *climate change* e COVID-19 foi realizada na Scopus, Web of Science e Dimensions, no período 2010-2020 e no Diretório de Grupos de Pesquisa na Plataforma Lattes do Conselho Nacional de Desenvolvimento Científico e Tecnológico do Brasil (CNPq). A análise de dados realizada na amostra demonstrou aproximação semântica entre termos, conceitos e palavras-chave que se refletem nos temas entre as áreas de conhecimento sobre a mudança climática e COVID-19.

Revisão de literatura

O conhecimento científico organizado por disciplinas é discutido por Beghtol (1998) que distingue os Sistemas de Classificação Bibliográfica (SCB) dos Sistemas Especiais de Classificação (SEC). Em primeiro lugar, os SCB abarcam todo o universo do conhecimento analisado através de facetas, classes e subclasses; assim como selecionam uma ou mais disciplinas acadêmicas com base em aulas ou em categorias. Nesse sentido, os sistemas de classificação nascem ao mesmo tempo em que surgem disciplinas como uma maneira para estruturar o conhecimento acadêmico. Em segundo lugar, a autora acrescenta que a estrutura disciplinar impõe uma distinção entre documentos de ficção e não-ficção, esta distinção perdeu parte de sua força na Biblioteconomia. Em terceiro lugar, a autora destacou que a Tecnologia da Computação permite recuperar unidades de informação de qualquer tamanho, em particular, unidades menores do que aquelas encontradas na maioria das bases de dados que recuperam textos completos.

Hjørland adverte que quando os documentos são mesclados em grandes bancos de dados, as informações sobre significados implícitos dos contextos anteriores são perdidas: “Quanto maior o grau de fusão, maior também a perda de informações

implícitas. Os Sistemas de Organização do Conhecimento (SOC) devem ser desenvolvidos para lidar com essa perda de informações implícitas, tornando-as explícitas através da semântica de bancos de dados.” (Hjørland, 2002, 445).

As 11 abordagens de Hjørland (2002) apontam para a possibilidade de integração dos sistemas de informação através da compreensão das distâncias semânticas entre termos, gêneros de documentos produzidos nos domínios de conhecimentos e as comunidades epistêmicas. Gramini (2020) discute as limitações dos canais de comunicação sobre mudança climática para o público em geral, pois se de um lado os sistemas são abertos para transferência de dados entre instituições, por outro são de difícil acesso e compreensão para o público. Neste sentido, cabe destacar a proposta de implementação de Sistemas de Informação de Pesquisa Corrente (CRIS) e das iniciativas em prol do formato comum de comunicação para recuperação de informações referenciais, no caso o Common European Research Information Format (CERIF).

Em relação aos sistemas complexos de organização de conhecimento baseados em modelagem de dados sobre o clima e o planeta terra, Boldrini et al (2014) descrevem como exemplo: The Global Earth Observation System of Systems (GEOSS); the European Infrastructure for Spatial Information in the European Community (INSPIRE); e ontologia do meio ambiente na web (ENVO).

Estes sistemas computacionais complexos expressam fórmulas e métricas geolocalizadas chamadas de geométricas que instauram poderes práticos e simbólicos sob a percepção do clima (Elden, 2013). Hulme (2009, 2013) discute a formação de um regime global de informação (Braman, 2004) sobre os fenômenos da mudança climática, por meio da redução estatística dos eventos do clima em diversas partes do planeta. Na visão de Hulme (2009), os aspectos sociais e culturais da percepção do clima e suas mudanças se aproximam da ciência pós-normal proposta por Funtowicz e Ravetz (1994), o qual também critica a perspectiva disciplinar do conhecimento científico separado por áreas que não conversam entre si. Ravetz (1995) e Hulme (2009) admitem a necessidade de uma convergência das Ciências Exatas e da Terra com as Ciências Humanas e Sociais, nas análises de problemas complexos entre os quais a percepção dos fenômenos do clima em diferentes culturas. As relações entre as ontologias dos fenômenos físicos, biológicos e humanos conduzem a reflexão sobre a teoria de sistemas e sua importância para a organização do conhecimento.

Clément, com base na teoria geral de sistemas de Ludwig Bertalanffy (1993), define um sistema aberto como “um sistema cuja existência e estrutura dependem de uma energia externa (como a chama de uma vela ou os redemoinhos de um rio ao redor do cais de uma ponte). No caso de sistemas vivos, esse alimento não é apenas material ou energético, mas também é organizacional e informativo.” (Clément, 2000, 55). Para o autor, o hipertexto pode ser considerado um sistema aberto que marca o início da nova era do livro em constante transformação. Desta forma, termos, palavras-chave e conceitos são “desterritorializados”, ou seja, os conceitos se tornam “nômades”, termo emprestado do campo epistemológico da complexidade e da teoria do caos e aplicado ao conceito de hipertexto (Clément, 2000).

A representação de informação sobre problemas globais como mudança climática e COVID-19 nos convidam a refletir estes termos como conceitos nômades, que

atravessam diferentes comunidades epistêmicas e domínios de conhecimento. Nessa relação de conceitos nômades, é possível citar a perspectiva holística de conhecimentos de Bogdan et al (2014) que sugerem a emergência do paradigma das teorias de Bioeconomia e Ecoeconomia como um aporte na direção da “Bio-eco-geoconomia”.

Kleinberg (2017) destaca a importância das ciências da complexidade como exemplo os estudos em genética que se aproximam da teoria dos níveis integrativos e dos níveis de complexidade de sistemas orgânicos e inorgânicos propostos no ensaio de Feibleman (1954). O princípio de organização dos níveis integrativos pode ser expresso em termos de ordem evolutiva com base nos princípios combinados de gradação por especialidade, na relação gênero-espécie e na relação parte-todo (Dousa, 2009).

A abrangência dos temas da mudança climática e COVID-19 reforçam a importância da teoria dos níveis integrativos no contexto dos novos sistemas de classificação (Keshet, 2010), como analisados por Kleinberg (2017) e Gnoli (2008) que ampliam as questões sobre a interação dos fenômenos naturais, culturais e sociais que entrelaçam o mundo material, o universo orgânico dos seres vivos e as representações mentais e culturais dos seres humanos. Manzano (2020), Klein (2014) e Saloranta (2001) sinalizaram que mudança climática e COVID-19 são problemas globais com diferentes percepções sobre a avaliação de risco e graus de incerteza. Nessa perspectiva, Latour (2004) já criticava a crença em torno da objetividade suprema em relação ao mundo natural como resultado de uma epistemologia baseada em duas formas ontológicas distintas, uma que compreende os seres humanos e a outra composta por não-humanos.

Metodologia

A presente pesquisa tomou como referência a análise de domínio, uma abordagem sociocognitiva que observa o contexto das organizações a partir das comunidades discursivas com o objetivo compreender um domínio através das 11 abordagens sugeridas por Hjørland e Albrechtsen (1995). Neste trabalho foram escolhidas as abordagens 7 e 8, bem como as ferramentas cientométricas VOSviewer e CitNetexplorer utilizadas para analisar uma amostra de 125 termos recuperados em novembro de 2020, a partir das expressões de busca “*climate change* e COVID-9”, nas bases de dados Scopus, Web of Science e Dimensions, no período 2010-2020.

A fim de verificar os princípios de encontrabilidade e acessibilidade, o experimento foi replicado na Plataforma Lattes e no Diretório de Grupos de Pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico do Brasil (CNPq).

O estudo empírico foi dividido em cinco etapas: 1) Identificar os termos e palavras-chave do domínio nas bases de dados; 2) Identificar a rede de autores mais frequentes na Web of Science; 3) Identificar “entidades, pessoas e organizações” do núcleo CERIF na recuperação das bases de dados; 4) Comparar os termos e áreas de conhecimento das bases de dados; 5) Testar as etapas anteriores na Plataforma Lattes e Diretórios de Grupos de Pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

As análises de domínio 7 (estudos de documentos e gêneros em domínios do conhecimento) e 8 (estudos epistemológicos e críticos de diferentes paradigmas, pressupostos, e interesses em domínios) foram utilizadas de maneira a auxiliar nas interpretações dos dados a fim de permitir uma visão multidisciplinar e abrangente de forma a perceber as propostas das ODS 14 e 15 da Agenda 2030, assim como as limitações das bases de dados analisadas.

Discussão dos resultados

A primeira etapa do estudo foi a análise das três bases de dados internacionais, tendo como resultado da recuperação da informação o seguinte: 378 documentos na Scopus, 423 na Web of Science e 775 na Dimensions. A partir de 125 termos recuperados nas bases, foi utilizado o programa VOSviewer para análise, no qual a Figura 1 representa o grafo dos clusters de termos e palavras-chave do domínio “*climate change* e COVID-19”.

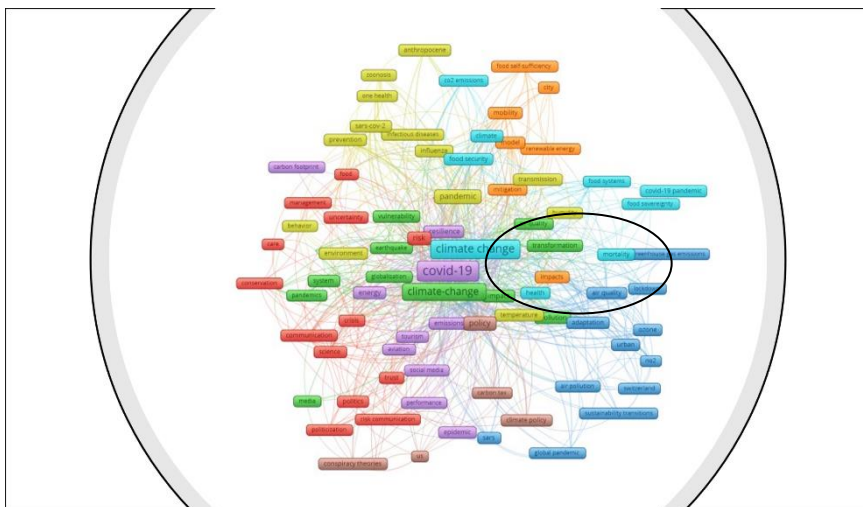


Figura 1. Mapa de frequência e correlação dos termos *climate change* e COVID-19 na Web of Science (2021).

A sintaxe realizada pelo VOSviewer resultou no grafo a centralidade dos termos COVID-19 e *climate change*. Os demais termos representados por diferentes clusters sinalizados pelas etiquetas com cores e tamanhos que simbolizam a aproximação e as distâncias entre os termos, em relação ao núcleo destacado. Cabe destacar que entre os termos mais próximos estão *impact*, *globalization*, *risk*, *earthquake* e *resilience*, conforme o grafo da Figura 1.

Na segunda etapa, foram analisados os dados recuperados na Web of Science no programa CitNetexplorer, através do acoplamento dos dados, citação e co-citação, com intuito de perceber a rede de autores que compartilham a temática da mudança climática e COVID-19, retratada na Figura 2.

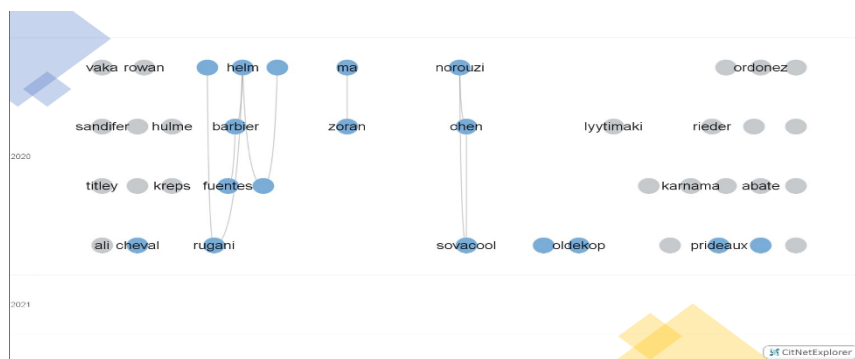


Figura 2. Rede de autores de *climate change* e COVID-19 na Web of Science (2021).

O grafo da Figura 2 destaca quatro agrupamentos autores, a saber: 1) Helm, Barbier, Fuentes e Rugani; 2) Ma e Zoran; 3) Norouzi, Ohen e Savacool; 4) Cheval, Oldekop e Prideaux.

Na terceira etapa, os dados das da recuperação nas três bases de dados internacionais foram tratados e analisados no programa VOSviewer a fim de perceber os itens “pessoas e instituições” do núcleo CERIF, através da amostra dos dez autores mais citados por base, tendo como resultado a Tabela 1.

Tabela 1. Resultado da análise dos autores mais citados nas três bases de dados.

N.º	Base de Dados	Autor	Afiliação	País	Periódico
1	Scopus	Gössling S.	Western Norway Research Institute	Norway	Journal of Sustainable Tourism
2	Scopus	Markard J.	ETH Zurich	Switzerland	Sustainability: Science, Practice, and Policy
3	Scopus	Rosenbloom D.	University of Toronto	Canadá	Science
4	Scopus	Zarazua De Rubens G.	Aarhus University	Denmark	Energy Research and Social Science
5	Scopus	Bodrud-doza	Bangladesh Rural Advancement Committee	Bangladesh	Environmental Research
6	Scopus	Rahman M.M.	Begum Rokeya University	Bangladesh	Frontiers in Public Health
7	Scopus	Shammi M.	Jahangirnagar University	Bangladesh	Frontiers in Public Health
8	Scopus	Higham J.	University of Otago	New Zealand	Journal of Travel Research
9	Scopus	D’adamo I.	Sapienza University of Rome	Italy	Energy Policy
10	Scopus	Meo A.S.	Army Medical College	Pakistan	Pakistan Journal of Medical Sciences
N.º	Base de Dados	Autor	Afiliação	País	Periódico
1	WOS	Fu, Shihua	Lanzhou University	China	Science of The Total Environment
2	WOS	He,	Lanzhou University	China	Science of The Total

		Xiaotao			Environment
3	WOS	Liu, Jiangtao	Lanzhou University	China	Science of The Total Environment
4	WOS	Luo, Bin	Lanzhou University	China	Science of The Total Environment
5	WOS	Niu, Jingping	Lanzhou University	China	Science of The Total Environment
6	WOS	Wang, Bo	Lanzhou University	China	Science of The Total Environment
7	WOS	Yan, Jun	Lanzhou University	China	Science of The Total Environment
8	WOS	Zhou, Ji	Shanghai Meteorological Bureau	China	Science of The Total Environment
9	WOS	De Rubens, Gerardo Zarazua	Aarhus University	Denmark	Energy Research & Social Science
10	WOS	O'callaghan, Brian	University of Oxford	United Kingdom	Environmental & Resource Economics
N.º	Base de Dados	Autor	Afiliação	País	Periódico
1	Dimensions	Bodrud-doza, Md	Bangladesh Rural Advancement Committee (BRAC)	Bangladesh	Asian Journal of Psychiatry
2	Dimensions	Rahman, Md Mostafizur	Jahangirnagar University	Bangladesh	Heliyon
3	Dimensions	Barello, Serena	Università Cattolica del Sacro Cuore	Italy	European Journal of Epidemiology
4	Dimensions	Bashir, Muhammad Farhan	Central South University	China	Air Quality, Atmosphere & Health
5	Dimensions	Djalante, Riyanti	United Nations University-Institute for the Advances Study of Sustainability (UNU-IAS)	Japan	Progress in Disaster Science
6	Dimensions	Flahault, Antoine	University of Geneva	Switzerland	International Journal of Environmental Research and Public Health
7	Dimensions	Galea, Sandro	Harvard Global Health Institute,	United States	Disaster Medicine and Public Health Preparedness
8	Dimensions	Graffigna, Guendalina	Università Cattolica del Sacro Cuore	Italy	Frontiers in Psychology
9	Dimensions	Haines, Andy	London School of Hygiene and Tropical Medicine	United Kingdom	Cardiovascular Research
10	Dimensions	Lelieveld, Jos	The Cyprus Institute/ Max Planck Institute for Chemistry	Cyprus/ Germany	Proceedings of the National Academy of Sciences of the United States of America

Referência: criação nossa com base na análise dos dados recuperados nas bases de dados (2021).

Na Scopus, em relação ao item “pessoas”, foram identificados três autores de Bangladesh (Bodrud-Doza, Rahman e Shammi), sendo que dois destes autores também constam na recuperação da Dimensions (Bodrud-Doza e Rahman). O autor Zarazua de Rubens da Dinamarca foi detectado na Scopus e Web of Science, mas com grafias de nomes diferentes. Além disso, o autor mais citado foi Stefan Gössling da instituição Western Norway Research Institute.

Na Web of Science, para “pessoas”, oito dos dez autores mais citados foram chineses, sendo todos autores do artigo intitulado “*Impact of meteorological factors on the COVID-19 transmission: a multi-city study in China*”. Portanto, a instituição que mais se destacou foi a Lanzhou University.

Na Dimensions, quanto a “pessoas”, o autor mais citado foi Bodrud-Doza da instituição Bangladesh Rural Advancement Committee (BRAC).

É interessante observar os títulos dos periódicos em relação às palavras-chave que indicam uma diversidade temática de produção de conhecimento: Turismo, Sustentabilidade, Energia, Ciências Sociais, Saúde Pública, Ciências Médicas, Ciência do Meio Ambiente Total, Economia Ambiental, Psiquiatria e Psicologia. Nesse sentido, fica evidenciada a complexidade da questão da mudança climática e COVID-19.

Na quarta etapa, os dados foram analisados no programa VOSviewer para identificação dos termos mais frequentes por base de dados e áreas de conhecimento para comparação diante da complexidade da representação da informação da mudança climática e COVID-19, tendo como resultado a Tabela 2.

Tabela 2. Comparativo de termos e áreas de conhecimento das bases de dados.

Comparativo dos principais termos por frequência nas bases de dados			
Nº	Scopus	Web of Science	Dimensions
1	<i>Climate Change</i>	<i>Covid-19</i>	<i>Cov</i>
2	<i>Covid-19</i>	<i>Climate Change</i>	<i>Crisis</i>
3	<i>Pandemic</i>	<i>Coronavírus</i>	<i>Lockdown</i>
4	<i>Human</i>	<i>Pandemic</i>	<i>Patient</i>
5	<i>Coronavirus Disease 2019</i>	<i>Policy</i>	<i>Transmission</i>
6	<i>Coronavirus Infection</i>	<i>Sustainability</i>	<i>Challenge</i>
7	<i>Coronavirus Infections</i>	<i>Resilience</i>	<i>China</i>
8	<i>Pneumonia Viral</i>	<i>Impact</i>	<i>March</i>
9	<i>Virus Pneumonia</i>	<i>Risk</i>	<i>Period</i>
10	<i>Betacoronavirus</i>	<i>Sars-Cov-2</i>	<i>Need</i>

Comparativo das áreas de conhecimento nas bases de dados			
N°	Scopus	Web of Science	Dimensions
1	<i>Social Sciences</i>	<i>Environmental Sciences</i>	<i>Studies in Human Society</i>
2	<i>Medicine</i>	<i>Environmental Studies</i>	<i>Medical and Health Sciences</i>
3	<i>Environmental Sciences</i>	<i>Public Environmental Occupational Health</i>	<i>Public Health and Health Services</i>
4	<i>Energy</i>	<i>Green Sustainable Science Technology</i>	<i>Economics</i>
5	<i>Business, Management and Accounting</i>	<i>Economics</i>	<i>Applied Economics</i>
6	<i>Multidisciplinary</i>	<i>Multidisciplinary Sciences</i>	<i>Commerce, Management, Tourism and Services</i>
7	<i>Earth and Planetary Sciences</i>	<i>Meteorology Atmospheric Sciences</i>	<i>Political Science</i>
8	<i>Economics, Econometrics and Finance</i>	<i>Medicine General Internal</i>	<i>Policy and Administration</i>
9	<i>Agricultural and Biological Sciences</i>	<i>Geography</i>	<i>Sociology</i>
10	<i>Arts and Humanities</i>	<i>International Relations</i>	<i>Environmental Sciences</i>

Referência: criação dos autores com base no VOSviewer versão 1.6.15 (2021).

O resultado das áreas de conhecimento mostrou que o termo *Environmental Sciences* aparece nas três bases de dados. O termo *Economics* está presente nas três bases, sendo que na Scopus é aparece como um termo composto *Economics, Econometrics and Finance* e na Dimensions também é percebido *Applied Economics*. O termo *Medicine* é identificado na Scopus e Web of Science, já na Dimension é encontrado *Medical and Health Sciences*.

No que diz respeito às palavras-chave, é possível inferir que existiu uma maior aproximação entre os termos recuperados pela Scopus e Web of Science, não sendo identificado nenhum termo na Dimensions. Os termos em destaque foram: *climate change, covid-19, pandemic* e *coronavírus*, além das versões de coronavírus como: *coronavírus disease 2019, coronavirus infection, e betacoronavirus*.

A quinta etapa, o resultado da análise da Plataforma Lattes recuperou o total de 14.448 currículos pelo termo COVID-19; 5.064 currículos pelo termo de busca COVID-19 AND meio ambiente; 679 currículos pelo termo de busca COVID-19 AND mudança climática.

O resultado no Diretório de Grupos de Pesquisa do CNPq para o termo “COVID-19 AND mudança climática” não foi encontrado nenhum grupo. Para o termo

“COVID-19” foram encontrados 233 grupos, sendo possível perceber a gama de áreas de conhecimento que englobam a pesquisa sobre COVID-19 com os respectivos percentuais em ordem decrescente: Ciências da Saúde 44%, Ciências Biológicas 27,47%, Ciências Sociais Aplicadas 9,01%, Ciências Humanas 7,73%, Ciências Exatas e da Terra 6,87%, Engenharias 2,57%, Outras 1,29%, e Letras, Linguística e Artes 0,85%.

O resultado do termo “COVID-19 AND meio ambiente” foi apenas identificado um grupo: Grupo de Economia do Meio Ambiente e Desenvolvimento Sustentável (GEMA), criado em 1990 (última atualização em julho de 2020), na área Economia nas Ciências Sociais Aplicadas e cujas linhas de pesquisa do grupo são: 1) COVID-19, Políticas Públicas e Meio Ambiente; 2) Economia verde e pagamento por serviços ecossistêmicos no Brasil; 3) Elaboração, análise e coordenação de políticas públicas ambientais; 4) Financiamento para o desenvolvimento sustentável; 5) Meio ambiente, comércio internacional e matrizes de insumo-produto; 6) O *trade-off* entre políticas de transferência de renda e conservação florestal.

É relevante assinalar que existe uma aproximação entre as linhas de pesquisa do GEMA com o escopo temático do periódico *Science of the Total Environment*, revista multidisciplinar internacional que aborda o conceito de ambiente total: atmosfera, litosfera, hidrosfera, biosfera e antroposfera. Além disso, os resultados indicaram áreas multidisciplinares que sugerem o crescimento da pós-disciplinaridade na organização do conhecimento, o que foi percebido também na revisão de literatura quanto ao desenvolvimento de sistemas complexos de representação do conhecimento.

Conclusão

A pesquisa revelou que a mudança climática e COVID-19 são domínios multidisciplinares de conhecimento, o que indica uma possível mudança de paradigma em relação aos processos de representação da informação. Ao invés da demarcação de áreas de conhecimento disciplinares, percebeu-se que os termos se tornam nômades diante da demanda dos efeitos sociais, políticos e econômicos. É possível inferir a emergência de um domínio de conhecimento como a Ciência do Meio Ambiente Total, demonstrando a perspectiva pós-disciplinar.

As bases de dados recuperam itens documentais representados por diferentes áreas de conhecimento e com temas não necessariamente recorrentes nas três bases analisadas. Os títulos dos periódicos reforçam o aspecto multidisciplinar dos fenômenos de mudança climática e COVID-19 que envolvem diferentes abordagens como Economia, Turismo, Ciências da Saúde, Ciências Ambientais, Psicologia, entre outros. A amostra da recuperação dos termos das bases indicou a relação de proximidade entre Ciências Ambientais e Saúde Coletiva, em virtude de pesquisas científicas sobre os efeitos do clima na saúde. Convém sinalizar que o SARS já tem sido estudado desde 2003 e o termo mudança climática foi indexado na base Scopus em 1910.

As abordagens 7 e 8 de Hjørland, conjugadas às ferramentas de visualização de dados VOSviewer e CitNetexplorer reforçaram as questões terminológicas que

envolvem a mudança climática e a COVID-19 na discussão dos aspectos ecológicos, econômicos, sociais, entre outros. Neste sentido, é observado o duplo desafio estrutural para os sistemas de organização do conhecimento, em especial em relação às bases de dados internacionais. Em primeiro lugar, a representação da informação para fins da recuperação eficiente de itens documentais enfrenta o desafio da necessidade de padronização diante da pluralidade dos diversos domínios de conhecimentos, suas comunidades específicas e na produção de suas fontes documentais. Em segundo lugar, o desafio é também epistemológico no contexto na organização do conhecimento marcado pelo debate sobre as ciências empíricas que analisam a complexidade dos sistemas naturais diante do quadro de risco global que geram níveis de incerteza nas demandas por respostas imediatas.

Os resultados mostraram que se de um lado a perspectiva pós-disciplinar ressalta os aspectos epistemológicos dos estudos multidisciplinares sobre mudança climática e COVID-19, por outro lado a profusão de termos, domínios e comunidades acrescentam novas dificuldades técnicas na representação do conhecimento, que dificultam a aplicação dos princípios de encontrabilidade e acessibilidade nas bases de dados. O surgimento de novos domínios de conhecimentos e de comunidades no contexto destes desafios globais ressalta a importância dos sistemas de referenciamento de pesquisas (CRIS) e formatos de comunicação enquanto agentes facilitadores na identificação dos fluxos de informação entre projetos, pessoas e instituições previstas no núcleo do CERIF.

A análise quali-quantitativa permitiu situar *climate change* e COVID-19 como um possível domínio emergente na organização do conhecimento, o que implica no redimensionamento de sistemas de organização de conhecimento. Nesta direção, a comparação dos grafos e a distribuição dos clusters nas bases de dados internacionais, sinalizaram relações conceituais entre redes de autores, instituições, fontes documentais (periódicos) e áreas transversais de conhecimento que merecem ser aprofundadas como continuação do estudo.

Em sentido amplo, os resultados da amostra destacaram a necessidade de infraestruturas convergentes que promovam a integração das fontes de dados de pesquisa com a produção científica dos pesquisadores nas mais diversas fronteiras de conhecimentos. No sentido específico da organização do conhecimento, as abordagens 7 e 8, as ferramentas cientométricas e aos dois princípios FAIR demonstraram que o desafio da interoperabilidade semântica entre conceitos, termos e palavras-chave não se resolvem ou terminam apenas com a implementação de sistemas com recursos de tecnologia digital e adoção de formatos. A solução para estes desafios depende também da cooperação internacional e interoperabilidade intercomunitária representada por ações de comunicação entre comunidades geolocalizadas dentro ou fora das academias que discutem soluções sustentáveis para os problemas emergentes globalizados.

A emergência de problemas globais sinaliza a possibilidade da passagem para o modelo dialógico e pós-disciplinar na busca por maior simetria entre projetos de pesquisa, governança de dados e troca de informações. As iniciativas e os debates da Agenda 2030 e dos resultados do trabalho coletivo dos pesquisadores do IPCC já indicam o caminho das comunidades abertas entre pesquisadores e não acadêmicos que compartilham ecossistemas interativos de conhecimento.

Por fim, acreditamos que o trabalho sirva como base para ampliação dos estudos sociais, políticos e econômicos da mudança climática e COVID-19, pois a metodologia proposta é replicável para outras temáticas que cobrem os 17 ODS da Agenda 2030.

Referências

- Agenda 2030. (2021). Acesso: 27 junho 2021. Disponível em: <http://www.agenda2030.org.br/sobre/>
- Beghtol, C. (1998). Knowledge Domains: Multidisciplinary and Bibliographic Classification Systems. *Knowl. Org.* 25, N.º 1/2
- Boldrini, E; Luzi, D.; Nativi, S.; Pecoraro, F. (2014). Integrating CERIF entities in a multidisciplinary e-infrastructure for environmental research data. In: CRIS2014: 12th International Conference on Current Research Information Systems, Rome May 13-15.
- Braman, S. (2004). The emergent global information policy regime, In: (Ed.) The emergent global information policy regime. Hampshire, Palgrave, cap. 2. 12-37.
- Buckland, M (2012). What kind of science can information science be? *Journal of the American Society for Information Science and Technology*, v. 63, n.º 1, 1-7.
- Clément, J. (2000). Hypertexte et complexité. *Études françaises*, 36(2), 39-57. <https://doi.org/10.7202/005256ar>
- Dousa, T. M. (2009). “Evolutionary Order in the Classification Theories of C. A. Cutter and E. C. Richardson: Its Nature and Limits.” In *Proceedings from North American Symposium on Knowledge Organization 2*, Syracuse, New York, ed. Elin K. Jacob and Barbara Kwasnik. ISKO-CUS, 76-90.
- Edwards, P. (2010). *A vast machine: computer models, climate data, and the politics of global warming*. Massachusetts: The MIT Press.
- Elden, S. (2013). *Secure the volume: vertical geopolitics and the depth of power*. *Political Geography*, v. 34, 35-51.
- Feibleman, J. K. (1954). Theory of Integrative Levels. *The British Journal for the Philosophy of Science* 5 n.º 17: 59-66.
- Foskett, D. J. (1978) The theory of integrative levels and its relevance to the design of information systems. *Aslib Proceedings*. Vol. 30, n.º 6, 202-208.
- Funtowicz S. O., Ravetz J. R. (1994). Uncertainty, complexity and post-normal science, *Environmental Toxicology*, 13(12): 1881-1885.
- Giddens, A (2008). The politics of climate change. National responses to the challenge of global warming. <http://www.policy-network.net/>
- Gnoli, C. (2008). Categories and Facets in Integrative Level. *Axiomathes* 18:177-92.
- Hjørland, B.; Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*. Vol.46, 400-425.
- Hjørland B. (2002). *Domain analysis in information science: Eleven approaches – Traditional as well as innovative*. *Journal of Documentation*, (58) 4, 422-462.
- Hulme, M. (2009). *Why We Disagree About Climate Change: Understanding Controversy, Inaction and Opportunity*. New York: Cambridge University Press.
- Hulme, M. (2013). How climate models gain and exercise authority. In: Hastrup, K.; Skrydstrup, M. (Eds.). *The social life of climate change models: anticipating nature*. Nova York: Routledge, 30-44.
- Ipcc. AR6 Climate Change (2021): The Physical Science Basis Acesso: 27 junho 2021. Disponível em: <https://www.ipcc.ch/report/ar6/wg1/>

- Klein, N. (2014). *This Changes Everything: Capitalism vs. the Climate*. New York: Simon & Schuster.
- Kleinberg, M. (2017). Reviews of Concepts in Knowledge Organization: Integrative levels. *Knowl. Org.* 44, N.º 5.
- Latour, B. (2004). *Politics of Nature*, Harvard University Press, Cambridge, MA.
- Lockyer, W. J. S. (1910). Does the Indian Climate Change? *Nature*, Volume 84, Issue 2128, 178.
- Ludwig von Bertalanffy (1993). *Théorie générale des systèmes*, Paris, Dunod.
- Manzano, B. (2020) Covid-19 and Climate Change: A Tale of Two Global Problems. GREEN-Bocconi WORKING PAPER SERIES Number 09 ISSN 1973-0381
- Ravetz, J. R. (1995). *Scientific knowledge and its social problems* (First ed. 1971), New Brunswick (NJ): Transaction.
- Saloranta T. M. (2001). Post-normal science and the global climate change issue. *Climatic Change* 50: 395-404.
- Wilkinson, M. et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Nature*.

ORGANIZAÇÃO DO CONHECIMENTO NA CÂMARA DOS DEPUTADOS BRASILEIRA: ESTUDO TEÓRICO-METODOLÓGICO PARA CATEGORIZAÇÃO DAS PROPOSIÇÕES SOBRE DIREITOS EM CONSTRUÇÃO DAS MULHERES

Carla Maria Martellote Viola¹, Luana Farias Sales²

¹Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) /
Universidade Federal do Rio de Janeiro (UFRJ), viola.carla@gmail.com,
ORCID ID 0000-0002-0895-8163

²Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) /
Universidade Federal do Rio de Janeiro (UFRJ), luanafsales@gmail.com,
ORCID ID 0000-0002-3614-2356

Resumo

Objetivo: Investigar os aportes teórico-metodológicos da Organização do Conhecimento que possam servir de base para a construção de novo modelo de categorização das proposições em tramitação na Câmara dos Deputados brasileira, que abordam os direitos em construção das mulheres, para viabilizar uma recuperação da informação comprometida com apreciações contemporâneas críticas e pluralistas, que sejam relevantes para atender as verdadeiras necessidades das mulheres.

Métodos: Pesquisa exploratória e qualitativa fundada na busca teórico-metodológica da Organização do Conhecimento, que possa assentir o desenvolvimento humano das mulheres rumo aos Objetivos de Desenvolvimento Sustentável da Agenda 2030. Métodos bibliográfico, a partir do levantamento e análise da literatura alicerce da Organização do Conhecimento, e documental, com base no levantamento e análise das categorias até então utilizadas para organizar as proposições em tramitação na Câmara dos Deputados brasileira, voltadas para os direitos em construção das mulheres.

Resultado: A pesquisa teórico-metodológica revelou que, para atender as reais necessidades das mulheres, a categorização das proposições deve ser flexível, criativa e desenhar associações, não baseadas em um conjunto de princípios predeterminados, mas, sim, num simples reconhecimento de semelhanças necessárias para demonstrar a qualidade e a eficiência das proposições em relação à Agenda 2030. Portanto, devem se valer de um agrupamento de baixo para cima, contemplando relações funcionais que indiquem as relevâncias motivacional, temática e interpretativa.

Conclusões: O governo brasileiro precisa implementar políticas públicas e criar normativas eficazes, requerendo técnicas da Organização do Conhecimento que auxiliem na sistematização das propostas apresentadas na Câmara dos Deputados,

para aferição da qualidade dos dados com vistas à Agenda 2030. Necessita-se de categorização que reconheça as vulnerabilidades das mulheres apontadas nos Objetivos do Desenvolvimento Sustentável, assegurando a recuperação eficaz das proposições em tramitação. A partir do arcabouço teórico-metodológico revelado, uma proposta de nova categorização para a interpretação das proposições, no que tange aos direitos em construção das mulheres, poderá possibilitar a aferição do grau de eficácia dessas propostas em razão da Agenda 2030, em especial do objetivo 5 e de suas metas e interligações com os outros objetivos.

Palavras-chave: Organização do Conhecimento, direitos em construção das mulheres, categorização, Câmara dos Deputados, Agenda 2030.

Introdução

Atualmente, as proposições em tramitação na Câmara dos Deputados brasileira são classificadas e organizadas em macrotemas. Os macrotemas têm relação com todas as comissões temáticas da Câmara dos Deputados, são as áreas da Consultoria Legislativa da Câmara, que levam em consideração todos os domínios do conhecimento (Câmara dos Deputados, 2016).

A classificação adotada pela Câmara dos Deputados atende a recuperação dos conteúdos que fomentam os trabalhos da Casa. Contudo, ao se considerar o documento internacional denominado Agenda 2030 – que reúne os 17 Objetivos do Desenvolvimento Sustentável (ODSs) e do qual o Brasil é signatário – e as necessidades evidenciadas em pesquisas institucionais, científicas e jornalísticas sobre as mulheres, percebe-se que o modelo de organização atualmente adotado para a sistematização das proposições não permite a aferição da relevância e da eficiência dos conteúdos legislativos recuperados, em razão das verdadeiras vulnerabilidades que abarcam a vida das mulheres. Nesse sentido, a questão que se coloca é: aportes teórico-metodológicos da Organização do Conhecimento (OC) podem representar um caminho para a elaboração de novas categorias que sistematizem de forma mais adequada as proposições que abordem os direitos em construção das mulheres, apresentando assim interpretações responsivas para a verificação da qualidade das propostas legislativas brasileiras acerca de políticas públicas e normativas para a concretização da Agenda 2030?

O ODS 5, com suas 9 metas e interligações da Agenda 2030, visa “alcançar a igualdade de gênero e empoderar todas as mulheres e meninas” (United Nations, 2015). Para que as intenções sejam conquistadas, são indispensáveis esforços e medidas governamentais para a criação de estruturas legais que promovam, façam cumprir e monitorem a igualdade de gênero. Portanto, faz-se necessário elaborar nova categorização responsiva para a verificação da qualidade das proposições legislativas brasileiras, visando a concretização da Agenda 2030.

Posto isto, enquanto recorte de pesquisa de uma tese em andamento, este artigo objetiva investigar os aportes teórico-metodológicos da OC que possam servir especialmente à construção de novo modelo de categorização para a interpretação associativa das proposições em tramitação na Câmara dos Deputados, que abordam os direitos em construção das mulheres, para viabilizar uma recuperação da infor-

mação comprometida com apreciações contemporâneas críticas e pluralistas, que sejam relevantes para atender as verdadeiras necessidades das mulheres, em consonância com o ODS 5 e suas metas e interligações com os demais objetivos da Agenda 2030.

A partir da concepção de OC de Hjørland (2008), o estudo se fundamentou basicamente nas propostas teóricas de autores como Dahlberg (1978a, 1978b; 2006), Smith e Medin (1981), Jacob (2004), Olson (1998, 2007) e nas concepções de relevância de Shutz (1970).

Metodologia

Adota-se 3 etapas para alcançar os objetivos propostos: na primeira, realiza-se pesquisa bibliográfica para abordagens das teorias que expressem conformidade para desenvolver argumentos sobre a Organização do Conhecimento, classificação, categorização e relevância. Este tipo de averiguação possibilita um amplo alcance de informações, além de permitir a utilização de dados dispersos em inúmeras publicações, auxiliando também a construção, ou a melhor definição do quadro conceitual que envolve o objeto de estudo proposto (Gil, 1994).

Na segunda etapa, complementa-se a pesquisa bibliográfica com a documental, a partir de levantamento, relatório e de enunciados provenientes dos Portais da Câmara dos Deputados, United Nations e ONU Brasil, visando buscar argumentos satisfatórios para a proposta de uma nova categorização das proposições legislativas. Quanto aos livros e documentos, a primeira fase correspondeu à leitura do texto; a segunda fase consistiu em selecionar, no texto lido, o conteúdo informacional importante para os usos determinados. E na terceira fase, procedeu-se a indexação das informações selecionadas, de modo a torná-las documentariamente manipuláveis (Kobashi, 1996).

Na terceira etapa, a pesquisa segue os parâmetros da *sequential exploratory strategy*, explicada por Creswell e Creswell (2018) como um tipo de abordagem especialmente adequado para desenvolver novos instrumentos de coleta e/ou aperfeiçoar instrumentos já existentes, em que a abordagem qualitativa precede a quantitativa. Nesta etapa, a partir da exploração dos macrotemas utilizados pela Câmara dos Deputados e dos objetivos e metas do desenvolvimento sustentável, sugere-se um percurso metodológico.

Mulheres e os dados intangíveis

O documento “Transformando Nosso Mundo: A Agenda 2030 para o Desenvolvimento Sustentável”, das Nações Unidas, registra intensões de estimular e apoiar ações em áreas de importância crucial para a humanidade: Pessoas, Planeta, Prosperidade, Paz e Parcerias (United Nations, 2015).

Razavi (2016) reconhece que a Agenda 2030 tem a pretensão de suprir muitas das deficiências que estavam registradas nos objetivos e metas dos Objetivos de Desenvolvimento do Milênio (ODM) quanto à sua visão estreita e da maneira de cima para baixo em que foram definidos. A autora assinala que o processo de definição da nova agenda foi mais inclusivo e de escopo universal, além de seus ODSs e metas chamarem atenção para um conjunto muito mais variado de preocupações estruturais e de estarem ancorados nos princípios dos direitos humanos.

Fato este que está evidente no escopo notavelmente amplo das metas do ODS 5, específico de gênero – desde leis discriminatórias, práticas prejudiciais e violência contra mulheres e meninas, à saúde sexual e reprodutiva e direitos reprodutivos, à desigualdade de gênero na distribuição de trabalho de cuidado não remunerado, igualdade de acesso aos recursos produtivos e participação das mulheres na tomada de decisões (Razavi, 2016; Azcona e Bhatt, 2020).

Mesmo que a agenda apresente progresso e maior completude na definição de seus objetivos e metas, para que os países demonstrem efetivamente avanço rumo ao desenvolvimento sustentável ambiental, econômico e social e a intenção de “não deixar ninguém para trás”, lema publicizado pela ONU, precisam de alguma forma contabilizar suas ações e apresentar dados que comprovem a melhoria de condições em amplo espectro.

O recente relatório *Implementation of sustainable development goals: addressing the data disaggregation challenge in the GCC*, publicado pela Oliver Wyman, enfatiza a existência de várias barreiras nos países para a coleta de dados em torno dos Objetivos de Desenvolvimento Sustentável. O relatório indica a necessidade dos governos de realizarem um processo sistemático de verificação de dados, aumentarem as capacidades das Organizações Nacionais de Estatística (ONEs) e ressaltarem a importância do uso de dados de terceiros (Sammakieh; Youssef; Salem, 2021).

O destaque desta questão-chave – uso de dados de terceiros – para a recuperação e monitoramento do cumprimento dos ODSs, apresenta-se como fonte de informação que permite complementar e até mesmo suprir a não disponibilização de informações pelas ONEs, para o desenvolvimento de políticas baseadas em evidências e orientadas para resultados, visando a realização dos ODSs.

Corroborando esta asserção o V Relatório Luz da Sociedade Civil Agenda 2030 de Desenvolvimento Sustentável Brasil (Grupo de Trabalho da Sociedade Civil para a Agenda 2030, 2021), que utiliza para a execução de suas análises dados oficiais e dados de terceiros.

Os dados de terceiros são entendidos como aqueles que provêm de toda fonte que observe os critérios de série histórica e apresente metodologia contextualizada e consolidada em sua comunidade sociocientífica, podendo ser decorrentes de estudos produzidos pela sociedade civil, organizações não governamentais e pesquisas acadêmicas catalogadas em bibliotecas, como SciELO e Portal Capes.

Esse relatório elaborado pelo Grupo de Trabalho da Sociedade Civil para a Agenda 2030 (GTSC A2030) – coalizão que atualmente reúne 57 organizações não governamentais, movimentos sociais, fóruns, redes, universidades, fundações e federações brasileiras, é a única pesquisa de monitoramento de dados existente no Brasil.

Fatos estes, demonstram a importância das pesquisas acadêmicas que resultem em novos modelos e apontem metodologias inovadoras para o monitoramento de dados sobre propostas e ações brasileiras para alcançar os ODSs.

No tocante à falta de monitoramento e de valorização adequadas dos dados que abordam os índices de desenvolvimento das mulheres, Azcona e Bhatt (2020) alertam para a dificuldade de se conhecer as ações relevantes iniciadas que possam vir a sanar as vulnerabilidades apontadas no *Global Gender Gap Report* (World Economic Forum, 2021) e alcançar as metas do ODS 5, até 2030 (United Nations, 2015).

Os governos e as ONEs têm sido tradicionalmente o padrão e os únicos guardiões dos dados para as questões relacionadas ao desenvolvimento de políticas e acompanhamento do progresso. Some-se a esta asserção, o monitoramento de dados sobre as mulheres, que exigem critérios interseccionais. Desta forma, para o levantamento e aferição de dados e resultados sobre os ODSs e as metas da Agenda de Desenvolvimento Sustentável, são necessários a participação de toda a sociedade e o envolvimento de todas as partes interessadas, em particular dos atores não governamentais.

Toda ação de coalisão, como o reconhecido *'advocacy* feminino' e estudos científicos acadêmicos ou da sociedade civil, são bem-vindos para relatar a dialética real entre a vulnerabilidade e o desenvolvimento das mulheres, que precisa ser representada na OC.

Portanto, segue-se com a discussão metodológica para uma categorização flexível das proposições em tramitação na Câmara dos Deputados sobre os direitos em construção das mulheres, com a base de formação estrutural desenhada nas relevâncias apresentadas pelo ODS 5, suas metas e seus correlatos.

A organização do conhecimento e a categorização baseada na relevância estrutural das mulheres

Fundamentos teórico-metodológicos indicam que a OC trata de atividades como descrição, indexação e classificação de documentos, existindo várias abordagens históricas e teóricas que se relacionam com diferentes visões de conhecimento, cognição, linguagem e organização social (Hjørland, 2008).

Enquanto campo de estudo, a Organização do Conhecimento é reconhecida por se encarregar da natureza e qualidade dos processos organizacionais, bem como dos sistemas de organização do conhecimento usados para estruturar documentos e documentar representações e conceitos.

Miranda (1999) acrescenta que a Organização do Conhecimento é o estudo dos recursos e instrumentos para a identificação, extração e descrição da informação registrada nos documentos, a partir da análise do seu conteúdo ou de seu modo de produção, visando classificá-los e ordená-los de modo que permita a recuperação mais eficaz e satisfatória inerente às demandas dos usuários.

Embora a classificação e a categorização na OC procurem representar o mapa de assuntos e disciplinas de forma objetiva e neutra, seus esquemas são histórica e culturalmente condicionados, pois refletem o pensamento social, político e religioso, bem como o estado da evolução científica e das mentalidades de profissionais da informação de seu tempo (Shera 1961; González Casanova 1996).

A organização desses assuntos se dá através de uma categorização formal dos conceitos que abarcam objetos, fenômenos, processos, propriedades; relações que têm importância na formação e combinação de sistemas de conceitos (Dalbergh, 1978a).

No entanto, Smith e Medin (1981), no livro *Categories and Concepts*, definem um modelo clássico de categorias que aponta para uma categorização arbitrária, podendo itens serem agrupados de inúmeras maneiras para formar categorias, além de pessoas poderem aprender a identificar ou construir essas categorias definidas pela sua cultura, pois nada no mundo ou em nosso sistema nervoso determina como devemos repartir as nossas observações. Nesse modelo, as categorias possuem atributos definidores ou críticos e todos os membros de uma categoria compartilham destes atributos definidores, nenhum não-membro compartilha deles, não há sobreposição entre membros e não-membros, e a intensão (ou conjunto de atributos) determina a extensão de uma categoria (quais itens são membros), sendo um método de agrupamento que se realiza de baixo para cima, isto é, que vai do significado que o objeto tem dentro de seu contexto cultural, em um agrupamento constante por semelhança, até se constituir uma categoria mais ampla. Dessa maneira, não faz sentido falar que uma categoria tem uma estrutura interna, com alguns itens se destacando como membros melhores do que outros itens.

Apesar de Dahlberg (1978b) também considerar esse tipo de metodologia de baixo para cima, sua proposta de agrupamento se dá em um nível lógico mais rígido, reunindo os conceitos apenas por seus tipos ou partes, deixando de fora da hierarquia o que ela chama de relações ônticas, ou seja, as relações que expressam outro tipo de associação entre conceitos.

Sob uma perspectiva flexível, Olson (1998, 2007) assevera que os estudos em OC devem ser pautados em valores que orientam a educação para uma catalogação crítica e flexível sobre as questões das mulheres e sugere um modelo alternativo que rejeita um modelo universal, aceita um conceito singular da verdade, foca nos relacionamentos, adequação de uma estrutura semelhante a uma teia em oposição a uma hierarquia piramidal, considera a situação do contexto e a experiência, envolve comunidades conhecedoras e reconhece o poder como fator de conhecimento.

Um modelo alternativo para a organização da informação está diretamente relacionado ao ato de dar voz às necessidades do “subalterno” (Morris; Spivak, 2010) e poder às desprovidas de conhecimentos.

Independente do método adotado, o processo de tratamento e classificação realizado pela OC “revela seu caráter estratégico decisivo, tanto na cooperação quanto no conflito que perpassam o jogo político, dado o papel capital da informação nas disputas de poder, o que engloba conhecimentos adequados das forças em disputa [...]” (Schneider, 2013, p. 63).

Importante destacar que Jacob (2004) assinala que a classificação e a categorização são processos diferentes, embora ambos sejam mecanismos para estabelecer a ordem por meio do agrupamento de fenômenos relacionados, no qual as diferenças fundamentais entre eles influenciam como essa ordem é efetuada nos contextos de informação. Embora a classificação tradicional seja rigorosa no sentido de que exige que uma entidade seja ou não membro de uma classe específica, o processo de categorização é flexível e criativo e pode desenhar associações que não são baseadas

em um conjunto de princípios predeterminados, mas no simples reconhecimento de semelhanças não vinculativas que existem em um conjunto de entidades.

Na busca de parâmetros para esse processo flexível apontado por Jacob (2004), são apresentados fundamentos para uma categorização, proposta por estas autoras, baseada nos três tipos de “concepção de relevância” apregoados por Shutz (1970), no contexto cognitivo do mundo da vida, mundo este que agrega especificidades das diversas vulnerabilidades encontradas pela mulher no seu dia a dia, em busca de equidade e direitos.

A partir de tal tipologia, o primeiro tipo encontra-se na relevância motivacional, e está regido por interesses encontrados no ODS 5, metas e interligações da Agenda 2030, que prevalecem na situação específica de mesclar, de forma equilibrada, as três dimensões do desenvolvimento sustentável – a econômica, a social e a ambiental – para alcançar a igualdade de gênero e empoderar todas as mulheres e meninas. Esses elementos presentes nesta situação servem para definir a relevância da categorização à luz de seus propósitos. O princípio revelado pela relevância motivacional opera satisfatoriamente somente na situação volitiva escolhida.

O segundo tipo está pautado na relevância temática. É o princípio que vai determinar os elementos que definem o problema em dada situação, em razão de interesses específicos. A problemática disposta nos ODSs se torna relevante somente na medida em que representa um processo de categorização das proposições legislativas.

O terceiro registra-se na relevância interpretativa, que se apresenta como extensão do segundo princípio. Ao reconhecer o problema (como as proposições legislativas que tratam dos direitos em construção das mulheres se qualificam com vistas a alcançar o ODS 5, metas e interligações da Agenda 2030), a formulação, enquanto tal, necessita de interpretação mais aprofundada. A partir deste princípio da relevância interpretativa, é possível verificar a necessidade de uma categorização que atenda a medição da eficácia das proposições em razão dos ODSs da Agenda 2030.

A partir dessas três noções de relevância, pretende-se construir, em um método *bottom-up*, a categorização dos conceitos que representam as proposições, considerando um agrupamento não hierárquico, mas funcional e flexível, atendendo, assim, as reais necessidades das mulheres brasileiras de acordo com os ODSs da Agenda 2030.

A figura 1 foi construída com base no *The Zendesk Triple Diamond*, desenvolvido por Mike Chen, Kim Lenox e Jennifer Chang, a fim de facilitar a visualização do fluxo para a construção da categorização proposta. A opção pelo modelo dos diamantes foi útil por ajudar na compressão das convergências que possibilitam a descoberta de recursos, processo e procedimento para a solução de problemas.

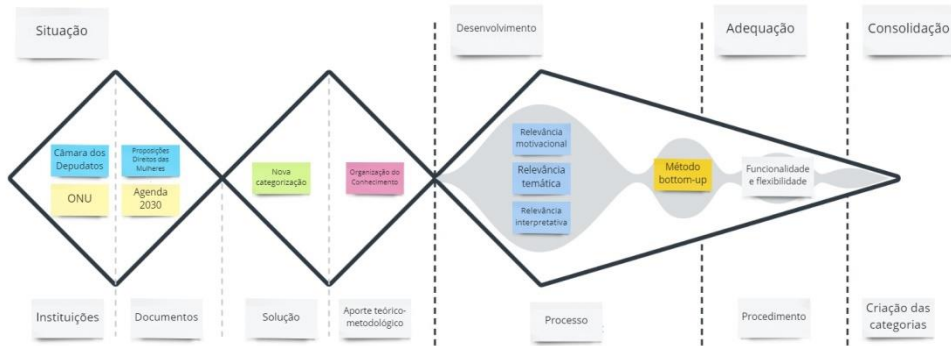


Figura 1. Situação, desenvolvimento, adequação e consolidação da criação das categorias (adaptada do The Zendesk Triple Diamond – Chen, 2020)

Assim, em uma adaptação para esta pesquisa, do modelo triplo Zendesk (Chen, 2020), apresenta-se na 1.^a fase, o contexto da situação: instituições e documentos; Câmara dos Deputados e as proposições que tramitam sobre os direitos em construção das mulheres + a ONU com a Agenda de 2015, que tem objetivos e metas a serem cumpridos até 2030. Segue-se para a solução encontrada – nova categorização baseada no aporte teórico-metodológico da Organização do Conhecimento. Na 2.^a fase, revela-se o desenvolvimento do processo baseado nas relevâncias motivacional, temática e interpretativa e o encaminhamento para uso do método *bottom-up*, que permite procedimentos funcionais e flexíveis para consolidar a criação das categorias na 3.^a fase.

Conclusões

Ao se buscar fundamentos teórico-metodológicos para alicerçar a categorização das proposições em tramitação na Câmara dos Deputados, pretendeu-se encontrar uma OC que apreciasse a aferição da qualidade e da eficiência das iniciativas legislativas atendendo à justiça e inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade prescritos na Constituição Federal (Constituição do Brasil, 1988), que se coaduna com os ODSs da Agenda 2030.

As referências teórico-metodológicas pesquisadas representam um caminho para a elaboração da categorização pretendida e para a sistematização mais adequada das proposições que abordam os direitos em construção das mulheres.

As categorias devem observar a tipologia das relevâncias motivacional, temática e interpretativa, ser flexíveis, criativas e desenhar associações não baseadas em um conjunto de princípios predeterminados, mas, sim, num simples reconhecimento de semelhanças necessárias para demonstrar a qualidade e a eficácia das proposições em tramitação em relação à Agenda 2030.

Nessa esteira, as diferenças a ser produzidas na formação das novas categorias têm associações situadas nos ODSs que respondem pelas necessidades das mulheres. É uma organização que se baseia em um conhecimento situado no momento em que a premência da utilidade exige.

Desta forma, a OC passa a englobar apreciações críticas e pluralistas contemporâneas das verdadeiras necessidades e vulnerabilidades das mulheres, registradas na Agenda 2030.

Como já mencionado no V Relatório Luz (Grupo de Trabalho da Sociedade Civil para a Agenda 2030, 2021), o Brasil precisa apresentar resultados mais contundentes que demonstrem suas reais ações para alcançar os ODSs, para tanto, precisa reunir dados confiáveis e efetivos, podendo, estes, vir de iniciativas e fontes de terceiros. O estudo teórico-metodológico para a nova categorização das proposições sobre direitos em construção das mulheres se apresenta como estratégia eficaz para demonstrar se o Legislativo brasileiro está formulando proposições que representem avanços ou retrocessos para as mulheres.

Fato agravante que demonstra a necessidade da proposta de construção de categorias, que possa aferir a eficácia das proposições, ou seja, dados organizados para a verificação de toda a sociedade e pesquisadores, é que o governo brasileiro não apresentou o *Voluntary National Review* (VNR) na última edição, em julho de 2021, do fórum da ONU com o tema “Recuperação sustentável e resiliente da pandemia COVID-19 que promove as dimensões econômica, social e ambiental do desenvolvimento sustentável: construindo um caminho inclusivo e eficaz para o cumprimento da Agenda 2030 no contexto da década de ação e entrega para o desenvolvimento sustentável”. Diferentemente do Brasil, 42 países realizaram as análises nacionais voluntárias da implementação da Agenda 2030, demonstrando resultados e intenções pós-pandemia no alcance dos ODSs (United Nations, 2021).

A este fato, acrescenta-se a informação da Secretaria das Mulheres da Câmara dos Deputados, de que até o dia 28 de agosto de 2021, havia 1.350 “Proposições sobre a Mulher em Tramitação” (Câmara dos Deputados, 2021), *corpus* que será categorizado para prover dados organizados e disponíveis para acesso com vistas à Agenda 2030

Também motivaram esta pesquisa os princípios de “promover instituições fortes, inclusivas e transparentes, a manutenção da paz e o respeito aos direitos humanos baseados no Estado de direito são a base para o desenvolvimento humano sustentável” (United Nations, 2015) que sustentam as metas do ODS 16. A transparência corrobora a intenção de sistematização de nova categorização para as proposições legislativas que pretende medir o grau de eficácia, em razão dos ODSs.

A meta 10 do ODS 16, que visa “assegurar o acesso público à informação e proteger as liberdades fundamentais, em conformidade com a legislação nacional e os acordos internacionais”, em conjunto com a meta 5 do ODS 9, que tem como uma das premissas “fortalecer a pesquisa científica [...]”, incentivam estudos que tornem mais acessíveis as informações públicas.

Os achados registrados neste artigo serão aprofundados na pesquisa de tese em andamento, com vistas a apresentar como resultado uma nova categorização, para conformar os dados sobre as proposições que tratam dos direitos em construção das mulheres, que tramitam na Câmara dos Deputados, aos ODSs da Agenda 2030.

Referências

- Azcona, G., & Bhatt, A. (2020). Inequality, gender, and sustainable development: measuring feminist progress, *Gender & Development*, 28(2), 337-355.
- Câmara dos Deputados [do Brasil] (2016). *Manual de indexação de proposição legislativa*. Brasília: Edições Câmara.
- Câmara dos Deputados [do Brasil] (2021). *Proposições sobre a Mulher em Tramitação*. Recuperado em 28 agosto, 2021, de <https://www2.camara.leg.br/a-camara/estruturaaadm/secretarias/secretaria-da-mulher/proposicoes-sobre-a-mulher/todas>.
- Chen, Mike (2020). *The Zendesk Triple Diamond*. Recuperado em 26 agosto, 2021, de <https://medium.com/zendesk-creative-blog/the-zendesk-triple-diamond-process-fd857a11c179>.
- Constituição da República Federativa do Brasil de 1988* (1988). Brasília: Senado Federal.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: qualitative, quantitative, and mixed methods approaches* (5nd ed.). Thousand Oaks: SAGE Publications.
- Dahlberg, I. (1978a). Teoria do conceito, *Ciência da Informação*, 7(2), 101-107.
- Dahlberg, I. (1978b). *Ontical structures and universal classification*. Bangalore: Sarada Ranganathan Endowment.
- Dahlberg, I. 2006. Knowledge organization: a new science? *Knowledge Organization*, 33(1), 11-19.
- Gil, A. C. (1994). *Métodos e técnicas de pesquisa social*. São Paulo: Atlas.
- González Casanova, P. (1996). Clasificaciones y Definiciones: Nota para un Bibliotecario, *Investigación Bibliotecológica*, 10(20), 3-8.
- Grupo de Trabalho da Sociedade Civil para a Agenda 2030 (2021). *V Relatório Luz da Sociedade Civil Agenda 2030 de Desenvolvimento Sustentável Brasil*. Recuperado em 26 agosto, 2021, de https://brasilnaagenda2030.files.wordpress.com/2021/07/por_rl_2021_completo_vs_03_lowres.pdf.
- Hjørland, B. (2008). What is knowledge organization (KO)? *Knowledge Organization*, 35(3), 86-101.
- Jacob, E. K. (2004). Classification and categorization: A difference that makes a difference, *Library Trends*, 52(3), 515-540.
- Kobashi, N. Y. (1996). Análise documentária e representação da informação, *Informare: Cadernos do Programa de Pós-Graduação em Ciência da Informação*, 2(2), 5-27.
- Miranda, M. L. C. de (1999). A organização do conhecimento e seus paradigmas científicos: algumas questões epistemológicas, *Informare: Cadernos do Programa de Pós-Graduação em Ciência da Informação*, 5(2), 64-77.
- Morris, R. C., & Spivak, G. C. (2010). *Can the subaltern speak? Reflections on the history of an idea*. New York: Columbia University Press.
- Olson. H. A. (1998). Education For Cataloging Is/As Women's Studies, *The Serials Librarian*, 35(1/2), 153-166.
- Olson. H. A. (2007). How We Construct Subjects: A Feminist Analysis, *Library Trends*, 56(2), 509-541.
- Razavi, S. (2016). The 2030 Agenda: challenges of implementation to attain gender equality and women's rights, *Gender & Development*, 24(1), 25-41.
- Sammakieh, S., Youssef, J., & Salem, J. (2021). *Implementation of Sustainable Development Goals: addressing the data disaggregation challenge in the GCC*. Nova York: Oliver Wyman. Recuperado em 26 agosto, 2021, de <https://www.oliverwyman.com/our-expertise/insights/2021/jan/implementation-of-sustainable-development-goals.html>.

- Schneider, M. (2013). Ética, política e epistemologia: interfaces da informação. In Albagli, S. (Org.). *Fronteiras da Ciência da Informação* (pp. 57-77). Brasília: Ibict.
- Schutz, A. (1970). *Reflections on the Problem of Relevance*. New Haven: Yale University Press.
- Shera, J. H. (1961). Social Epistemology, General Semantics, and Libraries, *Wilson Library Bulletin* 35(3), 767-70.
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge: Harvard University Press.
- United Nations (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*. New York: United Nations. Recuperado em 28 agosto, 2021, de http://www.un.org/ga/search/view_doc.asp?symbol=A/70/L.1&referer=/english/&Lang=E.
- United Nations (2021). *High-level Political Forum on Sustainable Development*. Recuperado em 28 agosto, 2021, de <https://sustainabledevelopment.un.org/hlpf/2021#vnrs>.
- World Economic Forum (2021). *Global Gender Gap Report 2021*. Recuperado em 28 agosto, 2021, de http://www3.weforum.org/docs/WEF_GGGR_2021.pdf.

(Página deixada propositadamente em branco)

O OBJETIVO JUSTIÇA E PAZ DA AGENDA 2030/ONU POR MEIO DE DADOS DO PORDATA

Camila Mattos da Costa¹

¹ PPGCI IBICT-UFRJ, camilamctagmail.com, ORCiD iD
<https://orcid.org/0000-0002-5738-3087>¹

Resumo

Esta comunicação objetiva apresentar e analisar o portal divulgação de dados estatísticos PORDATA. A análise exposta está centrada nos dados disponíveis sobre justiça, paz e fortalecimento das instituições na página do portal. Isso decorre do interesse em perceber como tal objetivo está diretamente relacionado com a segurança pública, inserindo-o em uma perspectiva que coloca o conhecimento e a informação como indispensáveis para o alcance da sustentabilidade econômica, ambiental e social. O PORDATA é um sítio na internet destinado a recolher e organizar a informação de modo a torná-la mais acessível, além de produzir metainformação, o que permite a apropriada interpretação dos dados de modo gratuito. O portal disponibiliza informações provenientes de diversas entidades oficiais, com especial colaboração do Instituto Nacional de Estatística português. Desde março de 2021, o PORDATA exhibe informações estatísticas relacionadas aos indicadores de desenvolvimento sustentável. O objetivo 16 da Agenda 2030, objeto de nossa de análise, é a promoção de sociedades pacíficas e inclusivas para o desenvolvimento sustentável, proporcionando o acesso à justiça para todos e construção de instituições eficazes, responsáveis e inclusivas a todos os níveis. A metodologia tem caráter exploratório e documental. Conclui-se que a ampliação do debate a respeito da segurança pública e da organização do conhecimento favorece o reconhecimento de que só é possível realizar políticas públicas eficientes baseadas em dados e estatísticas sólidos e, muitas vezes, de longo prazo para evitar decisões distantes das necessidades reais dos indivíduos. A criação de indicadores para o monitoramento do progresso para o alcance de cada um dos objetivos da Agenda 2030 prescinde de dados de qualidade, acessíveis, atualizados, confiáveis e desagregados, provenientes de fontes oficiais oriundos dos países e coletados em parceria com organizações regionais e internacionais, do modo como são apresentados pelo portal.

Palavras-chave: *PORDATA, Informação em Segurança Pública, Desenvolvimento Sustentável.*

¹ O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Introdução

O paradigma custodial dissolve-se cada vez mais rápido e dá lugar a um paradigma pós-custodial, científico e informacional (Malheiro da Silva & Ribeiro, 2012). A informação é gerada por agentes públicos e privados. Depois, usuários apropriam-se dela e apresentam novas leituras para problemas da sociedade por meio do cruzamento de dados extraídos dos registos. O fenómeno ocorre com muita velocidade, afetando diversas áreas tais como o planeamento urbano e a segurança pública, por exemplo.

Este trabalho objetiva analisar de que maneira ocorre transformação de registos produzidos por entidades públicas com vistas a produzir conhecimento e de que maneira os dados sobre justiça e paz são apresentados aos usuários no que se refere a segurança pública por meio do PORDATA. Além de compreender como o modo pelo qual são exibidos tais dados é capaz de remeter ao processo de gênese da informação. A exposição de estatísticas criminais e de justiça ressalta práticas informacionais, percebidas como fenómenos complexos, logo, politicamente inseridos. Portanto, o desenvolvimento sustentável deve ser considerado a partir de suas três dimensões, a saber, a social, a económica e a ambiental.

Há estudos que apontam as mudanças nas formas de representação visual dos modelos da organização do conhecimento, indicando que as mudanças nos volumes de dados influenciam os modelos de organização. A disponibilização para um grupo maior de usuários exige novas formas de representação da informação e do conhecimento (Clavier, 2019). São incipientes os estudos que apontam as implicações políticas do campo da organização do conhecimento e a busca por práticas profissionais que abracem as diferenças socioculturais em busca de sociedades mais justas (Leazer & Montoya, 2020).

No âmbito da gestão do conhecimento, destaca-se que é um terreno em disputa em contextos socioculturais mediados por dispositivos sócio-técnicos (Moura, 2017). Foi possível observar a relação entre o conhecimento e os objetivos da Agenda 2030 (Geraldo & Pinto, 2019. ver também Ochôa & Pinto, 2019). Estudos anteriores relacionam organização e gestão do conhecimento com a sustentabilidade (Martins et. al., 2019), referindo-se com frequência aos objetivos relativos a sustentabilidade ambiental e ao planeamento urbano, por meio da discussão a respeito das cidades inteligentes (ver Barros, 2017; Mikalauskiene & Atkociuniene, 2019; Souto & Pizzol, 2019; Martins & Santos, 2019; Baracho, 2020; Balaji & Dhanamjaya, 2020; Israilidis et. al, 2021). Destaca-se que as cidades inteligentes ainda devem ser vistas como espaços políticos centrados no cidadão e na transparência (Keymolena & Voorwinden, 2020). O crescimento da adoção de tecnologias da informação e comunicação tem aumentado a digitalização de serviços prestados aos cidadãos, mas é preciso lidar com complexidade e integração dos dados (Bokolo, 2021). Isso inclui governos e agentes externos como instituições, cidadãos e sistemas sócio-técnicos (Yahia et. al, 2021), como é o caso da criação do PORDATA por uma instituição extra-governamental. A importância dos dados abertos em cidades inteligentes para a elaboração de políticas públicas (Zuiderwijk & Hinnantb, 2018) também já foi anteriormente analisada.

Neste trabalho, analisaremos o portal Pordata (<https://www.pordata.pt/Home>) e os dados disponíveis sobre justiça e segurança pública. A Base de Dados de Portugal Contemporâneo ou PORDATA é organizada e desenvolvida pela Fundação Francisco Manuel dos Santos, criada em 2009. A página recolhe, organiza, sistematiza e divulga informações sobre diversas áreas da sociedade de Portugal e de seus municípios, além de outros países da Europa. São divulgadas estatísticas provenientes de fontes oficiais certificadas e com a devida competência para produzir informações nas áreas específicas. O PORDATA destina-se a recolher, organizando a informação para que se torne mais acessível, além de produzir metainformação, o que permite uma interpretação mais correta dos dados. A disponibilização das estatísticas não tem custo e apresenta informações provenientes de diversas entidades oficiais, com especial colaboração do Instituto Nacional de Estatística português.

Em março de 2021, o PORDATA passou a disponibilizar informações estatísticas relacionadas aos indicadores de desenvolvimento sustentável. Restringiremos nossa análise aos dados relativos à segurança pública, relacionando-a diretamente relacionada ao objetivo 16 da Agenda 2030. O objetivo 16 busca *promover sociedades pacíficas e inclusivas para o desenvolvimento sustentável, proporcionar o acesso à justiça para todos e construir instituições eficazes, responsáveis e inclusivas a todos os níveis*.

A elaboração deste tipo de ferramenta tem crescido, exigindo posicionamento científico e métodos apropriados diante da informação socialmente produzida por parte dos gestores e usuários da informação.

Este trabalho objetiva analisar o modo como ocorre a transformação de registos produzidos por entidades públicas, sua transformação em conhecimento na segurança pública e sua disponibilização aos usuários do PORDATA e em que medida esse tipo de iniciativa relaciona-se com os objetivos de desenvolvimento sustentável proposto pela Agenda 2030.

As infraestruturas de comunicação e informação utilizam-se de dados públicos e/ou privados a respeito da criminalidade registada para permitir novos modos de visualização de ocorrência criminal, assim como apresenta o PORDATA.

Metodologia

Esta pesquisa possui caráter exploratório e documental. Relacionaremos a abordagem teórico-conceitual da organização do conhecimento, correlacionando-a com o tema da segurança pública para, à luz do referencial teórico, analisarmos em que medida iniciativas como PORDATA favorecem o acompanhamento e implementação dos objetivos da Agenda 2030, conseqüentemente, um mundo mais sustentável. Pretendemos identificar de que fonte obtém suas informações, a disponibilidade dos dados para o usuário, a navegabilidade, a usabilidade, o organismo responsável e a frequência de atualização das mesma com foco na análise das seções *Justiça e Segurança e Paz, Justiça e Instituições Eficazes*.

Destaca-se que a opção metodológica de restringir a análise no que se refere ao objetivo da justiça e da paz não desconsidera que o problema da violência é multifatorial e está conectado com os demais objetivos da Agenda. No entanto, o recorte

objetiva analisar com maior profundidade a questão das estatísticas a respeito da violência e das instituições de justiça criminal por considerarmos que há uma lacuna a ser preenchida com relação aos aspetos da justiça social e da sustentabilidade ambiental.

Resultados

O alcance de um mundo mais sustentável exige preocupação com as questões relativas a defesa da vida e do património. O resultado alcançado é a reflexão teórica a respeito da promoção da justiça e da paz enquanto objetivos da Agenda 2030 da Organização das Nações Unidas a partir dos dados apresentados e de que maneira portais como o PORDATA funcionam como mecanismos auxiliam o alcance destes objetivos na medida em que favorecem que as informações sejam de conhecimento público.

O portal extrai os dados estatísticos dos países membros da UE de outro portal: o EUROSTAT². Também podem ser extraídos de outras entidades como o UNODC, entidades nacionais e a OMS, por exemplo. A navegabilidade e usabilidade do portal simplificam o entendimento das estatísticas. O portal também oferece cursos presenciais e online, além de tutoriais para melhorar a experiência do usuário.

A partir da observação do PORDATA, foi possível identificar que os campos relativos aos objetivos da justiça e da paz apresentam dados da seguinte maneira: os dados sobre Justiça e Segurança estão divididos em quatro secções³. A saber, 1 Crimes, que apresenta as estatísticas de criminalidade; 2 Processos, que apresenta as estatísticas de processos cíveis e/ou comerciais e criminais entrados, findos e pendentes, os processos criminais entrados, findos e pendentes, o saldo processual por tipo de processos, taxa percentual de congestão por tipo de processos e a de eficácia por tipo de processos; 3 Prisões com estatísticas a respeito da população em situação de cárcere; e 4 Recursos Humanos empregados na magistratura e nas forças policiais. Nessa secção são apresentados dados de Portugal e dos demais países que formam a União Europeia.

A secção Crimes divide-se em absolvidos, arguidos por agressão sexual por sexo, arguidos por homicídio intencional por sexo, arguidos por nacionalidade, arguidos por sexo, arguidos por violação por sexo, percentual de condenados no total de arguidos, condenados por 100 mil habitantes, condenados por agressão sexual por sexo, condenados por homicídio internacional por sexo, condenados por nacionalidade, condenados por sexo, condenados por violação por sexo, por categoria, vítimas de agressão por sexo e vítimas de violação por sexo. Em Portugal, podem ser observadas as seguintes taxas. O número de pessoas absolvidas saiu de 22.817 indivíduos em 2008 para 12.516 indivíduos em 2019, o que pode ser explicado pela redução de crimes e indivíduos arguidos, conforme pode ser visto a seguir. O total de indivíduos arguidos por agressão sexual foi 446 (432 homens e 14 mulheres) em 2008 e 542 (526 homens e 16 mulheres) em 2019. Foram 192 (180 homens e 12

² <https://ec.europa.eu/eurostat>

³ A análise foi realizada em agosto de 2021.

mulheres), os indivíduos arguidos por homicídio intencional em 2008 e 74 (66 homens e 8 mulheres) em 2019, representando uma diminuição. O número de indivíduos arguidos caiu de 141.109 (125.716 portugueses e 15.302 estrangeiros) para 67.030 (59.103 portugueses e 7.927 estrangeiros), o que representa uma redução significativa. O número de indivíduos condenados foi de 87.861 (77.172 nacionais e 10.689 estrangeiros) em 2008 para 45.053 (39.544 nacionais e 5.509 estrangeiros). O total de arguidos pelo crime sexual de violação foram 141 (140 homens e 1 mulher) indivíduos em 2008 e 149 (140 homens e 1 mulher) indivíduos em 2019. Do total de arguidos, 62,3% foram condenados em 2008 e 67,2% em 2009, aumentando a proporção de condenados. Os condenados por 100 mil habitantes em 2008 representavam 832,2% e em 2019 esse número caiu para 438,0. Os condenados por agressão sexual foram 271 (267 homens e 4 mulheres) indivíduos em 2008 e 398 (388 homens e 10 mulheres) indivíduos em 2019, o que representa um aumento. O número de indivíduos condenados por homicídio intencional caiu de 121 (113 homens e 8 mulheres) para 62 (56 homens e 6 mulheres). Os indivíduos condenados por violação eram todos do sexo masculino e representaram o total de 90 em 2008 e 88 em 2019. Foram 1.662 (318 homens e 1.344) vítimas de agressão sexual em 2008 e 1.761 (259 homens e 1502 mulheres) em 2019. Foram 209 (188 do sexo feminino e 21 do sexo masculino) vítimas de violação em 2019 e 251 (227 do sexo feminino e 24 do sexo masculino) e em 2018.

O crime é entendido como a infração penal de violação legal atribuída no registo pela polícia. O crime de agressão apresentou 661 registos em 2019. Os homicídios intencionais reduziram de 143 em 1994 para 73 em 2019. Foram registados 338 crimes de sequestro em 2019 e não há dados anteriores. Registaram-se 2.701 crimes de violência sexual em 2019. Em 1994, foram registados 14.322 roubos, decaindo para 10.926 em 2019. Foram registados 21.432 assaltos em 2019. Registaram-se 82.120 furtos em 2019. Por último, os registos de atos ilícitos que envolvem drogas e precursores controlados em 1994 foram 4.517, subindo para 5.749 em 2019.

Em processos, são apresentados dados sobre Processos cíveis e/ou comerciais nos tribunais de 1.ª instância: entrados, findos e pendentes, Processos criminais nos Tribunais de 1.ª instância: entrados, findos e pendentes, saldo processual dos Tribunais de 1.ª instância: tipo de processos, taxa de congestão dos Tribunais de 1.ª instância: tipo de processos (%), taxa de eficácia dos Tribunais de 1.ª instância: tipo de processos (%), taxa de resolução dos Tribunais de 1.ª instância: tipo de processos (%).

Os processos cíveis e/ou comerciais nos tribunais de 1.ª instância correspondem ao número casos cíveis e/ou comerciais novos que estão aguardando decisão ou são resolvidos pelos Tribunais de 1.ª instância. Houve um aumento de 314.467 em 2008 processos iniciados para 323.236 em 2019. O número de processos findos também teve um aumento de 312.007 para 339.370. Já o número de processos pendentes teve uma redução significativa de 362.897 para 186.351. Os processos criminais nos Tribunais de 1.ª instância representam o número de casos criminais novos, que aguardam decisão ou são resolvidos pelos Tribunais de 1.ª instância. Os números processos iniciados correspondem a 114.960 em 2008 e 64.339 em 2019. Os processos findos caíram de 129.737 em 2008 para 65.269 em 2019. O número de processos pendentes reduziu de 102.911 em 2008 para 38.673 em 2019. O saldo processual que corresponde a diferença entre o número de processos criminais entrados e resol-

vidos nos Tribunais de 1.^a instância e superiores foi de -14.777 em 2008 e -930 em 2019. O saldo processual de cíveis e/ou comerciais foi 2.460 em 2008 e -16.134 em 2019. A taxa de congestão dos Tribunais de 1.^a instância que corresponde a percentagem de processos pendentes face aos processos resolvidos pelos Tribunais de dessa instância. Eram 90,7% de processos criminais pendentes em 2008 contra uma redução para 60,7% em 2019. Quanto aos processos cíveis e/ou comerciais, havia uma taxa de 115,5% em 2008 contra 59,7%. A taxa de eficácia dos Tribunais de 1.^a instância, que equivale percentual de processos resolvidos pelos Tribunais de 1.^a instância, diante do conjunto dos processos novos e que aguardam decisão. Em 2008, a taxa de processos criminais resolvidos era de 55,8%, passando a 62,8%. Já as taxas relativas aos processos cíveis e/ou comerciais foram 46,2% em 2008 e 64,6%. A Taxa de resolução dos Tribunais de 1.^a instância, que significa a percentagem de processos resolvidos pelos Tribunais de 1.^a instância, frente aos novos processos. Em 2008, a 112,9% dos processos criminais face a 101,4% em 2019. Já dos processos cíveis e/ou comerciais foi 99,2% em 2018 e 105,0% em 2019.

Os números relativos às prisões apontam a quantidade de pessoas presas em relação à lotação das cadeias. O dado lotação oficial que corresponde ao número previsto de lugares disponíveis a 31 de dezembro sem sobrelotação, excluindo locais/capacidade utilizados para a detenção de pessoas com base no estatuto de imigração. Eram 12.464 indivíduos em 2008 e passaram a 12.934 em 2019. O número de reclusos era de 10.807 em 2008 e 12.947 em 2019, significando uma diferença entre o número de vagas no sistema prisional e o número de ocupantes. Havia em 2019 mais ocupantes do que vagas. A ocupação efetiva em percentual era de 87%, passando a 100% em 2019. O pessoal ao serviço em estabelecimentos prisionais para adultos era de 5.808 indivíduos em 2008, passando a 5.346 indivíduos em 2019, o que significa uma diminuição, apesar do aumento de reclusos. Havia em 2008 1.246 indivíduos a serviço em estabelecimentos prisionais para menores, passando 193 em 2019. Eram 227 indivíduos reclusos por agressão sexual em 2008 – não há dados sobre o sexo dos reclusos – e 209 indivíduos em 2019 (208 homens e apenas 1 mulher). Os reclusos por homicídio intencional eram 340 (327 homens e 13 mulheres) indivíduos em 2008 e passaram a 249 (234 homens e 15 mulheres) indivíduos em 2019. Os reclusos por nacionalidade já foram apresentados anteriormente.

Em 1993, havia 11.079 reclusos em Portugal e o portal não informa quantos eram do sexo masculino e quantos eram do sexo feminino. Em 2019, eram 12.947 reclusos, sendo 12.071 homens e 876 indivíduos do sexo feminino. O número de reclusos que aguardavam julgamento em 2008 era de 2.108 e 2.271 indivíduos em 2019. O número de reclusos com sentença proferida em 2008 era de 8.699 e 10.522 indivíduos em 2019. O total de reclusos por violação era de 209 indivíduos em 2008 e 148 indivíduos em 2019. Todos os inclusos por violação nos dois anos apresentados eram do sexo masculino. Por meio desses dados é possível observar o tempo de tramitação dos processos na justiça. A redução do número de processos afeta diretamente o objetivo 16, pois é capaz de influenciar a confiança na população nas instituições de justiça. Os recursos humanos apresentam estatísticas sobre o quantitativo de Magistrados Judiciais por 100 mil habitantes, ou seja, o número de juizes e procuradores dos Tribunais de 1.^a instância e superiores por cada 100 mil residentes. Houve um pequeno aumento percentual de 16,2% em 2008 para 16,9% em 2019.

Em 2008, Portugal tinha 1.712 magistrados judiciais. Destes, 824 indivíduos do sexo masculino e 888 do sexo feminino. Em 2019, esse número subiu para 1.734 indivíduos, sendo 663 homens e 1.071 mulheres. O percentual de pessoas a serviço das polícias por 100 mil habitantes era de 434,2% em 1996 e 441,7% em 2019. O número total de indivíduos a serviço das polícias em 1996 era de 43.695 indivíduos. Não há indicadores sobre o sexo das forças policiais no período. Em 2019, eram 45.431 indivíduos, sendo 41.613 indivíduos do sexo masculino e 3.818 do sexo feminino.

Não há dados sobre a União Europeia consolidados em muitas das estatísticas apresentadas, apesar de ser possível visualizar os números de outros países da UE. Percebe-se a ausência de sistematização de muitas informações relativas à UE.

Os indicadores apresentados pelo PORDATA dentro do objetivo 16, ou seja, diretamente ligados à meta de desenvolvimento sustentável, em área intitulada Justiça e Paz dividem-se em seis secções. A taxa padronizada de mortalidade por homicídio corresponde às mortes causadas por outra pessoa, com intenção de ferir ou matar, por 100 mil residentes. Em 2000 era de 0,9%, reduzindo para 0,8% em 2018. A proporção da população que reporta ocorrências de crime, violência e vandalismo na sua área era de 13,6% em 2004 e 6,7% em 2019 em Portugal. Houve uma diminuição que pode ser explicada pela diminuição de números de crimes. Na UE, a taxa estimada era de 11%. As despesas em tribunais *per capita* era de 34,8€ em 2000. Tinha o valor provisório 68,1€ em 2019. Já na UE, o valor correspondia a 100,4€.

Houve um aumento percentual na perceção de independência do sistema judicial, subindo de 33,0% em 2016 para 48,0% de pessoas que consideravam boa ou muito boa em Portugal. A perceção má ou muito má caiu de 44,0% para 43%. O percentual de pessoas que desconheciam reduziu de 23,0% em 2016 para 10% em 2019, o que pode ser explicado pela maior disponibilidade dos dados, pois auxilia no processo de formação de opinião, mas não apenas. Na UE, o percentual tinha o valor estimado de 50% em 2016 e passou a 54%⁴ em 2021. O índice de perceção de corrupção em Portugal reduziu de 63,0 para 61,0, sendo 100 o mais confiável e 0 o menos confiável. Já na UE não foram apresentados índices.

A proporção da população com confiança nas instituições da União Europeia, ou seja, que afirmam confiar no Parlamento Europeu corresponde a 56,0% em 2016 e 80% em 2021 em Portugal e 50% na UE. Quanto à Comissão Europeia, 52,0% em 2016 e 83,0% em 2021 em Portugal; na UE, 47,0% em 2021. A confiança no Banco Central Europeu era de 51,0% em 2016 e subiu para 77,0% em 2021; na UE, o índice era de 47,0%. É possível observar que tem sido crescente a melhora nos índices relativos ao objetivo 16 da Agenda 2030 da ONU.

Foi possível observar uma discrepância nas informações fornecidas pelos países membros da UE ao longo da análise, o que dificulta a sistematização de indicadores para a avaliação das metas da Agenda 2030.

As metainformações sobre os dados em ordem de apresentação incluem o âmbito geográfico, operação estatística, tipo de operação estatística, período ou momento de referência, periodicidade, entidade responsável, notas, data do último acesso, con-

⁴ Houve uma quebra de série.

ceitos, série e notas de âmbito geográfico. A simbologia dos dados também é apresentada. Essas informações auxiliam o usuário a compreender o processo de extração e organização das informações, contextualizando-as. É possível também entender as categorias de classificação utilizadas que muitas vezes estão baseadas nos códigos penais e/ou outros códigos de classificação como, por exemplo, a classificação de crimes que se dá pela categorização de acordo com a Classificação Internacional do Crime com Fins Estatísticos (International Classification of Crime for Statistical Purposes – ICCS). Já as causas de morte utilizam a “Lista Sucinta Europeia” que se orienta pela Classificação Estatística Internacional de Doenças e Problemas Relacionados com a Saúde (CID).

Segundo os metadados do portal, diversos fatores alteram as estatísticas de criminalidade. Em especial, os diferentes sistemas jurídicos e de justiça penal; diferentes fases de registo da infração; regras diferentes para a contagem da infração penal; tipos diferentes de infrações; diferença nas taxas reportadas e diferentes formas de organização e eficiência da justiça criminal nos países que formam a UE. Portanto, para melhorar a análise dos indicadores seria interessante a padronização das informações no momento da criação dos registos, o que não é fácil. Afimal, são dados oriundos de administrações da justiça nos diferentes países do bloco europeu.

A partir dos dados apresentados, observa-se uma diminuição dos crimes em Portugal e um aumento da confiança dos indivíduos nas instituições de justiça. O crescimento da confiança também fortalece que crimes sejam denunciados na medida que o cidadão sente-se mais confortável para procurar ajuda da justiça como, por exemplo, acontece nos crimes de agressão sexual que registaram aumento.

Os dados podem ser visualizados em gráficos em ranking ou utilizados em ranking (embed) e podem ser exportados em excel ou pdf.

A atualização da base de dados é periódica. Observou-se que a base possui dados atualizados em 2020 e 2021.

O EUROSTAT, portal de onde são extraídos diversos dados encontra-se em inglês, francês e alemão. A tradução para o português auxilia o usuário que não domina os idiomas citados anteriormente.

Além das estatísticas apresentadas anteriormente, também podem ser encontradas outras informações sobre Justiça e Segurança de Portugal, são disponibilizadas informações sobre 1 Atos Notariais e Registros; 2 Crimes; 3 Menores e Adoção; 4 Prisões; 5 Processos; 6 Rede Judiciária e; 7 Recursos Humanos. Sobre os municípios são apresentados dados a respeito atos notariais e seus registos, crimes e processos. No que se refere aos dados apresentados sobre Justiça e Segurança de Portugal, são disponibilizadas informações sobre 1 Atos Notariais e Registros; 2 Crimes; 3 Menores e Adoção; 4 Prisões; 5 Processos; 6 Rede Judiciária e; 7 Recursos Humanos.

Sobre os municípios portugueses são apresentados dados a respeito atos notariais e seus registos, crimes e processos.

Para que governos e a sociedade civil consigam acompanhar a execução das metas da Agenda 2030 é preciso que sejam disponibilizadas informações de todos os países signatários. É possível que as estatísticas que não puderam ser observadas nesse estudo sejam posteriormente inseridas. A continuidade da apresentação das análises é necessária para que usuários conheçam e acompanhem os indicadores.

A apresentação de dados precisa estar acompanhada de engajamento por parte dos cidadãos e complementados por mecanismos que garantam este envolvimento e auxiliem a interpretação dos dados, a governança e arranjos institucionais (Matheus et. al, 2020), o que pode ser verificado no PORDATA na medida em que os dados apresentam metadados que auxiliam sua interpretação por parte de pessoas interessadas.

Conclusões

O monitoramento do progresso para o alcance de cada um dos objetivos da Agenda prescinde de dados procedentes de países envolvidos com o pacto global e devem ser coletados em parceria com organizações regionais e internacionais. Portanto, a produção de indicadores depende que esses dados sejam disponibilizados com qualidade, acessíveis, atualizados, confiáveis e desagregados, provenientes de fontes oficiais, assim como são apresentados pelo PORDATA.

Os dados sobre justiça que apontam para a diminuição de crimes ajudam a inferir também o modo como outros objetivos da Agenda 2030 tais como a erradicação da pobreza e da fome, a promoção da equidade de gênero e da educação de qualidade, o crescimento econômico e a promoção de trabalho digno e a redução das desigualdades impactam nas estatísticas de criminalidade, reduzindo crimes.

Conclui-se que a análise do PORDATA com foco na promoção da justiça e da paz possibilita promover o entendimento a respeito do caminho a ser percorrido para o alcance do objetivo 16 e a garantia do respeito a vida e a propriedade, apesar de reconhecermos que apenas uma base de dados é incapaz de demonstrar toda a complexidade do problema da segurança pública. O trabalho amplia o debate a respeito da segurança pública e da organização do conhecimento, favorecendo o reconhecimento de que só é possível realizar políticas públicas eficientes baseadas em dados e estatísticas sólidas e, muitas vezes, de longo prazo. Recomenda-se a busca por iniciativas de bases de dados públicos ou privadas em outros países que se encontram fora da União Europeia de modo a comparar as iniciativas e demais realidades no que se refere ao alcance das metas relativas ao objetivo 16 da Agenda 2030, isto é, a busca por um mundo mais justo e em que a paz seja soberana.

Referências

- Baracho, R. (2020). Representação e gestão do conhecimento: Aplicações em Cidades Inteligentes – Smart Cities. *Perspectivas em Ciência da Informação*, 25 (número especial), 252-279.
- Barros, L. (2017). Sustentabilidade ambiental e direito de acesso à informação verdadeira: de Estocolmo aos dias atuais. *Revista Brasileira de Biblioteconomia e Documentação*, 13 (Esp), 2923-2940.
- Bokolo, A. (2021). Managing digital transformation of smart cities through enterprise architecture – a review and research agenda. *ENTERPRISE INFORMATION SYSTEMS*, 15(3), 299-331.
- Clavier, V. (2019). Knowledge Organization, Data and Algorithms: The New Era of Visual Representations. *Knowl Org*, 46 (N.º 8), 615-621

- Esther Keymolen & Astrid Voorwinden (2020) Can we negotiate? Trust and the rule of law in the smart city paradigm, *International Review of Law, Computers & Technology*, 34:3, 233-253, DOI: 10.1080/13600869.2019.1588844
- G. H. Leazer and R. Montoya. The Politics of Knowledge Organization: Introduction to the Special Issue. *Knowl. Org.* 47(2020) N.º 5. 367-371
- Geraldo, G & Pinto, M. (2019). Percursos da Ciência da Informação e os Objetivos Do Desenvolvimento Sustentável da Agenda 2030/Onu. *Revista ACB: Biblioteconomia em Santa Catarina*, 24(2), 373-389.
- Israilidis et al.. (2021). Exploring knowledge management perspectives in smart city research: A review and future research agenda. *International Journal of Information Management*, 56(1), 1-10.
- Malheiro da Silva, A & Ribeiro, F. (2012). Documentation / Information and Their Paradigms: Characterization and Importance in Research, Education, and Professional Practice. *Knowl Org*, 39(2), 111-124.
- Martins et al.. (2019). Knowledge management in the context of sustainability: Literature review and opportunities for future research. *Journal of Cleaner Productio*, 229, 489-500.
- Martins, F & Santos, P. (2019). Informação e conhecimento na vivência social das cidades: um caminho para a sustentabilidade ambiental. *Iris – Informação, Memória e Tecnologia*, 5(jan/dez), 55-74.
- Matheus, R. et al (2020). Data science empowering the public: Data-driven dashboards for transparent and accountable decision-making in smart cities. *Government Information Quarterly*, 37(1), 1-9.
- Mikalauskiene, A. (2019). Knowledge Management Impact on Sustainable Development. *Montenegrin Journal of Economics*, 15(4), 149-160
- Moura, M.A. (2017). INTERSECTIONALITY AND THE SOCIAL CONSTRUCTION OF KNOWLEDGE ORGANIZATION. 15th INTERNATIONAL ISKO CONFERENCE, 1-2.
- Ochôa, P & Pinto, L. (2019). Agenda 2030 e o campo de intervenção da Ciência de Informação: dinâmicas de aprendizagem, envolvimento e desenvolvimento de competências. IX Encontro Ibérico EDICIC, 1-7.
- Preedip balaji, B. & Dhanamjaya, M. (2020). Towards a Model of Urban Studies Classification. *Knowl Org*, 47(7), 574-581.
- Souto, L & pizzol, R. (2019). Sustentabilidade e Gestão do Conhecimento: Perfil de Autoria e Análise Temática das Publicações do Km Brasil no Período de 2002 a 2016. *Revista Digital Biblioteconomia e Ciência da Informação*, 17(1), 1-22.
- Yahia, N. et al (2021). Towards sustainable collaborative networks for smart cities co-governance. *International Journal of Information Management*, 56(1), 1-16.
- Zuiderwijk, A & Hinnant, C. (2019). Open data policy-making: A review of the state-of-the-art and an emerging research agenda. *Information Polity*, 24, 117-129.

ORGANIZAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO EM AMBIENTES DIGITAIS: UMA APRESENTAÇÃO DO HERBÁRIO VIRTUAL IAN DA EMBRAPA

Cristiane Pantoja de Moraes¹

¹Discente do Programa de Pós-graduação em Ciência da Informação da Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, São Paulo, Brasil, Email: Crikapj@gmail.com, ORCID iD 0000-0002-3872-2717

Resumo

Diante da relevância dos estudos sobre meio ambiente, principalmente de aspectos referentes à Amazônia, abordar temas que norteiam a sua sustentabilidade, sobretudo do ponto de vista da biodiversidade à luz da ciência da informação num contexto de organização do conhecimento e informação, faz-se necessário e importante. **Objetivo:** Pensando nisso, este trabalho tem como objetivo observar e analisar as informações da biodiversidade existentes no Herbário IAN da Embrapa Amazônia Oriental (EMBRAPA), no qual há ampla diversidade de amostras de várias aplicações, como na medicina, recuperação de áreas degradadas, serviços ambientais, alimentação animal, madeira, entre outras atividades. Dessa forma, buscou-se fazer um levantamento de informações e disponibilização dos conteúdos e investigar como a representação da informação contribui para o desenvolvimento do domínio da botânica. **Metodologia:** O método utilizado neste trabalho foi de uma pesquisa essencialmente bibliográfica com uma abordagem qualitativa voltada para a produção do conhecimento em herbários, em especial àquelas relacionadas aos dados e imagens de exsicatas e coleções botânicas. A princípio, foi feito um levantamento científico bibliográfico para o referencial teórico em livros, periódicos científicos e relatórios na Biblioteca digital da Embrapa e no Google acadêmico relacionados à respectiva pesquisa, ou seja, seleção, definição e delimitação do tópico as palavras-chaves Herbário, Botânica, organização da informação e do conhecimento como principais termos para a busca. Após isso, utilizando-se de pesquisa empírica para o levantamento dos conteúdos contidos no sistema *Brahms* no qual está o Herbário virtual IAN por meio de uma navegação exploratória para verificar quais matérias estão presentes no sistema e quais são os principais tipos de informações se podem ser encontradas no Herbário. **Resultados:** Os resultados mostraram uma grande massa de informações e a importância da biodiversidade e o desenvolvimento sustentável. **Conclusão:** Assim, quando se fala no processo de informatização e disponibilização das coleções botânicas no Herbário IAN, notou-se não somente a qualidade e forma de organização dos conhecimentos nessa área, como também, a gama de informações contida nesse sistema, que vai desde

uma simples pesquisa por um material específico até às inúmeras informações, que podem ser geradas e a disponibilização de outras fontes de pesquisa de informação, gerando conhecimento para a população mundial.

Palavras-chave: Herbário, EMBRAPA, Botânica, Organização do conhecimento

Introdução

A Organização do conhecimento e Organização da informação ainda geram muitas interrogativas sobre suas definições, aplicações e usos nas diversas áreas do conhecimento. Elas podem ser compreendidas, para Brascher e Café (2008), como um procedimento que modela o conhecimento colaborando com a elaboração das representações do conhecimento; a qual, por sua vez, pode ser entendida como um modelo de abstração do mundo real, elaborado para determinado fim.

Assim, a preservação e compartilhamento de dados de pesquisa atualmente pode ser um fator decisivo para a produção científica no mundo, pois permite que a ciência evolua em sua diversidade, organizando, armazenando e compartilhando o conhecimento.

Paralelamente a isso, o progresso que as tecnologias de informação e comunicação (TICs) promovem no cenário científico e tecnológico é bastante significativo, principalmente porque envolvem a visibilidade dada pelas organizações informacionais como as bibliotecas, repositórios, base de dados. Essas organizações, por sua vez, disponibilizam em suas plataformas digitais a organização de informações, que por fim tem o intuito de disseminação e produção do conhecimento; contemplando, assim, vasta amplitude de informações. Como exemplo, temos o objeto deste trabalho, o Herbário Virtual IAN:

Disponibiliza informações e imagens de exsicatas (amostras de plantas desidratadas), e coleções de madeiras, flores, frutos, sementes e plântulas da Amazônia. É o primeiro herbário do Brasil a adotar o sistema *Brahms – Botanical Research and Herbarium Management System*, desenvolvido pela Universidade de Oxford para manejar dados botânicos. O sistema foi introduzido no Brasil pela Embrapa Amazônia Oriental (Empresa Brasileira de Pesquisa Agropecuária, 2013).

Nesse cenário, temos o Herbário do Instituto Agrônomo do Norte (IAN), fundado em 1945. Atualmente ele está sob o domínio da Embrapa Amazônia Oriental, mais conhecido como Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), localizado em Belém do Pará. O acervo do herbário está composto por milhares de exsicatas, sendo esta uma das unidades básicas de um herbário. A xiloteca, por exemplo, apresenta muitas amostras de madeira e fotografias de tipos e coleções de flores, frutos, sementes e plântulas. Essas coleções de tipos são organizadas em sala independente do acervo geral a fim de melhorar sua conservação (Aranda & Thiengo, 2008).

Contudo, é necessário que compreendamos o conceito de herbários. Para Araújo et al. (2019, p. 131), “Os herbários servem como base para pesquisa em muitas áreas do conhecimento e são de grande importância para estudos de sistemática de plan-

tas”. Já para Wiggers e Stange (2008, p. 41), o “herbário é uma coleção de plantas inteiras ou ramos com folhas, flores e frutos desidratadas, montadas, geralmente, em cartolina padrão com etiquetas contendo informações de coleta e número de registro, recebendo o nome de exsicata e conservadas de acordo com técnicas específicas”

Para Lucas Vieira et al. (2018) o Herbário é uma coleção botânica destinada à catalogação de materiais botânicos. Na perspectiva de João Pires (1995, p. 102), Herbário é o museu onde são guardadas, em perfeita ordem consultáveis, as amostras secas de plantas que são o produto das coleções botânicas.

Em relação a materialização desses conteúdos da forma natural para a forma digital, só poderá ser consultada, conforme Manhique e Casarin (2019), se os sistemas de organização do conhecimento adotados pelas unidades de informação refletirem as necessidades de informação de usuários em contextos específicos.

As coleções botânicas têm por objetivo documentar, catalogar e subsidiar estudos referentes a diversas áreas, mas em particular, a botânica. Com enfoque na área de morfologia e taxonomia vegetal, essas coleções possuem reconhecida relevância para profissionais de diversas áreas e, principalmente, para especialistas botânicos, tendo via de regra o herbário como coleção principal. O herbário é uma coleção botânica destinada a catalogação de materiais botânicos e auxiliar estudos referentes a estes. Nos dias atuais, faz-se necessário que essas instituições se modernizem, de modo a fazer fluir o conhecimento concentrado nas coleções para o resto da sociedade. Uma das maneiras é a digitalização das coleções, transformando o material antes apenas em formato físico, agora também em formato digital (Vieira, 2018, pp. 20-21).

De acordo com Souza et al. (2013), as coleções de plantas são inestimáveis e são muito importantes para todo e qualquer trabalho de pesquisa relacionado à diversidade, estrutura, classificação e distribuição dos organismos vegetais. Desse modo, os herbários apresentam um papel muito relevante, pois são depositários de parte dos testemunhos de fortuna patrimonial e ecológica. Eles têm desempenho inquestionável, único que reforça e contribui principalmente aos estudos da biodiversidade vegetal. Diante disso, é notável a gama de materiais e informações contidas no herbário IAN. Provavelmente essas amostras já cresceram em números e espécimes que, de acordo com Souza et al. (2013, p. 10), apresenta como grande fonte de informação:

Sua principal missão é gerar conhecimento e contribuir para o manejo e conservação da flora amazônica. O acervo científico do Herbário IAN constitui uma valiosa fonte de pesquisa científica, documentação e informação sobre a Amazônia, composto por 191 mil exsicatas, abrigando ainda uma xiloteca com 8 mil amostras de madeira, uma carpoteca com 700 frutos desidratados e 289 em meio líquido, uma fototeca com cerca de 30 mil fotografias de tipos e uma coleção de tipos nomenclaturais com aproximadamente 2 mil amostras, devidamente acompanhadas da cópia de sua respectiva obra princeps (publicação contendo a primeira descrição de um táxon). Coleções menores, mas não menos importantes, fazem parte desse patrimônio científico, como flores em meio líquido (321 amostras), sementes desidratadas (191), plântulas em meio líquido (54) e desidratadas (65).

Sendo assim, este trabalho tem como objetivos demonstrar como são organizados, caracterizadas e apresentadas as amostras depositadas no Herbário IAN. Isso facilita nas pesquisas dos profissionais da área, sejam na localização desses exemplares, como em qualquer tipo de pesquisa de conteúdos. Como por exemplo, as imagens de exsicatas, que contenham informações específicas para os pesquisadores, primeiramente porque incorporam seu papel social como fonte mediadora e cultural de leitura, conhecimento e informação da biodiversidade. Desse modo, em decorrência das novas tecnologias da informação em plena sociedade da informação ou era do conhecimento, essas fontes de informações enfrentam vários desafios de incorporação e adaptação de seus materiais a essas novas tecnologias e os meios de disseminação desses conhecimentos.

O Herbário IAN da Embrapa Amazônia Oriental armazena milhares de dados acerca da biodiversidade vegetal, com destaque para a flora amazônica (Souza et al., 2013). Assim, este trabalho tem por objetivo descrever o ambiente de armazenamento da base de dados do Herbário IAN, bem como seu gerenciamento, aplicações e ferramentas, com o objetivo de subsidiar pesquisas em várias áreas do conhecimento, tornando, dessa forma, o acesso à informação mais rápido, eficiente e seguro.

Em se tratando de pesquisa, os dados são materiais reconhecidos pela comunidade científica que foram registrados durante uma pesquisa e servem para certificar os resultados obtidos. Em suma, são documentos que registram criações intelectuais ou artísticas de interesse para a pesquisa, divulgados por meio de unidades acadêmicas, sob a supervisão de indivíduos ou grupos de pesquisa, ou como coleções de indivíduos ou instituições (Pando & Almeida, 2015).

Com o advento, desenvolvimento e intensificação do uso das novas tecnologias eletrônicas, a atualização do processamento e comunicação de informações dos chamados conteúdos digitais, tornaram-se muito comuns essa constante evolução. Esse processo vem facilitando a maneira de como adquirir informação, possibilitar finalmente um estreitamento entre usuário e conteúdo disponibilizado em qualquer lugar no mundo.

Porém, os desenvolvimentos das pesquisas na Amazônia devem ser embasados com informações científicas que possam contribuir no planejamento do uso de seus recursos naturais, tendo em vista o manejo e conservação do meio ambiente e da biodiversidade para as futuras gerações. Em via de regra, o desenvolvimento das regras ecológicas e as inter-relações do meio ambiente são constantemente violadas, resultando num declínio rápido e irreversível da diversidade vegetal. Como resultado, ocorre a diminuição das opções de utilização dos recursos naturais pelas comunidades locais, indústria e científica, em nível nacional e internacional (Secco et al., 2003).

No Brasil verifica-se uma enorme preocupação com a proteção e o uso adequado da biodiversidade vegetal da Amazônia, desenvolvimento sustentável e meio ambiente. Afinal, são inúmeras as discussões que têm surgido, especialmente aquelas que envolvem recursos vegetais. Com a colaboração de alguns países do mundo os herbários vêm sendo apontados ultimamente como elemento essencial e único dessa biodiversidade, uma vez que detêm coleções dos mais variados ecossistemas regionais, contendo em seus acervos informações sobre local e data de coleta de espécies, assim como seus aspectos vegetativos e reprodutivos, bem como, seus usos, diversi-

dade e conservação da mais variadas espécimes vegetais já encontradas (Martins-Da-Silva & Ferreira, 1998).

Metodologia

Considerando que esta pesquisa constitui o caminho para se conhecer e descobrir a realidade do Herbário IAN da Empresa Brasileira de Pesquisa Agropecuária. Este trabalho visa demonstrar a organização, caracterização e divulgação dos tipos de materiais depositados no Herbário IAN a fim de facilitar pesquisa em muitas áreas do conhecimento que são de grande importância para estudos das plantas. Por meio de uma pesquisa bibliográfica com uma abordagem qualitativa voltada para a produção do conhecimento, sobretudo produções relacionadas aos dados e imagens das exsicatas e coleções botânicas em Herbários. A princípio, foi feito um levantamento científico bibliográfico para o referencial teórico em livros, periódicos científicos e relatórios na Biblioteca digital da Embrapa e no Google acadêmico relacionados à respectiva pesquisa, ou seja, seleção, definição e delimitação do tópico ou problema as palavras-chaves Herbário, Botânica, organização da informação e do conhecimento como termos para a busca. Após isso, utilizando-se de pesquisa empírica para o levantamento dos conteúdos contidos no sistema *Brahms* no qual está o Herbário virtual IAN por meio de uma navegação exploratória para verificar quais matérias está presente no sistema e quais são os principais tipos de informações se podem ser encontradas no Herbário.

Pretendeu-se mostrar como o herbário contribui com o desenvolvimento da representação da informação do domínio da Botânica, demonstrando como as informações biológicas são reorganizados e recuperados em um sistema digital, de modo que possa gerar listas, etiquetas, relatórios, mapas, tabelas, gráficos e muitas outras informações a partir do seu sistema de busca, cujas principais funções do sistema é contribuir para as pesquisas como referências para o meio ambiente e a biodiversidade mundial. Dessa forma, procurou explorar os recursos disponíveis no Herbário virtual IAN, bem como esclarecer a relevância de estudo da organização dos conhecimentos em botânica a recuperação da informação na prática.

Resultados

De acordo com a pesquisa, nota-se que o Herbário virtual IAN apresenta uma infinita lista de conteúdos, que são disponibilizadas para a pesquisa a qualquer hora em qualquer lugar de acordo com a (ver Figura 1). A imagem das páginas iniciais do Herbário IAN é possível notar grande quantidade de informações, variedades de abas e filtros para busca. Além disso, há apresentação das imagens logo na página principal, apesar de parecer um pouco complexas para quem é leigo em se tratando de tecnologia.

The screenshot displays the Herbario Virtual IAN interface. On the left, there are navigation panels for 'Collectors', 'Geography', 'Taxa', and 'Species Search'. The main area features a 'Species List' table with columns for TaxStat, Family, Name, Genus, Author, Accepted Name, and Images. The table lists various plant species such as *Licania heteromorpha*, *Heliconia elegans*, and *Quilicaria oppositifolia*. On the right side, there is a 'Record Images' section showing a grid of thumbnail images of plant specimens.

TaxStat	Family	Name	Genus	Author	Accepted Name	Images
	Chrysobalanaceae	<i>Licania heteromorpha</i> var. <i>heteromorpha</i>	<i>Licania</i>	Berth.		4
	Moraceae	<i>Heliconia elegans</i>	<i>Heliconia</i>	(J.F. Macbr.) C.		2
	Cappariaceae	<i>Capparis sola</i>	<i>Capparis</i>	J.F. Macbr.		0
	Violaceae	<i>Viola maculata</i>	<i>Viola</i>	H. Rob.		4
	Leptocarpaceae	<i>Couratari obtusa</i>	<i>Couratari</i>	A.C. Sm.		5
	Polygalaceae	<i>Polygala spessa</i>	<i>Polygala</i>	Berth.		0
	Asaraceae	<i>Ichthyophora rufa</i>	<i>Ichthyophora</i>	Gardner		8
	Asclepiadaceae	<i>Marsdenia macrophylla</i>	<i>Marsdenia</i>	(Humb. & Bon.)		1
	Sapotaceae	<i>Pouteria faha</i>	<i>Pouteria</i>	T.D. Penn.		0
	Leguminosae	<i>Aldina insignis</i> var. <i>insignis</i>	<i>Aldina</i>	(Berth.) Endl.		4
	Amesaceae	<i>Quercus reticulata</i>	<i>Quercus</i>	Maas		4
	Menispermaceae	<i>Cissampelos laeflora</i>	<i>Cissampelos</i>	Maléna		5
	Quilicaceae	<i>Lacunaria oppositifolia</i>	<i>Lacunaria</i>	Pires		15
	Solanaceae	<i>Solanum schlechtendalium</i>	<i>Solanum</i>	Walp.		3
	Boraginaceae	<i>Cordia kingstoniana</i>	<i>Cordia</i>	J.S. Millar		1
	Cyperaceae	<i>Kyllinga vaginata</i>	<i>Kyllinga</i>	Lam.		2
	Gentianaceae	<i>Chorizandra ovatum</i> var. <i>spapanum</i>	<i>Chorizandra</i>	Gleason		2
	Leguminosae	<i>Astragalus lentiginosus</i> var. <i>australis</i>	<i>Astragalus</i>	Douglas		2
	Euphorbiaceae	<i>Adioscimenon verticillatum</i>	<i>Adioscimenon</i>	(Rottb.) Bail.		2
	Mahoeaceae	<i>Pavonia dactyloides</i>	<i>Pavonia</i>	H. Montaino		4
	Leguminosae	<i>Cassia insularis</i>	<i>Cassia</i>	Berth.		2
gen	Piperaceae	<i>Piper</i>	<i>Piper</i>	L.		6
	Leguminosae	<i>Tibellum liturium</i>	<i>Tibellum</i>	Phil.		1
	Rutaceae	<i>Casimiroa edulis</i>	<i>Casimiroa</i>	La Lièvre & Lex.		3
	Manispermaceae	<i>Odotocarya sylvatica</i>	<i>Odotocarya</i>	Barnesby		5

Figura 1. print screen Herbario virtual IAN

Fonte: <http://brahms.cpatu.embrapa.br/herbarioian/Explore>

É possível obter pelas pesquisas no herbário imagens das exsicatas e de suas respectivas etiquetas, que representam as descrições e características de coleta do material (ver Figura 2). Em cada exsicata são colocadas as escalas fotografadas. Além disso, observa-se que as amostras possuem código de barras, o que facilita o controle dos materiais que são inseridos no herbário. Para Souza et al. (2013, p. 15), o sistema do herbário IAN vem sendo desenvolvido para dar suporte, principalmente, em três atividades: primeiro seria o manejo de coleções científicas em herbários, produzindo etiquetas para exsicatas, administrando o processo de intercâmbio científico de material botânico efetuado entre herbários, produzindo etiquetas para especialistas visitantes, organizando base de dados, fornecendo consultas via internet e atualizando identificações realizadas por especialistas. Em segundo lugar, a elaboração de produtos para taxonomia, como revisões, monografias, checklists, dentre outros. E, por fim, a elaboração de produtos com orientações geográficas, como checklists de localidades, flora, levantamentos de biodiversidade e outros.

De acordo com as pesquisas disponíveis da Empresa Brasileira de Pesquisa Agropecuária (2013) e no sistema IAN, o herbário permite que se façam diferentes tipos de pesquisas nos milhares de exsicatas digitalizadas. Assim, pode-se pesquisar e obter diferentes tipos de informações (ver Quadro 1). Por exemplo, a primeira coluna do quadro separou-se os diversos tipos de informações sobre amostras quanto às informações botânicas, ou seja, em suas várias categorias como: família, gênero, espécie, produto (imagens de lâminas ou fotos). É possível, também, pesquisar as informações com a descrição de determinado conteúdo do herbário. A segunda coluna exibe informações disponíveis para pesquisa e filtros de pesquisa que se pode utilizar para a busca de informações como: coletores ou colecionadores (especialis



Figura 2: print screen Exsicata

Fonte:

<http://brahms.cpatu.embrapa.br/IANDB/image/IAN199546.JPG/Zoom?width=800px&height=800px>

tas responsáveis pela coleta das amostras) organizadas em ordem alfabéticas agrupados letra a letra. Geograficamente se pode pesquisar também por países que estão agrupados em ordem alfabética. Já na pesquisa por taxa, a busca é feita de um termo geral que são as famílias. As subdivisões são destinadas para os termos mais específicos como gênero e espécies. Além disso, é possível gerar mapas com a localização das coletas das amostras e exibir relatórios delas. A terceira coluna trata de algumas fontes de informações contidas no herbário, através de botões específicos, em que se pode acessar diversos conteúdos como: banco e base dados nas áreas de ciências biológicas, referentes principalmente às plantas, botânica, meio ambiente e as mais diversas biodiversidades existentes no mundo.

Quadro 1. Tipos de informações contidas no Herbário virtual IAN

PRINCIPAIS INFORMAÇÕES DO HERBÁRIO IAN		
INFORMAÇÕES BOTÂNICAS	INFORMAÇÕES DO HERBÁRIO	FONTES DE INFORMAÇÕES
Famílias	Colecionadores	IPNI (Índice Internacional de Nomes de Plantas)
Gêneros	Geográfica (países)	POWO (Plants of the World Online)
Espécies	Taxa de pesquisa	The Plant List
Produtos	Localização	Global Plants database
Descrição	Relatórios	CJBG (Conservatoire et Jardin botaniques Genève)
		HBL (Biodiversity Heritage Library)
		GOOGLE
		Banco de dados da Trópicos
		GBIF (Global Biodiversity Information Facility)
		Kew (Royal Botanic Gardens)
		CRIA (Centro de Referência em Informação Ambiental)

No momento atual, o herbário possui um controle de registro de suas amostras que são feitas por meio de listas elaboradas a partir da base de dados, contendo registro, coletor e número, data da coleta, dados dos espécimes e de local de coleta. Dessa forma, os relatórios gerados são fundamentais para quem trabalha na área e deseja pesquisa sobre a temática. Esse controle é fundamental para se obter informação de forma rápida, pois agiliza o processo de busca e verificação do material, o que facilita na recuperação de dados pelos pesquisadores e usuários do herbário.

Conclusões

Nota-se que o herbário apresenta em seus sistemas amostras bem organizadas e preservadas de acordo com os sistemas já determinados. Assim, o herbário apresenta grande importância como material de pesquisa nas áreas da ciência, tecnologia e economia que necessitam de estudos relacionados ao meio ambiente e a biodiversidade.

Ressalta-se que o herbário virtual não foi criado para substituir completamente o herbário tradicional, porém pode ser solução para muitos pesquisadores, uma vez que uma imagem simples possa ser suficiente para identificar ou mesmo resolver certos problemas de classificação. Sendo assim, a informatização das coleções não veio para substituir suas cópias, mas para protegê-las, pois as informações contidas

no banco de dados preservam as coleções presentes no herbário convencional de danos e desgastes nas amostras.

O acervo botânico do herbário virtual IAN é uma ferramenta com vasto significado tanto para o ensino, pesquisa e extensão, quanto para a produção de conhecimento. Esse sistema virtual abriu o caminho acesso e recuperação da informação, sendo primordial para a sociedade, pois se acredita que há grande necessidade de aproximar a informação contidas nos sistemas disponíveis ao público, principalmente para instituições de pesquisas, profissionais que necessitam da disponibilização desses conteúdos on-line como uma forma obter conhecimento sobre a conservação e utilização da biodiversidade.

Em meio a tudo isso, observou-se que a digitalização do material pertencente aos herbários são ferramentas indispensáveis e capazes de capturar as imagens de diversos materiais que incentivam a informatização. Elas são de suma importância por fornecerem suporte tecnológico dos herbários para otimização do processo de informatização dos acervos botânicos. É indiscutível a importância das plantas, elas desempenham papel fundamental na manutenção das funções dos ecossistemas e são mundialmente reconhecidas como uma parte vital da biodiversidade que busca em sua essência a Convenção do meio ambiente, conservação de Plantas, isto é, que objetiva ao passar dos anos longo o controle da perda da diversidade de plantas, que é fundamental compreensão da diversidade para os estudos das plantas, animais e microrganismos.

As coleções, sejam elas de qual natureza forem, são fundamentais para a preservação e o conhecimento da biodiversidade de um local. A digitalização dessas coleções é um importante passo para auxiliar nesses processos, pois ao transmitir as informações que estão organizadas e contidas nos herbários no meio digital, por meio de fotografias e dados, é uma forma de salvaguardar permanentemente os materiais e também de fazer fluir com muito mais eficiência a gama de conhecimentos dessas instituições para a sociedade. No entanto, apesar dos benefícios advindos da digitalização dos herbários, a análise do material físico se mostra ainda indispensável para as análises. Isso porque, além de aumentar a visibilidade e relevância das coleções e de pesquisas nelas realizadas, a exposição pública dos dados permite melhora na sua qualidade. Porém, talvez o benefício preponderante do compartilhamento de dados seja a aceleração no avanço científico com o acesso aos dados e a possibilidade de maior integração e trocas entre diferentes áreas de conhecimento.

Referências

- Aranda, A. T., & Thiengo, S. C. (2008). Anais do segundo simpósio nacional de coleções científicas. In *Anais do segundo simpósio nacional de coleções científicas* (pp. 95-95).
- Araújo, K. L. G. de, Luz, J. C. L., Souza, P. S. S., SOUZA, G. D. S., & Kiill, L. H. P. (2019). Avaliação do banco de dados do acervo do Herbário do Trópico Semiárido. In *Embrapa Semiárido-Artigo em anais de congresso (ALICE)*. In: JORNADA DE INICIAÇÃO CIENTÍFICA DA EMBRAPA SEMIÁRIDO, 14., 2019, Petrolina. Anais... Petrolina: Embrapa Semiárido, 2019. <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1114644/avaliacao-do-banco-de-dados-do-acervo-do-herbario-do-tropico-semiarido>.

- Café, L., & Brascher, M. (2008). Organização da informação ou organização do conhecimento. *Encontro Nacional de Pesquisa em Ciência da Informação*, 9, 5. <http://repositorios.questoeseemrede.uff.br/repositorios/bitstream/handle/123456789/809/17.pdf?sequence=1>
- Empresa Brasileira de Pesquisa Agropecuária (2013). Herbário virtual disponibiliza informações e imagens sobre a flora amazônica. <https://www.embrapa.br/busca-de-noticias/-/noticia/1874058/herbario-virtual-disponibiliza-informacoes-e-imagens-sobre-a-flora-amazonica>.
- Pando, D. A., & Almeida, C. C. D. (2015). Organização da informação e do conhecimento no contexto da Ciência da informação: da análise terminológica à reflexão epistemológica. In *II Congreso ISKO España-Portugal/XII Congreso ISKO España (2015). Organización del conocimiento: sistemas de información abiertos. Actas. Murcia: Facultad de Comunicación y Documentación, Universidad de Murcia. ISBN ISBN* (pp. 978-84). http://www.iskoiberico.org/wp-content/uploads/2015/11/54_Pando.pdf
- Pires, J. (1958). Instruções sumárias sobre a organização de um herbário ou museu botânico. *Embrapa Amazônia Oriental-Artigo em periódico indexado (ALICE)*. <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/978224/instrucoes-sumarias-sobre-a-organizacao-de-um-herbario-ou-museu-botanico>
- Manhique, I. L. E., & Casarin, H. D. C. S. (2019). Abordagem cultural da organização do conhecimento na ciência da informação brasileira. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 24(56), 01-20. <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2019.e65174/41030>
- Martins-Da-Silva, R., & Ferreira, G. (1998). Acervo do herbário IAN. 1. Meliaceae Ventenat. *Embrapa Amazônia Oriental-Documentos (INFOTECA-E)*. <https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/388116/1/CPATUDoc120.pdf>
- Secco, R. D. S., Martins-da-Silva, R. C. V., Ferreira, C. A. C., Sarquis, R., Filer, D., & Duarte, J. R. M. (2003). Informatização dos herbários amazônicos: gerenciamento, integração e utilização da informação. In *Embrapa Amazônia Oriental-Artigo em anais de congresso (ALICE)*. In: CONGRESSO NACIONAL DE BOTÂNICA, 54.; REUNIÃO DE BOTÂNICOS DA AMAZÔNIA, 3., 2003, Belém, PA. Desafios da botânica no novo milênio: inventário, sistematização e conservação da diversidade vegetal. Belém, PA: Sociedade Botânica do Brasil: Universidade Federal Rural da Amazônia: Museu Paraense Emílio Goeldi: Embrapa Amazônia Oriental, 2003. <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/404453/informatizacao-dos-herbarios-amazonicos-gerenciamento-integracao-e-utilizacao-da-informacao>.
- Souza, H., Martins-da-Silva, R. C. V., Filer, D., Xavier Junior, S. R., & Fouro, A. (2013). Base de dados do Herbário IAN da Embrapa Amazônia Oriental. *Embrapa Amazônia Oriental-Documentos (INFOTECA-E)*. <https://www.infoteca.cnptia.embrapa.br/bitstream/doc/973872/1/DOC392.pdf>
- Vieira, L. L. A., Souza, H., & Furtado, C. M. (2018). A digitalização de quatro herbários da Amazônia e seus equipamentos, processos e atual situação. *Embrapa Amazônia Oriental-Artigo em periódico indexado (ALICE)*. <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1106458/a-digitalizacao-de-quatro-herbarios-da-amazonia-e-seus-equipamentos-processos-e-atual-situacao>
- Wiggers, I., & Stange, C. E. B. (2008). Manual de instruções para coleta, identificação e herborização de material botânico. *Programa de Desenvolvimento Educacional-SEED-PR UNICENTRO*, 45p. <http://www.diaadiaeducacao.pr.gov.br/portals/pde/arquivos/733-2.pdf>

THE CULTURAL ASPECTS OF KNOWLEDGE ORGANIZATION TOWARD A SUSTAINABLE DEVELOPMENT

Asmaa Bouaamri¹, Ágnes Hajdu Barát²

¹Library and Information Science institute, Faculty of Humanities, Eötvös Loránd University (ELTE), Budapest, Hungary, asmaa.bouaamri@gmail.com,
<https://orcid.org/0000-0003-3340-9673>

²Head of department professor of Library and Information Science Institute (LIS), Faculty of Humanities, Eötvös Loránd University (ELTE), Budapest, Hungary, hajdu.agnes@btk.elte.hu, <https://orcid.org/0000-0002-6774-5440>

Abstract

The awareness of the changing environment that we are living in, and the way our daily interactions and communications are changing, and swapping to different kinds of cultures worldwide, pauses us to re-think our perceptions of Knowledge Organization (KO) and databases as the new forms of the world's culture. The digitization of our world changed massively our behaviours to seeking and accessing information.

Knowledge Organization reflects the societal and cultural behaviours of the population shaping it. The action that takes place when information is recognized and organized based on human understanding is, therefore, an interaction between a social and cultural context with individuals, institutions, and the way of interpretation and comprehension of that particular information. Human action, the background of societal and cultural perception, which interferes in the KO of certain information, is crucially reflecting a particular heritage of human activities.

As a result of the development of technological tools and environments, the perception of KO is also changing to practically and exclusively a digitized one. Databases and other forms of virtual and digitized platforms became the main activity of human behaviour, and which is mirrored in the way that Knowledge is organized within a digitized culture.

It is in this purpose, and as a result of the aforementioned, that this research paper is about to explore and perceive how can Knowledge Organization be one of the aspects of culture in the digitized world, and how can its preservation maintain a Sustainable Development (SD) of countries around the globe.

Keywords: Knowledge Organization, Knowledge Management systems, Sustainable Development, Technological development, digitization.

Introduction

The development and the rapidness of the circulation of information, in the modern technological world that we are living in, became indispensable. The need of information by individuals is, therefore, linked directly to the systems utilized by them to retrieve and localize the sought information. Hence, this leads to the different ways of rethinking how Knowledge is organized.

Knowledge Organization (KO) is the field of research and study concerned with organizing and classifying knowledge in order to make it accessible to different entities. KO is also defined in the narrow sense, as the “activities such as document description, indexing and classification performed in libraries, bibliographic databases, archives and other kinds of “memory institutions” by librarians, archivists, information specialists, subject specialists, as well as by computer algorithms and laymen” (Hjørland, 2008).

This perception of KO is only the attenuated one, to simplify its understanding from the surface level. KO should not be mixed with the field of Information Retrieval (IR), as there are two separated fields but concerned with a similar purpose that is “to facilitate the findability of documents, knowledge, and information” (Hjørland, 2021).

In this respect, this paper’s goal is to provide and analyse the roles of Knowledge Organization in a broader sense of the term, which is related to the social, cultural settings of human actions. Using different approaches from different disciplines such as Humanities and Social Sciences, this article argues the importance of KO in organizing the world and how it can be perceived as a way for sustainable development.

Methodology

The methodology followed in this paper was based on the literature review of articles in three main domains: Knowledge Organization, linguistic analysis and social sciences. Combining different methods and perspectives was the main purpose of this study, in order to decorticate the main goals of KO and how it is perceived socially and culturally by humans.

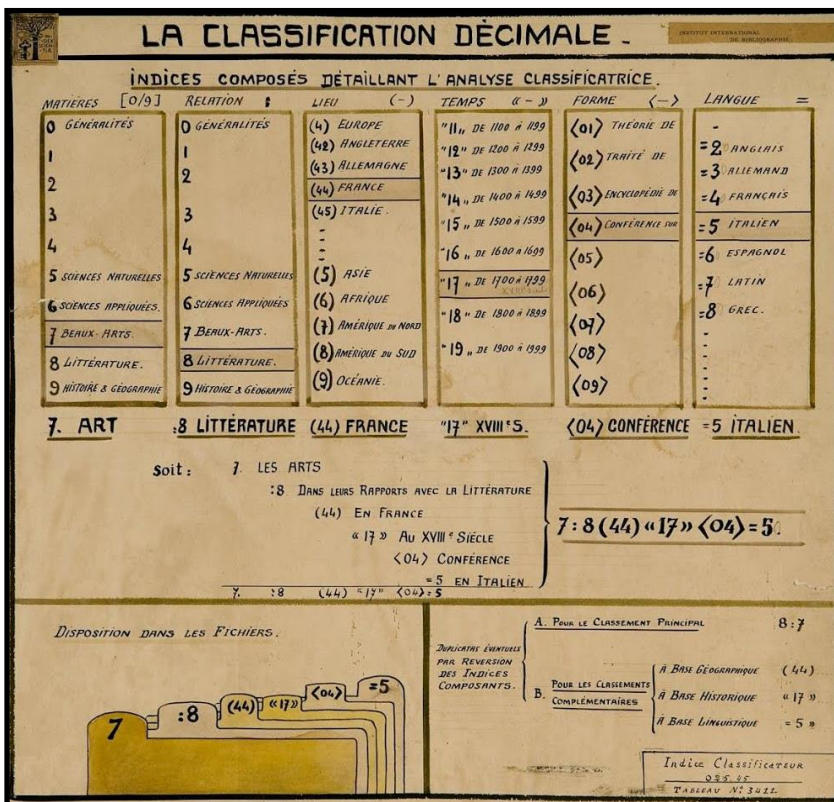
The studies adopted a selective literature approach that were concerned about the value of knowledge, the meaning of knowledge, how knowledge is perceived and acquired in societal and cultural settings. Using linguistic analysis, and interpretative methods to examine the different facets of KO within the society and how it is sustained. A thorough description is used to analyze the previous literature and results.

Knowledge Organization and its role

The definition of Knowledge Organization (KO) is to some extent difficult to contain. For a simplistic description, it can be defined as “Information Organization” (Salaba, 2020 as cited in Hjørland, 2021). It is considered as the science dedicated to

order knowledge (Smiraglia, 2018). The elaboration of KO is related to the evolution of sciences in the Renaissance period in purpose to make the discoveries less-ambiguous (Holzem, 2006). An evolution of Knowledge organization in the 17th century that took place, in order to organize knowledge based on a universal language system (see examples of Universal Decimal Classification (UDC) and its evolution from paper to digital versions), which permits the access of it. Therefore, a movement began to organize the new knowledge (of that period) and discoveries based on sciences (Holzem, 2006).

Figure 1: An old version of the Universal Decimal Classification (UDC). An explanatory schema of the Universal Decimal Classification index formation in French, 1920. Retrieved from Wikipedia



Hjørland (2008) describes KO in a restricted manner as the activities dedicated to describe a document by trying to classify it and index it referring to classification systems allocated for such purposes, and which are mostly done (but not only) by librarians, archivists, documentalists, and information scientists (p. 86). It is a field of study that focuses on the systems used for organizing knowledge known as Knowledge Organization Systems (KOS), and which will be detailed later on in this paper for more explicit understanding.

From a broader perspective, Hjørland explained it very clearly by stating that “the broader meaning KO is about the social division of mental labor, i.e. the organization of universities and other institutions for research and higher education, the structure of disciplines and professions, the social organization of media, the production and dissemination of “knowledge” etc.[...] We may distinguish between the social organization of knowledge on one hand, and on the other hand the intellectual or cognitive organization of knowledge. The broad sense is thus both about how knowledge is socially organized and how reality is organized” (2008, pp. 86-87).

La CDU : Classification Décimale Universelle									
0	1	2	3	4	5	6	7	8	9
INFORMATION COMMUNICATION	PHILOSOPHIE	RELIGION	SCIENCES SOCIALES	La classe 4 est actuellement inoccupée	SCIENCES	TECHNIQUES	ARTS LOISIRS	LINGUISTIQUE LITTÉRATURE	HISTOIRE GEOGRAPHIE
Bibliographies Catalogues	Métaphysique	Religion naturelle	Statistiques		Mathématiques	Médecine hygiène santé	Urbanisme	Littérature américaine	Géographie
01	11	21	31		51	61	71	81	91
Bibliothèque et lecture Documentation	Connaissance	Bible Evangiles	Politique L'Etat		Astronomie	Techniques industrielles	Architecture	Littérature anglaise	Biographies
02	12	22	32		52	62	72	82	92
Encyclopédies générales	Parapsychologie	Théologie doctrinale chrétienne(dogme)	Economie		Physique	Agriculture, forêt, pêche	Sculpture	Littérature allemande	Histoire ancienne archéologie
03	13	23	33		53	63	73	83	93
Essais en collection	Systèmes philosophiques	Théologie spirituelle (vie religieuse personnelle)	Droit (justice)		Chimie Minéralogie	Arts ménagers	Dessin Arts décoratifs	Littérature française	Histoire de l'Europe
04	14	24	34		54	64	74	84	94
Périodiques généraux Annuaires	Psychologie	Théologie pastorale (paroisse)	Administration de l'Etat		Sciences de la terre géologie météorologie	Entreprise	Peinture	Littérature italienne	Histoire de l'Asie
05	15	25	35		55	65	75	85	95
Organisations générales congrès, musées	Logique	L'Eglise chrétienne et la société	Aide Assistance Secours		Paléontologie (fossiles)	Industries chimiques et alimentaires	Arts graphiques	Littérature espagnole et portugaise	Histoire de l'Afrique
06	16	26	36		56	66	76	86	96
Presse Edition	Morale (éthique)	Histoire de l'Eglise chrétienne	Education Enseignement		Sciences de la vie biologie génétique	Fabrications industrielles	Photographie	Littérature latine et grecque	Histoire de l'Amérique du Nord
07	17	27	37		57	67	77	87	97
Recueils généraux (discours)	Philosophes anciens et orientaux	Autres confessions chrétiennes	Commerce Transports Communication		Botanique (plantes)	Articles manufactures	Musique	Littérature slave	Histoire de l'Amérique du Sud
08	18	28	38		58	68	78	88	98
Manuscrits Livres rares	Philosophes modernes (depuis le XVIe s)	Autres religions et mythologie	Coutumes et folklore		Zoologie (animaux)	Bâtiment Construction	Loisirs	Autres littératures	Histoire de l'Océanie
09	19	29	39		59	69	79	89	99

**Figure 2: The Schema contained in the UDC.
Retrieved from the internet on the 16/04/2021**

The screenshot shows the UDC Online interface. At the top, there is a navigation bar with 'UDC online' and 'AENOR' logos, and a menu with 'Guía CDU', 'Noticias', 'Contactos', 'Fuentes', and 'Desconectar'. Below this is a secondary menu with 'Búsqueda', 'Constructor de número', 'Índice alfabético', and 'Validación / Análisis'. The main content area is titled 'CDU Tablas' and features a search box with 'Número CDU' set to '34' and a 'Buscar' button. There are also checkboxes for 'Búsqueda exacta', 'Incluir ejemplos', and 'Incluir números cancelados'. A message states '1830 se han encontrado registros coincidentes'. The search results are listed as follows:

Número CDU	Descripción
34	Derecho. Jurisprudencia
34.01/09	Auxiliares especiales para derecho y jurisprudencia
34.01	Naturaleza del derecho. Teoría del derecho. Alcance del derecho
34.018	Codificación del derecho. Aplicación de la ley ■ Promulgación, coacción, sanciones
34.02	Condiciones de existencia del derecho
34.021	Precedencia, prioridad de las leyes
34.023	Condiciones relativas al sujeto de la ley
34.025	Condiciones relativas al objeto de la ley
34.028	Condiciones relativas a la forma de la ley ■ Formalidades prescritas oficialmente. Apertura al público de los procedimientos y actos legales
34.03	Funcionamiento del derecho. Efectos del derecho
34.033	Formación de sentencias
34.034	Pronunciamiento de la sentencia
34.035	Incorporación de las decisiones judiciales a las actas

Figure 3: UDC Online. Accessed in 2017 from “La Biblioteca Ana Maria Matute, in El instituto Cervantes de Casablanca”

The wider meaning of Knowledge Organization thus is the relation of how knowledge is perceived in the first place and how it is interpreted in the real world. However, this raises several questions and queries in relation to the ways used to organize knowledge, and if there is a reliability in the systems used and the accuracy of controlled vocabulary employed within the field of Knowledge Organization.

Hence, Knowledge organization is more than a metadata, or what it is called information about information, that is free from concepts, but rather, it contains more social, cultural, political and scientific aspects in it, that are not directly recognized, and which poses more ambiguity to usability of KOS in competency with the field of Computer science and Information Retrieval (IR).

The screenshot displays the Library of Congress Subject Heading (LCSH) for 'Rousseau, Jean-Jacques, 1712-1778'. The interface is divided into two main sections: a left sidebar and a right main content area.

Left Sidebar:

- LIBRARY OF CONGRESS** logo.
- Breadcrumb: [The Library of Congress](#) > [Linked Data Service](#) > [LC Name Authority File \(LCNAF\)](#)
- Rousseau, Jean-Jacques, 1712-1778**
- URI(s)**
 - <http://id.loc.gov/authorities/names/n79008220>
- Instance Of**
 - MADS/RDF PersonalName
 - MADS/RDF Authority
 - SKOS Concept
- Scheme Membership(s)**
 - Library of Congress Name Authority File
- Collection Membership(s)**
 - Names Collection - Authorized Headings
 - LC Names Collection - General Collection
- Variants**
 - Rousseau, Jan Jakub, 1712-1778
 - Russo, Zhan Zhak, 1712-1778
 - Rousseau, John James, 1712-1778

Right Main Content Area:

- found:** Khayāḥ[parvarihā-yi tafarruj'gar-i inziḩā]ū, 1966:t.p. (ژان ژاک روسو = Zhān Zhāk Rōsū) t.p. ver Rousseau [in rom.]
- LC Classification**
 - ML410.R86
 - PQ2030-PQ2049
- Editorial Notes**
 - [Machine-derived non-Latin script reference project.]
 - [Non-Latin script references not evaluated.]
 - [Old catalog heading: Rousseau, Jean Jacques, 1712-1778]
- Change Notes**
 - 1979-01-26: new
 - 2020-07-21: revised
- Alternate Formats**
 - RDF/XML (MADS and SKOS)
 - N-Triples (MADS and SKOS)
 - JSON (MADS/RDF and SKOS/RDF)
 - MADS - RDF/XML
 - MADS - N-Triples
 - MADS/RDF - JSON
 - SKOS - RDF/XML
 - SKOS - N-Triples
 - SKOS - JSON
 - MADS/XML
 - MARC/XML

Figure 4: Library of Congress Subject Heading (LCSH) example. Screenshot from Library of Congress website on 16/04/2021, by the author.

Knowledge Organization Systems

Knowledge Organization Systems (KOS) are systems used to organize and manage knowledge based on semantic understanding of the natural and real world. Numerous KOS exist and their utilization is dependent on the way knowledge is understood and how it is perceived by humans, and which might have been adapted to different social and cultural environments. Organizations, institutions and governing entities, when organizing knowledge, choose either one of these KO systems, or to combine them along with IR systems, in order to facilitate the information and knowledge accesses. Hence, a brief presentation on three different existing KOS is dealt with in the following section: Thesauri, Classification Systems, and Ontologies. Other KOS exist for classifying and indexing knowledge and information, such as the organization of maps and geographical dictionaries called Gazetteers, Library of Congress Subject Headings (LCSH), but we will not go through them in this paper.

Thesauri system

Thesauri is one of the most used Knowledge Organization Systems. It is defined as the system that allows the control and structure of vocabulary in “which concepts are represented by terms, organized so that relationships between concepts are made explicit, and preferred terms are accompanied by lead-in entries for synonyms or quasi-synonyms” (ISO 25964-1; clause 2.62; as cited in Hjørland, 2021). Thus, thesauri systems help the structure of world knowledge based on synonyms that are closer to each other in terms of concepts. It can be explained as a limited vocabulary and “a predefined set of semantic relations” (Hjørland, 2021).

Thesaurus does not give an exact and precise meaning to the sought word or concept, but instead, it gives related terms and synonyms that are close in meanings to that concept, which provides a better and closer understanding. Hjørland (2021) gave an example of UNESCO’s thesaurus (see figures 5 and 6) that is constructed on alphabetical bases but it also gives hierarchical structures of terms based on related concepts. Another thesaurus example that we have considered while doing the research is, the Rare Books and Manuscripts section (RBMS) of the Association of College and Research Libraries (ACRL), a division of the American Library Association (ALA), and which provides also an alphabetical and hierarchical system to follow (see figure 7).

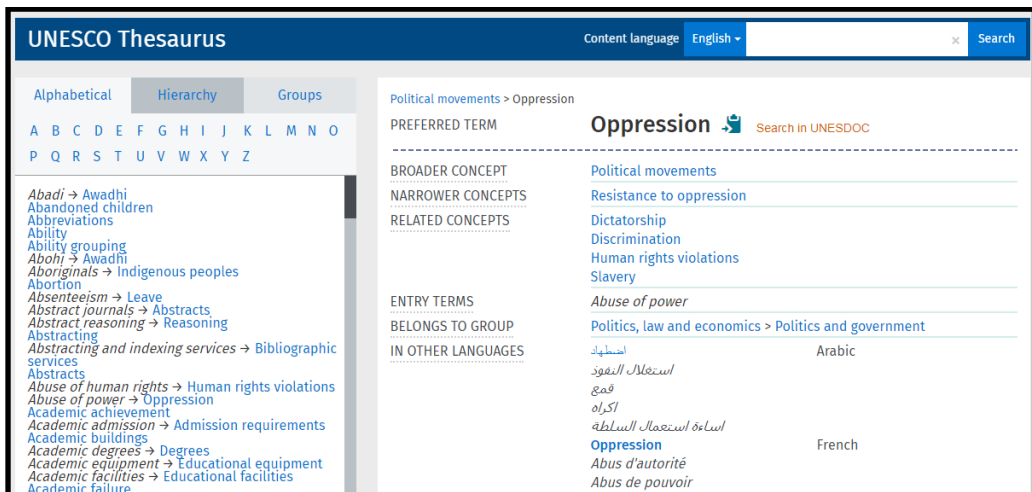


Figure 5: Example of Alphabetical search in the UNESCO Thesaurus.

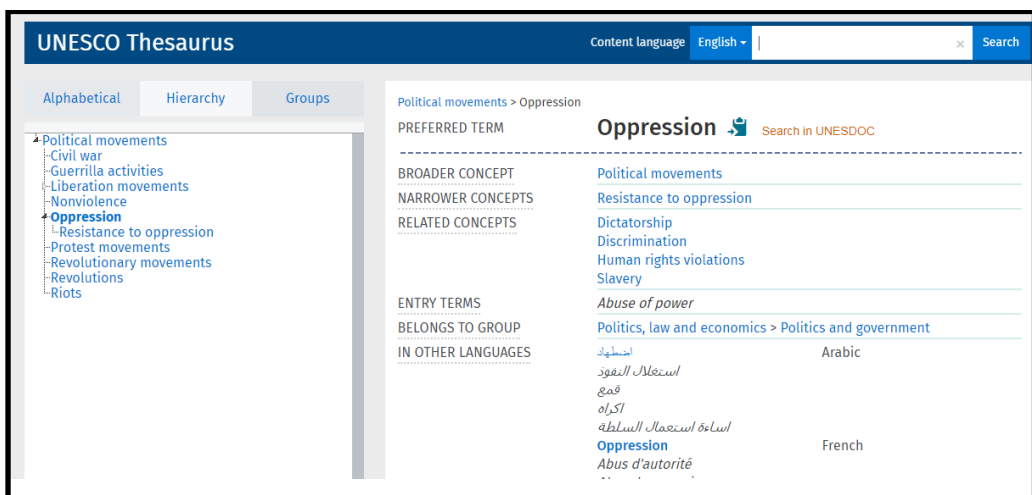


Figure 6: Example of the usage of UNESCO’s Thesaurus for a Hierarchical relations of the sought term ‘Oppression’ (see on the left the related hierarchical list)

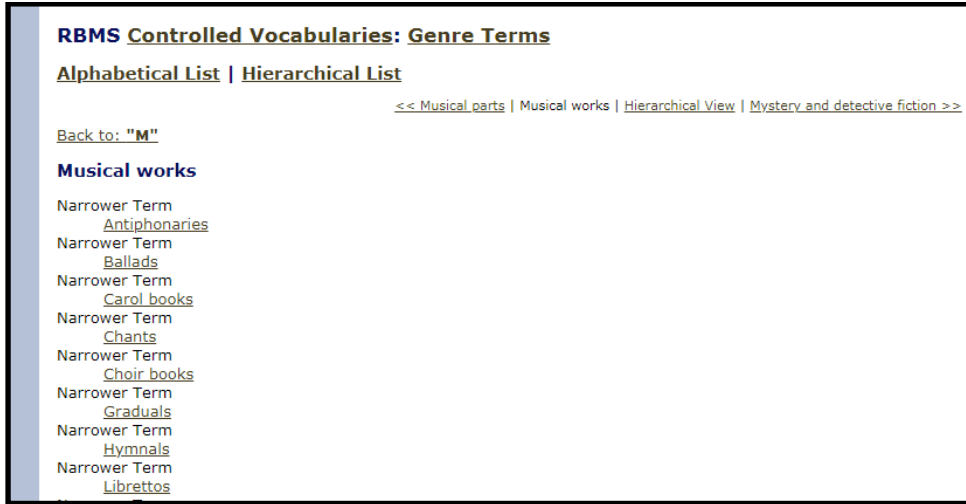


Figure 7: RBMS Thesaurus example, that is based on Alphabetical structure, and which also provides Hierarchical list

Notice that in both examples, shown above, there is a presentation of an alphabetical order of the thesauri, but which also provides a hierarchical list of the searched term, this way a conceptualization is narrowed down to the context that the user is seeking.

Classification systems

Classification systems are bibliographic systems invented for organizing and classifying documents (Books, essays, CD,...etc.), and/or document representations (Hjørland, 2021). They were created in order to “offer a theoretical and methodological ballast for solutions of problems concerning knowledge organization and thematic information treatment” (Ravazi, Moreira, 2017).

The most used and well known classification systems are Dewey Decimal Classification (DDC) and Universal Decimal Classification (UDC). These two systems were developed in the nineteenth century in order to facilitate the retrieval of documents and information. As bibliographical classification systems, they rely on the scientific classification and understanding of the world, in which they try to reflect based on scientific classification, and how information and knowledge are perceived and structured within each scientific field (Hjørland, 2021). However, as declared by Hjørland (2021), “library classification systems in particular often are very conservative in updating, and unfortunately often reflect outdated knowledge”. Bibliographical classification systems attempt to outline “knowledge in different areas featuring subjects based on hierarchical divisions from the general to the most specific” (Ravazi, Moreira, 2017). It is also worth mentioning, that these bibliographic classification systems were not designed particularly for digital resources, and therefore, their update and adaptation to the current circumstances is necessary.

Although, it has been argued by Hjørland (2008), that the traditional approaches followed when creating the bibliographic classification for example by Dewey, was hardly based on scientific approaches and academic foundations. As a businessman, Dewey's main concerns were not related to optimizing the systems for library users, but rather, to find an efficient way to handle and organize library's collections (Hjørland, 2008). Hence, Hjørland described DDC as "the dream of the library administrator rather than the dream of the library user" (2008).

Another example of classification systems is Taxonomy, which uses hierarchical models from the most general into more and more inclusive groupings. It is a scientific classification that finds its roots in biology, mainly, zoology, as it was created by the Swedish botanist, zoologist, and physician Carl Linnaeus. The scientific division used in the taxonomy classification segregates three main categories. The largest one is the domain category that has within it three domains (Bacteria, Archaea, and Eukarya). The second category, which is contained by the domain category, is the kingdom category and that includes many branches inside of it (animals, plants, fungi, bacteria,...etc.). The third and last subsequent categories are in order as such: phylum, class, order, family, genus, and species (see figure 8).

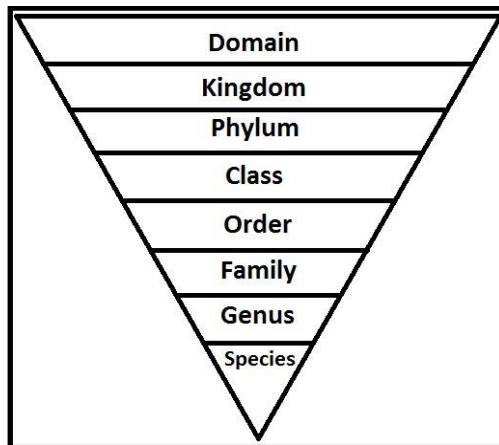


Figure 8: Taxonomy classification logic and structure.

Ontology

In the Knowledge Organization, there can sometimes be confusion between Thesauri and Ontologies. Hjørland (2021) described ontologies as systems that are "constructed by using formal languages, and many authors consider the use of such formal languages a necessary condition for a KOS to be an ontology". Similarly to thesauri, ontologies are based on simplified conceptualization of knowledge that is based on the world we are living in, and that we wish to represent (Hjørland, 2021). Therefore, Hjørland defines it as "an explicit specification of a shared conceptualization"(2021).

Smith et al. (2005) discussed the semantic relations in ontologies which are mainly expansion of Thesauri and classification systems (Hjørland, 2021). For biomedical ontologies, Smith et al. (2005) developed relations as such:

c instance_of *C* at *t* – a primitive relation between a continuant instance and a class which it instantiates at a specific time

p instance_of *P* – a primitive relation between a process instance and a class which it instantiates holding independently of time

c part_of *c1* at *t* – a primitive relation between two continuant instances and a time at which the one is part of the other

p part_of *p1*, *r* part_of *r1* – a primitive relation of parthood, holding independently of time, either between process instances (one a subprocess of the other), or between spatial regions (one a subregion of the other)

c located_in *r* at *t* – a primitive relation between a continuant instance, a spatial region which it occupies, and a time

r adjacent_to *r1* – a primitive relation of proximity between two disjoint continuants

t earlier *t1* – a primitive relation between two times

c derives_from *c1* – a primitive relation involving two distinct material continuants *c* and *c1*

p has_participant *c* at *t* – a primitive relation between a process, a continuant, and a time

p has_agent *c* at *t* – a primitive relation between a process, a continuant and a time at which the continuant is causally active in the process” (2005).

Therefore, ontologies are considered as systems that are concerned with the organization of concepts (Hjørland, 2021).

Information Retrieval (IR) in Computer Science

Since the rapid growth of technologies and technological materials in the last decades, the commonality of the usage of IR has become very absorbing for the field of KOs (Hjørland, 2021). The most famous and most popular in terms of usage worldwide is Google.

Recent statistics have shown that there is a complete domination of 92.41% by Google as a search engine that is utilized globally (see below figures 9 and 10). The reason for this popularity is that it gives to the users/searchers a precise results only by typing few key words on the search bar of Google (users' friendly), and which will be transformed the data that has been given into a response by retrieving all the related items where this data has been mentioned (either a title of a document, or a word within a document, a video/image link, and so on). This means that prior knowledge from the user is required to be able to retrieve the exact seeked information (Hjørland, 2014), and that can be a handicap to some extent for a searcher who does not hold this knowledge. An example of this, if a user would like

to search “Diplodocus” information, but he does not know the name of the dinosaur, he might only type dinosaur but he will get a bunch of information and many types of dinosaurs, which can become confusing for the searcher, or he might know that the dinosaur is a ‘Sauropod’ but does not know which one of them, as they have similarities in the shape. For non-experts, looking for such information can lead to confusion or even misinformation, and “Unknown documents of relevance for a given inquiry may have to be searched in a context and with concepts as well as symbolic systems unknown for the searcher. This is in particularly the case if there have been paradigm shifts in the field of enquiry” (Hjørland, 2021)

Recognizing the need of KOs to classify this type of information is crucial for better comprehension and precision. Not to forget that most search engines in most cases, specifically related to scientific confirmed data, use KOS, such as thesauri and ontologies, for information retrieval.

Let us broaden and try to simplify the vision of IR in the computer science field for a better comprehension of its utilization and utility.

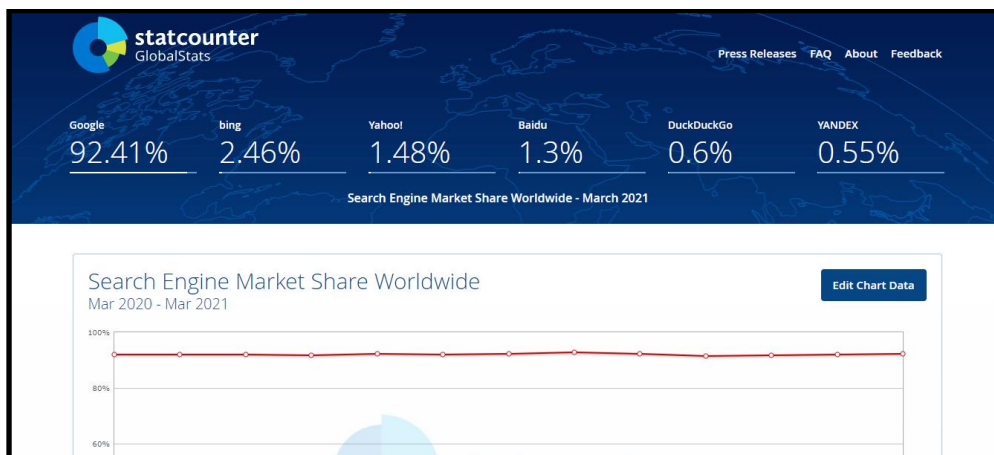


Figure 9: Statistics concerning famous search engines that show the domination of Google. Retrieved from ‘Statcounter GlobalStats’ on the 16/04/2021, main website: <https://gs.statcounter.com/search-engine-market-share>

Roberto Busa stated, in a foreword of the book “Text Encoding, A Companion to Digital Humanities” (2004), that “*Humanities computing is precisely the automation of every possible analysis of human expression (therefore, it is exquisitely a “humanistic” activity), in the widest sense of the word, from music to the theater, from design and painting to phonetics, but whose nucleus remains the discourse of written texts*”, it is indeed, no matter how much we try to deny it, computation still needs a humanistic action to be done, but let us first explain how does it work.

Computation, or what is known today by computer science, had started back in the 1940s, with it little by little during and after the second World War. Programs became available at universities related to the new field in the late of 1950s (Denning, 2010). Although, it was still argued if computation can be considered as a

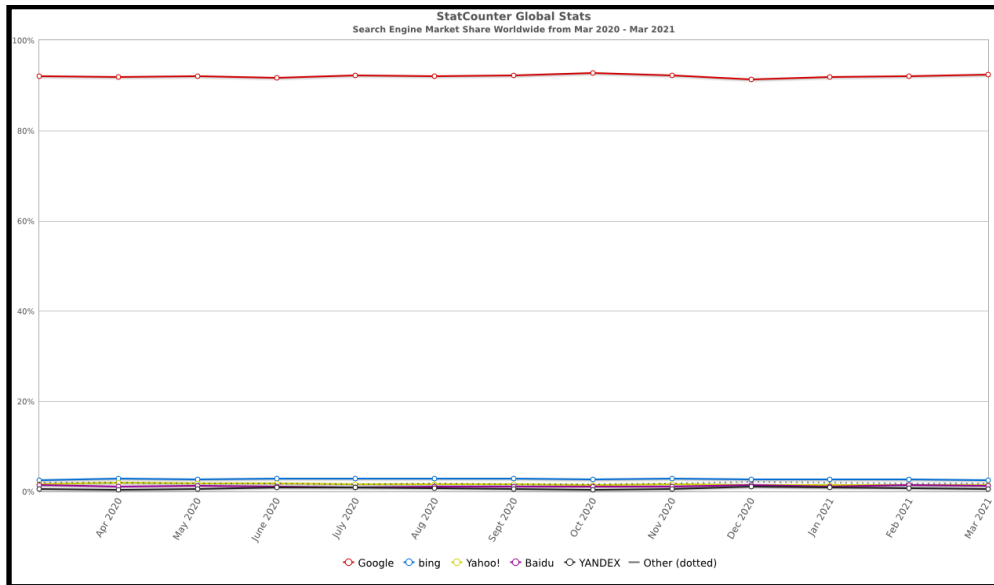


Figure 10: Search Engine Market Share Worldwide between Mar 2020 – Mar 2021.
 Source <https://gs.statcounter.com/> (consulted on 16/04/2021).

science as it is understood that “science as a way of dealing with the natural world”, however, “computers looked suspiciously artificial” (Denning, 2010). Indeed, the evolution of computation as a science and its projection of the natural world began in the late 1990s, “when leaders in the field of biology, [...] David Baltimore and echoing cognitive scientist Douglas Hofstadter, said that biology had become an information science and DNA translation in a natural information process. Many computer scientists have joined biologists in research to understand the nature of DNA processes and to discover what algorithms might govern them” (ibid, 2010).

In the present-day society, the usage of computation and new technologies became a must and no stupefaction to anyone, as it turned out that most of the world’s population is using new technologies one way or another. Thus, the retrieval of information depends primarily on the human information behaviour (Bawden, 2007), and which is alleged as “all aspects of behavior with respect to interaction with, and use of, all forms of information and knowledge, through all sources, channels and media, including informal and unrecorded communication” (Wilson, 2000, as cited in Bawden, 2007).

Accordingly, information-retrieval depends directly on the human information behaviour. Bawden pointed at the relation between three things that are connected and bound to each other when seeking information that are information seeking, human information behaviour and information retrieval (2007). It is the human action that defines the need of information, how the information is sought and how it is retrieved and finally the human interaction with the information found.

Wilson’s modelling of information human behaviour research (1999) is one of the most explanatory models (see below Figure 11), which shows how information

is compliant to human behaviour. Information on how it is sought, searched, retrieved and processed is a human activity, which is related to the social and cultural environment of the seeker. Information retrieval is defined as “the purposeful searching for information in a system, of whatever kind, in which information – whether in the form of documents, or their surrogates, or factual material (‘information itself), are stored and represented” (Willet and Jones, 1997, as cited in Bawden, 2007). This is not a new statement for the field of Library and Information Sciences (LIS) or computer science (CS) about the meaning of IR, however, the purpose of bringing it up is that IR systems do not only depend on information behaviour of the seeker, but also and importantly, they rely on the designer’s perceptions and ideas, in order to increase the user’s self-efficiency and self-efficacy, and reducing, as much as possible, the risks of failure by the user (Wilson, 1999), which means that human action is needed priorly to its introduction in any kind of systems, by all means, that IR systems are not social/cultural free nor independent from concepts, instead, they count on the adaptation from the designer whom himself depends on social, cultural, political, religious frames, and its awareness of the users’ surroundings, makes also the IR systems dependent and limited to their environment.

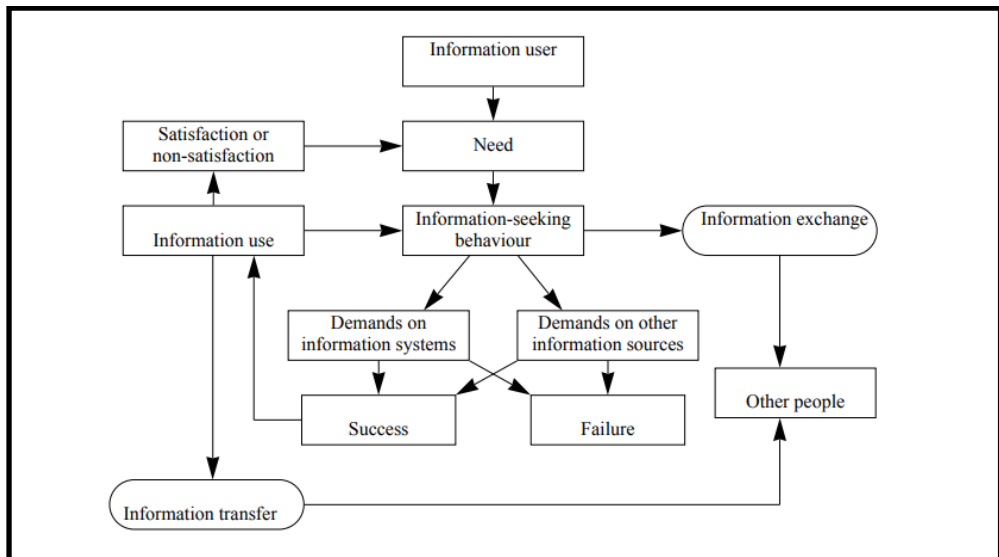


Figure 11: Wilson’s model of information behaviour, 1999.

Information Retrieval as a field of study that has been included within computer science, is not only concerned with systems and their superficial design and utilization, it is rather, the consideration of the designer’s mind-set and understanding of the users, which creates an indirect connection between the designer and the user on different levels. Wilson’s first model about information human behaviour (1996) project to some extent the circuit in which information is modelled by the designer and the user (see below figure 12). Hence, IR systems are

“an explanation of the system designer’s cognitive model of what the system should do and how it should function” (Wilson, 1999).

Nevertheless, the usefulness and efficacy of IR systems are not problems-free. Many issues and queries arise about their efficiency and trustworthiness, specifically for none-scientific users, who do not have any prior information about a sought-subject/information, and who might not have the necessary background to distinguish and evaluate information, and which can lead to mis-/dis-information. The question will be why and how can IR systems lead to confusion?

The latter that IR systems use to a document or information when a keyword is searched, is that it rely on the best match of the searched term and any type of documents containing that term (which can be a title, a full-text, a description...and so on) (Hjørland, 2021).

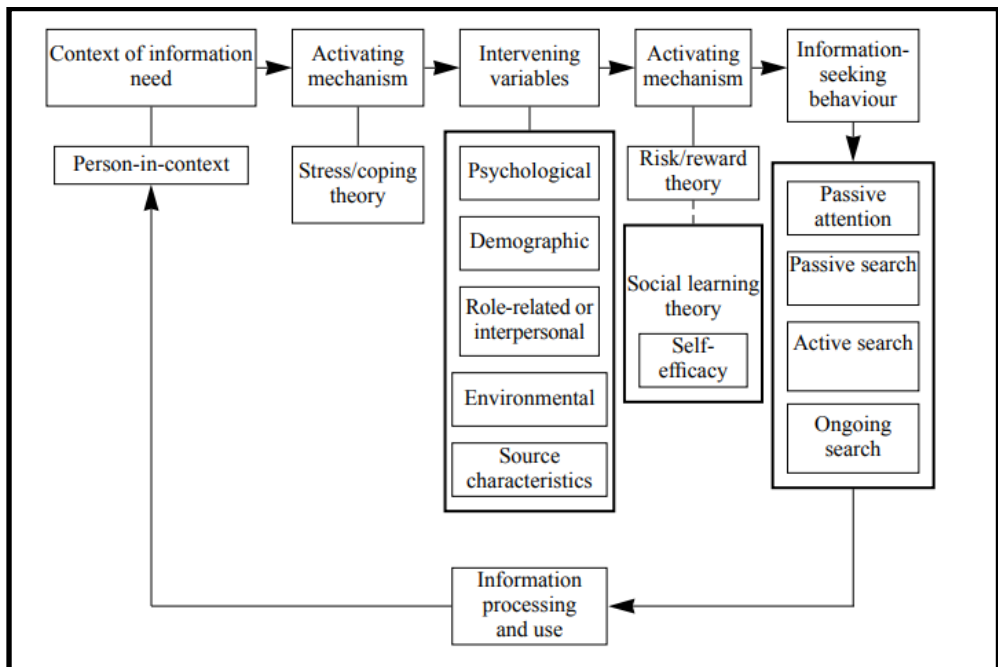


Figure 12: Wilson’s 1996, an earlier model of information behaviour

Therefore, the leading approach in IR systems “is about information from the document itself, not about value-added information” (Hjørland, 2021), and that is problematic as it contains no semantic structure such as the ones used in KOS, and which links concepts and information (ibid, 2021). This is very important when looking for unknown knowledge/information, as the latter can be clear to the user when using KOS. Of course in IR the search can be modified by excluding the irrelevant items, which will expand the query for more search. However useful it might be, this technique can be time and effort- consuming, especially, if the searcher does not have a clue of the structure of the searched item/information.

Compared to KOS in this matter, IR systems show weakness, as they do not hold a scientific description of the realities such as KOS do.

Most of the search engines used nowadays, such as Google, Yahoo, Bing...etc. are developed for economical reasons, and their main gain and benefit is based on the frequency of the usage and publicity. Therefore, they are not hundred percent trustworthy when delivering information, as the first links that appear in the first page are conserved to entities who pay more to maintain their commercial and economical status. The methods and approaches for finding information in IR systems are either based on the ‘classical IR model’ (Boolean, Vector and Probabilistic), ‘Non-classical IR model’ (Information logic model, situation theory model and interaction models), or ‘Alternative IR model’ (Cluster model, fuzzy model and latent semantic indexing (LSI) models)¹.

Hjørland discussed in his recent paper (2021) the different approaches of IR used by Google, underlying four principles that are “(a) “exact match”, (b) “best match”, (c) popularity measures, and (d) personalization”. Exact match of the query is a model based on similarities between the query and the document (of any type), it allows “building blocks search strategy” (Hjørland, 2021). This principle no matter how simple and helpful it can be, it holds many flaws, as it is unguarded, and the search is random as the matches can be anything (title, text, video...), and from anywhere (trusted and untrusted websites and databases).

On the flip side, Best match method consists in finding similarities of the query out of their context. An example of this, if a user is looking for “anthropology” as a keyword, the system will recall all the best matches to this query, that is relevant high but context-free, and which can be useful and fast, but at the same time a problematic issue as the best match principle is “statistical, there is an element of popularity measure in this” (Hjørland, 2021). This can create confusion to the users when looking for science-based information. Popular information can lead to misinformation.

The third method, used by Google (we take Google’s example as it is the most popular and used search engine), is the one based on popularity. As said earlier in this section, search engines in first place are commercial tools, their main goal is an economical one. We shall not deny the usefulness and efficacy of search engines such as Google, but at the same time, we should keep in mind that there are economical reasons behind their work.

Last but not least, Google’s personalization principle. This is more harmful than useful to any searcher, as first, when connecting to Google, it proposes you to accept the terms of usage, this includes the identification of the IP address of the user, and therefore his geographical location (see figure 13). By doing so, the search engines propose information (specifically on the first 10 pages) based on geographical characteristics (Hjørland, 2021). This, by definition; becomes a subjective search rather than an objective one, which restrain the ability of the users to modify it, as it is automatically recognized by the system and set as a default option. The search

¹ More information about these models can be found: https://www.tutorialspoint.com/natural_language_processing/natural_language_processing_information_retrieval.htm (consulted on 19/04/2021).

becomes more exclusive to the user, instead of an inclusive one, which does not give free access to knowledge and information equally, and can be discriminative, as it depends on the availability of content on geographical traits. The free-will of the searcher to self-selection is then limited and reduced to this feature that can block the expansion and exploration of general and global knowledge. The personalization aspect is not limited to this issue, but moreover, the search engine recalls previous searches made by the user and presents information based on prior memory. It is, indeed, useful in some cases to look back at history and location when searching a known information or previous forgotten knowledge, but most of the time it is more harmful, as it contradicts with the right of the users to access global information with no discrimination of any sort.

On that account, we should not deny the usefulness and popularity of search engines such as Google, but a further improvement and amelioration is needed in order to deliver an easy, structured, efficient and equal information and knowledge to the users, and for this, more inclusion of KOS and combination in IR systems should be made.

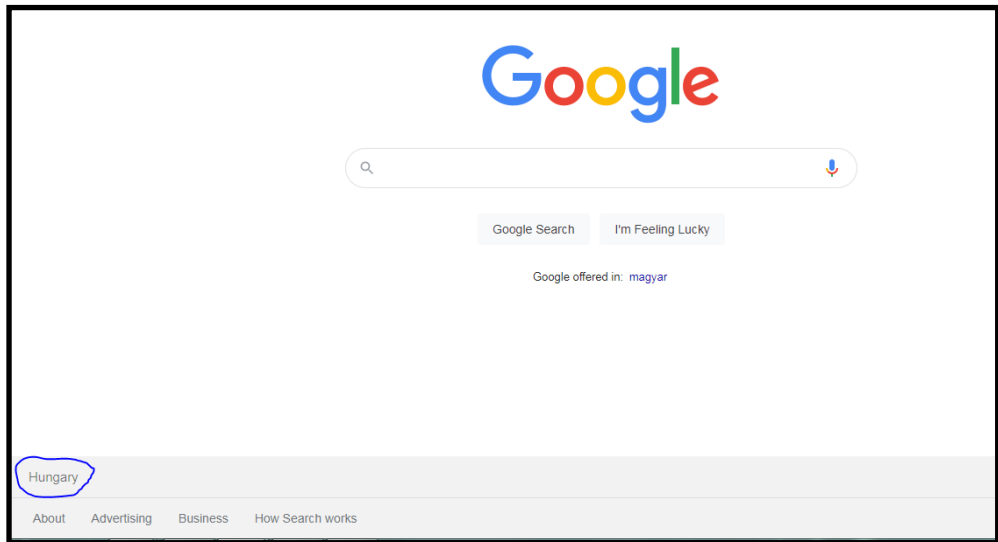


Figure 13: Google search engine locating the IP address of the searcher. By the author (20/04/2021)

Language and KO

The relationship between language and KOS is indispensable. Language changes continuously, by developing different meanings each time; either by including new terms or erasing others.

The terms and synonyms used in each field get modified all the time. Which depends on scientific research and new discoveries, but as well, the environment, the cultural and societal settings. This leads directly to organizing not only “neutral”

knowledge, but as well as organizing concepts that can have either syntagmatic relations, or paradigmatic relations (De Saussure, 1916).

There is an ongoing debate on how concepts are organized, how language plays a significant role in shaping these concepts, and how they can relate (or not) to each other. Hjørland states that “Homonymity is not an inherent quality, but a human choice, the usefulness of which may depend on the size and characteristics of the literature represented in a given database and thus is relative” (2021).

A good example to illustrate this is; Red is to be considered as a color, which is the hyponym (Red) of the hypernym (Color); integrating a concept that humans use mainly in everyday life worldwide is the example of the Road code. Red as a sign of ‘stopping’, which can have two different hypernyms such as “road code” and/or also “color”. Therefore, a word or a concept can have syntagmatic relation, but also paradigmatic relation (Chiu and Lu, 2015).

In this sense, Knowledge Organization and language usage are indispensable to each other which necessitates human action. Language can be free from concepts, but it can also contain important paradigmatic approaches as it is used to express the human actions and which are used to categorize knowledge that is stored and organized.

This idea of organizing concepts using language will be expanded in the following section that is related to structural linguistics, which is based on binary oppositions from Ferdinand de Saussure’s point of view.

Concepts organization from the Structural linguistics approaches

Considered as the father of modern linguistics, Ferdinand de Saussure came up with new thinking of linguistics based on a structural approach, which is based on contrast and equivalence systems, and what have been mentioned earlier in the last section of this article, syntagmatic relations and paradigmatic relations.

De Saussure argued that the existence of words consists in the contrast with other objects. He stated that there is a sign that contains the signifier and the signified; let us take the previous example as before for a better explanation of the key concept of structural linguistics. Taking the word “Red”, on the one hand, is considered as the signifier (signifié), that is the sound of the letters used to express what we are discussing; on the other hand, the signified (signifiant) is the concept of the thing that we are talking of, which is in this case “color” or sign “Stop”, this helps the human perception to relate the word ‘red’ with the concept color or the road sign ‘stop’, and which is a way of organizing the world for a general human understanding. Therefore, there is a referent in our case that is ‘Color’ or “‘Stop”, which is an internationally recognized token.

The construction of a language and how it is organized is very interesting when looking at it from the structural approach. Languages are used to restore and conserve knowledge, but how it is done depends on different aspects of life such as the social, cultural, religious and political matters of the life of each community.

The linguistic structural analyses also found their ways to social sciences, and they have been applied to explain also the social/cultural structures of communities.

One of the leading anthropologists in structuralism approaches is Claude Levi-Strauss, who had applied the structural linguistics theories of De Saussure in several of his analyses. He had discussed and analysed the existence of myths, and the way they are created and stored as a collective knowledge based on structural linguistics approaches.

Levi-Strauss argued that myths are, firstly, an individual creation, utterance that has a ascertain function which finds its way to an individual act (1981), he puts it as such *“it may be asked, should one have such reservations with regard to the subject when dealing with myths, that is, with stories which could not have come into being unless at some moment-even though, in most cases, that moment is beyond the reach of enquiry-each of them had been conceived and narrated in the first instance by a particular individual? Utterance is a function confined to subjects, and every myth, in the last resort, must have its origin in an individual act of creation. This is no doubt very true, but, in order to achieve the status of myth, the created work must cease precisely to be individual and, in the process of generalization, must lose the essential part of those factors determined by probability with which it was infused at the outset, and which could be attributed to the particular author’s temperament, talent, imagination, and personal experience”* (1981, p. 64).

Based on the structural binary opposition analyses, and from the last mentioned from Levi-Strauss’ ideas about myths, organizing knowledge can be seen to have similar patterns. Knowledge at first can be described as an individual discovery (that can also be an idea that has been developed by a group of people: scientists, politicians, artists...etc.). This discovery is not restored or saved if it is not recognized by the community in which the individual belongs to. Recognition is the key to the implementation of that knowledge inside the society. Therefore, an individual human action is always in play. Expanding this idea much further, and taking the KOS and IR tools, there is always a certain duality between the human interference when setting these systems. The contrast between how humans perceive things and how technology analyzes data is, however, complementary. Technological tools cannot exist without human actions, whether we admit it or not, but also these technological systems show the weaknesses and limitations of human perceptions and inputs.

It also leads us to the question of how “languages” are constructed and how they are understood by people with different backgrounds and different educations and knowledge. Language is defined as the ‘ability’ to speak and understand (Chomsky, 1997, p. 13), which can be confused with knowledge. Chomsky argued in his article ‘Language and Problems of Knowledge’, that Knowledge cannot be considered and limited as an ‘ability’ when referring to languages (1997), he explains that knowledge is more than that, as he gives an example of a person suffering from Parkinson’s disease that hasn’t forgotten the P-ability to speak (Philosopher’s ability), but who lost the ability to do so (Chomsky, 1997, p. 13).

Chomsky’s idea can be quite confusing to comprehend, but it is more explanatory than it might seem to us. Giving another significant example of this, a person who had an accident and lost his abilities to walk, drive or run; but he definitely has not lost his P-ability for these activities, as he will certainly recover them giving the time and exercises needed to do so. The question that poses itself is:

what is knowledge then? Is it the P-ability or the ability to perform certain activities? Or does knowledge combine them both?

Chomsky (1997) points out to the unsolved knowledge in language, by addressing the absurdity of what he called the "E-language", and "*where "E" is intended to suggest "extensional" and "externalized". The definition is "extensional" in that it takes language to be a set of objects of some kind, and it is "externalized" in the sense that language, so defined, is external to the mind/brain. Thus a set, however chosen, is plainly external to the mind/brain.*" (1997, p. 7).

This was also identified as "the totality of utterances that can be made in a speech community" by the American linguist Leonard Bloomfield (as cited in Chomsky, 1997, p. 7). De Saussure referred to this by separating 'Parole' and 'Langue' in the form of the use of language 'Parole', and the system of language 'Langue'. He indicated that, on the one hand, 'Parole' is what an individual uses when speaking; while on the other hand, 'Langue' is the system of the language and speech shared by the community (ex. syntax, semantic, phonology...etc.) (De Saussure, 1916).

Comparing the two ideas from Chomsky and De Saussure, we come to understand that what Chomsky referred to as E-language is the so-called 'Langue' for De Saussure. While what 'Parole' meant for De Saussure might be interpreted as the 'I-language' for Chomsky, and which the "I" is defined as "*"intensional" and "internalized". The "I-language" is what the grammar purports to describe: a system represented in the mind/brain, ultimately in physical mechanisms that are now largely unknown, and is in this sense internalized; a system that is intensional in that it may be regarded as a specific function considered in intension -that is, a specific characterization of a function- which assigns a status to a vast range of physical events*" (1997, p. 10). In a more recent paper (2006), Chomsky added to his definition the "I" as 'individual' (2006, p. 175).

This I-language that Chomsky calls attention to is what he named as the 'universal grammar' that is the initial state of a common human possession, that is already composed within the mind/brain prior to the acquisition of any language (Chomsky, 1997, p. 11). Based on this idea, the I-language/universal grammar might be interpreted as the state 0 in which all humans are at when born? Or prior to that, inside the womb of the mother? As it has been proven already by scientists that babies start their learning process from the wombs of their mothers.

Either way, there is a certain initial state of the brain/mind as Chomsky mentioned, which organizes the acquisition of any language by appropriating concepts to the spoken and written words. Therefore, there is always an internalized human action that comes into play, and which is necessary to organize knowledge, and which the source of it is still unknown. The way of how we attribute certain words to certain concepts and in different languages that are grammatically related but also different is very gripping. It also pauses us to question the neutrality of knowledge organizations, which knowledge is thus constructed based on social and cultural features.

Hence, a more discussion about Knowledge Organization and its linkage to cultural-heritage is to be more developed later on in this paper.

Categorization and power

The main objective of Knowledge Organization is to put things into categories that are based either on scientific discoveries or social, cultural, and political settings. The manners of how categories are perceived and placed are necessary to be discussed in order to comprehend how knowledge is constructed and put together.

Sacks (1995) explains that categories help in organizing the world. He gave an example of putting women into the female category which is a description of a person, and which reflects its social position. He goes on adding that once the category is known, we do not need to know anything else except for the category. This is a very interesting point in Knowledge Organization, and how knowledge is managed for the human perceptions of the world. If we take the example of food, and we say that an apple is in the category of fruits, then our understanding and usage of the apple goes accordingly to that by making a dessert (which is also a category for itself and which is in the category of sweets) out of apples. While our perception would be very much different for an apple if it was perceived as a legume or a seed, we would have to make something out of it that makes it a salty dish presented for lunch for example (this is not a rule, but just an example for a better comprehension, apples can still be used for salty dishes).

The relation between Knowledge and categorization of the world is very powerful. Categorization helps in stabilizing Knowledge within our world, and vice versa, "Knowledge is stable for any category" (Sacks, 1995, p. 46). If knowledge is established and acknowledged by all the members of the society (or just the leading members of the society e.g. politicians, scientists, historians... etc.) to a certain category, then an image is created and applied to all members of that category. It projects a certain description of any member (persons, things...etc.) belonging to that category.

Therefore, the stability of Knowledge depends on the categories made by humans in order to control the world, and once the knowledge is established, acknowledged and confirmed by all members, then the categorization is sought as true and depends on that knowledge that was implemented within the society and culture.

We shall expand extensively and in detail, this idea of categorization of the world and its relation with knowledge in different contexts.

Boolean data type and Categorization

The Boolean system is one of the categorization systems that is based on an algebraic system of logic that considers things either true or false, black or white. Its concept is utilized in computer science that considers two values X and Y, which one is related to the other either $X=Y$, or $X\neq Y$. If X is true then Y is true (case $X=Y$), or X is true then Y is false (case $X\neq Y$), or X is not true then Y is true ($X\neq Y$) or X is not true then Y is false ($X=Y$). These are most of the cases with these two values that are related to each other and depend on one another, as they are categorized, the options of what X and Y can be narrowed down to four cases, which makes it easier to retrieve or to extract later on.

This type of categorization is also used in our social life, 'we' humans might not be aware of it in our daily usage, but 'we' unconsciously tend to apply it in our conversational or our everyday behaviours. Sacks (1995) has had analyzed this kind of categorizations in social conversational contexts, and found it interesting how this 'We' is utilized and applied by humans to categorize themselves, but at the same time to prove a kind of contradiction in their behaviour with members of certain categories that they belong to or consider themselves part of (pp. 333-340).

Sacks had expanded the idea of the 'We', and its categorization by saying: "'We' can be used to refer to all members of the category that have ever lived and may ever live... once you've gotten the plural pronoun replacing the categorial term, there is no movement, necessarily, to any set-of- members reference, i.e., reference via a list of particular names. The categorial reference, once established, holds." (1995, p. 335).

This is a very appealing note from Sacks about categorization and usage of plural forms that are necessary to humans for organizing the world and knowledge in particular. For example, if a person is looking at what is 'Sauropod', once, the category is known that it is a type of dinosaur, the name is no longer important to remember (maybe only for specialists like paleontologists or just fans of dinosaurs..etc.). But for a person who is just seeking to discover for the first time the meaning of the word, it is maybe enough to know in which category to put it, such as the dinosaurs category, and therefore, all the characteristics known about dinosaurs are also applied to that particular one, even if for example, Sauropod is herbivorous dinosaur compared to Theropoda that is carnivorous, this detail would not be of importance once the holding category is known, unless the person is looking for more detailed information. This procedure organizes the sense of the person's understanding of the world.

To resume the idea of this section, which purpose to project the similarities between the Boolean system and the concept of categorization, once the categorization is done and acknowledged for a certain category, the characteristics that this category hold are applied to all its members, which makes it easier to extract information about the category or member if only the characteristics are known. This is only the first case, the second case is to prove the contradiction of a member belonging to a certain category for which the characteristics are known, but which makes that certain member be the exception of that particular category for which he/she/it belongs. Sacks gave a good example of this referring to a black fellow who might be categorized as a Negro pertaining to Negroes/black/African.., category, and who might be an exception when the word 'but' is expressed, so as such "he is a Negro, but...". Therefore, a categorization is either an expression of similarities of two entities/ values, or of contradiction of them.

Categorization in social contexts

From the last example discussed in the previous section, it is necessary to enlarge the discourse about social categorization, and how it is done from an anthropological point of view that concerns the categories applied upon people, and how they are captured and implemented within societies.

In his article “*Ethnicity without Groups*”, Brubaker points out the differences between groups and categories which also makes it clear to find a relation between them. Brubaker discussed that categorization “includes limiting access to scarce resources or particular domains of activity by excluding categorically distinguished outsiders (Weber 1968 [1992]: 43ff, 341ff; Barth, 1969; Brubaker, 1992; Tilly, 1998), but it also includes more mundane actions such as identifying or classifying oneself or others (Levine, 1999) [...] We can analyze the organizational and discursive careers of categories—the processes through which they become institutionalized and entrenched in administrative routines (Tilly, 1998) and embedded in culturally powerful and symbolically resonant myths, memories and narratives (Armstrong, 1982; Smith, 1986).” (Brubaker, 2002, p. 169).

Therefore, categorization is important in the sense of holding control of the world by making sense of it (Brubaker, 2002). It gives an agreed explanation (Sacks, 1995) to certain actions or phenomena. The political aspects of categories are to be analyzed from above and below to comprehend how these categories are forced and inculcated at certain points in social and cultural settings (Brubaker, 2002). Categories are made and “used by individuals to make sense of the social world; linked to stereotypical beliefs and expectations about category members; invested with emotional associations and evaluative judgments; deployed as resources in specific interactional contexts; and activated by situational triggers or cues.” (Brubaker, 2002, p. 170).

Schegloff (2001) goes beyond this explanation of the importance of categories in constructing a culture by describing how categories consist of major repositories for building a knowledge that becomes a common-sense by members within a particular society. He pointed out that “*The importance of these category terms (and their organization) goes far beyond the role they play in persons’ practices for referring to other persons (Schegloff, 1996d), important though that be. The categories of these collections are one major repository, perhaps the major repository, for commonsense knowledge of the society by members of the society as members of the society. “Knowledge” of what different “sorts” of people are like, what they do, how they behave, and so on—one key element of what is often termed “culture”—is organized and stored by reference to these categories (Sacks, 1992, 1:40-49)*” (Schegloff, 2001, p. 308).

A noteworthy point about categorization, when building knowledge for commonsense, is that it is ‘bound’ by applied activities, and as Sacks (1972, 1992) referred to it as “category-bound activities” (as cited in Schegloff, 2001). These activities are the keys of “labelizing” members into certain categories. In this matter, Schegloff argues that “One could provide a transparent account for someone having done some action or behaved in some way by involving their membership in that category, for example, by referring to them with a reference form—a “category label” (Moerman, 1988), which names a “sort” of person who does that sort of thing. Indeed, one could figure who had done some action, especially a problematic one, by seeking out persons who were members of a category to which actions of that sort were bound, who were “known” to do “things like that.” As Sacks put it in his initial discussion of this matter in an early one of his lectures (1992, p. 1:180): “*The fact that some activities are bound to some categories is used, then, in a tremendous*

variety of ways, and if somebody knows an activity has been done, and there is a category to which it is bound, they can damn well propose that it's been done by such a one who is a member of that Category. What's important, in part, is that it's not the case that deviant activities are especially problematic, but there are categories of persons who do deviant activities and you've got a solution to a deviant activity if you've got a member of a category which is known to do this."² (Schegloff, 2001, p. 308).

Therefore, there are activities in social context that are done by individuals that might be incomprehensible to the society, and for a better comprehension, Humans need to seek explanations for these activities or actions, hence, they seek to put labels on them in order to categorize them in a certain way that make them easier to absorb and maybe accept. These categories are what comprise the assemblance of 'Knowledge'.

The role of categorization is more than preserving a certain knowledge for a better explanation of the world, it is, as many researchers have discussed, a way to apply power of one category over another. The Conversational Analysis (CA) studies done concerning "interruption" in discourses between two parties, have had great upshots regarding the implication of categories when the interruption takes place as an 'action-bound category' (Schegloff, 2001).

Schegloff expanded this point based on several previous empirical studies that found the relation between categorization and interruption to be indispensable. In this sense, Schegloff explains that "*Commonly the participants to interaction where interruption figures are not understood or described in such anonymous and activity-specific terms as candidate-interruptor or interrupted. Rather, they are as a matter of cultural practice (both putatively on their part qua participants, and on the part of observers) understood to be members of deeply grounded categories of societal membership, the categories composing the membership categorization devices, but those categories reconfigured in a particular way. In this configuration, the categories come in coordinate pairs, as in: candidate interruptor is male, candidate-interruptee is female; candidate-interruptor is employer, candidate-interruptee is employee; candidate-interruptor is teacher, candidate-interruptee is student; candidate-interruptor is professional, candidate-interruptee is client; and, generically, candidate-interruptor is superordinate (in power, status, class, income, wealth, knowledge, skill, prestige, legal entitlement, etc.), candidate-interruptee is subordinate (in the same resource)*" (2001, p. 310)².

On this account, categorization is dependent on the relation between categorizer and the categorized, which implies a certain power that favored the categorizer over the categorized. It is noteworthy to mention that in linguistics this relationship between the two parties is designated as a "Binary-Oppositions".

² In his article, Schegloff refers to the work of Smith-Lovin and Brody (1989, p. 425), where more implicit results showed how interruption takes place based on categorical organization and relation between the interruptor and interruptee is more controlled by these categories within society.

Knowledge Organization for Sustainable Development

Before discussing the impact and the importance of Knowledge Organization in Sustainable Development (SD), we find it necessary to introduce the meaning and approaches of SD.

As the technologies are developing fast each minute, human activities follow a similar rhythm. This means the consumption of environmental resources (natural and unnatural ones), and which increases the sustainability challenges caused by their heavy deterioration. It is no secret nowadays that human activity has caused more damage than good to the Earth planet and to all its creators. Robèrt et al. (2013, 2) emphasized that “the gradual decline of the biosphere’s potential to sustain civilization in the face of growing global societal needs” as a funnel (Aldabaldetrekú et al., 2016). The need for change of human behaviours in regard to unsustainable actions is necessary, in order to prevent the vanishing of resources and restore the Earth environment for future generations. It is indeed very challenging for humanity to maintain its technological and consumptive development, and at the same time to preserve its surrounding and material/ immaterial resources for those who come after us.

Hence, Sustainable Development is globally needed and necessary for humanity. The well-known definition of SD was given by the World Commission on Environment and Development’s (WCED) in the 1987 Brundtland report for a “development that meets the needs of the present without compromising the ability of future generations to meet their own needs”. Although, this definition is very optimal and unrealistic to the technological world in which we are living today, as every development has its two coin sides. The normative aspect of SD (Aldabaldetrekú et al., 2016) sets obligations of how humans act based on moral attitudes, however, it does not establish bases for the consequences of each action (Hedenus et al., 2016; as cited in Aldabaldetrekú et al., 2016).

Therefore, Sustainable Development is a complex matter, as it is linked to the complex issues related to economy, natural resource systems and society (Aldabaldetrekú et al., 2016). The interconnection between these life aspects within a social system aggrandize their complexity, as one cannot be understood or separated from the other, which result in interference of actions, one advantage economical action is not necessarily in favor of natural resources or social life. The unsustainable challenges are thus complicated to erase at once and without compromises, but it is crucial to try to reduce the harmful consequences and raise people’s awareness about the outcomes of our controversial activity.

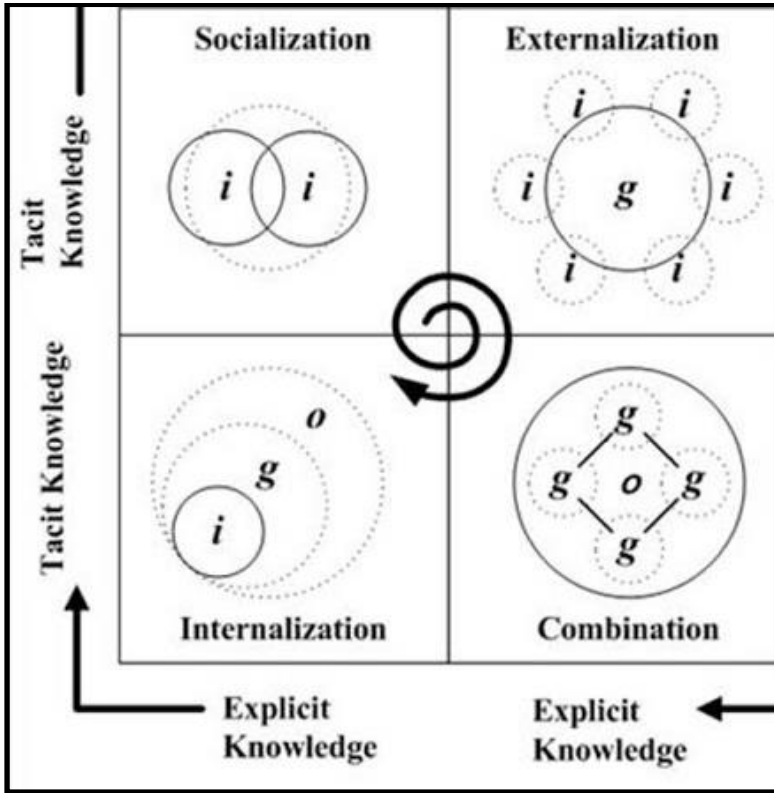
It is necessary for us to see Sustainability as a spiderweb to make it (or at least part of it), we have to consider all silks inside the web and their connection and interrelations with each other. Aldabaldetrekú et al. (2016) argued that “in order for governments, businesses, and research institutions to innovate for a better future, they need to be able to examine the full system with clarity and organisation, together with scientific boundaries and opportunities”, instead of addressing issues independently, it is preferable to treat the issue in its whole context for indelible results. So where and how Knowledge organization comes to play?

Based on linguistics analyses discussed earlier in this paper, and which was related to I-language and E-language (Chomsky, 1997), and also the structural linguistic approaches from Ferdinand De Saussure (1916) that are concerned with “Parole” and “Langue”, we realize that there are two kind of knowledge, tacit knowledge and explicit knowledge (Nonaka, 2007; Nonaka, Konno, 1998). Nonaka and Konno (1998) describe explicit knowledge as the knowledge that “can be expressed in words and numbers and shared in the form of data, scientific formulae, specifications, manuals. and the like [...] This kind of knowledge can be readily transmitted between individuals formally and systematically”. As for tacit knowledge, it is recognized as non-verbalized, and inborn knowledge (Polanyi, 1967). Tacit knowledge as Nonaka (1998) describes it “is highly personal and hard to formalize, making it difficult to communicate or share with others”. Spender’s (1996) definition to tacit knowledge that is well-defined as “knowledge that has not yet been abstracted from the practice” (as cited in Gonzalez, Martins, 2017).

Organizing knowledge helps in the preservation of tacit and explicit knowledge that are necessary for sustainability, KO “facilitate the flow of tacit knowledge” (Leistner 2010; Geromin 2015; as cited in Wu et al., 2019). The significance of Knowledge organization in regard to tacit knowledge, is that it helps not only to restore the knowledge but to develop it in a way that makes it expand to create and know-how (Nonaka, 2007) to process it, “The characteristics of tacit knowledge enable it to be a source of sustainable advantage of organizations because of its immobility and inimitability (Ambrosini and Bowman 2001)” (Wu et al., 2019). However, since Tacit knowledge is internal in each one of us, and it is equal, but at the same time uniquely individual, therefore, it cannot be systematized nor duplicated (Wu et al., 2019). Instead, it can reach a certain level of external-transformation that enhances creativity and environment change. An example of that Einstein’s tacit knowledge and his perception of the world could never be restored (cause of death and loss), but his mathematical discoveries and relativity theories are indeed restorable, this helps to the continuity and the transfer of knowledge, that is the base in many fields and sciences in today’s world. This explicit knowledge, which is the outcome of interactions between tacit knowledge and the exterior world, is the foundation and root for structural knowledge which is not limited and can be very complex.

Organizing explicit knowledge, which “is managed using the codification strategy which makes knowledge easy to store, transfer, and communicate with others through electronic means” (Wu et al., 2019), helps in the development and expansion of tacit knowledge, which also goes into explicit knowledge as it generates and creates new information and knowledge for future generation, which is the main principle of sustainable development, circulation of information/knowledge in none-ending circuit.

Nonaka and Konno (1998) have developed a ‘SECI’ model that shows how the combination and interaction between tacit and explicit knowledge generates a new knowledge (see figure 14), which is conceptualized into four patterns (Nonaka, Konno, 1998) that are Socialization, Externalization, Combination, and Internalization.

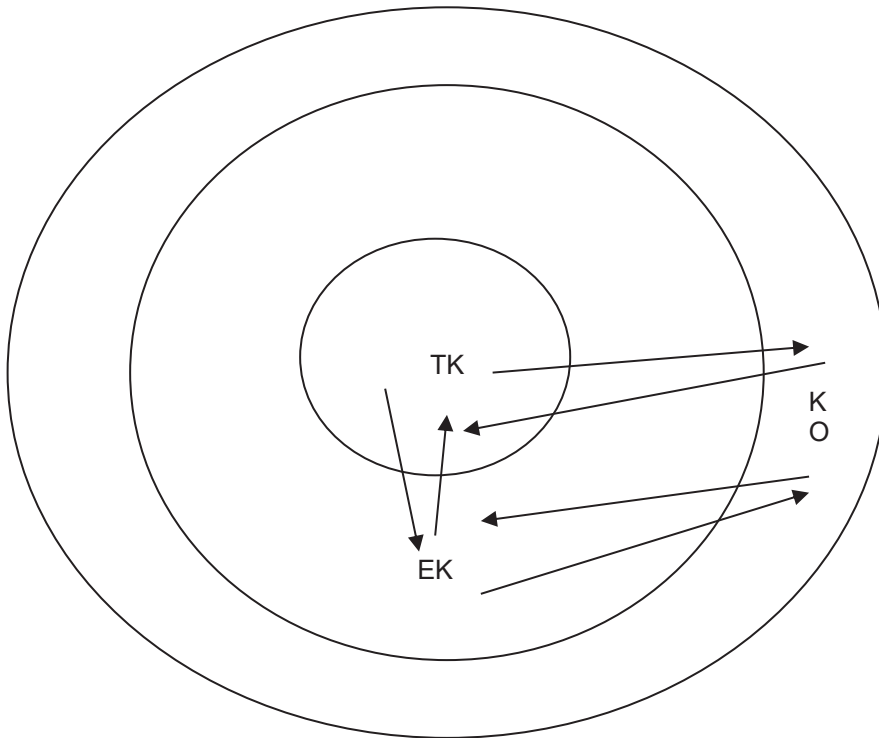


i: individual; g: group; o: organization

Figure 14: Spiral Evolution of Knowledge Conversion and Self-transcending Process by Nonaka and Konno

(1998).

It is important to comprehend the flow of knowledge and how it is created, in order to put in place the necessary tools and systems to enhance its development. The interaction between tacit and explicit knowledge as shown in the SECI model (Konaka and Konno, 1998) is infinite and unlimited in time and space. Thus, KO manages the explicit knowledge with the aim to prepare it for tacit knowledge, a simplified model based on the interconnection relating tacit, explicit knowledge and KO is drawn in figure 15. This circuit is very crucial for the sustainability of knowledge.



TK= Tacit Knowledge
 EK= Explicit Knowledge
 KO= Knowledge Organization

Figure 15: Simplified model of the interrelations between TK, EK and KO.
 By the author (22/04/21)

Therefore, sustainability is directly dependent on how knowledge is created, stored, transferred and managed (Nonaka; 1994). In the economical field, many studies have been concerned by the Knowledge Management (KM) and its impact on companies' businesses. Nonaka (1989, 1991, 1994) stated that "core competence" "(Prahalad and Hamel 1999) of any organization is a result of capability of its management in creating relevant organizational knowledge, which induces sustainable competitive advantage. To achieve sustainability, organizations need to "accumulate extensive trial and error learning and to build up an organizational knowledge base for problem-solving (Metcalf and Benn 2013)" (as cited in Wu, et al., 2019).

In social systems, KO provides a trusted social system for information and knowledge based on scientific discoveries from the real world. The main purpose of this is to create stability for individuals inside the society by ensuring the satisfaction of one of the human fundamental needs that is the right to access information and

knowledge based on trusted structure, and an equal social system. KO insures equal access to information and knowledge with no discrimination, which is one of the main goals of SD. Not only that, but also, it helps in the development of structural, human and social capital which are the basis of social, cultural, economical and political development.

Colnar et al. (2019) argued that there is an unequalized attention about sustainability that primarily focused on economical and environmental aspects of SD while omitting the social facet of it. Related to the social work field, Colnar et al. add that the mutual development of Knowledge organizations could potentially strengthen the overall performance levels of social work organizations, if well implemented, and which may have positive influences on the entire social systems (2019). Fruitful long-term functions of any organization are critically dependent on how knowledge and intellectual assets are organized and managed (Colnar et al., 2019).

However, there is a lack of studies that are concerned with public sectors compared to private economic-based fields that is due to the lack of financial support (Colnar et al., 2019). Colnar et al. pointed out that “dominant discussions related to knowledge management in general are not adapted to fit the context of the public sector and social work (Leung,2007)” (Colnar et al., 2019), as public organizations function uniquely and differently in comparison to private entities. Hence, more studies and research need to focus on the importance of KOs in regard to public sectors, in order to adapt adequate ways for its implementation. In addition to this, Colnar et al. posit that the relationship between knowledge organization and social work should be examined as it is very important (2019). It has been discussed by Kianto et al. (2016) that knowledge-creation, which is defined as “the organisation’s ability to develop new and useful ideas and solutions regarding various aspects of organisational activities”, is “a key factor in enabling sustained performance in turbulent environments (Teece et al., 1997; Eisenhardt and Martin, 2000)” (as cited in Kianto et al., 2016), and that depends on the job satisfaction among individuals (Kianto et al., 2016). In their study, Kianto et al. (2016) demonstrated the impact of knowledge organization within companies and economic-sectors on the job satisfaction of employees, which is crucial to sustain the human resources and retain them within organizations.

In the context of higher education, knowledge organization’s help is translated by the enhancement of tacit knowledge of individuals (Wu et al., 2019), which results in the creation of new ideas that improve the social environment by encouraging regular change and adaptation to new challenges. It is also translated with the facilitation of the circulation of the flow of information and knowledge between students and academic members.

In the past decade, there has been an abandoned focus on the social features of sustainability and sustainable development (Colnar et al., 2019). This might be due to the lack of directed resources and knowledge, but also to the misunderstanding of the complexity of the interconnection of sustainability with all other aspects of life. As discussed earlier in this paper, the attention about sustainability has been only concerned with economical and ecological attitudes. Hence, the necessity to organize knowledge in regard to sustainable development is crucial.

Poverty exists worldwide, and not only refers to the economic poverty here, but instead, to the intellectual poverty. Bourdieu (1986) argued that there are Three kinds of capitals, the economical capital, which is a financial one related to the social status and class levels, and cultural capital, which is an intellectual possession of knowledge that individuals acquire in time and space, and which needs sacrifices and dedication of time to achieve (Bourdieu, 1986), Bourdieu states that “the work of acquisition is work on oneself (self-improvement) an effort that presupposes a personal cost” (1986). Cultural capital becomes an embodied capital that is part of the individual “habitus’ (Bourdieu, 1986), and it cannot be instantly transmitted but needs time, unlike the economical capital. The third capital is the social capital; Bourdieu defined it as “the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition—or in other words, to membership in a group—which provides each of its members with the backing of the collectively owned capital, a “credential” which entitles them to credit, in the various senses of the word” (1986). It is thus the relationships that individuals have with their environment that allow them to acquire and transfer knowledge which is infinitely reproduced and exchanged. The social capital, as described by Bourdieu, is a network of connections which is not naturally nor socially given, but rather, “constituted once and for all by an initial act of institution” (1986).

For that being said, sustainability can be attained only when we consider it in every act and in all aspects of life, and “the reproduction of social capital”, which “presupposes an unceasing effort of sociability” (Bourdieu, 1986). Hence, it is not the economical capital that we should seek to improve at first levels, but rather, the social and cultural capital of individuals as it is the base for social actions. For that reason, a well structured, and organized knowledge is momentous.

Conclusion

Knowledge organization consists not only in a scientific field that is concerned about how knowledge is organized and how it should be represented for an efficient retrieval (Anderson, 2015). It is instead a way of perceiving the world and it should be recognized “as a social and cultural practice shaping us in our daily encounters with the worlds of social and cultural communication” (Anderson, 2015). With the immense growth of technological tools and assets, we witness and observe as well the expansion of a digital culture which is a reflection of our current world. Therefore, knowledge organization is a consequence of a diverse social human experience (Anderson, 2015) as it is based on human activity.

It is affirmed that “classifying is human” (Bowker and Star, 1999, as cited in Anderson, 2015). Our social interaction is based on classification and hierarchical inputs, our use of languages and their structure is based on classified information and knowledge, whether it is a tacit or explicit knowledge we tend to structure and find logical patterns to be able to simplify our communications and understanding depending on our environment. However, we need to consider knowledge organization as a social action (Anderson, 2015) that has sustainable outcomes

within our societies. If knowledge organization is used in the favor of individuals in order to enhance and encourage their creativity and the growth of their tacit knowledge, a social change can be generated out of this as there will be a maintain of the circulation of knowledge and the creation of new one, which is the main purpose of sustainable development, as it provides the fundamentals for changes that ensures the production of new livable environment that is sustained for the next generation. To conclude, knowledge organization should not be seen only as a theoretical and practical way to systemize knowledge, but rather, as part of human behaviours and actions in daily life. There is a need to change the way of perceiving and applying knowledge organization, knowledge is created, restored, transformed and transmitted depending on the culture of members of societies. Thorough studies on the connection and impact of knowledge organization, and the way it should be implemented depending on cultural settings, are needed in order to contribute to sustainable development.

References

1. Adams, G. L., & Lamont, B. T. (2003). Knowledge management systems and developing sustainable competitive advantage. *Journal of Knowledge Management*, 7(2), 142-154. <https://doi.org/10.1108/13673270310477342>
2. Aldabaldetrek, Rita; Lautiainen, Juuso; Minkova, Alina (2016). The role of knowledge management in strategic sustainable development- Comparing Theory and practice companies applying the FSSD; School of Engineering Blekinge Institute of Technology Karlskrona, Sweden.
3. Andersen, J. (2015). Re-Describing Knowledge Organization—A Genre and Activity-Based View. In *Genre Theory in Information Studies* (Vol. 11, pp. 13-42). Emerald Group Publishing Limited. <https://doi.org/10.1108/S2055-537720140000011003>
4. Bendix, R.F., Eggert, A., Peselmann., A. (ed.) (2012) *Heritage Regimes and the State* (Göttingen: Universitätsverlag Göttingen) 11-20.
5. Biemann, C., Crane, G. R., Fellbaum, C. D., & Mehler, A. (2014). *Computational Humanities—Bridging the gap between Computer Science and Digital Humanities (Dagstuhl Seminar 14301)* [Application/pdf]. 32 pages. <https://doi.org/10.4230/DAGREP.4.7.80>
6. Bourdieu, P. (1986). The forms of capital. In: Richardson, J., *Handbook of Theory and Research for the Sociology of Education*. Westport, CT: Greenwood: 241–58. Available: <https://www.socialcapitalgateway.org/content/paper/bourdieu-p-1986-forms-capital-richardson-j-handbook-theory-and-research-sociology-educ>
7. Brubaker, R. (2002). Ethnicity without groups. *European Journal of Sociology / Archives Européennes De Sociologie / Europäisches Archiv Für Soziologie*, 43(2), 163-189.
8. Chomsky, N. (1997). Language and Problems of Knowledge. *Revista Internacional de Filosofía*, Vol. 16, No. 2, pp.5-33. Retrieved on (06/04/2021), available at: <https://www.jstor.org/stable/43046198>
9. Chomsky, N. (2006). *Language and Mind*, III ed., Cambridge, C.U.P.
10. Colnar, S., Dimovski, V., & Bogataj, D. (2019). Knowledge Management and the Sustainable Development of Social Work. *Sustainability*, 11, 6374. <https://doi.org/10.3390/su11226374>

11. David A. Walden (2011) *From Policy to Praxis: Cultural Diversity, Human Rights and the Challenges of International Standard-Setting* (York: York Centre for Public Policy and Law) 3-12.
12. Denning, P. (2010). *Computing Science: The Great Principles of Computing*. *American Scientist*, 98(5), 369-372. Retrieved April 25, 2021, from <http://www.jstor.org/stable/27859559>
13. Djeflat, A. (2010). Sustainable knowledge for sustainable development: Challenges and opportunities for African development. *World Journal of Science, Technology and Sustainable Development*, 7(2), 131-149. <https://doi.org/10.1108/20425945201000009>
14. Duncan Light (2015) *Heritage and Tourism* in E. Waterton, S. Watson (eds.), *The Palgrave Handbook of Contemporary Heritage Research* (New York: Palgrave Macmillan) 144-158.
15. Dundes, A. (1997). Binary Opposition in Myth: The Propp/Lévi-Strauss Debate in Retrospect. *Western Folklore*, 56(1), 39-50. doi:10.2307/1500385
16. Gonzalez, R., & Martins, M. (2017). Knowledge Management Process: a theoretical-conceptual research. *Gest. Prod.*, São Carlos, v. 24, n. 2, p. 248-265. <http://dx.doi.org/10.1590/0104-530X0893-15>
17. Hajdu Barat, A. (2005). *The relationship between human perception and knowledge organization*.
18. Hallowell, A. (1947). Myth, Culture and Personality. *American Anthropologist*, 49(4), new series, 544-556. Retrieved April 26, 2021, from <http://www.jstor.org/stable/662894>
19. Hjørland, B. (2015). Classical databases and knowledge organization: A case for boolean retrieval and human decision-making during searches. *Journal of the Association for Information Science and Technology*, 66(8), 1559–1575. <https://doi.org/10.1002/asi.23250>
20. Hjørland, B. Information Retrieval and Knowledge Organization: A Perspective from the Philosophy of Science. *Information* 2021, 12, 135. <https://doi.org/10.3390/info12030135>
21. Holzem, M. (2006). L'organisation des connaissances à l'ère du village planétaire. *La contemporaine*, (N° 82), pp. 82 à 86. <https://www.cairn.info/revue-materiaux-pour-l-histoire-de-notretemps-2006-2-page-82.htm>
22. John Schofield (2015) 'Thinkers and Feelers' A Psychological Perspective on Heritage and Society in E. Waterton, S. Watson (eds.), *The Palgrave Handbook of Contemporary Heritage Research* (New York: Palgrave Macmillan) 417-425.
23. Jones, Karen Sparck; Willett, Peter (1997). "Overall Introduction" in *Readings in Information Retrieval*, ed. Karen Sparck Jones and Peter Willett. San Francisco: Morgan Kaufmann, pp. 1-7.
24. Kianto, A., Vanhala, M., & Heilmann, P. (2016). The impact of knowledge management on job satisfaction. *Journal of Knowledge Management*, 20(4), 621-636. <https://doi.org/10.1108/JKM-10-2015-0398>
25. Lévi-Strauss, C. (1981). Structuralism and Myth. *The Kenyon Review*, 3(2), 64-88. Retrieved April 26, 2021, from <http://www.jstor.org/stable/4335186>
26. *Managing Organisational Knowledge for Sustainability and Competitive Advantage: View as single page*. (n.d.). Retrieved 20 April 2021, from <https://www.open.edu/openlearncreate/mod/oucontent/view.php?id=5768&printable=1>
27. Marie-Theres Albert (2012) *Perspectives of World Heritage: towards future-oriented strategies with the five 'Cs'* in Albert, M.-T., Richon, M., Viñals, M. J., Witcomb, A. (eds) *Community Development through World Heritage*. (Paris: UNESCO) 32-38.
28. Moerman, Michael. (1988). *Talking culture: ethnography and conversation analysis*. Philadelphia: University of Pennsylvania Press

29. Mohamed, M., Stankosky, M., & Mohamed, M. (2009). An empirical assessment of knowledge management criticality for sustainable development. *Journal of Knowledge Management*, 13(5), 271-286. <https://doi.org/10.1108/13673270910988105>
30. Michel Batisse (2005) Nature and Culture. Recollections of a (conventional) marriage in Batisse, M., Bolla, G., *The Invention of "World Heritage"* (Paris: AFUS) 13-43.
31. Naomi Deegan (2012) The local-global nexus in the politics of World Heritage: space for community development? in Albert, M.-T., Richon, M., Viñals, M. J., Witcomb, A. (eds.) *Community Development through World Heritage*. (Paris: UNESCO) 77-83.
32. Noel B. Salazar (2010) The glocalisation of heritage through tourism, in Labadi, s., Long., C., *Heritage and Globalization* (London: Routledge) 145-161.
33. Nonaka, I. (1994a). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science*, 5(1), 14-37.
34. Nonaka, I. (1994b). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science*, 5(1), 14-37.
35. Nonaka, I. (2007, July 1). The Knowledge-Creating Company. *Harvard Business Review*. <https://hbr.org/2007/07/the-knowledge-creating-company>
36. Nonaka, I., & Konno, N. (1998). The Concept of "Ba": Building a Foundation for Knowledge Creation. *California Management Review*, 40(3), 40-54. <https://doi.org/10.2307/41165942>
37. Nonaka-1994-A Dynamic Theory of Organizational Knowledge Creat.pdf. (n.d.). Retrieved 24 April 2021, from https://www.svilendobrev.com/1/Nonaka_1994-Dynamic_theory_of_organiz_knowledge_creation.pdf
38. Paul Betts, Corey Ross (2015) Modern Historical Preservation – Towards a Global Perspective in *Past and Present*, Supplement 10, 7-26.
39. Polanyi, M. (1967). The tacit dimension. London: Routledge & Kegan Paul Ltd.
40. Ravazi, Franciscnaira Cristina; Moreira, Walter (2017). Documentary languages and audiovisual resources treatment: theoretical and methodological interfaces. In *Knowledge Organization and Cultural Diversity*, Sociedade Brasileira de Organização do Conhecimento (ISKO-Brasil). Marília, São Paulo, Brasil
41. RBMS Thesaurus. Available online: <https://rbms.info/vocabularies/> (accessed on the 25th of April, 2021).
42. Renear, A. (2004). *Text Encoding, A Companion to Digital Humanities*. Ed. Susan Schreibman, Ray Siemens, John Unsworth. Oxford: Blackwell. Available at: <http://www.digitalhumanities.org/companion>
43. Robért, Karl Henrik; Broman, Göran I.; Basile, George (2013). Analyzing the Concept of Planetary Boundaries from a Strategic Sustainability Perspective: How Does Humanity Avoid Tipping the Planet? *Ecology and Society* 18 (2).
44. Robyn Buskell (2015) Heritage and Sustainable Development in E. Waterton, S. Watson (eds.), *The Palgrave Handbook of Contemporary Heritage Research* (New York: Palgrave Macmillan) 492-507.
45. Sacks, Harvey (1995). *Lectures on conversation*. Edited by Gail Jefferson; with introduction by Emanuel A. Schegloff. The Estate of Harvey Sacks, Blackwell Publishing: Oxford, UK.
46. Schegloff, E. A. (2001). Accounts of Conduct in Interaction: Interruption, Overlap, and Turn-Taking. *Handbook of Sociological Theory*, 287–321. doi:10.1007/0-387-36274-6_15
47. Smiraglia, Richard P. (2018). Introduction: Theory, Knowledge Organization, Epistemology, Culture

48. Smith, B., Ceusters, W., Klagges, B., Köhler, J., Kumar, A., Lomax, J., Mungall, C., Neuhaus, F., Rector, A. L., & Rosse, C. (2005). Relations in biomedical ontologies. *Genome Biology*, 6(5), R46. <https://doi.org/10.1186/gb-2005-6-5-r46>
49. Sophia Labadi (2013) UNESCO, Cultural Heritage and Outstanding Universal Value, (Plymouth: Altamira Press) 11-58.
50. Spender, J. (1996). Organizational knowledge, learning and memory: three concepts in search of a theory. *Journal of Organizational Change Management*, 9(1), 63-78. <http://dx.doi.org/10.1108/09534819610156813>.
51. Sustainable Development Solutions Network. (2020). (Rep.). Sustainable Development Solutions Network. Retrieved April 25, 2021, from <http://www.jstor.org/stable/resrep25832>
52. Tim Winter (2010), Heritage tourism. The dawn of a new era? In Labadi, s., Long., C., *Heritage and Globalization* (London: Routledge) 131-144.
53. UNESCO Thesaurus. Available online: <http://vocabularies.unesco.org/browser/thesaurus/en/> (accessed on the 25th of April, 2021).
54. Wilson, T.D. (2000), "Human information behaviour", *Informing Science*, Vol. 3 No. 2, pp. 49-56
55. Wu, J., Lo, M. F., & Ng, A. W. (2018). Knowledge Management and Sustainable Development. In W. Leal Filho (Ed.), *Encyclopedia of Sustainability in Higher Education* (pp. 1-9). Springer International Publishing. https://doi.org/10.1007/978-3-319-63951-2_175-1
56. XU, S. (2015). A STUDY ON KNOWLEDGE MANAGEMENT CAPABILITIES TOWARDS NEW PRODUCT INNOVATION TYPE AND DEVELOPMENT PERFORMANCE OF CHINESE BUSINESSES. *Acta Oeconomica*, 65, 145-157.

(Página deixada propositadamente em branco)

APLICAÇÃO DE METADADOS NA PADRONIZAÇÃO DE REGISTROS DE OCORRÊNCIA DE ESPÉCIES NO CONTEXTO DA CIÊNCIA CIDADÃ PARA A BIODIVERSIDADE: UM ESTUDO DE CASO

Filipi Miranda Soares¹, Raíssa Yuri Hamanaka²

¹Universidade de São Paulo, Center for Artificial Intelligence (C4AI),
filipisoares@usp.br, ORCID 0000-0002-0674-7960

²Universidade Estadual de Londrina, raissa.hamanaka@uel.br,
ORCID 0000-0001-9516-5825

Resumo

A ciência cidadã pode ser definida como a condução da ciência por cientistas amadores ou não-profissionais. O aplicativo iNaturalist é um exemplo de projeto de ciência cidadã, no qual cidadãos podem compartilhar dados de observações de organismos da biodiversidade. O objetivo do estudo é investigar a dinâmica da representação do conhecimento, com foco na representação descritiva feita por meio de metadados, adotada pelo iNaturalist. Esta pesquisa possui objetivo exploratório e descritivo, abordagem qualitativa, natureza aplicada e utilizou a técnica de estudo de caso de um banco de dados. Foram analisados dois registros de ocorrência da espécie popularmente conhecida como monarca-do-sul e os recursos disponíveis no iNat para descrevê-los. A correta aplicação de padrões de metadados e vocabulários controlados permitiu que os dados inseridos no iNat fossem importados de maneira automática para outros repositórios de dados. Este estudo foi realizado com um conjunto de dados pequeno que não expressa todos os problemas que a interoperabilidade entre sistemas pode apresentar. Acredita-se que uma pesquisa mais abrangente, com diferentes bancos de dados, possa levantar questões-problema relacionadas ao tema que não foram aventadas por este trabalho.

Palavras-chave: padrão de metadados, biodiversidade, ciência aberta, iNaturalist, GBIF.

Introdução

A ciência aberta pode ser definida como o “[...] conhecimento transparente e acessível que é compartilhado e desenvolvido por meio de redes colaborativas” (Vicente-Saez & Martinez-Fuentes, 2018, p. 428). Também pode ser entendida como um termo guarda-chuva no qual estão presentes o acesso aberto, a educação aberta, dados científicos abertos, ferramentas e materiais científicos abertos, a ciência cidadã, avaliação científica aberta, políticas de ciência aberta e os cadernos de

pesquisa abertos (Albagli et al., 2014; Pontika & Knoth, 2015). Nesta direção, a presente pesquisa utiliza como objeto de estudo um aplicativo de ciência cidadã para a biodiversidade intitulado iNaturalist (iNat), desenvolvido por uma iniciativa conjunta da California Academy of Science e da National Geographic Society (<https://www.inaturalist.org/>).

A ciência cidadã objetiva aumentar a participação da sociedade na ciência e pode ser definida como a “[...] pesquisa científica conduzida, no todo ou em parte, por cientistas amadores ou não-profissionais” (Albagli et al., 2014, p. 443). Entretanto, este trabalho realizado por cientistas não-profissionais pode apresentar diversos problemas, dentre eles a qualidade dos dados produzidos e sua representação (Lukyanenko et al., 2016). É possível distinguir cinco tipos de ciência cidadã: a computação compartilhada, a inteligência distribuída, o sensoriamento voluntário, o diálogo com a sociedade e a pesquisa direta (Albagli et al., 2014). O iNat se enquadra no sensoriamento voluntário, no qual cidadãos podem observar organismos da biodiversidade e compartilhar os dados das observações por meio de fotos e ou áudios associados a metadados descritivos. Metadados, de maneira resumida, podem ser definidos como termos que descrevem dados de recursos por meio de pares propriedade-valor, como por exemplo <title>Mona Lisa</title>, que designa o título da pintura Mona Lisa, de Leonardo da Vinci (Soares, 2019).

Os registros de ocorrência de biodiversidade podem ser relevantes para estudos científicos com diferentes propósitos, como o monitoramento de espécies ameaçadas possibilitar o entendimento das dinâmicas espaço temporais da biodiversidade, para promover sua gestão sustentável, dentre diversas outras possibilidades de uso desses dados (Petersen et al., 2021). Os registros de dados do iNat podem ser importados por outros repositórios de dados científicos e acessados livremente pelos usuários interessados.

A essa comunicação de dados entre sistemas de maneira automática ou semiautomática, atribui-se o nome de interoperabilidade (Almeida, 2020; Moura, 2009). Para que haja interoperabilidade entre sistemas de informação, acredita-se que uma das ações fundamentais é o uso de vocabulários controlados para descrição e representação dos dados. Os vocabulários controlados assumem papel descritivo quando são utilizados como metadados, ou seja, assumem a função de descrever dados. Assumem o papel de objeto quando são utilizados como valores para campos de metadados (é comum que os padrões de metadados recomendem um conjunto de termos de vocabulário controlado como opções de preenchimento de determinados campos de metadados).

No contexto da ciência da biodiversidade, que tem dentre uma de suas atribuições o registro de ocorrências de espécies nos diferentes ecossistemas da Terra, comunidades científicas internacionais têm se organizado em prol de desenvolver ferramentas para a organização do conhecimento produzido sobre a biodiversidade. O grupo de pesquisa mais conhecido pela sigla TDWG (<https://www.tdwg.org>) – Biodiversity Information Standards, anteriormente denominado Taxonomic Databases Working – destaca-se por seu proeminente esforço na criação de padrões de metadados para a descrição de diversas dimensões dos dados produzidos acerca da biodiversidade. Um de seus padrões de metadados mais consolidados, o Darwin Core (Wieczorek et al., 2012), é utilizado por instituições no mundo todo para a

publicação de dados de biodiversidade. Somente no repositório da Global Biodiversity Information Facility (GBIF) há atualmente mais de 1,6 bilhões (<https://www.gbif.org/>) de registros de metadados indexados em formato Darwin Core. Parte dos registros indexados no GBIF são provenientes do iNat, que também utiliza o padrão de metadados Darwin Core na descrição de seus registros. A importação de dados é feita por meio de um API do GBIF que reconhece arquivos em formato DarwinCore Archive, que são arquivos zip compostos por arquivos CSV de registros de metadados Darwin Core (<https://www.inaturalist.org/pages/developers>). Além disso, cada registro de dados recebe um identificador persistente do tipo Uniform Resource Identifier (URI), que facilita o referenciamento dos recursos entre as bases de dados e garante sua disponibilidade a longo prazo.





A partir da contextualização exposta, o presente estudo pretende responder ao problema de pesquisa: como os metadados podem contribuir para a padronização de registros de ocorrência de espécies no contexto da ciência cidadã para a biodiversidade? Para elucidar o problema, o objetivo do estudo é descrever e analisar os metadados utilizados em registros de ocorrência de espécie do iNat.

Material e métodos

De acordo com Gil (1994), uma pesquisa pode ser caracterizada conforme seus objetivos, abordagem, natureza e técnicas. Este estudo possui objetivo exploratório e descritivo ao analisar e descrever metadados utilizados em registros de ocorrência de espécies do iNat. Foi utilizada abordagem qualitativa, realizando-se inferências a partir de dados particulares. A natureza da pesquisa é aplicada, ao contribuir para a padronização dos metadados utilizados por naturalistas ao compartilharem seus dados de observação de organismos da biodiversidade. E foi utilizada a técnica de estudo de caso na análise de um banco de dados com registros de ocorrência de espécies.

Para realizar as análises necessárias para se atingir o objetivo deste trabalho, foi criado um banco de dados com registros de ocorrência da espécie *Danaus erippus* Cramer, 1775 (conhecida popularmente como monarca-do-sul), realizados em campo por um dos proponentes deste estudo (Soares, 2021). Dois registros de ocorrência foram efetuados tendo fotografias como evidências, conforme o Quadro 1.

Quadro 1. Registros de ocorrências da espécie *Danaus erippus*

	<p>Ocorrência #01 Nome científico: <i>Danaus erippus</i> Cramer, 1775 Data da observação: 2020-09-05T15:54:00 Localidade: Moeda – MG, Brasil</p> <p>© Filipi Miranda Soares</p> 
	<p>Ocorrência #02 Nome científico: <i>Danaus erippus</i> Cramer, 1775 Data da observação: 2020-12-24T13:16:00 Localidade: Virginópolis – MG, Brasil</p> <p>© Filipi Miranda Soares</p> 

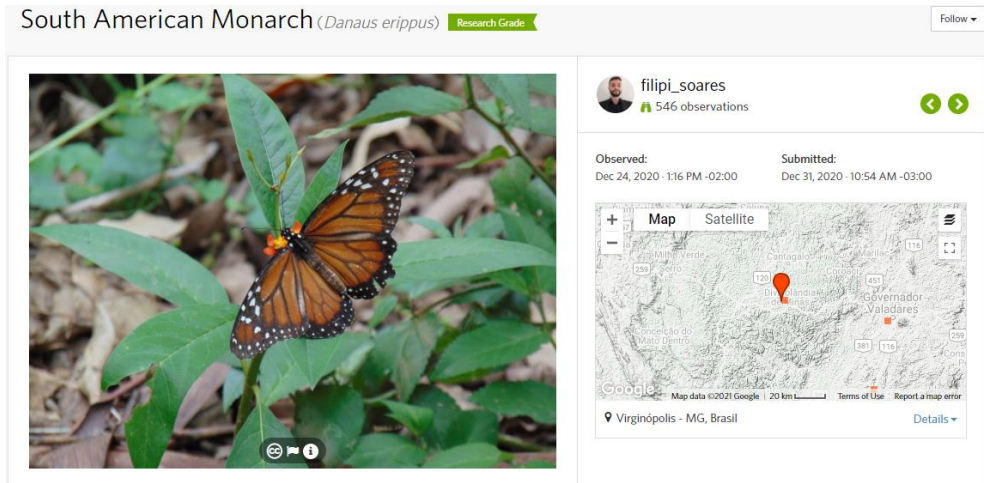
Fonte: Soares (2021)

As informações completas e detalhadas das duas ocorrências representadas no Quadro 1 podem ser recuperadas de Soares (2021). As fotografias em ambas as observações foram feitas com uma câmera digital Sony, modelo DSC-H300.

As fotografias dos organismos foram submetidas na base do iNat, juntamente com dados de localização e data de observação. Utilizou-se também a inteligência artificial presente no iNat para identificar a espécie observada. Além do conjunto *default* de metadados oferecido pelo iNat, foram adicionados outros campos ao registro por meio dos Campos de Observação, que são metadados adicionais que o iNat oferece para melhorar a descrição do registro. Por fim, foram utilizadas etiquetas em linguagem natural como uma das formas de indexação dos registros de dados.

Resultados

O registro da ocorrência #01 (vide Quadro 1) gerou a URI <https://www.inaturalist.org/observations/59048126> e o registro da ocorrência #02 gerou a URI <https://www.inaturalist.org/observations/67352810> no iNat. A Figura 1 apresenta capturas de tela de parte de ambos os registros de dados indexados no iNat.



Fonte: dados da pesquisa (2021)

Ambos os registros da Figura 1 receberam o selo de qualidade Grau de Pesquisa do iNat, o que significa que esses registros atenderam a alguns critérios de qualidade, a saber (iNat, 2021):

- a) data específica;
- b) localização específica;
- c) possuem fotografia;
- d) possuem uma identificação suportada por dois ou mais membros da comunidade iNat no nível de espécie (em ambos os casos, adotou-se a identificação taxonômica sugerida pela inteligência artificial, que foi confirmada por outros membros da comunidade: francierlem e henriqueandrades);
- e) data correta (a comunidade concorda que a data parece ser acurada);
- f) localização correta (a comunidade concorda que a localização parece ser acurada);
- g) o organismo observado é selvagem (não é um organismo cultivado ou em cativeiro);
- h) apresenta evidência de organismo (é possível observar um organismo nas fotografias, em vez de apenas paisagens, rios, rochas etc);
- i) a comunidade concorda que há evidências recentes do organismo observado (menos de 100 anos aproximadamente);

Os detalhes do significado de cada um dos critérios de qualidade podem ser conferidos no site do iNat (iNat, 2021). Em ambos os registros de dados, foram utilizados metadados das classes ocorrência, evento, identificação, taxonomia, localização provenientes do padrão Darwin Core e metadados do padrão Dublin Core para descrição dos arquivos das fotos. Na seção Anotações de ambos os registros, foram preenchidos os atributos: a) <estágio de vida>, que recebeu o valor “adulto”, indi-

cando que os organismos observados encontravam-se na fase adulta; b) <sexo>, que recebeu valor “indeterminado” por não ser possível indentificar o dimorfismo sexual das espécies apenas por foto; c) e indicou-se que o organismo estava vivo no momento da observação. Foram adicionadas as tags “borboleta” e “butterfly”, em linguagem natural, como ferramenta alternativa de indexação dos registros. Por fim, foi utilizado o termo “Feeding on” para registrar que os organismos observados em ambas as ocorrências estavam se alimentando em plantas da espécie *Asclepias curassavica*, popularmente conhecida como oficial-de-sala. Os Campos de Observação do iNat são campos adicionais que permitem registrar dados de interações entre o organismo observado com outras espécies e uma infinidade de outros dados. Entretanto, foi observado no repositório https://www.inaturalist.org/observation_fields que há alguns problemas com a padronização desses termos adicionais, que aparecem em múltiplos idiomas, em linguagem natural, com definições conceituais incompletas ou ausentes e em formas tipográficas variadas (palavras totalmente em caixa alta, outras em caixa baixa). Existe um esforço para padronizar os termos adicionais mais utilizados pela comunidade, cujas discussões acontecem no Wiki <https://forum.inaturalist.org/t/observation-field-standardization-wiki/380>. O termo <Feeding on> faz parte de um vocabulário controlado para representação de interações ecológicas, que são relações que acontecem na natureza entre dois ou mais organismos. Os registros de dados com o selo grau de pesquisa que aplicam metadados de interações são indexados automaticamente pelo repositório Global Biotic Interactions (Globi), que indexa dados de interações ecológicas do mundo todo. Desta forma, as ocorrências #01 e #02 foram automaticamente indexadas pelo Globi (Soares, 2020).

Após receberem o selo Grau de Pesquisa no iNat, os registros de ocorrências também foram importados pelo GBIF, que gerou as URIs <https://www.gbif.org/occurrence/2873905520> para a ocorrência #01 e <https://www.gbif.org/occurrence/3013943235> para a ocorrência #02. O GBIF importou todos os campos de metadados Darwin Core dos registros feitos no iNat, o que não inclui os campos adicionais de observação. O algoritmo do GBIF fez algumas inferências automáticas sobre os dados, por isso é possível notar quatro elementos em cada par propriedade-valor no registro, em vez dos elementos habituais: 1) *term*, coluna que designa o nome do termo (propriedade); 2) *interpreted*, que apresenta o dado importado e processado pelo algoritmo do GBIF (valor processado); 3) *original*, que apresenta o dado em sua forma original, ou seja, da maneira como está formatado no iNat (valor original); 4) *remarks*, que indica o tipo de alteração que o dado sofreu após o processamento. A Figura 2 mostra parte do registro de dados da URI <https://www.gbif.org/occurrence/2873905520>, em que é possível observar as diferenças entre os dados originais e os dados interpretados.

Figura 2. Valores originais e interpretados em um registro de dados no GBIF

Record			
Term	Interpreted	Original	Remarks
Dataset name	iNaturalist research-grade observations	iNaturalist research-grade observations	
Institution code	iNaturalist	iNaturalist	
Basis of record	Human observation	HumanObservation	
Collection code	Observations	Observations	
Occurrence			
Term	Interpreted	Original	Remarks
Establishment means		wild	Excluded
Life stage	Adult	adult	
Catalogue number	59048126	59048126	
Occurrence ID	https://www.inaturalist.org/observations/59048126	https://www.inaturalist.org/observations/59048126	
Occurrence status	PRESENT		Inferred
Recorded by	Filipi Miranda Soares	Filipi Miranda Soares	
Recorded by ID	 https://orcid.org/0000-0002-0674-7960	https://orcid.org/0000-0002-0674-7960	
Sex		undetermined	Excluded

Fonte: dados da pesquisa (2021)

Em termos de representação da informação, a funcionalidade apresentada pelo GBIF (vide Figura 2) é fundamental para garantir a integridade dos dados após a exportação. A troca de dados entre diferentes bases de dados pode levar à perda de semântica quando não é feita adequadamente, porém constatou-se que o rigor no tratamento dos dados, que inclui aplicação precisa de padrões de metadados, tanto no iNat quando no GBIF foi capaz de preservar a semântica dos dados originais e acrescentar significado de maneira automática: no caso do iNat, a inteligência artificial identificou a espécie taxonômica do organismo da foto automaticamente, o que agregou valor aos dados; e no GBIF, o algoritmo fez algumas inferências que resultaram em dados interpretados.

No Quadro 2 é possível observar algumas equivalências entre metadados e dados associados do iNat e do GBIF para o registro de dados da ocorrência #01 do Quadro 1. As equivalências foram indicadas a partir da observação e comparação dos registros de dados das duas bases.

Quadro 2. Correspondência entre alguns metadados utilizados no iNat e GBIF

iNat metadados	iNat valores	GBIF metadados	GBIF valores
id	59048126	catalogNumber	59048126
url	https://www.inaturalist.org/observations/59048126	occurrenceID	https://www.inaturalist.org/observations/59048126
observed_on_string	2020-09-05 3:54:00 PM GMT-03:00	verbatimEventDate	2020-09-05 3:54:00 PM GMT-03:00
time_observed_at	2020-09-05 3:54:00 UTC	eventDate	2020-09-05T15:54:00
observed_on	05/09/2020	day	5
observed_on	05/09/2020	month	9
observed_on	05/09/2020	year	2020
created_at	2020-09-09 15:37:22 UTC	dateIdentified	2020-09-09T15:37:22
latitude	-20.290608	decimalLatitude	-20.290608
longitude	-43.992762	decimalLongitude	-43.992762
place_state_name	Minas Gerais	stateProvince	Minas Gerais
scientific_name	Danaus erippus	scientificName	Danaus erippus Cramer, 1775
taxon_kingdom_name	Animalia	kingdom	Animalia
taxon_phylum_name	Arthropoda	phylum	Arthropoda
taxon_class_name	Insecta	class	Insecta
taxon_order_name	Lepidoptera	order	Lepidoptera
taxon_family_name	Nymphalidae	family	Nymphalidae
taxon_genus_name	Danaus	genus	Danaus
taxon_species_name	Danaus erippus	species	Danaus erippus

Fonte: dados da pesquisa (2021)

Conforme o Quadro 2 permite observar, os nomes atribuídos aos metadados e o formato dos valores (dados) nem sempre são iguais entre o iNat e o GBIF. Porém, o mapeamento dos metadados entre as duas bases é eficiente de tal forma que permite

a comunicação entre elas sem perda de semântica dos dados. Dessa forma, as transformações sofridas pelos metadados e pelos dados não afetam o significado deles. Por exemplo, o termo que expressa o nome científico da espécie no registro do iNat é <scientific_name> e contém o valor associado <Danaus erippus>; já o registro do GBIF apresenta o termo <scientificName> e o valor associado <Danaus erippus Cramer, 1775>. No caso do GBIF, o nome científico apresenta também o nome do autor e a data de identificação da espécie, conforme recomendações do International Code of Zoological Nomenclature (ICZN), que é um “[...] sistema de nomes científicos aplicados a unidades taxonômicas [...] de animais existentes e extintos” (Ride et al., 1999). Ambas as representações estão corretas e permitem a identificação inequívoca da espécie, apesar de a representação do GBIF ser aparentemente mais completa que a do iNat neste caso específico.

Em geral, considera-se que ambas as bases de dados apresentam recursos sólidos de representação da informação que permitem a interoperabilidade entre os dois sistemas sem prejuízo de significado dos dados e metadados.

Conclusões

O iNat demonstrou ser uma aplicação de ciência cidadã que apresenta ferramentas bem desenvolvidas para garantir a qualidade de dados, o que ajuda a garantir o uso desses dados pela comunidade científica. Esses recursos para qualidade de dados, que incluem a aplicação de padrões de metadados, ajudam a transpor o problema da falta de padronização dos dados coletados por cientistas cidadãos, que muitas vezes não possuem conhecimentos para o correto tratamento dos dados, conforme observado por Lukyanenko et al. (2016).

Este estudo verificou que o uso de padrões de metadados permite que diferentes ambientes informacionais compartilhem dados entre si por meio da interoperabilidade entre seus metadados (Castro & Santos, 2014). A análise realizada evidenciou que a interoperabilidade semântica entre o iNat-GBIF ocorre por meio do uso de padrões de metadados entre os dois ambientes informacionais, o que permitiu estabelecer uma correspondência de metadados, e o compartilhamento de significado dos dados entre esses ambientes. Corroborou-se o estudo de Alves e Souza (2007), que destacam o uso de metadados, normas e modelos informacionais na integração e na troca de dados entre sistemas, ao verificar-se que as duas plataformas estudadas utilizam padrões de metadados em comum. Além disso, utilizam APIs que reconhecem esses metadados e fazem a extração automática dos dados de uma plataforma para outra. Também foi possível evidenciar que o uso de URIs e de ontologias na explicitação formal de relacionamentos no registro de ocorrência de espécie do iNat permitiu que este fosse rastreado e indexado pelo GBIF, esse acesso a bases de dados de forma integrada e com esforço reduzido é o desafio que a interoperabilidade pretende cumprir, segundo Marcondes e Sayão (2001).

A correta aplicação de padrões de metadados e vocabulários controlados permitiu que os dados inseridos no iNat fossem importados de maneira automática para outros repositórios de dados. Soma-se a isso a geração de URIs para cada registro de observação feito no iNat, que facilita o referenciamento da fonte de dados por outros repositórios que importam os dados com Grau de Pesquisa, como o GBIF. Este

estudo foi realizado com um conjunto de dados pequeno, acredita-se que uma pesquisa mais abrangente, com diferentes bancos de dados, possa levantar questões-problema relacionadas ao tema que não foram aventadas por este trabalho.

Espera-se que este estudo sirva como um incentivo para o engajamento da comunidade científica da organização do conhecimento na Ibero-América na seara de trabalho da informação sobre biodiversidade, que apresenta-se como campo de trabalho promissor para a comunidade de Knowledge Organization (KO).

Referências

- Albagli, S., Clinio, A., & Raychtock, S. (2014). Ciência Aberta: correntes interpretativas e tipos de ação. *Liinc em Revista*, 10(2), 434-450.
- Almeida, M. B. (2020). *Ontologia em Ciência da Informação: teoria e método*. CRV. <https://doi.org/10.24824/978655578679.8>
- Alves, M. D. R., & Souza, M. I. F. (2007). Estudo da correspondência de elementos metadados: Dublin Core e MARC 21. *Revista Digital de Biblioteconomia e Ciência da Informação*, 4(2), 20-38.
- Castro, F. F., & Santos, P. L. V. A. C. (2014). Elementos de interoperabilidade na perspectiva da catalogação descritiva. *Informação & Sociedade: Estudos*, 24(3), 13-25.
- Gil, A. C. (1994). *Métodos e técnicas de pesquisa social* (4. ed.). Atlas.
- iNaturalist. (2021). What is the data quality assessment and how do observations qualify to become "Research Grade"? <https://www.inaturalist.org/pages/help#quality>
- Lukyanenko, R., Parsons, J., & Wiersma, Y. F. (2016). *Conservation Biology*, 30(3), 447-449. <https://doi.org/10.1111/cobi.12706>
- Marcondes, C. H., & Sayão, L. F. (2001). Integração e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: a proposta da Biblioteca Digital Brasileira. *Ciência da Informação*, 30(3), 24-33.
- Moura, M. A. (2009). Informação, ferramentas ontológicas e redes sociais ad hoc: a interoperabilidade na construção de tesouros e ontologias. *Informação & Sociedade: Estudos*, 19(1), 59-73.
- Petersen, T. K., Speed, J. D. M., Grøtan, V., & Austrheim, G. (2021). Species data for understanding biodiversity dynamics: The what, where and when of species occurrence data collection. *Ecol Solut Evidence*, 2. <https://doi.org/10.1002/2688-8319.12048>
- Pontika, N., & Knoth, P. (2015). *Open science taxonomy*. <http://oro.open.ac.uk/47806/>
- Ride, W. D. L., Cogger, H. G., Dupuis, C., Kraus, O., Minelli, A., Thompson, F. C., Tubbs, P. K. (1999). *International Code of Zoological Nomenclature*. International Trust for Zoological Nomenclature, London. <https://www.iczn.org/the-code/the-code-online/>
- Soares, F. M. (2021, 4 Abril). *GBIF Occurrence Download*. <https://doi.org/10.15468/dl.5jye6p>
- Soares, F. M. (2019). *Princípios para a criação de uma extensão de metadados sobre interações ecológicas na agrobiodiversidade para o padrão Darwin Core* [Dissertação de Mestrado, Universidade Federal de Minas Gerais]. Biblioteca Digital de Teses e Dissertações da UFMG. <https://repositorio.ufmg.br/handle/1843/33387>
- Soares, F. M. (2020). *Danaus erippus feeding on Asclepias curassavica*. <https://github.com/globalbioticinteractions/inaturalist/archive/1706a6dc719a6033d6cea1bebf438fee9871fd43.zip>.
- Vicente-Saez, R., & Martinez-Fuentes, C. (2018). Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research*, 88, 428-436.
- Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., & Vieglais, D. (2012). Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLoS ONE*, 7(1). <https://doi.org/10.1371/journal.pone.0029715>

(Página deixada propositadamente em branco)

UN MAPEO TERMINOLÓGICO DEL DOMINIO COVID-19 CON BASE EN BIBLIOMETRÍA Y GARANTÍA ACADÉMICA

Mario Barité¹, Exequiel Fontans²

¹Universidad de la República, mario.barite@fic.edu.uy, ORCID iD 0000-0002-2992-6582

²Universidad de la República, exequiel.fontans@fic.edu.uy, ORCID iD 0000-0002-4627-4678

Resumo

El término ‘dominio’ ha adquirido creciente importancia en la terminología de la organización del conocimiento (OC), con el sentido de ‘field of work’, cualquier campo de trabajo que involucre a actores diversos de una comunidad de discurso (organizaciones profesionales, idóneos, investigadores, responsables de bibliotecas y diseñadores de sistemas de organización del conocimiento). Como tal constituye una expresión genérica comprensiva de todas las disciplinas, macrodisciplinas, subdisciplinas, interdisciplinas, microdisciplinas y áreas de estudio. Los análisis de dominios involucran modalidades de mapeo de dominios para diversos propósitos, entre ellos la creación y revisión de sistemas de organización del conocimiento (tesauros, listas, ontologías, taxonomías). A la clasificación dicotómica tradicional entre dominios consolidados y dominios emergentes, en este trabajo se agrega una categoría adicional, la de los dominios de emergencia súbita (DES), para nominar distintas áreas del saber que en los últimos cuarenta años han irrumpido de situaciones de la realidad no anticipadas por estudios proyectivos (VIH/sida, ébola, derribo de las torres gemelas, y actualmente la dolencia COVID-19 ocasionada por el virus SARS-CoV-2). Este trabajo se propone como objetivo general, fundamentar la idea de que el espacio de conocimiento creado en torno al SARS-CoV-2 y la COVID-19 es un dominio desde la perspectiva de la OC, particularmente un dominio de emergencia súbita, que ha ido conformando, por lo menos durante la pandemia, su propia comunidad de discurso y práctica. Sobre esa premisa se aspira a determinar el origen disciplinario de la investigación que ha generado, identificando los veinte dominios preexistentes cuyos autores más han aportado a la producción científica para conceptualizar, categorizar y tratar a la enfermedad. Para ese fin se utilizan técnicas bibliométricas de análisis temático por granularidad gruesa (WC, WOS Category), y se hace una primera aproximación utilizando una granularidad fina (KWp, Key Word Plus) tomando como referencia el universo documental alcanzado por la Web of Science (WOS) en el período 2019-2020. Se identificaron 42.175 artículos en el dominio analizado, el 51% de los cuales se concentran en 20 WC. Los resultados primarios apuntan a una concentración de la investigación en áreas directamente emparentadas o subordinadas a la medicina: 13 disciplinas en 20. En cuanto a WC vinculados a ciencias sociales, no aparecen en el top 20; recién en el lugar 22 comparece Economics. En cuanto a las humanidades, la primera WC que comparece, Religión, lo hace recién en el lugar 61. La terminología

WOS, tanto en granularidad gruesa como fina, por su formalidad y afinidad con el lenguaje científico tradicional, parece constituir una forma de validación que se corresponde con la garantía académica. De entre las conclusiones se destaca la necesidad de ajustar los métodos de diseño de los sistemas de organización del conocimiento, para elaborar en corto plazo herramientas de urgencia de bajo costo que faciliten la indización y la organización temática del caudal documental que genera en poco tiempo un DES. Asimismo, se consideran adecuados los análisis bibliométricos de granularidad gruesa para ofrecer una cartografía inicial y provisional de un dominio de emergencia súbita.

Palabras clave: COVID-19, Análisis de dominio, Garantía académica, Bibliometría, Web of Science.

Introducción

El término ‘dominio’ ha adquirido creciente importancia en la terminología de la organización del conocimiento (OC). Coadyuvan a ello la necesidad de mayor precisión conceptual para conjugar en una, distintas expresiones utilizadas a veces en forma indistinta, sin ser equivalentes; y la sistematización de métodos y técnicas que exijan, para su mejor aplicación, consensos sobre el alcance del término.

Respecto al primer punto, se acusaba la falta de una expresión genérica comprensiva de todas las disciplinas, macrodisciplinas, microdisciplinas, subdisciplinas, interdisciplinas, áreas y campos de estudio que constituyen un núcleo particular y dinámico de conocimientos. La expresión ‘dominio’, por otra parte, tiene una larga existencia asociada a la historia de la propiedad privada de la tierra, y refiere también a la posibilidad de ejercer el gobierno sobre un territorio. Si bien Corominas no le asigna una entrada independiente en su diccionario crítico etimológico (1991), ya en el primer diccionario publicado por la Real Academia se registra con esta definición: “El mando, imperio y señorío que tiene uno sobre alguna cosa, Lugar, o Provincia, del qual puede usar libremente. Latín. *Dominium*” (Real Academia Española, 1732). Por extensión de sentido, la noción de dominio comprende también al territorio intelectual de una especialidad, en la que conviven especialistas, idóneos, practicantes y aprendices, que mantienen y desarrollan las ideas, los métodos y los productos que son de su incumbencia.

En lo que hace a la OC, la sistematización de métodos y técnicas relacionadas con el mapeo de dominios para fines de representación temática y recuperación temática de información, y para la creación de sistemas de organización del conocimiento (SOC), tuvo su inflexión en 1993 cuando se extrapola la noción de análisis de dominio (‘domain analysis’) desde la informática a la OC (Albrechtsen, 1993; Hjørland & Albrechtsen, 1995). Ello implicó también la mutación de su contenido semántico, pues se entendió que “the most fruitful horizon for Information Science is to study knowledge domains as thought or discourse communities” (Hjørland & Albrechtsen, 1995, p. 400).

Así, el término ‘dominio’ comenzó a utilizarse regularmente con el sentido que le asignan Schmidt and Wagner (2004): a ‘field of work’, un campo de trabajo que, según Albrechtsen (2015) involucra a actores diversos de una comunidad de discurso (organizaciones profesionales, investigadores, responsables de bibliotecas y

diseñadores de SOC), pues todos ellos están involucrados en “the development of specific knowledge fields or tools” (p. 559).

Hjørland confirmó la naturaleza metodológica del análisis de dominio al proponer once modalidades de mapeo: producción de guías de literatura; construcción de clasificaciones especializadas y tesauros; indización y recuperación de información en áreas especializadas; estudios de usuarios; estudios bibliométricos; estudios históricos; estudios documentales y de tipos documentales; estudios epistemológicos y críticos; estudios terminológicos; estructuras e instituciones en la comunicación científica; y, cognición científica, conocimiento experto e inteligencia artificial (Hjørland, 2002). Si bien esta lista ha sido contestada en parte (Smiraglia, 2015) y ampliada (Guimarães y Tognoli 2015; Smiraglia, 2015), se sigue aceptando pacíficamente la inclusión de los estudios terminológicos y los estudios bibliométricos, que están en el centro metodológico del estudio que se desarrollará, como formas de análisis de dominio.

En la literatura se ha prestado una atención lateral a la clasificación de los dominios, y no se ha percibido con claridad que esa categorización podría ser necesaria, por ejemplo, para seleccionar las metodologías de análisis y mapeo más adecuadas. Se ha distinguido entre dominios consolidados y menos consolidados en función de las características de la investigación en cada disciplina (Mote, 1962). También hay referencias esporádicas a las características de los dominios emergentes (Al-Thubaity and Ahmad, 2002; Lu, 2019), incluso respecto a la OC (Smiraglia, 2012).

La pandemia por coronavirus que viene afectando las condiciones sociales, culturales, económicas, políticas y educativas de la humanidad desde fines de diciembre de 2019, volvió a poner sobre la mesa la cuestión de los dominios emergentes, pues en torno al virus y la enfermedad surgió un cúmulo de información y documentación desmesurado en su cantidad, y sospechado en su fiabilidad. Todas las disciplinas de la salud, pero también muchas de las disciplinas biológicas, sociales y humanas fueron convocadas a identificar los nuevos problemas creados por la pandemia, y buscar soluciones de urgencia. La fiabilidad de la información que se fue acumulando estuvo y está bajo la lupa tanto por la continua aparición de *fake news* (especialmente en las redes sociales pero también en los medios tradicionales de comunicación), como por las certezas rápidamente sustituidas unas por otras, en función de los paulatinos avances de la investigación y la experimentación.

Puede afirmarse que la documentación y la investigación alrededor del virus SARS-CoV-2- y de la enfermedad COVID-19, cuenta con tantos vectores como disciplinas son afectadas, y fueron conformando un dominio emergente que por sus características y dinámicas, exige respuestas originales por parte de la OC. Sin embargo, lo sucedido con la pandemia no es enteramente nuevo. Ya desde los años 80 se han venido delineando -a partir de hechos de la realidad como el surgimiento del VIH y del ébola, el derribo de las torres gemelas- dominios emergentes atípicos, que recientemente se ha propuesto denominar ‘dominios de emergencia súbita’ o DES (Barité, 2020). En particular, los DES que se forman en torno a un virus y una enfermedad cuentan, en su estructura nocional con dos términos tope (justamente los nombres del virus y la enfermedad) y no con uno, lo que conlleva la toma de decisiones metodológicas no previstas si, por ejemplo, se quisiera construir un tesoro sobre estos tópicos. Por ello estos dominios podrían denominarse con justicia, binarios.

Varios aspectos distinguen a los DES de otros dominios emergentes:

- su irrupción intempestiva, no anticipada por estudios prospectivos;
- la magnitud de su impacto en términos económicos, políticos y sociales;
- la producción de un gran volumen de información y documentación en poco tiempo, con fiabilidad diversa;
- una terminología inicial inestable y ambigua, que complica la tarea de profesionales mediadores (periodistas, divulgadores, diseñadores de campañas, gestores de contenidos, bibliotecólogos, traductores);
- por la naturaleza de los problemas que generan, se verifica la contribución multi- e interdisciplinaria de numerosas especialidades;
- en consecuencia, es posible que los mapeos de estos dominios requieran varios niveles de aproximación con base en la procedencia de su terminología y la categorización temática de la investigación que generan.

En este trabajo se propone como objetivo general, realizar un análisis del dominio binario SARS-CoV-2 / COVID-19 (en adelante COVID-19), considerando su naturaleza de dominio de emergencia súbita, a fin de determinar en primera instancia, aquellos dominios preexistentes cuyos autores más han aportado desde la investigación a la producción científica, y por tanto a la creación de su perfil como comunidad de discurso.

Para cumplir ese objetivo se utilizan técnicas bibliométricas de análisis temático, tomando como referencia el universo documental alcanzado por la Web of Science (WOS).

Metodología

Las preguntas que disparan este estudio son: ¿en qué dominios del saber se concentró la investigación científica en torno al DES COVID-19 en el primer año de la pandemia? ¿Cómo establecer una primera cartografía del nuevo DES, a efectos de favorecer su inserción en los mapas del conocimiento preexistentes, y la generación de SOC de urgencia, que permitan clasificar e indizar la documentación correspondiente?

Dieron sustento metodológico para la investigación, dos de las modalidades de análisis de dominio propuestas por Hjørland (2002): los estudios terminológicos y los bibliométricos. Se tomó como justificación de la terminología utilizada la garantía académica, entendiendo por tal la justificación de la inclusión/exclusión de términos de indización en base a los acuerdos alcanzados por los científicos o especialistas en cada área del saber. Esos acuerdos se respaldan en los consensos, los disensos y la opinión experta de los especialistas de un dominio (Barité & Rauch, 2019).

En este trabajo la garantía académica procede de las categorías temáticas (WC) propuestas por Web of Science (WOS) para agrupar las publicaciones y las Key Word Plus (KWp) para indizar y agrupar temáticamente los documentos.

Los estudios terminológicos pueden asumir dos tipos de abordajes:

i) abordaje microestructural: estudio término a término, que indaga sobre la etimología, acuñación, conceptualización, definición e identificación de sinónimos de cada término, así como las cuestiones gramaticales, sintácticas y ortográficas que correspondan (lo que es conocido en terminología como investigación puntual).

ii) abordaje macroestructural: busca como resultado el mapeo de un dominio, tan genérico o tan específico, tan primario o exhaustivo como sea necesario a los objetivos de la investigación. Este abordaje -llamado en terminología investigación sistemática- es el que mejor se ajusta a este trabajo, y permite visualizar además los vínculos con otros dominios, facilitando su inserción en los mapas preexistentes del conocimiento.

Para el abordaje bibliométrico se utilizó el análisis temático a través de la WOS Category y de las KWp. Las WC ofrecen una tematización de ‘granularidad gruesa’, pues mediante un método heurístico se asignan a los títulos de revistas (y por tanto a los artículos publicados) algunas de las 251 categorías que refieren a dominios preestablecidos (Pudovkin & Garfield, 2002). Si bien esta tematización es perfectible, como señala incluso uno de sus autores (Garfield, 2006), es muy utilizada en estudios bibliométricos y se ha demostrado más robusta que la tematización de Scopus, por ejemplo (Wang & Waltman, 2016).

Para el estudio de las temáticas desde una granularidad más fina se optó por utilizar el campo ID, que contiene los KWp. Estos descriptores son asignados por la Web of Science mediante un algoritmo que analiza los títulos de las referencias contenidas en los documentos (Garfield, 1990; Garfield y Sher, 1993). Cantos-Mateos y colaboradores han señalado que la ventaja sobre otros métodos como el de palabras clave asignadas por el autor es que son más específicos y aluden directamente al contenido conceptual de los documentos (Cantos-Mateos et al., 2014). De hecho, un estudio comparado entre vocabulario libre (palabras clave asignadas por los autores), vocabulario controlado (descriptores del Medical Subject Headings (MeSH) asignados manualmente), y los KWp ha permitido determinar que estos constituyen la unidad de análisis más adecuada a la hora de realizar un análisis temático sobre un dominio científico (Cantos-Mateos et al., 2013).

Para agrupar las WC en nivel de granularidad más gruesa se utilizó la metodología propuesta por Milojevic (2019), que consiste en la agrupación de las 251 WC en 14 grupos mediante un algoritmo que toma en cuenta la cocitación de toda la base WOS.

En suma, como técnica de exploración se decidió utilizar un abordaje terminológico de corte macroestructural, combinado con un análisis temático, a efectos de identificar –a partir del universo documental relacionado- los veinte dominios en los que se concentró la investigación sobre COVID-19 en el período 2019-2020. Se es consciente de que lo correspondiente al año 2019 se concentra sobre fin de ese año, ya que el primer caso de COVID-19 se originó en diciembre. Estos dominios constituyen la garantía académica del nuevo DES.

La fuente es WOS, una plataforma en línea de información científica de acceso por suscripción considerada, junto a Scopus, una de las principales bases bibliográfi-

cas multidisciplinares. Se utilizó su Colección Principal (CP), que indiza unos 12.700 títulos de revistas y más de 160.000 actas de congreso. Su corpus supera los 62 millones de registros y el billón de referencias citadas.

Se cumplieron tres fases metodológicas:

1. Búsqueda en CP de las expresiones SARS-coV-2 OR COVID-19. Se recuperaron todos los registros de la base que contuvieran “SARS-coV-2” y/o “COVID-19” en el título, las palabras clave o el resumen de cada documento. El tipo de documento considerado fue “Article”.

2. Organización de los datos extraídos de la CP dentro de las 251 categorías (dominios) WOS, en orden cuantitativo decreciente, hasta un total de veinte.

3. Ubicación de los veinte dominios obtenidos, en alguna de las categorías establecidas en la jerarquía “macrodisciplinas (MD), disciplinas (D), subdisciplinas (SD), y tópicos (T)”. Para la ubicación de los términos en las categorías, en caso de duda se tomó como referencia la Clasificación Decimal Universal (AENOR, 2004).

4. Agrupamiento de las WC en las 14 categorías más generales propuestas por Milojevic (2019).

5. Análisis de las 20 KWp más frecuentes por categoría.

Se reconocen antecedentes de investigaciones similares aunque desde otras perspectivas o por períodos de varios meses de 2020. Por ejemplo, Farook y colaboradores en una búsqueda realizada en la WOS el 1ro. de agosto de 2020 identifican 6694 registros, y se centran en el análisis de países e instituciones más productivos (Farook et al., 2021). Wang y Tian, (2021) trabajan con WOS y varias bases de preprints (bioRxiv, medRxiv, Preprints y SSRN), pero tematizan con categorías de la OMS.

Se identifican algunos trabajos que analizan la producción en el área social. En junio de 2020 Samir y colaboradores (2020) identifican 395 artículos (en WOS, Scopus y Google Scholar) en donde, en lo que respecta a la tematización, describen KWp y palabras clave de autores, títulos y resúmenes.

Resultados

La búsqueda en WOS arrojó un total de 42.175 registros (o sea, artículos) distribuidos en 6.235 títulos de revistas. Los datos extraídos se distribuyeron en 7 columnas:

1. Lugar en el ranking de ocurrencias.
2. Atribución asignada en la jerarquía “macrodisciplinas (MD), disciplinas (D), subdisciplinas (SD), y tópicos (T)”.
3. Categoría asignada por WOS.
4. Cantidad de registros.
5. Dato porcentual sobre el total de registros.
6. Dato acumulado de registros.
7. Dato porcentual sobre los registros acumulados.

La siguiente tabla exhibe el primer análisis de los resultados. Cabe consignar que los títulos de revistas pueden estar categorizados en más de una WOS Category

Tabla 1. Distribución de la producción sobre SARS-coV-2 y COVID-19 en las WC de WOS 2019-2020.

Rank	Clas	WOS Categories	# Art. Con ponderador	% / 42175	Acumulado	% Acumulado
1	D	Medicine, General & Internal	3.642,25	8,64%	3.642,25	8,64%
2	SD	Public, Environmental & Occupational Health	2.915,15	6,91%	6.557,40	15,55%
3	MD	Multidisciplinary Sciences	1.794,25	4,25%	8.351,65	19,80%
4	SD	Infectious Diseases	1.461,83	3,47%	9.813,48	23,27%
5	SD	Psychiatry	924,78	2,19%	10.738,27	25,46%
6	MD	Environmental Sciences	912,72	2,16%	11.650,98	27,63%
7	SD	Pharmacology & Pharmacy	903,42	2,14%	12.554,40	29,77%
8	SD	Surgery	863,92	2,05%	13.418,32	31,82%
9	SD	Medicine, Research & Experimental	838,85	1,99%	14.257,17	33,80%
10	SD	Cardiac & Cardiovascular Systems	794,75	1,88%	15.051,92	35,69%
11	SD	Health Care Sciences & Services	760,40	1,80%	15.812,32	37,49%
12	D	Education & Educational Research	758,75	1,80%	16.571,07	39,29%
13	SD	Virology	752,75	1,78%	17.323,82	41,08%
14	SD	Immunology	746,98	1,77%	18.070,80	42,85%
15	SD	Radiology, Nuclear Medicine & Medical Imaging	666,65	1,58%	18.737,45	44,43%
16	SD	Pediatrics	603,33	1,43%	19.340,78	45,86%
17	SD	Microbiology	602,25	1,43%	19.943,03	47,29%
18	SD	Oncology	585,40	1,39%	20.528,43	48,67%
19	SD	Clinical Neurology	583,02	1,38%	21.111,45	50,06%
20	D	Law	548,67	1,30%	21.660,12	51,36%

Referencia: Los autores.

Los resultados indican que el 51% de los artículos científicos sobre COVID-19 se concentra en las 20 categorías temáticas seleccionadas. Si se considera la clasificación por dominios, se concluye que la producción científica se ha concentrado en dos macrodisciplinas, tres disciplinas y quince subdisciplinas. Las quince subdisciplinas están subordinadas o vinculadas a la medicina y a la biología. Solo dos de los veinte dominios (educación y derecho) guardan relación con las ciencias sociales y humanas, pese al alto impacto que la pandemia ha tenido sobre la economía, el trabajo y la sociedad. Así, puede determinarse que durante el primer año de pandemia la investigación sobre COVID-19 se ha concentrado claramente en el área de medicina y/o en alguna de sus subdisciplinas, o en dominios fuertemente relacionados con ella.

En la tabla 2 se agrupan las WC según la propuesta de Milojevic (2019) en 14 categorías más generales. Luego nos centramos en las 20 WC más importantes, que aportan el 50% de los artículos al dominio estudiado. De estas 20 WC analizamos las 20 KW Plus de mayor frecuencia en cada WC. Diseñamos una matriz de WC (20) por KW Plus (123) que nos permitió identificar las temáticas que se repetían en los diferentes dominios (WC). Finalmente asignamos la WC a una de las categorías cuando el 50% o más de los KWp cumplían con el criterio de clasificación.

Es importante señalar que no todos los registros tienen KWp asignados, lo que constituye un elemento a tener en cuenta a la hora de analizar los resultados. Para tener una aproximación de la incidencia que podría tener esta situación en las WC seleccionadas en el top 20, realizamos un análisis de cuartiles observando el porcentaje de registros sin KWp. Obtuvimos los siguientes valores: Q1= 27,5%; Q2= 44,0% y Q3= 54,7%. De los 20 dominios seleccionados resultaron subrepresentados tres: *Surgery* (61%); *Education & Educational Research* (59%); y *Law* (79%). Por su parte, resultó sobrerrepresentado un solo dominio: *Environmental Sciences* (24%). Por lo que, teniendo en cuenta que los 20 dominios seleccionados representan el 50% del corpus documental, se considera una buena aproximación exploratoria.

Tabla 2. Producción sobre SARS-coV-2 y COVID-19 en WOS 2019-2020

Rank	WOS Categories	Milojevic	KWp específicas		KWp compartidas		KWp otras		Total
			#	%	#	%	#	%	KWp
1	Medicine, General & Internal	Medical sciences	–	0%	13	65%	7	35%	20
2	Public, Environmental & Occupational Health	Medical sciences	–	0%	11	55%	9	45%	20
3	Multidisciplinary Sciences	Engineering	–	0%	12	60%	8	40%	20
4	Infectious Diseases	Medical sciences	1	5%	12	57%	8	38%	21

Rank	WOS Categories	Milojevic	KWp específicas		KWp compartidas		KWp otras		Total
			#	%	#	%	#	%	KWp
5	Psychiatry	Medical sciences	3	15%	7	35%	10	50%	20
6	Environmental Sciences	Geosciences	3	15%	9	45%	8	40%	20
7	Pharmacology & Pharmacy	Medical sciences	1	5%	9	43%	11	52%	21
8	Surgery	Medical sciences	4	16%	11	44%	10	40%	25
9	Medicine, Research & Experimental	Medical sciences	–	0%	11	55%	9	45%	20
10	Cardiac & Cardiovascular Systems	Medical sciences	6	27%	7	32%	9	41%	22
11	Health Care Sciences & Services	Professional fields	4	17%	11	46%	9	38%	24
12	Education & Educational Research	Professional fields	16	73%	1	5%	5	23%	22
13	Virology	Biological sciences	1	5%	10	48%	10	48%	21
14	Immunology	Medical sciences	2	10%	10	48%	9	43%	21
15	Radiology, Nuclear Medicine & Medical Imaging	Medical sciences	5	25%	11	55%	4	20%	20
16	Pediatrics	Medical sciences	4	18%	10	46%	8	36%	22
17	Microbiology	Biological sciences	1	5%	9	45%	10	50%	20
18	Oncology	Medical sciences	2	9%	11	50%	9	41%	22
19	Clinical Neurology	Medical sciences	1	5%	9	43%	11	52%	21
20	Law	Professional fields	12	57%	4	19%	5	24%	21

Rank	WOS Categories	Milojevic	KWp específicas		KWp compartidas		KWp otras		Total
			#	%	#	%	#	%	KWp

Notas:

KWp específicas: sólo aparecen en esta WC en el top de las 20 más frecuentes

KWp compartida: aparece en 10 o más de las WC seleccionadas

KWp otras, reúne a las KWp que no fueron clasificadas en ninguna de las dos categorías anteriores.

En color se identifica el porcentaje de la columna que es 50% o mayor.

Se identifica una cuarta categoría conformada por aquellas KWp que están distribuidas entre las 3 categorías de manera que ninguna alcanza al 50% o más (las que tienen toda la fila sin color).

Referencia: Los autores

Lo que sigue es una breve descripción sobre la manera en que se comportaron las KWp conforme a los datos presentados en la tabla 2.

KW Plus específicas: *Education & Educational Research* y *Law* son dos WC que pertenecen a dominios alejados de la medicina y parecen centrarse en temáticas propias de sus campos con algunos aspectos referentes al dominio general (COVID-19). Se destaca *Education & Educational Research*, que presenta el 73% de las KWp específicas de este dominio. También es importante recordar que estos dos dominios están subrepresentados, tienen 59% y 79% de registros sin KWp asignadas.

KW Plus compartidas: 7 WC. Con excepción de *Multidisciplinary Sciences*, todas las WC de este grupo pertenecen a la medicina. En este agrupamiento destacan *Medicine, General & Internal, Medicine, Research & Experimental, Multidisciplinary Sciences, Public, Environmental & Occupational Health*, que no tienen ninguna KWp específica. Las KWp compartidas son 13, y en su gran mayoría refieren a asuntos generales que es de esperar que aparezcan en el corpus establecido para este trabajo; por ejemplo, *ACUTE RESPIRATORY SYNDROME; CORONAVIRUS; COVID-19; DISEASE; HEALTH; IMPACT; INFECTION; OUTBREAK; PNEUMONIA; RISK; SARS*. Luego hay dos KWp que se explican por el contexto de aparición de la pandemia: *WUHAN; CHINA*.

KW Plus otras: 4 WC. En esta categoría se agrupan las WC que no cumplen con los dos criterios anteriores, tres vinculadas a *Medical Sciences: Psychiatry; Pharmacology & Pharmacy* y *Clinical Neurology*. Y una vinculada a *Biological Sciences: Microbiology*.

KW Plus sin categoría: 7 WC. Finalmente en esta categoría se agrupan las WC que no cumplen con los criterios establecidos para ser asignadas a las tres categorías anteriores. Se encuentra la única WC clasificada en *Geosciences* por las categorías de Milojevic: *Environmental Sciences*. También se destaca que para casi todas las WC las KWp que los integran tienen un peso importante en temáticas compartidas, mayor a 44%, excepto para *Cardiac & Cardiovascular Systems*. Dicho de otra manera, las temáticas más frecuentes en estas WC no refieren a tópicos de especialidad del dominio.

Ciencias Sociales y Humanidades: Estas dos grandes categorías de Milojevic no aparecen en el top 20 de dominios analizados para este trabajo (Tabla 3). Las cuatro WC que aparecen más cercanas al top 20 pertenecen a Ciencias Sociales: (en el lugar 22) *Economics*; (33) *Social Sciences, Interdisciplinary*; (39) *Political Science*; y (43) *Public Administration*. Para el caso de Humanidades, aparecen entre los

lugares 60 y 100: (61) *Religion*; (85) *Humanities, Multidisciplinary*; (88) *Ethics*; (94) *Philosophy*; y (96) *Medical Ethics*.

Estos datos son muy preliminares y constituyen una aproximación, no hay que perder de vista que se produce mucho menos en *Ethics*, por ejemplo, que en *Medical Sciences*.

Tabla 3. Producción sobre SARS-coV-2 y COVID-19 en Social Sciences y Humanities en WOS 2019-2020

Rank	WOS Categories	Milojevic
22	Economics	Social sciences
33	Social Sciences, Interdisciplinary	Social sciences
39	Political Science	Social sciences
43	Public Administration	Social sciences
61	Religion	Humanities
85	Humanities, Multidisciplinary	Humanities
88	Ethics	Humanities
94	Philosophy	Humanities

Referencia: Los autores

Conclusiones

La pandemia en curso ha contribuido a identificar un patrón común en el surgimiento de dominios en las últimas décadas, en torno a la irrupción imprevista de enfermedades o de hechos de la realidad (sida, ébola, torres gemelas). Por sus particularidades, estos campos temáticos podrían ameritar una categoría específica, aquí denominada ‘dominios de emergencia súbita’ (DES), que puede sumarse a las dos categorías clásicas correspondientes a los dominios consolidados y los emergentes.

La aceptación del concepto DES tiene consecuencias conceptuales y metodológicas en la OC. Las conceptuales refieren a la necesidad de revisar la noción de dominio, a la que se le ha prestado especial atención en la literatura de los últimos veinticinco años. Las metodológicas se orientan a las peculiaridades que los análisis de dominio deben atender, cuando se analizan dominios con elementos constituyentes de muchas disciplinas de base.

Los inevitables componentes multi- e interdisciplinarios que constituyen la red interna de conceptos de los DES forman también parte de la revisión, si se considera que aunque se ha avanzado mucho en la determinación de las particularidades de los dominios interdisciplinarios y sus consecuencias en la OC, no parece disponerse todavía de un cuerpo de trabajos canónicos en la materia.

En particular el mapeo realizado del DES COVID-19 sugiere que la investigación realizada en el primer año de la pandemia se ha enfocado claramente en distintas áreas, disciplinas y subdisciplinas de la salud, o en otras relacionadas con ella, en detrimento de la investigación originada en las ciencias sociales y humanas.

Los resultados pueden constituir una orientación para el diseño de sistemas de organización del conocimiento sobre el dominio COVID-19, así como para la preparación de glosarios y diccionarios de divulgación, en cuyos equipos responsables pueden insertarse profesionales de la información. Esto parece ir en concordancia y confirmar el pensamiento de Hjørland, cuando señala que “*in order to identify scientific concepts and terms and their relation to other terms, in the end one needs to inspect the primary literature and bibliometrics is an important tool for such an inspection*” (Hjørland, 2013, p. 1323).

Vista la recurrente aparición de nuevos dominios de emergencia súbita, quizás deba considerarse la necesidad de ajustar los métodos de diseño de los SOC, para elaborar en corto plazo herramientas de urgencia de bajo costo, que auxilien en su representación temática y la consiguiente recuperación de información. Los métodos basados en análisis bibliométricos de granularidad gruesa pueden ofrecer la primera aproximación para la cartografía inicial y provisional de un dominio de emergencia súbita por al menos dos razones: i) se apoyan en la investigación efectivamente realizada y documentada lo que asegura esquemas y clasificaciones ajustados al estado del arte del DES; y, ii) reavivan la idoneidad de la garantía literaria (criterio de selección de terminología de representación basado en cuerpos de documentación preexistente), para contribuir a soluciones, también de urgencia, en el diseño de sistemas de organización del conocimiento.

Referencias

- Asociación Española de Normalización y Certificación (AENOR) (2004). *CDU: Clasificación Decimal Universal* (2a edición abreviada). Madrid: AENOR.
- Al-Thubaity, A. M., & Ahmad, K. (2002). Tracking knowledge of emergent domains. In *Proceedings Sixth International Conference on Information Visualisation*. <https://doi.org/10.1109/iv.2002.1028849>
- Albrechtsen, H. (2015). This is Not Domain Analysis. *Knowledge Organization*, 42(8), 557-561.
- Albrechtsen, H. (1993). Subject analysis and indexing: from automated indexing to domain analysis. *The Indexer*, 18(4): 221-224, Oct. 1993.
- Barité, M. (2020). Terminología de urgencia y garantías para la representación temática: Elementos para el análisis de dominios de emergencia súbita (DES). *Informação & Informação*, 25(3), 239. <https://doi.org/10.5433/1981-8920.2020v25n3p239>
- Barité, M. & Rauch, M. (2019). *Apuntes para el estudio de la garantía académica: primera aproximación*. Ponencia presentada ante las III Jornadas de Investigación de la Facultad de Información y Comunicación (FIC). Montevideo: FIC.
- Cantos-Mateos, G., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., & Zulueta, M.-Á. (2013). Estudio comparativo sobre la visualización de redes de cowords a través de los descriptores del Science citation index y Medline. In F. Ribeiro & M. E. Cerveira (Eds.), *I Congresso ISKO Espanha e Portugal / XI Congresso ISKO España: Informação e/ou*

- Conhecimento: as duas faces de Jano* (pp. 173-189). Faculdade de Letras da Universidade do Porto – CETAC.MEDIA. <https://doi.org/10.1080/00387019808003304>
- Cantos-Mateos, G., Zulueta, M.-Á., Vargas-Quesada, B., & Chinchilla-Rodríguez, Z. (2014). Estudio evolutivo de la investigación española con células madre. Visualización e identificación de las principales líneas de investigación. *El Profesional de la Información*, 23(3), 259-271. <https://doi.org/10.3145/epi.2014.may.06>
- Corominas, Joan (1991). *Diccionario crítico etimológico castellano e hispánico*. Madrid: Gredós. 6 v.
- Farooq, R. K., Rehman, S. U., Ashiq, M., Siddique, N., & Ahmad, S. (2021). Bibliometric analysis of coronavirus disease (COVID-19) literature published in Web of Science 2019-2020. *Journal of Family & Community Medicine*, 28(1), 1-7. https://doi.org/10.4103/jfcm.JFCM_332_20
- Garfield, E. (1990). KeyWords Plus – ISI’s Breakthrough Retrieval Method. 1. Expanding Your Searching Power on Current-Contents on Diskette. *Current Contents*, 1(32), 5-9.
- Garfield, E., & Sher, I. H. (1993). KeyWords Plus Algorithmic Derivative Indexing. *Journal of the American Society for Information Science*, 44(5), 298-299.
- Garfield, E. (2006). The history and meaning of the journal impact factor. *JAMA*, 295(1), 90–93. <https://doi.org/10.1001/jama.295.1.90>
- Guimarães, J. A. Ch., and Tognoli, N.B. (2015). Provenance as a domain analysis approach in archival knowledge organization. *Knowledge Organization*, 42(8): 562-569.
- Hjørland, B. (2013). Citation analysis: A social and dynamic approach to knowledge organization. *Information Processing & Management*, 49(6), 1313-1325. <https://doi.org/10.1016/j.ipm.2013.07.001>
- Hjørland, B. (2017). Domain Analysis. *Knowledge Organization*, 44(6): 436-464.
- Hjørland, B. (2002). Domain analysis in information science: eleven approaches – traditional as well as innovative. *Journal of documentation*, 58: p. 422-462.
- Hjørland, B. & Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*, 46(6): 400-425.
- Lu, K. (2019). What Kind of Research Topics Emerged in the Biomedical Domain?: A Perspective from Newly Added Subject Terms in a Thesaurus. In *iConference 2020 Proceedings*. ISchools.
- Mote, L.J. (1962). Reasons for the Variations in the Information Needs of Scientists. *Journal of Documentation*, 18(4): 169-175.
- Pudovkin, A. I., & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technology*, 53(13), 1113-1119. <https://doi.org/10.1002/asi.10153>
- Real Academia Española (1726-1739). *Diccionario de autoridades*. Madrid: RAE.
- Schmidt, K. & Wagner, I. (2004). Ordering Systems: Coordinative Practices and Artifacts in Architectural Design and Planning. *Computer Supported Cooperative Work (CSCW)*, 13: 349-408.
- Smiraglia, R. P. (2015). *Domain Analysis for Knowledge Organization: Tools for Ontology Extraction*. Chandos information professional series. Oxford: Chandos Publishing
- Smiraglia, R. P. (2012). Knowledge organization: some trends in an emergent domain. *El profesional de la información*, 21(3): 225-227. <http://dx.doi.org/10.3145/epi.2012.may.01>
- Wang, Q., & Waltman, L. (2016). Large-scale analysis of the accuracy of the journal classification systems of Web of Science and Scopus. *Journal of Informetrics*, 10(2), 347-364. <https://doi.org/10.1016/j.joi.2016.02.003>
- Wang, P., & Tian, D. (2021). Bibliometric analysis of global scientific research on COVID-19. *Journal of Biosafety and Biosecurity*, 3(1), 4-9. <https://doi.org/10.1016/j.jobb.2020.12.002>

(Página deixada propositadamente em branco)

A NOVA PLATAFORMA PUBMED: ANÁLISE DOS PERITOS

Maria Luz Antunes^{1,2*}, Carlos Lopes^{2,3}, Licínio Roque⁴

¹Instituto Politécnico de Lisboa (ESTeSL), mluz.antunes@estesl.ipl.pt, ORCID iD 0000-0003-0942-7601

²APPsyCI-Applied Psychology Research Center Capabilities and Inclusion

³ISPA – Instituto Universitário, clopes@ispa.pt, ORCID iD 0000-0002-6440-4739

⁴Universidade de Coimbra, lir@dei.uc.pt, ORCID iD 0000-0002-1911-2788

Resumo

Introdução: A PubMed é a plataforma científica da área biomédica mais utilizada pelos profissionais de saúde em todo o mundo para realizar pesquisas de informação, validação de diagnósticos, procura da evidência científica e para efeitos académicos. **Objetivo:** Analisar a nova plataforma PubMed a partir do contributo de cinco peritos, contrastando experiências e averiguando como a visualizam face à sua prática profissional. **Métodos:** Metodologia qualitativa, com uma entrevista semiestruturada sobre as perceções de cinco peritos, representativos de um conjunto de utilizadores (bibliotecário, investigador, professor do ensino superior, médico e gestor editorial). **Resultados:** Constata-se que se perdeu a oportunidade de renovar de forma efetiva e válida a PubMed. Os perfis do investigador, do médico e do gestor editorial consideraram que a nova plataforma só é nova no visual, nada trazendo de novo para as suas práticas profissionais. Professor e bibliotecário, recorrendo à sua experiência pedagógica, consideraram-na um sinónimo de mudança de paradigma, assente na agilidade e na interoperabilidade de processos, mais amigável para os utilizadores, possibilitando a proximidade desta ferramenta ao cidadão comum. **Conclusão:** Diferentes utilizadores, com diferentes experiências, têm uma visão própria da PubMed. Os peritos esperavam também mais da nova plataforma, mais qualidade e mais maturidade tecnológica.

Palavras-chave: PubMed, Navegabilidade, Literacia digital, Contraste de experiências, Visualização.

Introdução

O desenvolvimento e a implementação de métodos de *design* centrados no Homem para criar infraestruturas de tecnologias da informação são vantajosos para os vários campos do saber. Considerando a área da saúde, o *design* centrado na relação homem-artefacto ajuda a criar valor para doentes e profissionais de

* Programa de Doutoramento em Ciência da Informação (Universidade de Coimbra, Portugal).

saúde, por exemplo, no apoio à tomada de decisão clínica e à coordenação de cuidados (Carayon, & Hoonakker, 2019). Estando o Homem ativamente envolvido no *design*, na conceção e na implementação das tecnologias da informação em saúde, será possível avaliar na prática o impacto do fator humano e dos seus princípios.

Qualquer que seja o seu contexto, o uso das tecnologias é influenciado em duas fases: no *design* (e.g., fornecedores de tecnologias da informação) e na implementação (i.e., configuração de *software* pelas organizações de saúde). De acordo com os seus princípios e com fatores de ordem variada, as organizações de saúde podem decidir sobre a personalização, a configuração e outros processos de implementação (e.g., formação, *inputs* e *outputs*) que podem influenciar significativamente a utilização final dos sistemas a disponibilizar aos profissionais de saúde. É de evitar que estes tenham a sensação de que dedicam muito tempo e esforço a apoiar a tecnologia, conquanto deva ser a tecnologia a apoiá-los nas suas atividades clínicas (Gawande, 2018).

O uso das tecnologias da informação pelos profissionais de saúde tem sido objeto de estudos diversos. Em 2016, Shanafelt *et al.* destacaram que o uso regular das tecnologias em ambiente clínico foi associado à baixa satisfação profissional do médico que, além da sua atividade clínica, despendia o seu tempo em pesquisas de informação (para atualização de conhecimentos e para resolução de dúvidas), mas também em tarefas administrativas (e.g., atualização de dados de registos clínicos), podendo ocasionar a exaustão profissional. Também o estudo de Arndt *et al.* (2017) mostrou que os médicos de medicina geral e familiar passam quase seis horas por dia interagindo com as tecnologias da informação após o horário das consultas, indicando uma possível sobrecarga de trabalho e a interferência entre a esfera profissional e a privada.

Poder-se-á esperar por resultados positivos resultantes da implementação de tecnologias da informação em saúde. Mas, entretanto, podem ocorrer episódios negativos, colocando em risco os doentes (e.g., riscos e erros médicos), frustrando profissionais de saúde e atrapalhando o seu desempenho – como o *burnout*. É necessária, por isso, a proatividade nas tecnologias da informação em saúde: fatores humanos e um *design* centrado no utilizador, garantindo que estas são bem implementadas e se encaixam com a área clínica, o fluxo de trabalho com os doentes e a sua permanente atualização de conhecimentos, em regime de autoformação (Carayon, & Hoonakker, 2019).

A PubMed é a plataforma científica da área biomédica mais utilizada pelos profissionais de saúde em todo o mundo – aproximadamente 2,5 milhões de utilizadores de todo o mundo acedem diariamente à PubMed, realizando cerca de três milhões de pesquisas e visualizando cerca de nove milhões de páginas (Fiorini, Lipman, & Lu, 2017). A PubMed é usada para realizar pesquisas de informação para a validação de diagnósticos e terapêuticas, para a procura da evidência científica, bem como para efeitos académicos (Hoogland, 2019). Foi criada e continua a ser coordenada pela National Library of Medicine, inserida nos National Institutes of Health (EUA). É de acesso livre, compreende cerca de 33 milhões de registos da base de dados MEDLINE, revistas científicas de ciências da vida, da saúde e de disciplinas conexas, como as ciências comportamentais, as ciências químicas e a bioengenharia. Não

inclui artigos em texto integral, mas o *link* está frequentemente presente quando disponível a partir de outras fontes, como o *site* da editora ou a PubMed Central (PubMed, 2020).

Entre 1971 e 1997, o acesso à base de dados MEDLINE era feito principalmente através de bibliotecas universitárias, que a subscreviam. A plataforma PubMed, lançada em janeiro de 1996, deu início à era da pesquisa privada e gratuita, no domicílio, no hospital, no consultório, no laboratório, na sala de aulas, da MEDLINE. Ao longo dos anos foram introduzidas funcionalidades específicas que contribuíram para melhorar o diálogo com os seus utilizadores que, na Primavera de 2019, foram alertados para uma versão beta em teste e para a implementação de uma nova plataforma (Fiorini et al., 2018), de acordo com padrões de inovação e de modernidade, a ser disponibilizada ainda em 2019 – o que aconteceu em outubro (Figura 1).

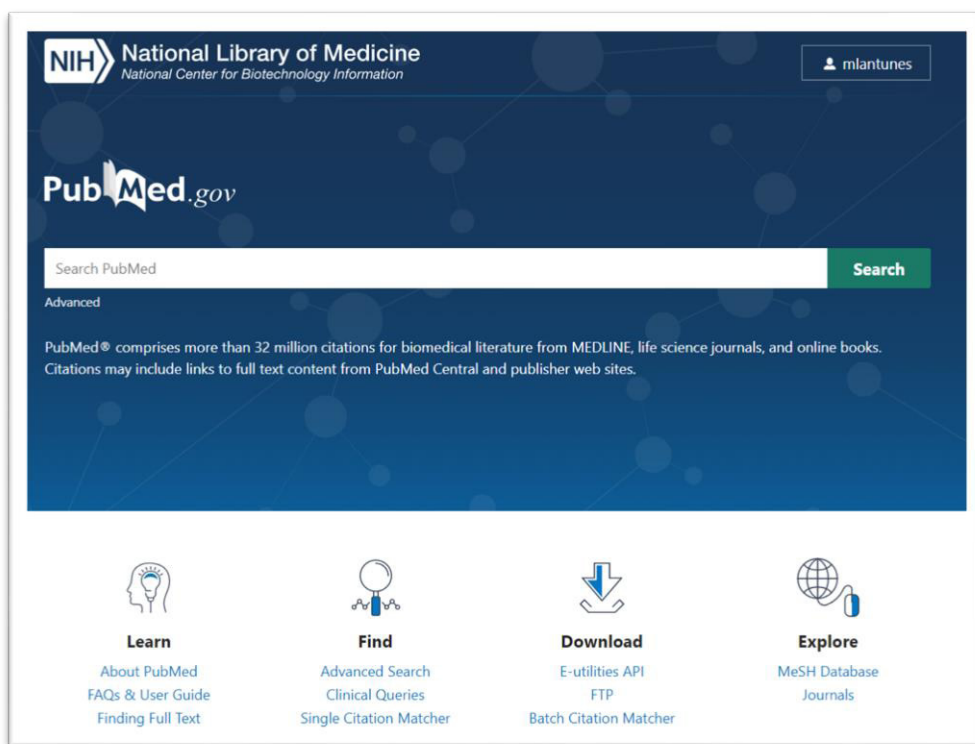


Figura 1. Homepage da nova plataforma PubMed.

Tendo sido objeto de uma evolução que ocorreu ao longo de décadas, acompanhando as transformações operadas pelas tecnologias da informação, para que a PubMed seja efetivamente utilizada, dever-se-á então considerar a sua interface, se é *friendly* e se apresenta funcionalidades de utilização de forma clara e objetiva, de modo que a interação homem-artefacto possa ocorrer com elevado grau de satisfação. Assim, o presente estudo tem como objetivo analisar a nova plataforma PubMed a partir do contributo de cinco personas ou peritos na matéria, num con-

traste de experiências e de perspectivas e averiguar como cada um visualiza os impactos do novo artefacto na sua prática profissional.

Metodologia

Foi usada uma metodologia qualitativa, através da realização de uma entrevista semiestruturada sobre as percepções de cinco peritos, representativos de um conjunto de utilizadores, sobre a nova plataforma da PubMed. Os entrevistados foram informados do objetivo da entrevista e tiveram acesso prévio aos tópicos. A conversa guiada entre o entrevistador e cada perito permitiu a obtenção de detalhes adicionais. As entrevistas foram transcritas, procurando assegurar alguma consistência aos elementos discursivos. As respostas às perguntas da entrevista semiestruturada foram objeto de uma análise qualitativa de conteúdos, de onde foram extraídas categorias (Denzin & Lincoln, 2017).

1. Amostra

Para este estudo foi usada uma amostra de conveniência, que procurou recolher as impressões de um conjunto diferenciado de utilizadores da PubMed da área da saúde: o bibliotecário, o investigador, o professor do ensino superior, o médico e o gestor editorial.

Foram identificados individualmente os cenários de uso de cada persona, os objetivos de utilização da plataforma, as ações que desenvolvem para as concretizar, avaliando se o novo sistema permite responder às suas necessidades enquanto profissionais.

2. Personas

- O bibliotecário. Possui uma experiência vasta adquirida ao longo de mais de trinta anos de atividade na área da informação em saúde.
- O investigador. Professor universitário há quinze anos. A sua produção científica é muito vasta, de que resultou uma dúzia de prémios ao longo de mais de vinte anos dedicados à investigação. A sua área de estudos são as ciências médicas e as ciências da saúde com ênfase nos media e na comunicação.
- O professor do ensino superior. Exerce a docência há mais de vinte anos nas áreas da psicologia e dos recursos de investigação inseridos nos processos de ensino-aprendizagem dos estudantes universitários.
- O médico. É investigador e professor universitário há sete anos. Apresenta uma vasta produção científica na área da epidemiologia, da medicina baseada na evidência e no rastreio oncológico.
- O gestor editorial. Gere a plataforma da edição de duas revistas científicas da área da saúde desde 2008: uma de uma associação profissional e a outra de uma instituição do ensino superior.

Resultados e discussão

Os resultados obtidos na análise de conteúdo às entrevistas semiestruturadas assinalaram respostas que se categorizaram, numeraram e analisaram, tópico a tópico, sendo a discussão apresentada no mesmo alinhamento (Tabela 1).

Tabela 1. Categorias e subcategorias analisadas

Categorias analisadas	Subcategorias
1. <i>Link</i> de acesso	
2. Estética	
3. Ícones de identificação	<i>Learn, Find, Download, Explore</i>
4. Ligação às redes sociais	
5. Perfil de utilizador	
6. Pesquisa avançada	
7. Apresentação dos resultados da pesquisa	<ul style="list-style-type: none"> • Proximidade temática vs. atualidade • Leitura temporal dos resultados (gráfico) • Seleção de artigos c/ dados associados • Navegabilidade entre resultados • Contagem de citações recebidas • Visualização das referências citadas • Referenciação (AMA, MLA, APA e NLM)
8. Salvaguarda dos resultados	
9. Opinião final e crítica das funcionalidades da nova plataforma	

De entre o conjunto de resultados obtidos destacam-se:

1. *Link* de acesso. Todos os entrevistados se aperceberam de que o *link* de acesso à plataforma é novo, mas consideram-no meramente cosmético. O investigador acrescenta que lhe é indiferente, porque um utilizador habitual guarda o URL nos favoritos.

2. Estética. Também todos os entrevistados se aperceberam da nova estética; um ou outro gosta, mas os outros consideram irrelevante. O investigador acha o tamanho de letra grande de mais, mas crê que essa é a nova teoria de usabilidade da *web*. O bibliotecário e o professor valorizaram a mudança da estética porque se mantém a cor azul a que o utilizador se habituou a identificar como marca da National Library of Medicine, mas também porque se registou uma melhoria no *browsing*, especialmente para aqueles que usam dispositivos com ecrãs mais pequenos, nomeadamente a versão *mobile* (Frisone & Micali, 2020). O médico está a adaptar-se, ainda não sabe exatamente onde estão todas as funcionalidades, pelo que sente mais dificuldades na utilização.

3. Ícones de identificação. O bibliotecário considera os ícones de identificação fáceis de utilizar e úteis. Sendo mais modernos, fogem um pouco aos tutoriais da PubMed existentes. Para o professor, o uso destes ícones evita a ambiguidade,

tornando a pesquisa fácil, amigável e intuitiva. Considera que os gráficos e a forma de exibição das figuras e tabelas aproximam a PubMed das melhores práticas, apresentando uma boa dinâmica na forma de classificar/organizar a informação. Mas é aqui que o investigador começa a recear a mudança. Considera que a PubMed viveu duas etapas muito diferentes: a primeira até 2012, a etapa do editor Sheldon Kotzin, em que a PubMed era demasiado fechada e excessivamente tradicionalista; e a segunda, a era pós-Kotzin, em que se procurou imitar o Google. O investigador defende a PubMed para uso profissional, não para uso amador. Acredita que estes ícones são a demonstração de que a estética para amadores passou a ser igual, ou até mais importante, que as funcionalidades para profissionais.

4. Ligação às redes sociais. O bibliotecário considera que ao permitir que os utilizadores partilhem citações no Facebook ou no Twitter a PubMed se aproxima de outras bases de dados agregadoras de conteúdos, na medida em que estas possibilidades são atualmente incontornáveis nestes recursos. O professor considera uma melhoria substancial a utilização do *link* permanente e a sua disseminação nas redes sociais. O médico não pensa usar e o investigador não comentou.

5. Perfil de utilizador. O bibliotecário não encontra alterações, crê que a nova versão remete para conteúdos já existentes. O investigador é da mesma opinião, mas acrescenta que se perdeu uma boa oportunidade de incluir indicadores unívocos de autor, como o ORCID ou, se necessário, poder-se-ia criar um PubMed ID como o Scopus criou o Author ID. Para o professor, o perfil do utilizador é mais simples e muito semelhante ao Mendeley.

6. Pesquisa avançada. O professor é de opinião que esta área se assemelha à da Scopus, Web of Science e Science Direct. Crê que a PubMed procurou um novo posicionamento, absorvendo uma arquitetura conhecida pelos estudantes, investigadores e bibliotecários, que torna a navegação e a pesquisa avançada mais simples e intuitiva, mais interoperável, mais amigável para os utilizadores, sendo um estímulo à inteligência na arte da pesquisa de informação científica e possibilitando a proximidade desta ferramenta ao cidadão comum (White, 2020). A esta perspetiva não é alheia a experiência de cerca de 20% dos utilizadores da PubMed que a consultam através dos dispositivos móveis (Fiorini, et al., 2017).

O investigador não identifica grandes diferenças com o *Advanced* antigo. Encontra mais alguns campos no identificador de campo (*collection*, *publisher* e pouco mais). Considera a relevância de um *field descriptor* de palavras-chave de autor, que não existia nem foi contemplado na nova versão. O médico, por seu turno, não conseguiu identificar nada de novo, ainda que esta seja a área que privilegia porque lhe permite visualizar a estratégia de pesquisa e aplicar filtros. Mas receia que as truncaturas possam não ter sido trabalhadas, preocupação também evidenciada por Vieira, Sousa-Pinto e Figueira (2021).

O bibliotecário manifesta alguma resistência à opção da pesquisa avançada, porque por hábito privilegia a pesquisa a partir dos termos MeSH (com acesso a partir do *MeSH Database*) e a construção da pesquisa usando o *Search Builder*. Lamenta que o histórico de pesquisas continue a ser eliminado após oito horas de inatividade, mas Frisone e Micali (2020) destacam que a nova PubMed consegue assegurar um histórico de 500 pesquisas, que pode ser eliminado na íntegra ou faseadamente.

7. Apresentação dos resultados de pesquisa. O investigador não identifica grandes diferenças, mas encontra mais alguns campos (Figura 2). Destaca as *display preferences*, que são agora mais visíveis. O bibliotecário assinala a atualização e automatização do mapa de termos, que acrescenta sinónimos, truncaturas, plurais, géneros, verbos, variantes britânicas e americanas da língua inglesa (e.g., behavior/behaviour; gynecology/ gynaecology, etc.) (Feld et al., 2020; García-Puente et al., 2020).

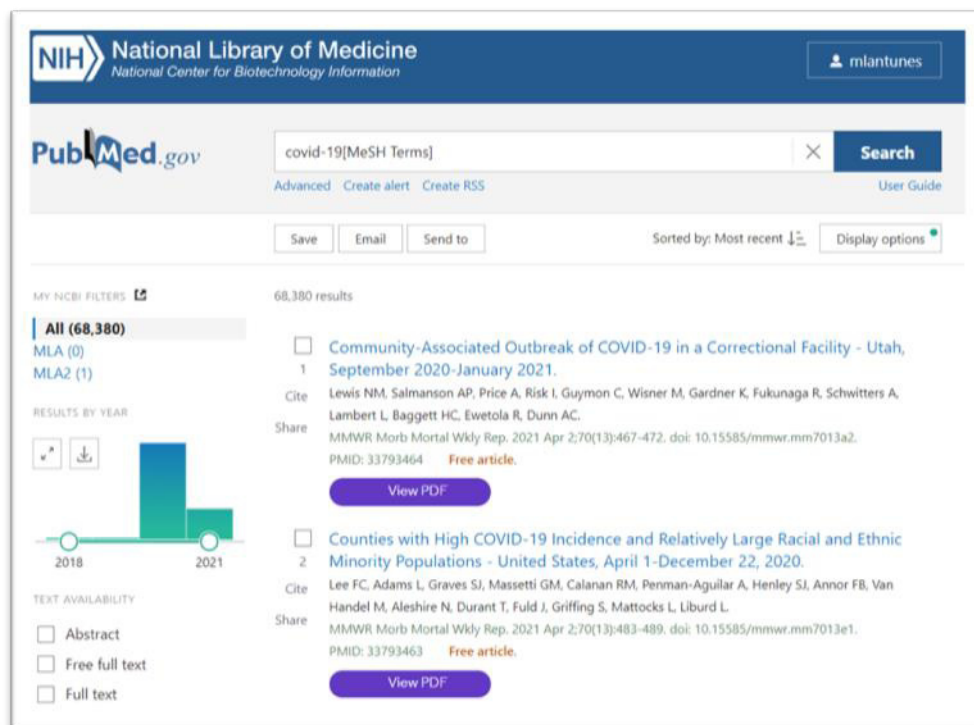


Figura 2. Apresentação de resultados na nova plataforma PubMed.

7.1. Proximidade temática vs atualidade. O investigador não aprecia o algoritmo do *best match*, porque não o conhece; prefere ser ele a definir a relevância dos resultados. Considera que esta é uma das dinâmicas que se está a copiar do Google e com a qual não concorda (Feld et al., 2020). Também o professor não usa. E pede aos seus estudantes que seja a primeira coisa a modificar aquando da apresentação dos resultados. O médico, que foi desde sempre um *early adopter* desta opção, não está convencido de que lhe permita localizar informação mais depressa. O bibliotecário, pelo contrário, considera razoável a apresentação dos resultados obtidos por *best match*, uma vez que considera que não compromete, antes acentua, a relevância (García-Puente et al., 2020).

7.2. Leitura temporal dos resultados (gráfico). O bibliotecário considera o gráfico dos *results by year* um indicador interessante, mas de utilidade relativa. O investigador reforça que já havia um, mas do lado direito. O gráfico atual permite

usar a régua do *timeline* como filtro, mas era mais importante manter uma *query* bem desenhada, com o identificador de campo *date of publication*. O médico não considera o gráfico particularmente útil porque todos os gráficos são exponenciais. O professor, por seu turno, avalia positivamente o dinamismo e a visibilidade da régua temporal; pode proporcionar a filtragem imediata através da data de publicação, ainda que as datas (de criação, de conclusão, de acesso eletrônico, de publicação, de entrada na PubMed) representem justamente o calcanhar de Aquiles da nova PubMed (García-Puente et al., 2020).

7.3. Seleção de artigos com dados associados. O investigador considera positivamente esta possibilidade, mas de pouca utilidade para ele. Afirma que pode querer consultar um artigo que não associa os dados, mas que o autor correspondente disponibiliza *on request*. O bibliotecário, por seu turno, é de opinião que pode ser uma boa funcionalidade a possibilidade de filtrar artigos pelo *Article Attribute*, mas ainda encontrou vantagens indiscutíveis. O médico não se tinha apercebido.

7.4. Navegabilidade entre resultados. O bibliotecário, o investigador e o gestor editorial não encontram vantagens em navegar lateralmente artigo a artigo; acham que o que existia resultava perfeitamente. O professor afirma que tem ensinado os seus estudantes a abrir os artigos mais relevantes num separador novo, de modo a nunca perderem a pesquisa, mas reconhece que maximiza a seleção de registos. O médico não se tinha apercebido.

7.5. Contagem de citações recebidas e visualização das referências citadas. Cada resultado apresenta o número de citações recebidas e quais. O investigador encara estas citações com algum ceticismo. É necessário analisar a origem dos dados dessas citações, porque os metadados que as revistas indexadas na PubMed enviam mensalmente não contemplam estes dados. O professor considera que esta funcionalidade está ainda longe do desejável.

O bibliotecário refere que as métricas de citação estavam ausentes na PubMed e que é importante que agora sejam contempladas. Mas parece-lhe insuficiente listar apenas citações de artigos indexados na base de dados, sem que haja o cálculo de um impacto. O médico considera que é uma função muito útil para fazer *citation tracking*.

7.6. Referenciação. Cada resultado oferece a possibilidade de referenciação de acordo com AMA, MLA, APA e NLM[†]. O bibliotecário é de opinião que é uma ferramenta prática e útil. O investigador, por seu turno, reage com alguma desconfiança, questionando quem usa essa ferramenta a partir da PubMed. Considera que pode ser uma estratégia para evitar que os estudantes menos aplicados continuem a citar erradamente, porque um mestrando ou um doutorando trabalha com um gestor de referências (Mendeley, EndNote, Reference Manager, ou outros). Professor e médico são da mesma opinião, ainda que o primeiro reconheça a sua eficácia. O editor editorial não confia na ferramenta, porque já encontrou exemplos que claramente não seguem a referenciação escolhida. Também destaca um lapso da PubMed: a NLM deveria ser identificada por *Vancouver Style*, porque essa é a sua identificação e é por ela que os profissionais de saúde a conhecem.

[†]AMA (American Medical Association style), MLA (Modern Language Association style), NLM (National Library of Medicine style guide).

8. Salvaguarda dos resultados. O investigador não identifica nada de novo na salvaguarda de artigos. Porém, quando verifica que o formato CSV é agora claramente identificado, confirma que a NLM está a tentar chegar ao público mais leigo, porque as potencialidades de um ficheiro CSV já eram permitidas na exportação em ficheiros [.txt] do formato MEDLINE. Na sua qualidade de investigador é esse tipo de ficheiro que lhe interessa quando precisa exportar grandes quantidades de artigos (mais de 25.000). Reconhece que nesta nova exportação se perdem imensos campos que a ele, investigador, lhe interessam (e.g., as datas) (García-Puente et al., 2020). Por esse motivo, não vai usar um ficheiro CSV porque o anterior ainda existe e permite-lhe trabalhar com os metadados que lhe interessam.

O editor apercebeu-se dos mesmos detalhes assinalados pelo investigador. Refere ainda que os ficheiros [.nbib], que permitem a exportação direta para os gestores de referências, estão agora mais arrumados, mas também menos visíveis. O professor considera positivamente as melhorias no campo da interoperabilidade com gestores. O bibliotecário não se apercebeu de quaisquer alterações. No seu entender, não se registou sequer qualquer inovação. Porém, nenhum dos entrevistados se apercebeu que o formato .ris foi substituído por um formato PubMed, que usa a mesma codificação da MEDLINE – Unicode UTF-8 –, reduzindo a possibilidade de erros aquando da exportação de ficheiros (Frisone & Micali, 2020).

9. Opinião final e crítica das funcionalidades da nova plataforma.

i. O investigador é de opinião que se perdeu uma ferramenta importantíssima: a caixa do *See Details*, em que o sistema colocava a *query* realmente pesquisada. Agora, se o investigador se esquece de algum operador booleano ou de algum identificador de campo, o sistema sugere o que lhe parece mais adequado. Face a uma *query* grande e complicada, às vezes cometem-se erros ou esquecem-se pormenores e o investigador quer o conteúdo dessa caixa visível para confirmar se a PubMed corrige, ou não, a sua *query*.

ii. O investigador considera ainda que a PubMed perdeu uma boa oportunidade para inserir as *keywords* de autor no identificador de campos da pesquisa avançada (Minguet, et al., 2014). O gestor editorial corrobora esta afirmação e acrescenta que estas só são visíveis quando os artigos são incorporados na plataforma, na versão EPUB AHEAD OF PRINT. Quando são efetivamente indexados, os artigos perdem as *keywords* de autor. Do mesmo modo, também no identificador de campos da pesquisa avançada, o investigador refere que era importante inserir a figura do *reviewer* – para a qual já chamou a atenção num dos seus trabalhos.

iii. Consideram ainda o investigador e o gestor editorial que se perdeu outra vez a oportunidade de rever os critérios de descrição da tipologia de artigos.

iv. O bibliotecário reconhece vantagens nesta nova plataforma, como a proximidade a modelos de bases de dados com interfaces modernos e mais próximos da pesquisa, que lhe parecem mais amigáveis para o utilizador final, contrariamente ao que costuma acontecer com as bases de dados académicas. Destaca, em especial, a indicação dos *trending articles*, da *latest literature* e das FAQ, que lhe parecem continuar a ser bem detalhadas. Mas o bibliotecário também realça as omissões, em

especial, a bibliometria e as métricas alternativas. Na sua opinião, os investigadores desejam aceder de imediato às métricas de citação dos seus artigos, o que passa a ser possível na nova PubMed, mas também à possibilidade de aceder à análise de resultados com outros indicadores bibliométricos, à semelhança do que existe na Web of Science ou na Scopus, por exemplo.

v. O professor considera a nova PubMed um sinónimo de mudança de paradigma: uma mudança mais ágil, mais interoperável, mais amigável para os utilizadores. Destaca a arquitetura semelhante às grandes plataformas de investigação (e.g., Scopus ou Web of Science), o equilíbrio com o passado através de outras bases de suporte (e.g., *bookshelf*), a linguagem mais adaptável a qualquer investigador, seja das ciências da saúde seja das ciências sociais. Na sua perspetiva, enquanto professor empenhado na transmissão de conteúdos aos seus estudantes, a PubMed tornou-se mais pragmática e mais intuitiva, sendo um estímulo à inteligência na arte da pesquisa de informação científica e possibilitando a proximidade desta ferramenta ao cidadão comum.

vi. O médico conclui não estar a usar a nova PubMed com regularidade. De acordo com a sua prática clínica, está a canalizar as suas pesquisas para bases de dados que categorizem os artigos científicos prioritariamente por *guidelines*, revisões sistemáticas e artigos primários. Reconhece, porém, que a PubMed disponibiliza uma área que lhe é essencial (a das *clinical queries*), que espera ter sido melhorada. A este respeito, Chan (2020, 2021) considera vantajosa a correção e atualização dos filtros das *clinical queries*, inalteráveis desde o ano 2000, que, devido à sua combinação única, garantem a efetividade das estratégias de pesquisa, respondendo a questões clínicas sobre terapia, diagnóstico, etiologia ou prognóstico.

O médico assinala, de igual modo, a incorporação de funcionalidades importantes, como um campo de pesquisa específico para a investigação publicada sobre o COVID-19 e a possibilidade de recuperar *preprints*. Esta evidência é corroborada por Flanagan, Fontanarosa e Bauchner (2020) e por Kang, Kalloniatis e Doig (2021), que afirmam que as novas funcionalidades da PubMed permitem aos profissionais de saúde uma pesquisa de literatura científica em três passos para a obtenção de respostas em tempo real, mas também para a realização de revisões sistemáticas.

Em síntese, ainda que os entrevistados apontem algumas reticências para a nova PubMed e identifiquem claramente as suas preocupações, não parece que resistam a continuar a usá-la enquanto base de dados de referência.

Conclusões

Analisar a nova plataforma PubMed a partir do contributo de cinco peritos na matéria, numa diversidade de experiências e de perspetivas, permitiu enriquecer a sua avaliação e considerar o seu desempenho em função de cinco diferenciados níveis de exigência, que a consideraram de acordo com as suas necessidades e padrões de prática profissional.

O presente estudo demonstrou que diferentes utilizadores, com diferentes experiências, têm uma visão própria da PubMed. Pela idade do artefacto e pela prática

profissional desenvolvida, os peritos entrevistados movem-se com razoável destreza e esperam dela um elevado grau de respostas relevantes e satisfatórias. Genericamente, esperavam mais da nova plataforma, mais qualidade e mais maturidade tecnológica. Constata-se que se perdeu a oportunidade de renovar de forma efetiva e válida a PubMed. Os perfis do investigador, do médico e do gestor editorial consideraram que a nova plataforma só é nova no visual, nada trazendo de novo para as suas práticas profissionais. Professor e bibliotecário, recorrendo às suas experiências pedagógicas, consideraram-na um sinónimo de mudança de paradigma, assente na agilidade e na interoperabilidade de processos, mais amigável para os utilizadores, possibilitando a proximidade desta ferramenta ao cidadão comum.

Face às alterações efetuadas e à renovação da maioria dos conteúdos da plataforma de acesso, permanecem, porém, algumas questões em aberto:

1) A quem se destina a nova PubMed? A utilizadores inexperientes (a quem a mudança registada será útil) ou a utilizadores mais avançados e profissionais (que esperavam por mudanças relevantes e que resolvessem os problemas existentes da PubMed)?

2) Para quando assegurar padrões de acessibilidade destinados a deficientes visuais de modo que o objetivo da PubMed seja realizado de forma integral, universalizando o acesso ao conhecimento científico da área da saúde?

Equaciona-se, no futuro, a continuidade deste projeto, desenvolvendo um teste de usabilidade com uma amostra representativa de profissionais variados da área da saúde. A escolha de itens em discussão poderá beneficiar dos contributos ora apresentados.

A resposta às questões apresentadas e o desenvolvimento do estudo projetado, alicerçados num sistema tecnológico consistentemente desenvolvido, renovará a imagem de uma plataforma que, ao longo dos anos, tem testemunhado o avanço da ciência em saúde com elevados níveis de evidência científica.

Referências

- Arndt, B. G., Beasley, J. W., Watkinson, M. D., Temte, J. L., Tuan, W. J., & Sinsky, C. A. (2017). Tethered to the EHR: Primary care physician workload assessment using EHR event log data and time-motion observations. *Annals of Family Medicine*, 15(5), 419-426.
- Carayon, P., & Hoonakker, P. (2019). Human factors and usability for health information technology: Old and new challenges. *IMIA Yearbook of Medical Informatics 2019*. p. 71-77.
- Chan, J. (2020). PubMed clinical queries update coming soon. *NLM Technical Bulletin*, 436, e8.
- Chan, J. (2021). PubMed update: Clinical queries usability study and interface updates. *NLM Technical Bulletin*, 441, e3.
- Denzin, N. K., & Lincoln, Y. S. (2017). *The SAGE handbook of qualitative research* (5th ed.). Sage.
- Feld, M., Neuhäuser, K., Esteban, R. M., Tolkmitt, F., & Brzoska, T. (2020). The new PubMed: Underestimated regulatory obstacles? *Medical Writing*, 29(3), 52-59.
- Fiorini, N., Canese, K., Bryzgunov, R., et al. (2018). PubMed labs: An experimental system for improving biomedical literature search. *Database* (Oxford). <https://doi.org/10.1093/database/bay094>
- Fiorini, N., Lipman, D. J., & Lu, Z. (2017). Cutting edge: Towards PubMed 2.0. *eLife* 6, e28801. <https://doi.org/10.7554/eLife.28801>

- Flanagin, A., Fontanarosa, P. B., & Bauchner, H. (2020). Preprints involving medical research. *JAMA*, 324(18), 1840-1843. <https://doi.org/10.1001/jama.2020.20674>
- Frisone, F., & Micali, R. (2020). The new-look of PubMed: A brief description of the changes in PubMed. *Mediterranean Journal of Clinical Psychology*, 8(2). <https://doi.org/10.6092/2282-1619/mjcp-2534>
- García-Puente, M., Pastor-Ramon, E., Agirre, O., Morán, J. M., & Herrera-Peco, I. (2020). Research note. Open letter to the users of the new PubMed: A critical appraisal. *Profesional de la información*, 29(3), e290336. <https://doi.org/10.3145/epi.2020.may.36>
- Gawande, A. (2018). *Why doctors hate their computers*. The New Yorker.
- Hoogland, M. A. (2019). How medical students discover and use medical information tools. *Medical Reference Services Quarterly*, 38(4), 347-357. <https://doi.org/10.1080/02763869.2019.1661197>
- Kang, P., Kalloniatis, M., & Doig, G. S. (2021). Using updated PubMed: New features and functions to enhance literature searches. *JAMA*, 326(6), 479-480.
- PubMed. (2020). *PubMed: About*. <https://pubmed.ncbi.nlm.nih.gov/about/> (visit on 05/12/2020)
- Shanafelt, T. D., Dyrbye, L. N., Sinsky, C., Hasan, O., Satele, S., & Sloan, J. (2016). Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction. *Mayo Clinic Proceedings*, 91(7), 836-848.
- Vieira, R., Sousa-Pinto, B., & Figueira, L. (2021). Beware of word truncation in PubMed: Unexpected problems and potential solutions. *BMJ Evidence-Based Medicine*, 26(1), 34-36.
- White, J. (2020). PubMed 2.0. *Medical Reference Services Quarterly*, 39(4), 382-387. <https://doi.org/10.1080/02763869.2020.1826228>

INDEXAÇÃO E REPOSITÓRIOS INSTITUCIONAIS: ENQUADRAMENTOS, DEFINIÇÕES E TRAÇOS DISCURSIVOS

Marta Filipa Pinheiro dos Reis Luro, Maria Cristina Vieira de Freitas

¹Doutoranda em Ciência da Informação, Faculdade de Letras, Universidade de Coimbra,
mfluro@gmail.com ORCID iD 0000-0002-7603-5168

²Centro de Estudos Interdisciplinares do Século XX, Faculdade de Letras, Universidade
de Coimbra, cristina.freitas@fl.uc.pt, ORCID iD 0000-0002-8849-8792

Resumo

Os repositórios institucionais oferecem um conjunto de serviços que permitem o armazenamento, a organização, a recuperação e uma ampla disseminação da informação científica produzida na instituição. O número de recursos publicados em formato digital cresce cerca de 4% ao ano, o que faz com que a indexação, pela sua capacidade de representar os conteúdos, seja um processo cada vez mais necessário à eficácia e efetividade na recuperação da informação. O objetivo principal deste estudo é revisar a literatura científica acerca da indexação, de modo geral, e da indexação em repositórios, de modo particular. Os objetivos específicos passam, então, por identificar e selecionar um conjunto de textos especializados e confrontar alguns dos seus traços discursivos relacionados com a indexação e com algumas das problemáticas em torno da indexação dos documentos depositados em repositórios em acesso aberto e votados à reunião da produção científica institucional. O estudo tem um caráter exploratório e descritivo e os textos são recolhidos intencionalmente em bases de dados e motores de busca. Os estudos analisados e sumariados permitem-nos inferir que, apesar de uma alegada morosidade do processo, os diferentes tipos de indexação usados em repositórios institucionais acarretam benefícios a aproveitar. Sobre o uso da linguagem natural na indexação em repositórios, entre outros, foram mencionados os seguintes benefícios: registo direto da informação, pelo/a utilizador/a (auto-depósito), agilização do processo (autonomia das tarefas), redução da distância na linguagem usada pelos/as utilizadores/as e pelo sistema. Assim, o uso de linguagem natural (palavras-chave de autor e etiquetas colaborativas), combinado com a linguagem controlada, embebida no sistema, surge como proposta a ser usada para potenciar a recuperação de informação. Conclui-se que as vantagens de um sistema podem colmatar as falhas de outros e que a amostra de textos deve ser alargada para a reunião e o confronto de casos concretos de uso recentes, em próximos estudos.

Palavras-chave: Repositórios institucionais; Revisão bibliográfica; Indexação; Representação temática; Organização do conhecimento.

Introdução

A organização é uma característica essencial do ser humano e, mesmo de forma inconsciente, está presente nos diversos tipos de atividades e tarefas, desde às mais simples às mais complexas (Vignoli, Almeida & Catarino, 2014). A pesquisa por informação científica não é mais realizada exclusivamente “intramuros”, mas na Web, onde podemos encontrar a informação armazenada e disponível em sítios de bibliotecas e em repositórios digitais.

A Ciência da Informação (de agora em diante, apenas CI) considera as atividades próprias da organização e da representação do conhecimento como meios para atingir o objetivo final de eficácia na recuperação da informação. Algumas destas atividades, especialmente as que concernem à representação temática da informação, materializam-se na indexação, que é o processo de análise documental realizado para determinar, em cada documento, um conjunto de palavras-chave ou assuntos e os traduzir em termos. Estas são áreas de toque da CI, porque a procura por informação de qualidade e compatível com as necessidades dos/as utilizadores/as tornou-se uma constante, face aos imperativos tecnológicos atuais (Bandim, & Correa, 2019; Souza, 2017; Vignoli, Almeida, & Catarino, 2014). O número de artigos científicos publicados na Web, atualmente, é elevadíssimo¹, o que torna a indexação um processo indispensável, embora moroso (Bandim, & Correa, 2019), dependendo dos meios a empregar e dos resultados a obter (Sousa, 2015) (*e.g.*, diferentes graus de precisão na recuperação da informação).

Com a proliferação dos repositórios digitais, muito em particular dos institucionais, surgem novas formas de registar e de disseminar a informação, tornando a sua identificação e o seu acesso cada vez menos lineares. Tornar as informações disponíveis em acesso aberto e melhorar a sua recuperação, nos repositórios, afiguram-se, nesse contexto, como preocupações centrais das atividades de indexação, especialmente, quando estas são realizadas pelos/as próprios/as autores/as dos textos, sendo eles/elas também os/as fiéis depositantes das suas publicações nos repositórios.

Com base nesse enquadramento, pretendemos, à luz de alguma literatura científica pertinente sobre essas matérias, visível nos meios de divulgação certificados, rever alguns dos temas predominantemente associados à indexação, de modo geral, e à indexação em repositórios institucionais, de forma particular. Sendo possível, também intentaremos identificar algumas das suas linhas privilegiadas de discussão.

Metodologia

Este estudo irá assumir uma orientação qualitativa, na medida em que iremos recolher dados em função das suas características e do enquadramento nas matérias pretendidas, bem como uma natureza exploratória e descritiva, já que iremos reco-

¹ Conforme dados reportados pela base de dados Scopus®(<https://www.scopus.com/home.uri>), os resultados de investigação vêm crescendo cerca de 4% ao ano. Entre 2008 e 2018, passou de 1.8 para 2.6 milhões de artigos publicados por ano.

lher textos da literatura científica recentemente publicada sobre a indexação e retirar elementos importantes para avivar a discussão em torno do seu uso em repositórios.

Os objetivos específicos passarão então por identificar e selecionar um conjunto de textos especializados e revisar e confrontar aspetos relacionados com a indexação e com a problemática em torno da indexação dos documentos depositados em repositórios em acesso aberto e votados à reunião da produção científica institucional.

Para atingir os objetivos propostos, faremos uso do método de revisão de literatura, utilizando a técnica de pesquisa bibliográfica e consultando as seguintes fontes de informação: B-On®, Google®, Google Académico®, Scopus® e RCAAP®. A partir dos resultados obtidos, excluindo-se as redundâncias, intentaremos realizar uma análise e sumarização de uma pequena fração da literatura científica descoberta.

Portanto, tendo em conta as funcionalidades disponíveis nas bases de dados consultadas, consideraremos as diversas tipologias de documentos e focar-nos-emos em critérios estabelecidos *a priori*, como: relevância do tema/assunto, autores/as mais citados/as, textos atuais e disponibilizados em texto integral. Os termos de pesquisa serão usados em três idiomas diferentes: português, inglês e espanhol. Não serão estabelecidas *a priori* barreiras temporais, uma vez que pretendemos manter o foco também em autores/as considerados/as expressivos/as no âmbito dos assuntos abordados.

No decorrer das buscas, admitindo-se os critérios antes explicitados, foram selecionados 31 textos, devidamente identificados nas referências bibliográficas deste estudo. Deste grupo, excluindo-se apenas um texto (3%), que atendeu aos critérios intencionais de notoriedade e não de atualidade, dado que foi publicado antes do ano 2000, bem como outros sete textos (23%), que se referem a publicações situadas em períodos igualmente mais recuados (2000-2005), a esmagadora maioria pode ser considerada de origem recente (12 textos; 39%; 2006-2015) ou mesmo muito recente (11 textos; 35%; 2016-2020) (figura 1).

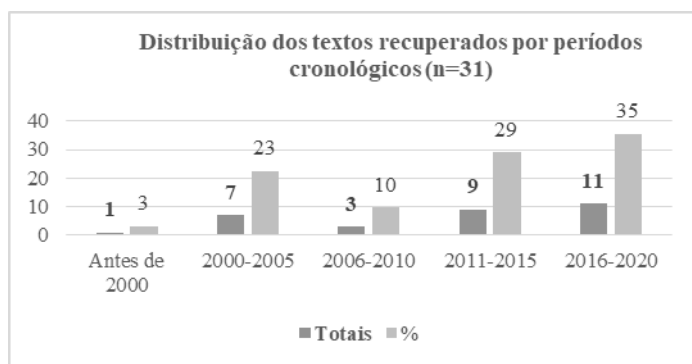


Figura 1. Distribuição dos textos recuperados (n=31).

Fonte: elaboração própria com base no *corpus* de dados de pesquisa.

A partir de uma pré-análise, seguida de alguma prospeção nos conteúdos, concluímos que os 31 textos selecionados versavam sobre problemáticas próprias da indexação, enquanto processo e atividade já consolidada na CI, por um lado, e sobre

a sua relação com os diferentes ambientes digitais, entre os quais os repositórios institucionais, por outro, sendo, naturalmente, estas perspetivas condicentes com as expressões de busca utilizadas. Também observamos que os conteúdos principais seguiam em três sentidos ou eixos principais, identificados na figura 2. As frequências de ocorrência dos referidos eixos não foram contabilizadas, quer pela sua transversalidade, quer pelo facto de não aportarem novidade ao estudo. Assim sendo, as análises e discussões centraram-se nos aspetos que nos pareceram ser mais pertinentes, no conjunto dos textos, à luz dos interesses de pesquisa.



Figura 2. Eixos de análise identificados na amostra de textos selecionada.

Fonte: elaboração nossa, com base com base no *corpus* de dados de pesquisa.

Após esta inicial distribuição dos textos nos respetivos eixos, verificamos detidamente os seus conteúdos, tentando alinhar os diferentes discursos que iam sendo identificados, para logo organizá-los, em função dos seus traços discursivos predominantes (figura 3).

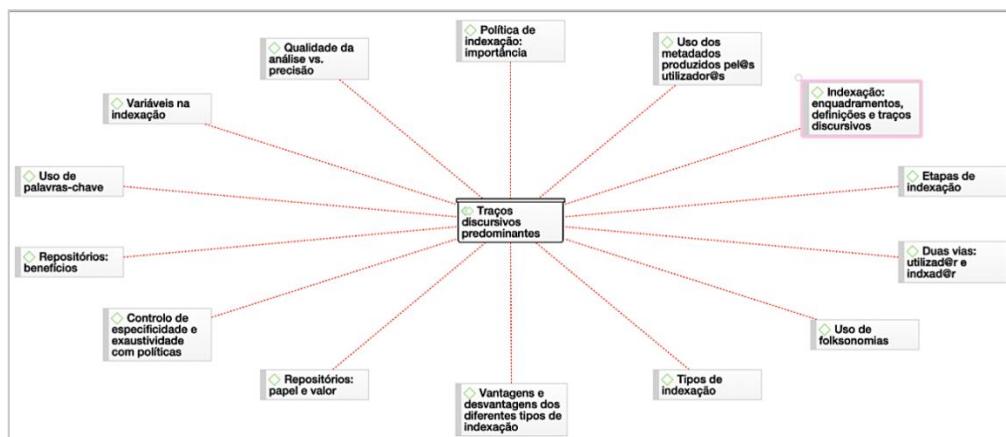


Figura 3. Conteúdos organizados em função dos seus traços discursivos predominantes.

Fonte: elaboração nossa, com base com base no *corpus* de dados de pesquisa.

A sumarização decorrente desse alinhamento discursivo encontra-se consubstanciada nos pontos que se seguem (cf. pontos 1, 2 e 3 dos resultados e discussão). Nela, são incorporados os discursos de autores/as presentes nos 31 artigos publicados, em autoria individual ou partilhada, num período cronológico compreendido entre 1989² e 2020. Uma nuvem de *tags* foi elaborada para representar esses resultados, bem como as frequências de ocorrência (autores/as, datas de publicação, traços discursivos) (figura 4).



Figura 4. Nuvem de *tags* contendo representações das codificações realizadas por autores/as, datas e traços discursivos predominantes.

Fonte: Obtenção automática, através do *software* Atlas.Ti@v.8.4.4, com base no *corpus* de dados de pesquisa.

Destas revisões em torno dos temas indexação, repositórios institucionais e indexação no âmbito dos repositórios institucionais, foram, ainda, retiradas algumas conclusões e recomendações a aproveitar em futuros estudos.

Resultados e discussão

1. Indexação: enquadramentos, definições e traços discursivos

Historicamente, a indexação aparece como uma atividade associada à elaboração de fontes bibliográficas especializadas em áreas do conhecimento dotadas de índices bibliográficos de autores/as, assuntos e títulos. A informatização leva à automatização desses índices e as bibliografias tornam-se bases de dados especializadas, dispo-

² Alguns textos podem conter dados que se desviam ligeiramente desses marcadores (*e.g.*, 1988; *e.g.*, citações de autores/as incluídas no quadro 1 deste estudo), o que se deve à existência de citações indiretas a alguma bibliografia que foi codificada nas etapas iniciais de análise, mas que não foi considerada diretamente no *corpus* final de dados de pesquisa.

níveis em plataformas e em sítios Web, na Internet. A indexação, enquanto processo de análise documental, passa a ser realizada, de modo realmente intensivo, desde a explosão das publicações periódicas e, conseqüentemente, da literatura técnico-científica, muito em função da necessidade de aplicação de mecanismos de controle bibliográfico nesses meios. A necessidade de recuperação cada vez mais imediata e especializada, fez com que o tratamento da informação passasse a contar com procedimentos e estratégias mais diversificadas e centradas no contexto de cada documento³ (Fujita, & Troitiño, 2018).

Segundo Vignoli, Almeida e Catarino (2014), a organização e a representação do conhecimento recorrem a instrumentos ou sistemas e têm como meta a eficácia na recuperação da informação. Neste sentido, Bandim e Correa (2019) complementam que a organização e a recuperação da informação materializam-se, por excelência, na indexação, sendo esta, conforme dissemos mais atrás, um processo de análise documental realizado com a finalidade de determinar, para cada documento, um conjunto de palavras-chave ou de assuntos capazes de os representar em sistemas de recuperação de informação. Oliveira (2017) também define a indexação como “uma atividade central nas unidades de informação, responsável por evidenciar os atributos temáticos dos documentos pela constituição de termos representativos de seus assuntos” (p. 40). Fujita e Gil-Leiva (2014), de igual modo, identificam-na (a indexação) como um processo crucial numa política de tratamento temático da informação e sujeito a variáveis tais como: seleção da linguagem, exaustividade, especificidade, revocação e precisão. Também complementam (Fujita, & Gil-Leiva, 2014), dizendo que a indexação transita por duas vias: uma, associada ao/à indexador/a e à identificação dos conceitos extraídos do documento com a sua conversão em termos de indexação; outra, associada ao/à utilizador/a e limitada aos conceitos que este/a considera que orientam a sua estratégia de pesquisa.

Ora, a indexação para ser bem-sucedida, tanto na representação como na recuperação, dependerá de uma política delineada e enquadrada no contexto institucional. Souza (2017) refere mesmo a necessidade de as instituições adotarem uma política de indexação assente num conjunto de decisões que determinem, de forma pormenorizada, os interesses e os objetivos a conseguir num sistema de informação. É importante, pois, que com essa política de indexação, para além dos requisitos, elementos e variáveis associadas, o sistema de informação controle os níveis de especificidade e de exaustividade pretendidos, o que se consegue pela aplicação de controlo do vocabulário (Fujita, & Gil-Leiva, 2014). Para Fujita e Gil-Leiva (2014), as decisões sobre a especificidade e a exaustividade da indexação são tomadas nas etapas de análise de assunto e de tradução dos conceitos em termos constantes do vocabulário controlado. A opção por uma ou por outra, ou por ambas, irá determinar o índice de precisão a obter na recuperação da informação. Podemos, assim, inferir que a política de indexação delineia as estratégias e as ferramentas ou técnicas a usar e concorre para o sucesso da atividade de indexação.

A este respeito, e num alinhamento convergente, Souza (2017) pondera que a eficácia na recuperação se encontra diretamente relacionada com a procura por infor-

³ Que no contexto da Web e dada a multiplicação dos suportes informacionais, passa a designar-se “recurso de informação” ou, imlemente, “recurso”.

mação de qualidade e compatível com as necessidades dos/as utilizadores/as do sistema. Mendes e Simões (2001, p. 17) complementam que “a qualidade da indexação depende, antes de tudo, da qualidade da análise, resultando esta da fidelidade com que [se] exprime o conteúdo total ou parcial do documento, ou seja, da fidelidade com que [se] exprime o pensamento do[a] autor[a]; resulta, ainda, de uma avaliação positiva da sua pertinência informativa em relação ao potencial utilizador[a]”. O foco, deste modo, situa-se em duas variáveis relevantes do processo – o/a utilizador/a e o sistema – permeadas por uma característica a perseguir: a precisão.

Dos aportes anteriores, inferimos que a indexação se processa em mais de uma etapa, que a qualidade deve ser uma constante do processo e que a precisão na recuperação será uma meta a alcançar. As diferentes etapas ou fases da indexação, embora partam de uma base comum, variam ligeiramente na forma como são expressas pelos/as diferentes autores/as consultados/as. Silva (2020), por exemplo, reúne algumas, a juntar a outras, que apresentaremos, a modo de resumo livre, no quadro 1.

Quadro 1. Identificação das etapas de indexação

Autor/a (data)	Etapas de indexação
Chaumier (1988)	1. Reconhecimento e extração dos conceitos informativos 2. Tradução dos conceitos numa linguagem documental
Pinto (2000)	1. Análise conceitual 2. Tradução em termos de indexação 3. Controlo de qualidade
Robredo (2005)	1. Análise conceitual do conteúdo do documento 2. Expressão num conjunto de frases ou palavras 3. Tradução em assuntos relevantes 4. Organização dos assuntos em função da sintaxe da linguagem de indexação
Gil-Leiva (2008)	1. Análise dos documentos e seleção dos conceitos explícitos ou implícitos 2. Armazenamento das palavras-chave na forma como se apresentam 3. Normalização por meio de um vocabulário controlado
Lancaster (2004); Fujita (2013); Oliveira (2017)	1. Análise conceitual do conteúdo do documento 2. Tradução em termos de indexação
NP 3715 (1989)	1. Análise do documento e definição do seu conteúdo 2. Identificação e seleção dos conceitos representativos do conteúdo 3. Representação desses conceitos por termos de indexação
ABNT 12676 (1992)	1. Análise do documento e definição do assunto do seu conteúdo 2. Identificação dos conceitos presentes no assunto 3. Tradução dos conceitos em termos usados por uma linguagem de indexação

Fonte: elaboração nossa, com base nos/as autores/as citados/as e em Silva (2020).

Do quadro anterior, inferimos que, independentemente da linha de interpretação, nas fases de indexação encontram-se presentes as atividades de análise de conceitos e de tradução em termos. Oliveira (2017) confirma que análise concei-

tual e tradução são, de facto, etapas fundamentais desse processo, dado que se destinam à compreensão do conteúdo do documento e à adequação dos termos a usar para o representar, em conformidade com a linguagem usada pela “unidade de informação”. Ainda, a NP 3715 (1989, p. 4), reconhece que apesar de estipular que o processo de indexação decorre em três fases fundamentais (cf. quadro 1), na prática, elas se sobrepõem.

No âmbito da tipificação dos processos de indexação, Sousa (2015) e também Simões, Machado, Souza e Lopes (2017), analisam os três tipos usualmente referidos na literatura científica: manual, automático e semiautomático⁴. Os traços discursivos identificados nos textos publicados por esses/as autores/as revelam alguma convergência no que tange às principais ideias. Por exemplo, quanto a que processo adotar, ponderam que a decisão fica na dependência dos objetivos a atingir, reconhecendo haver vantagens e desvantagens em cada opção. Simões, Machado, Souza e Lopes (2017) recordam que os prós e os contras estão usualmente relacionados, por um lado, com a consistência, em desvantagem nos processos de indexação manual e, por outro, com a semântica, prejudicada na indexação automática. No âmbito desta última, segundo referem, tem especial relevo a necessidade de compreender os critérios utilizados pelo *software*, que podem não ser exatamente iguais aos critérios estabelecidos, por exemplo, na indexação manual. Quanto a esta, referem que a morosidade do processo, que é devedora do grau de complexidade do assunto a representar, condiciona a sua adoção, especialmente, diante de uma necessidade de aceleração de procedimentos. Conforme evidenciam, estas condicionantes sentem-se especialmente nos ambientes Web, onde o exponencial crescimento dos acervos digitais tem favorecido, evidentemente, um maior desenvolvimento da indexação automática. Contudo, recordam que os sistemas automáticos ainda se deparam com constrangimentos⁵ que perturbam a indexação e, conseqüentemente, a recuperação da informação (Simões, machado, Souza, e Lopes, 2017). De modo semelhante, Sousa (2015) assevera que “a indexação automática é opção interessante quando se trata de um acervo digital que contém documentos textuais heterogêneos e que o objetivo é que a informação chegue mais rápido até os/as usuários/as” (p. 15). Também assevera que os avanços obtidos nos estudos sobre essas matérias levaram ao aparecimento de métodos que, no presente, conseguem combinar estatística, linguística e linguagem documental em favor da indexação automática⁶.

⁴ Na indexação manual, o processo é feito por humanos; na automática são as tecnologias computacionais que gerem as etapas e as decisões operacionais; e na semiautomática é repartido o trabalho entre humanos e máquinas, cabendo aos primeiros a decisão final sobre a atribuição dos termos apropriados.

⁵ Entre os quais, destacam: diferenças nos vocabulários e terminologias específicas de nichos e domínios; documentos em linguagens múltiplas; polissemia e sinonímia. A título de informação, Simões, Machado, Souza e Lopes (2017) defendem o uso dos sistemas de organização do conhecimento (vulgo SOC), na dotação dos sistemas computacionais de uma maior capacidade de inferência.

⁶ De facto, segundo Borges (2008 cit. por Sousa, 2015), a maioria dos procedimentos de indexação automática relaciona-se com a extração de termos pelo método da frequência ou análise estatística. O processo de extração dá-se, segundo Lancaster (2004), quando as palavras ou expressões presentes num documento são escolhidas para representar o seu conteúdo ou, ainda,

Outro dos traços discursivos recentes ou mais recentes⁷ que emergem dos textos consultados são as linguagens a usar no ato de indexação. Nesse contexto, há duas vertentes relevantes a analisar: o uso de palavras-chave e a adoção das *folksonomias*.

Fujita e Tartarotti (2020) dizem-nos que o uso de palavras-chave está relacionado ao uso de uma palavra significativa, *i.e.*, “chave” na representação de um recurso informacional, para que o mesmo possa ser acedido e recuperado em variadas outras fontes de informação. Como ponderam, essa “chave”, sistematicamente extraída ou atribuída, potencia a divulgação e a visibilidade do recurso informacional. Assim sendo, “[o] conceito de palavra-chave está condicionado à representação do significado de um determinado conteúdo verbal ou não verbal e tem diversas finalidades, mas, essencialmente, é utilizado para identificação de ideias e temas importantes” (p. 336). Reiteram que entre as diversas finalidades da palavra-chave está, evidentemente, o seu uso na indexação e na recuperação de informação.

As *folksonomias*, conforme sabemos, são o resultado de um processo de indexação social, realizado de forma colaborativa e partilhado entre utilizadores/as e indexadores/as. Santos e Corrêa (2015) referem que as *folksonomias* representam uma alternativa nos processos de organização e tratamento da informação em ambientes digitais e que “[...] nos sistemas colaborativos, a descrição dos objetos informacionais não fica apenas a cargo dos[as] seus gestores[as], mas também dos[as] usuários[as], com a realização da etiquetagem em recursos informacionais” (Santos, & Corrêa, 2015, p. 274). Para Medeiros (2014, p. 246), as *folksonomias* são uma forma de indexação cooperativa que faz uso de palavras-chave para representar a informação de modo espontâneo e livre, sendo, por isso, mais usadas no âmbito da Web 2.0, em que cada utilizador/a opta pelas palavras-chave (neste caso, etiquetas ou *tags*) que lhe pareçam ser convenientes para descrever os conteúdos⁸ com que se depara. Medeiros (2014) considera que a utilização do vocabulário pessoal dos/as utilizadores/as é um traço realmente vantajoso desse tipo de indexação, dado que contribui para a localização dos assuntos, na Internet, na medida em que os documentos se tornam recuperáveis pelos termos mais utilizados no momento. Como constrangimentos, refere, naturalmente, o seu contrário, ou seja, a ausência de controlo terminológico, que pode levar ambiguidade ao processo de representação e recuperação da informação, bem como a ausência de uma estrutura hierárquica ou associativa, para além da utilização indistinta do plural e do singular. Apesar disso, e como aspeto positivo [ou como solução para os constrangimentos apontados], refere que “as *folksonomias* servem para extrair léxico, a partir do qual se produzem os vocabulários controlados, após um processo de normalização” (Medeiros, 2014, p. 245). Dada a relevância destas questões, voltaremos à sua abordagem, num alinhamento mais à frente.

por atribuição de termos a um documento, feita por uma fonte que pode não ser o próprio documento.

⁷ Especialmente, no âmbito da indexação em repositórios institucionais, conforme adiante se verá.

⁸ Medeiros (2014) refere dois tipos de *folksonomias*: genéricas e específicas. As primeiras oferecem a possibilidade de etiquetagem do objeto, pelos/as utilizadores/as, com recurso ao próprio vocabulário; as segundas limitam o processo de etiquetagem a uma ou poucas pessoas.

2. Repositórios institucionais: enquadramentos, definições e traços discursivos

Na ótica de Freitas e Leite (2018), os repositórios institucionais são hoje poderosos canais de comunicação ao serviço das diversas Instituições de Ensino Superior (de agora em diante, IES). A sua existência acarreta diversos benefícios, entre os quais citam-se, nomeadamente: a visibilidade, a gestão e o controlo da produção científica, para além da ampliação das probabilidades de acesso à informação aí depositada.

Numa visão igualmente alargada, Lynch (2003, p. 2) pondera que esses repositórios são “[...] *a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members*”. Também considera que esses repositórios foram criados para fornecer uma infraestrutura propícia à disseminação da informação científica, tornando-a acessível a investigadores/as e outros públicos, para além de gerar um ambiente favorável à certificação da informação [e do conhecimento] aí depositada, devendo garantir a sua preservação e disponibilidade a longo prazo. Para Martins (2013) e Crow (2002), cujos discursos reproduzem as asserções de Lynch (2003), esses repositórios serão o arquivo digital da produção intelectual institucional, traduzida em uma ou várias comunidades de investigadores/as, estudantes e colaboradores/as, que desempenham a importante missão de promover a investigação que realizam. Neste sentido, os repositórios estimulam a inovação na comunicação entre as áreas académicas e servem como indicadores de qualidade, podendo esta ser medida pela visibilidade, pelo prestígio e pelo valor que representam para o público. Também este é o alinhamento discursivo que encontramos em Martins (2013) e em Viana, Márdero Arellano e Shintaku (2005), quando ponderam que os repositórios serão, de facto, os meios principais utilizados pelas IES para divulgar e dar visibilidade aos resultados das suas pesquisas científicas e para legitimar e estimular a publicação dos trabalhos produzidos pelos seus membros.

Os discursos anteriores alinham-se às reflexões de Miguéis (2012), que com a sua experiência na matéria também sumaria os vários benefícios dos repositórios institucionais⁹, acrescentando, ainda, um outro importante aspeto, que é o facto de estes serem sistemas muito viáveis, do ponto de vista económico, dados os relativos custos reduzidos que pressupõem. Nesse mesmo registo, cuja intenção será relacionar os repositórios aos diferentes valores que podem agregar, Kuramoto (2009) refere que os diversos países, ainda que indiretamente, tiram proveito da sua rede de repositórios institucionais nacionais, valendo-se dos indicadores de produção científica que daí podem extrair e que serão cruciais para o financiamento científico e para os investimentos em ciência.

⁹ Que, de resto, coincidem com os aspetos antes mencionados, ou seja: aumento da visibilidade, acesso, utilização e impacto dos resultados da investigação; melhoria e eficiência na monitorização, avaliação e gestão da atividade científica; criação de novos modelos de investigação; afirmação da identidade institucional; aumento da capacidade de armazenamento de informação; adoção de padrões e de formatos abertos.

Os repositórios institucionais configuram, assim, um novo modelo de organização do conhecimento e exercem um papel muito importante enquanto arquivos científicos digitais certificados. Não obstante, como recorda Wheatley (2004), nem sempre será esta a regra, pois instituições com necessidades de armazenamento, preservação, registo e fornecimento de acesso a outros tipos de recursos digitais podem lançar mão dessas tecnologias para configurar os seus próprios repositórios. De igual modo, Hockx-Yu (2006, p. 1) recorda que, embora os repositórios forneçam o acesso aberto aos resultados da investigação científica institucional, não serão esses os únicos conteúdos a incluir nesses sistemas, citando outros, como, por exemplo, os dados de pesquisa, os recursos de ensino-aprendizagem e as imagens. Martins (2013) também alega que um repositório institucional funciona, simultaneamente, como imagem de marca e objeto de aprendizagem em meio digital. Lembra que o conceito remete para um lugar onde se guarda algo. Assim sendo, os repositórios são criados para depositar e arquivar objetos de aprendizagem com qualidade, adequação, retroalimentação, adaptabilidade e com capacidade suficiente para motivar e gerar interesse em quem pesquisa. Tais características, aliada às anteriores, tornam ainda mais viável a opção por esses sistemas.

3. Indexação em Repositórios Institucionais: alinhamento discursivo

No âmbito da adoção da indexação de recursos informacionais depositados em repositórios institucionais, percebemos, em alguns dos discursos, argumentos relativos ao uso da linguagem natural e/ou da linguagem controlada, alinhando-se as discussões em torno dos benefícios que podem aportar ao processo de recuperação da informação, aliados às dificuldades eventualmente impostas (cf. análises realizadas no ponto 1, resultados e discussão). Vejamos, pois, como alguns desses discursos se constroem e se equilibram entre estas diferentes variáveis.

Lopes (2002), por exemplo, refere ser vantajoso o uso da linguagem natural nos repositórios, dado que permite o registo direto da informação, pelo/a utilizador/a, sem que tenha de recorrer à consulta de uma linguagem controlada para o efeito. A nosso parecer, este será um expediente especialmente útil, sobretudo, no âmbito do auto-depósito, dado que contribui para a agilidade do processo e também estimula a conclusão da tarefa por parte de quem a realiza (cf. no ponto 1, Medeiros, 2014; Santos, & Corrêa, 2015; Sousa, 2015; Simões, Machado, Souza, & Lopes, 2017). Lopes (2002) recorda, igualmente, que o processo de pesquisa nesses repositórios é simples e que os termos a usar, nesse caso, são extraídos diretamente dos recursos que os irão constituir (cf. no ponto 1, Lancaster, 2004; Borges, 2008 cit. por Sousa, 2015). Também alega que o uso da linguagem natural favorece o encontro dos temas de carácter específico, porque o seu uso reduz os conflitos de comunicação entre indexadores/as e utilizadores/as, dado que, nesse caso, ambos/as utilizam os mesmos termos (cf. no ponto 1, Fujita, & Gil-Leiva, 2014; Fujita, & Tartarotti, 2020). No entanto, não deixa de alertar para os possíveis problemas que o uso dessas linguagens acarreta, entre os quais: sinonímia, homonímia, ausência de padronização dos assuntos e perda de confiança dos/as utilizadores/as perante uma eventual resposta negativa do sistema (cf. no ponto 1, Medeiros, 2014; Simões, Machado, Souza, & Lopes, 2017).

A este respeito, Santos e Neves (2018) referem que a utilização das palavras-chave fornecidas pelos/as autores/as na indexação dos recursos, no ato de depósito, dada a carga de subjetividade inerente, poderá trazer complicações à recuperação da informação. Tal acontece, a seu parecer, porque ao selecionarem livremente as palavras que consideram que melhor representam os seus trabalhos, os/as autores/as nem sempre estarão a pensar nos prováveis interesses ou necessidades informacionais da comunidade ou grupo de pessoas que também irão utilizar o repositório e que poderão procurar pelo seu texto (cf. no ponto 1, (Fujita, & Gil-Leiva, 2014). Salientam, por isso, que o controlo do vocabulário poderá ser um recurso pertinente, na medida em que conseguirá mitigar as dispersões lexicais presentes na linguagem natural e auxiliar na escolha dos termos representativos dos conteúdos, atuando como um mediador na comunicação entre utilizadores/as e sistemas. Noruzi (2007), num alinhamento discursivo convergente com Santos e Neves (2018), com Lopes (2002), com Simões, Machado, Souza e Lopes (2017 e com Medeiros (2014), menciona algumas das implicações menos positivas, que podem advir da utilização da linguagem natural no processo de representação da informação nesses ambientes digitais, entre as quais destacamos: uso incorreto do singular ou plural; emprego de polissemia ou de sinonímia; desequilíbrio no nível de profundidade/especificidade das palavras ou expressões usadas para representar os conteúdos dos documentos. Efetivamente, concordamos com o facto de que todos esses problemas podem contribuir para reduzir a precisão com que se representa um recurso num sistema, dificultando a sua recuperação pelos/as utilizadores/as (cf. ponto 1, Fujita, & Gil-Leiva, 2014), sendo, pois, necessário lançar mão de uma política de indexação clara, firme e definidora das linhas de condução a usar na representação dos assuntos, no ato de “povoamento” desses repositórios (cf. no ponto 1, Souza, 2017).

Fujita e Tartarotti (2020, p. 335), no entanto, ponderam que “[os] produtos da representação documentária mais visíveis e utilizados para a recuperação da informação em bases de dados digitais, são o título, resumo e palavras-chave”. Assim, num alinhamento discursivo consistente com a importância que as palavras-chave assumem na representação da informação, tendo em conta o grau de expertise dos/as utilizadores/as e, ainda, o modo como os documentos são depositados nesses repositórios [referindo-se, naturalmente, ao auto-depósito], Fujita e Tartarotti (2020) discutem as vantagens de ser o/a próprio/a autor/a a representar os conteúdos do texto que produziu, justificando com o facto de esta pessoa – um/a especialista na matéria – ser quem tem o melhor domínio sobre o conteúdo do próprio texto e, portanto, do vocabulário a ser utilizado na sua representação. Ainda, ponderam sobre o facto de um/a autor/a, nestas condições, ter a devida consciência do seu papel – enquanto indexador/a – e refletir sobre a necessidade de adotar padrões que se sejam compatíveis com os possíveis e diferentes empregos dos termos que está a usar, no futuro.

Também, Miguéis e Neves (2013) acrescentam que o uso das palavras-chave apresenta vantagens, tais como potenciar o acesso ao conteúdo por outros pontos de acesso, que não apenas aqueles que se encontram expressos no título e/ou no resumo; traduzir as ideias dos/as autores/as; manter uma ligação com a realidade quotidiana, acompanhando a evolução científica dos recursos. Segundo analisam (Miguéis, & Neves, 2013), a pesquisa de artigos científicos nesses repositórios faz-

-se a partir dos metadados que incluem as palavras-chave e também dos conteúdos do texto integral, aplicando-se técnicas que representam e criam relações semânticas entre os recursos armazenados. Dizem-nos, também, que reconhecer a importância das palavras-chave para a representação e a recuperação da informação é uma forma de incentivar o auto-depósito e de estimular o preenchimento dos respectivos metadados. Direta ou indiretamente, a adoção destas medidas irá contribuir para uma maior visibilidade e probabilidade de aumento das citações e, conseqüentemente, do impacto da produção científica dos/as autores/as depositantes, bem como da própria instituição.

Numa outra linha discursiva, identificamos estudos que problematizam o uso dos modelos colaborativos de indexação social em repositórios institucionais, na forma livre ou, em alternativa, na forma controlada (cf. no ponto 1, Medeiros, 2014). No primeiro caso, conforme diz Santos (2017), respeita-se a linguagem adotada pelos/as utilizadores/as ao indexarem os documentos e, no segundo caso, restringe-se essa liberdade. Deste modo, o que está em causa é com que grau de liberdade poderá o/a utilizador/a – e autor/a – indexar os documentos que deposita e o quão consistente e pertinente será o resultado desta interação com o sistema, sendo esta escolha livre ou moderada (cf. no ponto 1, Medeiros, 2014; Santos, & Corrêa, 2015).

Sobre esse assunto, Santarém Segundo e Vidotti (2011) evidenciam a importância do uso das *folksonomias*, na modalidade assistida (ou seja, moderada), no contexto dos repositórios institucionais e destacam o seu potencial para a construção de inteligência coletiva e para o inter-relacionamento dos termos de indexação usados num sistema (o que, a seu parecer, não ocorre na forma livre, dado que os termos não se relacionam ou não possuem a precisão ou a consistência necessárias, enquanto vocabulário socialmente construído). Este tipo de “*folksonomia* assistida”¹⁰ pode, assim, melhorar a qualidade de descrição do documento, preservando, simultaneamente, a criatividade [liberdade] do/a utilizador/a na escolha das *tags*, fornecendo os elementos que irão permitir relacionar o recurso que está a depositar a outros depositados anteriormente, e/ou a uma estrutura de representação do conhecimento embebida no próprio sistema.

Fujita e Tartarotti (2020), também a este respeito, referem um estudo realizado por Kun e Kipp (2014), para averiguar a eficácia do uso de *tags* colaborativas e de palavras-chave de autor na recuperação de informação. Os resultados obtidos demonstraram que o uso de palavras-chave de autor melhorava a recuperação da informação e que, pelas suas diferentes valências, os sistemas deveriam incorporar ambas as formas na indexação dos recursos depositados.

Ainda, uma outra linha discursiva referenciada nos textos analisados, alude aos metadados comumente usados nos repositórios. Por esta via seguem Bacha e Almeida (2013), que relatam uma experiência realizada no repositório institucional da Fundação Getúlio Vargas (FGV), onde foram usadas, em simultâneo, a indexação por palavras-chave e a indexação por texto livre. Salientam Bacha e Almeida (2013) que o enriquecimento dos metadados descritivos permite uma maior recuperação da

¹⁰ Que, segundo Santarém Segundo e Vidotti (2011), surge da necessidade de auxiliar o/a utilizador/a a descrever a *tag* ou assunto do documento que está a ser depositado, em relação ao domínio do conhecimento a que pertence.

informação, devendo por isso usar-se ambos os processos. Ainda, recordam (Bacha e Almeida, 2013) que o DublinCore¹¹ é um esquema simples, que não permite a hierarquização de assuntos ou o estabelecimento de relações entre os termos; assim sendo, o uso da linguagem natural veio, neste caso implementado, para colmatar esta lacuna. Segundo Bacha e Almeida (2013), no repositório em causa, a combinação das duas linguagens afigurou-se útil para resolver duas questões: i) atender às reivindicações de autores/as, que sentiam a necessidade de reconhecer os termos utilizados no auto-depósito; ii) melhorar a eficácia do sistema. Esta experiência, a seu parecer, veio demonstrar que “a linguagem natural e a controlada podem e devem caminhar juntas, uma completando a outra” (Bacha, & Almeida, 2013, p. 7).

Finalmente e concordando com as alegações anteriores, Miguéis e Neves (2013) acrescentam que os metadados mais relevantes num repositório serão aqueles provenientes dos/as autores/as depositantes dos recursos, dado que espelham a informação precisa dos recursos visados. Ainda, sublinham Miguéis e Neves (2013) que as aplicações atuais conseguem explorar os dados provenientes dos sistemas de gestão dos repositórios, podendo-se, deste modo, efetuar uma análise criteriosa dos dados obtidos e escolher as soluções apontadas como mais profícuas. Também, com o mesmo propósito, Santos e Corrêa (2015) referem a importância de utilizar os metadados produzidos pelos/as utilizadores/as, podendo estes ser alvo de moderação, para aumentar a qualidade da descrição dos recursos.

Conclusão

Iniciamos este estudo com o objetivo principal de revisitar a literatura científica recuperada sobre a temática da indexação, de um modo geral, e da indexação em repositórios institucionais, em particular. Para o efeito, foram recuperados 31 textos publicados entre 1989 e 2020, verificando-se que uma esmagadora maioria, composta por 21 (74%) itens, foi publicada recentemente (2006-2015) ou muito recentemente (2016-2020).

Por esses textos, constatamos, sem surpresa, que os repositórios institucionais com um cariz científico são uma realidade, desde há vários anos, conferindo uma forma eficaz de arquivo, gestão, disseminação e visibilidade da informação em formato digital, além do que conferem prestígio e agregam valor à comunidade académica e científica na qual se inserem. Constatamos, também, que a tarefa de indexação continua sendo defendida como algo indispensável à gestão e ao tratamento temático da informação. Trata-se de uma área interdisciplinar, que recebe os contributos de diferentes áreas do saber, para além da Ciência da Informação (Linguística, Informática, Ciências Computacionais, Programação, Estatística, Gestão de Dados, entre outras).

¹¹ O DSpace é, reconhecidamente, o *software* mais usado internacionalmente no âmbito dos repositórios. Rubin (2016 citado por Santos, & Neves, 2018) salienta que o esquema Dublin Core, no qual se baseiam os campos descritivos deste *software*, possibilita que utilizadores/as inexperientes descrevam os seus recursos de forma autónoma e sem grandes dificuldades, o que facilita o alinhamento desse sistema com o uso, por exemplo, da linguagem natural.

Pelos textos, também foi possível descortinar alguns dos benefícios ou constrangimentos apontados pelos/as autores/as, no que toca aos diferentes tipos de indexação (manual, automática, semi-automática). O aumento da utilização da informação em meio digital tem levado a um grande desenvolvimento da indexação automática e dos *softwares* e aplicações que a suportam, mas este processo ainda comporta imprecisões. Simultaneamente, constatamos o interesse pelos modelos colaborativos de indexação ou indexação social, que poderão ser menos eficazes, mas, certamente, serão mais interativos, podendo ser realizados com ou sem assistência ou moderação profissional. Estes modelos surgem como alternativa e complemento à indexação controlada, mais eficaz, mas nem sempre mais eficiente. Os discursos alinham-se no sentido da sua utilização, de forma assistida, nos repositórios, alegando que podem contribuir para a melhoria na recuperação da informação e servir como ponto de partida para a construção de vocabulários controlados.

Relativamente ao uso destes tipos de vocabulários, livres ou controlados, verificamos que os benefícios de uns podem provavelmente ajudar a colmatar as falhas sentidas noutros, defendendo-se que as duas abordagens podem e devem mesmo caminhar juntas, retirando-se daí mais proveitos.

Quanto ao uso dos metadados, salientamos o seu papel indispensável na descrição dos recursos em repositórios, bem como a relativa simplicidade que implica o seu preenchimento, nos sistemas, sendo este um fator que veio para favorecer os modelos de gestão que funcionam recorrendo ao auto-depósito.

Outro aspeto a destacar é a necessidade da criação de políticas de indexação nas instituições que sejam norteadoras das decisões a tomar, orientando gestores/as e utilizadores/as dos repositórios para uma indexação consistente e uma recuperação da informação de qualidade que satisfaça as suas necessidades.

A Indexação continua a ser um tema atual e pertinente no meio digital e carece de mais estudos, muito em particular no que respeita aos resultados concretos do uso dos diversos tipos de linguagens ou sistemas de organização do conhecimento, bem como dos seus efeitos práticos na precisão no ato de recuperação da informação.

Por fim, serão precisas investigações futuras que alarguem o espectro da atual amostra, de reduzida dimensão, mas que serve a um fim qualitativo de evidenciar a problemática em torno da indexação nos repositórios científicos, nos últimos 10 anos. Convirá, igualmente, alargar as discussões para outros territórios – os casos concretos de uso – que este estudo, certamente, não terá conseguido alcançar, sendo esta, efetivamente, a sua principal limitação.

Referências bibliográficas

- Bacha, M., & Almeida, M. (2013). Vocabulário controlado e palavras-chave em repositórios digitais: relato de experiência do repositório institucional da FGV. In *Portal do Conhecimento FEBAB: Anais do XXV Congresso Brasileiro de Biblioteconomia, Documento e Ciência da Informação, Florianópolis, SC, Brasil, 07-10 de julho de 2013*. <https://portal.febab.org.br/anais/article/view/1361/1362>
- Bandim, M. & Correa, R. (2019). Indexação automática por atribuição de artigos científicos em português da área de Ciência da Informação. *Transinformação*, 31. <https://repositorio.ufpe.br/handle/123456789/25272>

- Crow, R. (2002). *The case for institutional repositories: a SPARC position paper*. Scholarly Publishing & Academic Resources Coalition. https://ils.unc.edu/courses/2014_fall/inls690_109/Readings/Crow2002-CaseforInstitutionalRepositoriesSPARCPaper.pdf
- Freitas, M. & Leite, F. (2018). Proposição de diretrizes para o depósito da produção científica em repositórios institucionais de acesso aberto baseada na visão de diferentes atores do sistema de comunicação científica. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, 23(53), 96-109. <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2018v23n53p96/37380>
- Fujita, M., & Gil-Leiva, I. (2014). Avaliação da indexação por meio da recuperação da informação. *Ciência da Informação*, 41(1). <http://revista.ibict.br/ciinf/article/view/1418>
- Fujita, M., & Tartarotti, M. (2020). Análise de palavras-chave da produção científica de pesquisadores: o autor como indexador. *Informação e Informação*, 25(3), 332-374. <http://www.uel.br/revistas/uel/index.php/informacao/article/view/41866/pdf>
- Fujita, M., & Troitino, S. (2018). Política de indexação em arquivos de instituições de saúde. *Informação em Pauta*, 3, 95-116. <https://doi.org/10.32810/2525-3468.ip.v3iEspecial.2018.39719.95-116>
- Hockx-yu, H. (2006). Digital preservation in the context of institutional repositories. *E-prints in Library & Information Science*, 40(3). <http://eprints.rclis.org/8189/>
- Kuramoto, H. (2009). Repositórios institucionais: políticas e mandatos. In L. Sayão, L. B. Toutain, F. G. Rosa, & C. H. Marcondes. *Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação*. EDUFBA. https://repositorio.ufba.br/ri/bitstream/ufba/473/3/implantacao_repositorio_web.pdf
- Lancaster, F. (2004). *Indexação e resumos: teoria e prática*. Briquet de Lemos. <https://bibliotextos.files.wordpress.com/2014/07/livro-indexac3a7c3a3o-e-resumos-teoria-e-prc3a1tica-lancaster.pdf>
- Lopes, I. (2002). Uso das linguagens controlada e natural em bases de dados: revisão da literatura. *Ciência da Informação*, 31(1), 41-52. <http://revista.ibict.br/ciinf/article/view/976>
- Lynch, C. (2003). Institutional repositories: essential infrastructure for scholarship in the digital age. *ARL 13: a bimonthly report*, 226, 1-7. <https://www.cni.org/wp-content/uploads/2003/02/arl-br-226-Lynch-IRs-2003.pdf>
- Martins, E. (2013). O repositório: imagem de marca e objeto de aprendizagem em meio digital. In *III Conferência do IPCB sobre o livre acesso ao conhecimento científico – O desafio da publicação em meio científico: como, onde, porquê? Livro de resumos*, Instituto Politécnico de Castelo Branco. https://repositorio.ipcb.pt/bitstream/10400.11/1739/1/Apresenta%c3%a7%c3%a3o_RCIPCB_1%5b1%5d.pdf
- Medeiros, F. (2014). *A historiografia medieval portuguesa na viragem do milénio: análise bibliométrica (2000-2010) e representação taxonómica*. (Tese de doutoramento). Universidade de Évora. http://dspace.uevora.pt/rdpc/bitstream/10174/11220/1/VOLUME_1.pdf
- Mendes, M. & Simões, G. (2001). *Indexação por assuntos: princípios gerais e normas*. Gabinete de Estudos a&b. <http://ojs.letras.up.pt/index.php/paginasaeb/article/view/3453/3162>
- Miguéis, A. (2012). *Atitudes e percepções dos autores depositantes do Repositório Científico da Universidade de Coimbra*. (Dissertação de Mestrado). Universidade de Coimbra. <https://estudogeral.uc.pt/handle/10316/21116>
- Miguéis, A., & Neves, B. (2013). Uma abordagem à linguagem de indexação dos artigos científicos depositados no Repositório Científico da Universidade de Coimbra. *Pontode-Acesso*, 7(1), 116-131. <https://estudogeral.sib.uc.pt/bitstream/10316/23450/1/Uma%20abordagem%20c3%a0%20linguagem%20de%20indexa%c3%a7%c3%a3o%20dos%20artigos%20cient%c3%adficos.pdf>

- Noruzi, A. (2007). Folksonomies: Why do we need controlled vocabulary? *Webology*, 4(2). <http://www.webology.org/2007/v4n2/editorial12.html>
- NP 3715 (1989). *Documentação. Método para a análise de documentos, determinação do seu conteúdo e seleção de termos de indexação*. Instituto Português da Qualidade. <http://ancacid.yolasite.com/resources/NP3715.pdf>
- Oliveira, L. (2017). Política de indexação: concepções acerca do conceito e percepções em torno de sua elaboração. *Ciência da Informação*, 4(2), 39-58. <https://brapci.inf.br/index.php/res/v/36466>
- Santarém Segundo, J. & Vidotti, S. (2011). Representação iterativa e folksonomia assistida para repositórios digitais. *Liinc em Revista*, 7(1), 283-300. <https://repositorio.unesp.br/handle/11449/114930>
- Santos, R. F. dos (2017). Indexação em repositórios digitais: uma abordagem sobre o metadado assunto da Biblioteca Digital de Monografias da UFRN. *Revista Informação na Sociedade Contemporânea*, 1, 1-22. <https://periodicos.ufrn.br/informacao/article/view/12279>
- Santos, R. e Corrêa, R (2015). Modelos colaborativos de indexação social e sua aplicabilidade em bibliotecas digitais. *Liinc em Revista*, 11(1). <http://dx.doi.org/10.18617/liinc.v11i1.768>
- Santos, R. & Neves, D. (2018). Práticas de indexação em repositórios digitais de acesso aberto: análise do metadado assunto do repositório institucional da UFRN. *In Encontro Nacional de Pesquisa em Ciência da Informação, XIX ENANCIB*. <https://brapci.inf.br/index.php/res/v/102563>
- Silva, S. (2020). *Sistemas de indexação automática por atribuição: uma análise comparativa* (Dissertação de Mestrado). Universidade Federal de Pernambuco. <https://repositorio.ufpe.br/handle/123456789/37626>
- Simões, M. G., Machado, L., Souza, R. R., & Lopes, A. T. (2017). Indexação automática e ontologias: identificação dos contributos convergentes na ciência da informação. *Ciência da Informação*, 46(1), 153-169. <http://revista.ibict.br/ciinf/article/view/4020/3720>
- Souza, I. O. (2017). *Proposta de implantação de uma política de indexação para a BEAGRI* (Dissertação de mestrado). Universidade Federal do Ceará. <http://www.repositorio.ufc.br/handle/riufc/32174>
- Sousa, J. A. G. de (2015). *A prática da indexação automática no DSpace pelas bibliotecas digitais e repositórios institucionais de Brasília*. (Monografia). Universidade de Brasília. <https://bdm.unb.br/handle/10483/13020>
- Viana, C., Márdero Arellano, M., & Shintaku, M. (2005). *Repositórios institucionais em ciência e tecnologia: uma experiência de customização do DSpace*. IBICT. <http://eprints.rclis.org/7168/1/viana358.pdf>
- Vignoli, R., Almeida, P., & Catarino, M. (2014). Folksonomias como ferramenta da organização e representação da informação. *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, 12(2), 120-135. <https://doi.org/10.20396/rdbci.v12i2.1606>
- Wheatley, P. (2004). *Institutional repositories in the context of digital preservation*. Digital Preservation Coalition. <https://www.dpconline.org/docs/technology-watch-reports/90-institutional-repositories-in-the-context-of-digital-preservation/file>

(Página deixada propositadamente em branco)

CIÊNCIA DA INFORMAÇÃO E INTELIGÊNCIA ARTIFICIAL: UM CAMINHO PARA ARQUIVOS E BIBLIOTECAS INTELIGENTES

Moisés Rockembach¹

¹Universidade Federal do Rio Grande do Sul, moises.rockembach@ufrgs.br,
ORCID iD 0000-0001-9057-0602

Resumo

Este trabalho traz uma investigação sobre as proximidades existentes entre a Ciência da Informação e o campo da Inteligência Artificial, com o objetivo de traçar possibilidades para a construção de arquivos e bibliotecas inteligentes. Para isto, procurou-se verificar na literatura existente, que converge a Ciência da Informação e a Inteligência Artificial, alguns dos principais pressupostos que apontem para modelos teóricos e aplicados no campo dos Arquivos e Bibliotecas. Utilizou-se como metodologia de investigação a pesquisa bibliográfica e documental, com uma abordagem exploratória, procurando evidenciar questões relevantes nos aspectos científicos e profissionais da Ciência da Informação. A partir dos resultados da investigação, observa-se que os métodos da inteligência artificial, sobretudo o tratamento dos dados, a implementação de uma cultura de dados e a construção de algoritmos e aprendizagem de máquina para os problemas dos arquivos e bibliotecas, podem levar a uma revolução em como percebemos estes espaços informacionais.

Palavras-chave: *Inteligência Artificial, Ciência da Informação, Organização do Conhecimento, Arquivos inteligentes, Bibliotecas inteligentes.*

Introdução

Estamos no decorrer de uma nova revolução tecnológica, que pode passar despercebida em uma visão geral ou superficial, mas aos poucos vai se instalando no âmago da Sociedade. Esta revolução, por ser tecnológica e digital, não nos causa estranhamentos, dado o hábito que temos da utilização dos computadores nas últimas décadas, a expansão do computador pessoal nos anos 1980, a concepção e globalização da world wide web desde os anos 1990 e a consequência da própria Sociedade em Rede e do Paradigma Tecnológico que nos auxilia a compreender questões atuais (Castells, 1999), tudo potencializado com o uso das redes sociais e a internet móvel nos anos 2000. Estes últimos fenômenos citados cresceram nos anos 2010, com uma infinidade de novas redes sociais e a melhoria da velocidade de conexão, saltando em sua capacidade até o atual 5G.

Entretanto, o crescimento dos grandes conjuntos de dados (Big Data) aliado às possibilidades de aplicação de análises estatísticas e de ciência de dados, traz consigo novas questões e o crescimento de uma área que não é propriamente nova, pois tem praticamente a idade da Ciência da Informação e compartilha com ela a procura pelo interesse em suas inúmeras aplicações. Esta área, ou campo de estudo, chamada Inteligência Artificial, leva nos a pensar no termo inteligência, usualmente compreendida como uma capacidade humana, e o termo adicional artificial, como uma capacidade computacional, ambas procurando realizar o mesmo objetivo, de possuir habilidades de aprender, planejar e solucionar problemas de forma autônoma. O problema consiste em compreender como a Ciência da Informação e a Inteligência Artificial, além do campo teórico e empírico nos Arquivos e Bibliotecas, podem interagir na concepção de novas aplicações.

Nosso escopo abrange uma investigação sobre Inteligência Artificial, Ciência da Informação, Arquivos e Bibliotecas Digitais, de modo a perceber algumas das principais investigações neste cruzamento de temas e pontos de convergência que possam levar a desdobramentos de pesquisa. Neste sentido, realizamos um trabalho com a aplicação de metodologia de pesquisa bibliográfica e documental, analisando a aproximação da Ciência da Informação e a Inteligência Artificial no contexto de Arquivos e Bibliotecas. De forma geral, percebeu-se que uma revolução tecnológica neste sentido, no campo da Ciência da Informação, Arquivos e Bibliotecas, tem a possibilidade de gerar diversos tipos de impactos, desde questões teóricas, aplicadas, de investigação científica e de práticas profissionais.

Metodologia

A metodologia adotada incluiu pesquisa bibliográfica e documental, relacionando publicações que abordam a Inteligência Artificial, Ciência da Informação, Arquivologia, Biblioteconomia, Organização do Conhecimento e pesquisas que usam o contexto dos Arquivos e Bibliotecas. Procedeu-se, deste modo, a buscas de referências em bases da WoS (Web of Science), Scopus e BrapCI (Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação), com o uso de termos em inglês e português nos temas listados (Inteligência Artificial, Ciência da Informação, Arquivologia, Biblioteconomia, Organização do Conhecimento, Arquivos, Bibliotecas).

A busca por fontes de informação relacionadas ao tema também procurou contextualizar a revolução tecnológica atual, ultrapassando o uso do computador como ferramenta, para uma abordagem de interação entre agentes humanos e não-humanos e uma continuidade da Quarta Revolução, que foi iniciada por Alan Turing, e definida por Floridi (2014).

De forma a indicar o problema de pesquisa, este consistiu na compreensão de como a Ciência da Informação e a Inteligência Artificial, além dos Arquivos e Bibliotecas como campo teórico e empírico, podem interagir na concepção de novas aplicações em IA. Para isso, foi necessário observar questões teóricas e projetos que vem sendo desenvolvidos, na delimitação proposta por esta pesquisa. Como um estudo exploratório do tema, e por ser uma temática relativamente nova no campo

dos Arquivos e Bibliotecas, a metodologia se demonstrou adequada a resolução do problema.

O estudo possui limitações, por não arrolar todos projetos e pesquisas acadêmicas relacionados a IA nos Arquivos e Bibliotecas e, sem querer esgotar o tema, levanta questões pertinentes debatidas a partir da pesquisa bibliográfica e documental. A relevância desta pesquisa demonstra-se na medida da adoção cada vez maior de projetos de IA em todos os campos do conhecimento. Desta forma, na próxima secção apresentaremos os resultados da investigação.

Convergências entre inteligência artificial e ciência da informação para arquivos e bibliotecas inteligentes

Floridi (2014) exemplifica a Quarta Revolução como o momento em que vivemos e em que aprofundaremos nos próximos anos, a partir de uma infosfera que produz a interação entre agentes humanos e não-humanos e em um contexto em que o ser humano não se encontra no centro da inteligência. Esta revolução sucede a transformação das três revoluções anteriores, a Primeira Revolução com Nicolau Copérnico e o Heliocentrismo, onde o Homem não estava mais no centro do Universo, a Segunda Revolução com Charles Darwin e a Teoria da Evolução, onde o Homem não se encontra mais no centro do Reino Biológico e a Terceira Revolução, com Sigmund Freud e a Psicanálise, onde o Homem não está mais no centro de sua própria Mente.

A Quarta Revolução tem início com o matemático Alan Turing, conhecido como o pai da Computação e da Inteligência Artificial. De uma forma geral, estas revoluções são de difícil compreensão no momento em que ocorrem, sem o devido distanciamento do tempo. Entretanto, os indícios que aparecem são de uma crescente adoção de aplicações de Inteligência Artificial nos mais diversos setores da Sociedade, o que certamente produzirá impactos permanentes no nosso dia a dia.

A Inteligência Artificial ou IA, de maneira geral, está presente em nosso inconsciente de muitas formas, positivas e negativas, essencialmente nas muitas representações em filmes e séries, algumas com visões positivas, outras com visões negativas sobre o impacto da Inteligência Artificial na Sociedade. A lista de obras que abordaram a Inteligência Artificial, com possíveis implicações éticas nas suas interações com humanos, é infidável, muitos exemplos podem ser citados, desde o Hall 9000, de 2001: a space odyssey (1968) e os Replicants de Blade Runner (1982), a exemplos mais recentes, como The Matrix (1999), I, Robot (2004), derivado do livro de Isaac Asimov (Asimov, 1950), Her (2013), Transcendence (2014), Ex Machina (2014) e a série inglesa Black Mirror (2011-). Podemos mencionar também como o tema Inteligência Artificial vem aparecendo de forma constante nos principais jornais e notícias da imprensa, com destaque para as inovações que vem promovendo em muitos campos do conhecimento.

Este imaginário alimenta as ideias da Sociedade em torno do tema Inteligência Artificial, o que, se por um lado, pode estimular a busca e adoção de soluções criativas a problemas reais, por vezes pode ser contraproducente quando gera equívocos e falta de compreensão dos usos da IA, se pensarmos em projetos

aplicados de Inteligência Artificial. Long e Magerko (2020) falam da possibilidade dos estudos em literacia em Inteligência Artificial. Desta forma, um dos pontos que destacamos aqui é a necessidade do desenvolvimento de uma literacia em Inteligência Artificial, com o objetivo de adquirir competências necessárias a colaboração em projetos interdisciplinares de IA, por sujeitos que não possuem uma formação específica nesta área, especialmente nas Humanidades, onde encontra-se nosso contexto de pesquisa.

Como campo de pesquisa, a Inteligência Artificial não se trata de uma área nova, pois seu conceito remonta 1956, na conferência Dartmouth Summer Research Project on Artificial Intelligence, realizada na Universidade de Dartmouth (Moor, 2006), com a participação de diversos pesquisadores dos campos das exatas, incluindo John McCarthy, que cunhou o termo Inteligência Artificial, e Claude Shannon, que propôs a conferência e é conhecido no campo da Ciência da Informação pela sua contribuição, juntamente com Warren Weaver, pela Teoria Matemática da Informação e Comunicação (Shannon, 1948). A título de exemplificação, mesmo na discussão brasileira em Ciência da Informação, a busca de referências na Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BrapCI), retornou 100 resultados com o termo “Inteligência Artificial”, com a primeira publicação de 1985 e com um crescimento de produção nos últimos anos.

As chamadas “máquinas pensantes” já faziam parte do escopo de diversas áreas de estudo, nomeadamente da cibernética, onde o trabalho de Norbert Wiener (2019) destaca-se, procurando compreender como os sistemas podem se autorregular e tomar ações com base em seu ambiente. Entretanto, é com o surgimento do termo Inteligência Artificial em que é dado o enfoque na reprodução do comportamento de inteligência humana.

O próprio conceito de Inteligência Artificial não é unânime e pode-se encontrar distinções entre uma inteligência artificial fraca, como são considerados os assistentes Alexa e Siri ou forte, algo mais próximo a inteligência humana. A Inteligência Artificial utilizada nos carros autônomos (Tesla, Waymo/Google, Zoox/Amazon), no jornalismo autônomo (Graefe, 2016), no reconhecimento de imagens e visão computacional (Wechsler, 2016) também são exemplos das atuais aplicações de sistemas inteligentes e do caminho a ser percorrido até atingir uma Inteligência Artificial forte.

A discussão pública tem se tornado cada mais séria e preocupante nos últimos anos, quando diversos cientistas, como Stephen Hawking (Cellan-Jones, 2014), alertaram para a posição que a Inteligência Artificial terá na relação com os seres humanos. Já na relação com a gestão do conhecimento, como aponta Rhem (2021) é preciso garantir que a base de conhecimento não tenha preconceitos e que as decisões tomadas por uma IA, com base nestes conhecimentos, sejam éticas. Ainda, questões relativas a substituição de profissionais por Inteligência Artificial (Webb, 2019) são problemas éticos e sociais que precisamos debater. Isto adiciona desafios ao atual cenário, em que já discutimos as implicações éticas sobre a coleta e uso do Big Data, as legislações de proteção de dados e as suas limitações (Ferreira, Rockembach, Krebs, 2017; Rockembach, Silva, 2018; Rockembach, Silva, 2021).

A Inteligência Artificial também pode ser exemplificada a partir de cinco grandes ideias (Long, Magerko, 2020): os computadores podem aprender com os dados;

perceber o mundo usando sensores; agentes mantêm modelos do mundo e os usam para raciocinar, fazer os agentes interagirem com humanos é um desafio substancial para os desenvolvedores de IA e as aplicações de IA podem impactar a sociedade de formas positivas e negativas. Estas ideias podem vir a ser alguns dos fundamentos que norteiem futuras aplicações de IA em Arquivos e Bibliotecas. Por sua vez, as questões da organização do conhecimento, das classificações, taxonomias e ontologias, encontram-se diretamente relacionadas as aplicações de aprendizagem de máquina, ou *Machine Learning*, com o uso de modelos que podem aprender com os dados, sendo que esta aprendizagem pode acontecer de forma supervisionada, semi-supervisionada e não supervisionada.

Para a construção de Arquivos e Bibliotecas Inteligentes torna-se necessário formar uma base a partir da Inteligência Artificial e de Sistemas Especializados Inteligentes. De forma geral, muitos termos novos se juntam no guarda-chuva da Inteligência Artificial e podem ser observados nos documentos analisados, como sistemas especialistas, mecanismos de regras ou sistemas baseados em regras, aprendizado de máquina, aprendizado profundo, redes neurais (Rolan, et al. 2019), mas também processamento de linguagem natural (NLP) e visão computacional. A aplicação de modelos estatísticos e processamento computacional, por exemplo, auxiliam no estabelecimento de formas de aprendizagem de máquina, que podem ser aplicadas alimentando o modelo com textos, imagens e diversos tipos de dados. A qualidade da informação ou IQ (Information Quality) pode ser compreendida por questões intrínsecas (precisão, objetividade, confiabilidade), de acessibilidade (acesso, segurança), contextuais (relevância, valor agregado, oportunidade, integridade, quantidade de dados) e representacionais (interpretabilidade, facilidade de compreensão, representação concisa e consistente) (Floridi, 2014).

Os conceitos de dataficação, competência em dados e de data-driven, ou de ações guiadas por dados, é algo presente em artigos da Ciência da Informação, da Arquivologia e Biblioteconomia (Song, Zhu, 2017; Tartarotti, Dal'Evedove, Fujita, 2019; Melo, Rockembach, 2019; Costal, Sales, Zattar, 2020) e vem desempenhando um papel de formação da cultura de dados em diversas organizações e todos os benefícios que a tomada de decisão guiada por dados pode trazer para os usuários dos sistemas informacionais.

Uma das abordagens metodológicas é o uso de aprendizagem de máquina em avaliação de documentos a partir de grande conjunto de e-mails (Rolan, et al. 2019), assim como a adoção de revisão assistida por tecnologia (Technology Assisted Review – TAR) e a função de classificação, a partir de estudo de caso sobre e-mails e análise de sentimentos (Liu, Lee, 2018). A Arquivologia Computacional (Computational Archival Science), abrangendo análises de materiais arquivísticos, mineração de dados, dados, análise de sentimento, análise de rede, escalabilidade dos serviços de arquivo, curadoria e preservação digital, incluindo a web, mídias sociais e blockchain (Marciano, et al., 2018; Payne, 2018) são algumas das possibilidades que se aproximam das aplicações de Inteligência Artificial, assim como o InterPARES Trust AI (www.interparestrustai.org) no desenvolvimento de projetos de IA que apoiem a disponibilidade de documentos públicos confiáveis.

Algumas considerações precisam ser realizadas na adoção de Inteligência Artificial em setores públicos, como colocado pelo Office for Artificial Intelligence

(2020), do Reino Unido, especialmente em relação a ética da informação e uso dos dados e referente a segurança da informação, compreendendo a qualidade dos dados; o uso de modelos justos, treinados e implementados tendo cuidados sobre viés e preconceitos; responsabilidade; privacidade e transparência.

Outra abordagem que consideramos relevante é o uso de deduplicação para a avaliação, procurando reduzir dados duplicados nos espaços de armazenamento (Shabou, Tièche, Gaudinat, 2019) e uso de aplicações como o DROID para a identificação de formatos de arquivos (The National Archives UK, 2021). Uma outra aplicação é o uso de modelos preditivos também é uma prática observada por Vickery (2017) em programas de aquisição de bibliotecas orientados por demanda. Além disto, as bibliotecas não se restringem as coleções, mas podem ter um papel fundamental na orientação de aquisição de ferramentas de inteligência artificial que colaborem com a mineração de dados e na função essencial de acesso ao conhecimento (Cox, Pinfield, Rutter, 2019).

Conclusões

É possível perceber um caminho sem volta na implementação e crescimento na adoção de soluções de Inteligência Artificial em diversos setores e a Ciência da Informação e os Arquivos e as Bibliotecas não estão fora deste escopo. Na própria concepção destas visões é possível verificar que há, ao menos, dois caminhos possíveis na relação entre as inteligências humana e artificial, de competição ou de colaboração. A automação e a substituição de mão de obra, seja por equivalência de capacidades ou por superação das capacidades humanas, por custo ou por melhor eficiência e eficácia na execução das atividades laborais leva a um contexto de competição e pode ser percebido como um dos pontos negativos, já que pode gerar consequente aumento nos índices de desemprego. Outras consequências incluem a necessidade dos profissionais em adquirir novas habilidades para o mercado de trabalho, tendo em conta um novo contexto tecnológico. Um ambiente de colaboração, por sua vez, estimularia a adoção de inteligência artificial como assistente no planejamento e execução de atividades profissionais, objetivando sempre melhorar a condição humana em todos os sentidos e em todas as pontas de uma cadeia de valor. Também melhoraria processos repetitivos, onde a automação tem sua função e auxilia na diminuição de erros provenientes de falha humana.

Entretanto, estas relações de competição e colaboração com a Inteligência Artificial, infelizmente, não são tão delimitadas, não trata-se somente de uma questão de escolha, mas do contexto social, econômico, geográfico e político no qual estão inseridos, pois podem reproduzir algumas das relações de competição e colaboração que já existem em relações de poder.

Por estar presente no imaginário popular, uma das questões que podem ser apontadas sobre a Inteligência Artificial e sua correlação com outros campos de atuação, nomeadamente, neste caso, a Ciência da Informação, os Arquivos e as Bibliotecas, diz respeito a uma aproximação entre as possibilidades atuais de projetos de IA e a observação dos problemas e lacunas do ambiente profissional e científico que podem ser beneficiados com estes projetos.

Neste sentido, assim como a literacia em informação nos traz condições de identificar, encontrar, avaliar, aplicar e reconhecer a informação e suas fontes, uma literacia em Inteligência Artificial pode nos auxiliar a adquirir competências necessárias para construir e colaborar e projetos de IA, de forma interdisciplinar.

Por isso, acompanhar de perto as investigações teóricas e aplicadas, a implementação de uma cultura de dados, a execução de Provas de Conceito (PoC) e projetos envolvendo Inteligência Artificial, Ciência da Informação e o contexto dos Arquivos e Bibliotecas é, para nós, um desafio que deve ser enfrentado continuamente e com a discussão necessária entre os pares. Reflexões das relações do desenvolvimento da Ciência da Informação contemporânea, da Organização do Conhecimento e da classificação, taxonomias e ontologias frente a aplicações como o uso da aprendizagem de máquina, do contexto dos Arquivos e Bibliotecas nos ambientes digitais e da Ética da Informação são, portanto, alguns dos pontos que a investigação futuramente poderá se desdobrar, assim como a necessidade de regulação da Inteligência Artificial que muitos países vem discutindo e adotando nacionalmente e nas suas relações internacionais. Arquivos e Bibliotecas, como espaços de informação confiável e promoção do conhecimento, podem se desenvolver e ampliar possibilidades com a implementação de abordagens de Inteligência Artificial.

Referências

- Assimov, I. (1950) *I, Robot*. Gnome Press
- Castells, M. (1999). A sociedade em rede. A era da Informação: Economia. Sociedade e Cultura, 1(5).
- Cellan-Jones, R. (2014). Stephen Hawking warns artificial intelligence could end mankind. BBC news, 2(2014), 10-10.
- Costal, M., Sales, L., & Zattar, M. (2020). Competência em dados: habilidades na atuação e formação do bibliotecário. BIBLOS, 34(2).
- Cox, A. M., Pinfield, S., & Rutter, S. (2019). The intelligent library. *Library Hi Tech*.
- Ferreira, L. B., Rockembach, M., & Krebs, L. M. (2017). Reflexões conceituais e éticas sobre Big Data: limites e oportunidades. In: XVIII Encontro Nacional de Pesquisa em Ciência da Informação (XVIII ENANCIB).
- Floridi, L. (2014). Big Data and information quality. In: *The philosophy of information quality* (pp. 303-315). Springer, Cham.
- Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. OUP Oxford.
- Graefe, Andreas (2016). *Guide to Automated Journalism*. New York City: Columbia Journalism Review
- Liu, S., & Lee, I. (2018). Email sentiment analysis through k-means labeling and support vector machine classification. *Cybernetics and Systems*, 49(3), 181-199.
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-16).
- Marciano, R., Lemieux, V., Hedges, M., Esteva, M., Underwood, W., Kurtz, M., & Conrad, M. (2018). Archival records and training in the age of big data. In *Re-Envisioning the MLS: Perspectives on the future of library and information science education*. Emerald Publishing Limited.

- Melo, J. F., & Rockembach, M. (2019). Arquivologia e Ciência da Informação na era do Big Data: perspectivas de pesquisa e atuação profissional em arquivos digitais. *Prisma. Com: Revista de Ciências e Tecnologias de Informação e Comunicação*. Porto: CIC. Digital. N. 39 (2019), p. 14-28.
- Moor, J. (2006). The Dartmouth College artificial intelligence conference: The next fifty years. *AI Magazine*, 27(4), 87-87.
- Office of Artificial Intelligence UK (2020). A guide to using artificial intelligence in the public sector. Disponível em: <https://www.gov.uk/government/publications/a-guide-to-using-artificial-intelligence-in-the-public-sector>. Acesso 20 mar. 2021
- Payne, N. (2018, December). Stirring the cauldron: redefining computational archival science (CAS) for the Big Data domain. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 2743-2752). IEEE.
- Rhem, A. J. (2021). AI ethics and its impact on knowledge management. *AI and Ethics*, 1(1), 33-37.
- Rockembach, M., & da Silva, A. M. (2018). Epistemology and Ethics of big data. In *Challenges and Opportunities for Knowledge Organization in the Digital Age* (pp. 812-819). Ergon-Verlag.
- Rockembach, M., & da Silva, A. M. (2018). Epistemology and Ethics of big data. In: *Challenges and Opportunities for Knowledge Organization in the Digital Age*. International ISKO Conference (pp. 812-819). Ergon-Verlag.
- Rockembach, M., & da Silva, A. M. (2021). Web Data and the Relationship Between the General Data Protection Regulation in Europe and Brazil. In *Digital Transformation and Challenges to Data Security and Privacy* (pp. 222-233). IGI Global.
- Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsouпова, T., & Stuart, K. (2019). More human than human? Artificial intelligence in the archive. *Archives and Manuscripts*, 47(2), 179-203.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423.
- Shabou, B. M., Tièche, J., Knafou, J., & Gaudinat, A. (2020). Algorithmic methods to explore the automation of the appraisal of structured and unstructured digital data. *Records management journal*.
- Song, I. Y., & Zhu, Y. (2017). Big data and data science: opportunities and challenges of iSchools. *Journal of Data and Information Science*, 2(3), 1.
- Tartarotti, R. C. D. E., Dal'Evedove, P. R., & Fujita, M. S. L. (2019). Biblioteconomia de dados em repositórios de pesquisa: perspectivas para a atuação bibliotecária. *Informação & Informação*, 24(3), 207-226.
- The National Archives UK (2021). DROID: file format identification tool. Disponível em: <https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>. Acesso 20 mar. 2021
- Webb, M. (2019). The impact of artificial intelligence on the labor market. SSRN. Disponível em <http://dx.doi.org/10.2139/ssrn.3482150>. Acesso 10 mar. 2021
- Wechsler, H. (2016). *Computational vision*. Elsevier.
- Wiener, N. (2019). *Cybernetics or Control and Communication in the Animal and the Machine*. MIT press.

MODELOS DE REQUISITOS PARA SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO

Luis Corujo¹, Jorge Revez²

¹Universidade de Lisboa, Faculdade de Letras, Centro de Estudos Clássicos, luiscorujo@campus.ul.pt, ORCID iD 0000-0003-4411-2453

²Universidade de Lisboa, Faculdade de Letras, Centro de Estudos Clássicos, jrevez@campus.ul.pt, ORCIDID: 0000-0002-3058-943X

Resumo

A necessidade de planeamento e desenvolvimento de Sistemas de Organização do Conhecimento (SOC) implica a identificação de requisitos que garantam a sua eficácia e eficiência. Pretende-se conhecer Modelos de Referência, e os seus requisitos, com potencial para aplicação em sistemas que tenham como finalidade ou sejam utilizados para a Organização do Conhecimento (OC). Nesse âmbito, procura-se: contextualizar a origem e o enquadramento de três Modelos de Referência: MoReq2010, Modelo de Referência OAIS, *Digital Library Reference Model*; analisar os modelos conceptuais dos Modelos de Referência indicados; comparar os requisitos constantes nos Modelos de Referência; identificar problemas e possíveis soluções para a definição de requisitos para SOC. Para tal, recorreu-se a uma análise comparativa entre os três modelos. Os resultados permitem perceber que cada Modelo representa um universo SOC específico e com enfoques diferentes. Verifica-se que o desenvolvimento de SOC deve considerar aspetos distintos, que são apresentados pelas diferentes ferramentas analisadas. Sugere-se que há espaço para o desenvolvimento de modelos de referência SOC, que utilizem os aspetos destacados nesta análise e, igualmente importante, apresentem critérios que levem em consideração uma camada superior da OC, que abranja as restantes.

Palavras-chave: Sistemas de Organização de Conhecimento, MoReq2010, OAIS, *Digital Library Reference Model*, Modelos de Requisitos

Introdução

Esta investigação versa o estudo de modelos de referência para o desenvolvimento de Sistemas de Organização do Conhecimento (SOC) para avaliar a sua eficácia/eficiência e sua capacidade de modelação, recorrendo-se a uma comparação de requisitos. Os SOC são “itens funcionais concebidos para organizar o Conhecimento e a Informação e tornar a sua gestão e recuperação mais fácil” (Mazzocchi, 2018), “destinados a abranger todos os tipos de esquemas para organizar a informação e promover a gestão do conhecimento” (Hodge, 2000). Sintetizando, o «sistema

de organização do conhecimento é um esquema que modela uma estrutura (ou seja, elementos e relações mútuas) de um conjunto organizado de conhecimentos» (Bratková & Kucerová, 2014, p. 8). Procura-se fazer uma comparação entre três modelos com a intenção de encontrar elementos comuns e/ou específicos, e que permita o futuro desenvolvimento de um documento único com requisitos para SOC. As ferramentas para o desenvolvimento de sistemas tecnológicos de gestão do conhecimento que serão comparadas são: *Reference Model for an Open Archival Information System* (OAIS) (CCSDS, 2012), *Modular Requirements for Records Systems* (MoReq) (DLM Forum, 2011) e *Digital Library Reference Model* (DLRM) (Candela et al., 2011b).

O MoReq foi originalmente concebido em 2001 pelo DLM Forum, um organismo europeu financiado pela Comissão Europeia. MoReq significava *Model Requirements for the Management of Electronic Records* e pretendia descrever um sistema eletrônico de gestão documental (SEGA) ótimo, fornecendo orientações que definiam as capacidades de um sistema genérico para toda a União Europeia e em qualquer setor ou ramo de atividade (DLM Forum, 2001, 2008, 2011). Posteriormente, foi publicada pela Comissão Europeia em 2008, e atualizada como MoReq 2, uma especificação formal de requisitos para um SEGA, acompanhada de documentação para testes e informação relacionada (DLM Forum, 2008, 2011).

O MoReq 2010 é uma especificação de gestão documental, publicada em 2011 pelo DLM Forum, que descreve “requisitos modulares para sistemas documentais” (DLM Forum, 2011). O seu objetivo é fornecer um conjunto abrangente de requisitos para um sistema documental que pretende ser adaptável e aplicável a informação e atividades, setores da indústria e tipos de organização diferentes. Tal evita uma abordagem “única” para implementar uma solução de gestão documental, definindo, em vez disso, um conjunto comum de serviços centrais, que são partilhados por tipos diferentes de sistemas de informação de arquivo, mas que também são modulares e flexíveis, permitindo que sejam incorporados em aplicações especializadas e dedicadas, que podem não ter sido anteriormente reconhecidos como sistemas de informação de arquivo (DLM Forum, 2011).

Este instrumento contém requisitos funcionais e não funcionais para sistemas de informação de arquivo. Na primeira parte, relativa ao serviço central e aos módulos de extensão, fornece informação de serviço, conceitos-chave e requisitos funcionais relativos aos serviços do sistema, ao serviço de utilizadores e grupos, serviço de perfis, serviço de classificação, serviço de registo de documentos, serviço de metainformação, serviço de seleção e eliminação, serviço de retenção, serviço de pesquisa e geração de relatórios, e serviço de exportação. Também fornece os principais conceitos e os aspetos não funcionais de um sistema de informação de arquivo e requisitos relativos a desempenho, escalabilidade, viabilidade (de gestão), portabilidade, segurança, privacidade, usabilidade, acessibilidade, disponibilidade, confiabilidade, recuperabilidade, capacidade de manutenção, suporte, garantia e conformidade. Finalmente, apresenta um glossário, seguido do modelo de informação, tipos de entidades, estruturas de dados, definições de elementos de metainformação do sistema e definições de funções. A segunda parte, referente apenas aos módulos de extensão, apresenta informação sobre o módulo, conceitos-chave, requisitos funcionais e não funcionais e glossário sobre a interface gráfica do utilizador, a interface

de programação da aplicação, as séries de classificação (classificação hierárquica) e as séries de componentes (componentes eletrônicos) (DLM Forum, 2011).

O OAIS é uma recomendação desenvolvida pelo *Consultative Committee for Space Data Systems* em 2002, e publicada como ISO 14721: 2003, tendo sido ambas atualizadas em 2012. Uma terceira versão está a ser elaborada (a última atualização da proposta é de setembro de 2020). Considera que o OAIS é um arquivo, que consiste numa organização, que pode fazer parte de uma organização maior, de pessoas e sistemas, e que aceitou a responsabilidade de preservar informação e disponibilizá-la a uma comunidade designada. A informação que é mantida foi considerada como necessitando de Preservação a Longo Prazo, mesmo que o OAIS em questão não seja permanente. Longo prazo é suficientemente extenso para se preocupar com os impactos das mudanças de tecnologias, bem como com o apoio para novos suportes e formatos de dados, ou com mudanças na Base de Conhecimento da Comunidade Designada ou dentro da Comunidade Designada ou a sua definição. O Longo prazo pode estender-se indefinidamente (CCSDS, 2002, 2012, 2020).

Desde a sua adoção, este Modelo de Referência foi saudado e amplamente adotado por praticamente todos os tipos de comunidades de preservação digital. A maioria das iniciativas de preservação digital recentes refere-se à norma do Modelo de Referência OAIS. Também tem sido amplamente utilizado por organizações para informar as suas implementações de sistemas de preservação novos ou atualizados. A recomendação pretende fornecer: uma estrutura para a compreensão e maior consciencialização dos conceitos arquivísticos necessários para a preservação e acesso de informação digital a longo prazo; os conceitos necessários às organizações não arquivísticas para serem participantes efetivos no processo de preservação; uma estrutura, incluindo terminologia e conceitos, para descrever e comparar arquiteturas e operações de arquivos existentes e futuros; uma estrutura para descrever e comparar diferentes estratégias e técnicas de preservação a longo prazo; uma base para comparar os modelos de dados de informação digital preservada pelos Arquivos e para discutir como os modelos de dados e a informação subjacente podem mudar ao longo do tempo; uma estrutura que pode ser expandida por outros esforços para abarcar a preservação a longo prazo de informações que não estão em formato digital (por exemplo, suportes físicos e amostras físicas). Pretende também: ampliar o consenso sobre os elementos e processos de preservação e acesso à informação digital a Longo Prazo; promover um mercado maior que os fornecedores podem apoiar; orientar a identificação e a produção de normas relacionadas com o OAIS (CCSDS, 2002, 2012, 2020).

A Recomendação apresenta os conceitos OAIS, no que respeita ao seu ambiente, informação, objetivo de preservação e interações externas de alto nível. Segue-se a identificação das responsabilidades obrigatórias e exemplos de mecanismos para o cumprimento de responsabilidades. Também detalha os modelos funcional, de informação e de transformação de pacotes de informação. Em seguida, apresenta as perspetivas de preservação, como Migração Digital, agregação de informação de representação e entrega para outro OAIS. Por fim, aborda a interoperabilidade do arquivo, em termos de níveis de interação entre os arquivos OAIS (CCSDS, 2020).

O modelo de referência aborda uma gama completa de funções de preservação de informação de arquivo, incluindo ingestão, armazenamento de arquivo, gestão de

dados, acesso e disseminação. Também aborda a migração da informação digital para novos suportes e formatos, os modelos de informação usados para representar a informação, o papel do software na preservação da informação e a troca de informação digital entre Arquivos. Identifica as interfaces internas e externas para as funções de arquivo e identifica vários serviços de alto nível para essas interfaces. Fornece vários exemplos ilustrativos e algumas recomendações de “boas práticas”. Define um conjunto mínimo de responsabilidades para um arquivo ser considerado OAIS e também define um arquivo máximo para fornecer um amplo conjunto de termos e conceitos úteis. Também considera que o termo ‘Aberto’ em OAIS é usado para indicar que a Recomendação e as normas que o seguem, são desenvolvidos em fóruns abertos, e não implica que o acesso ao Arquivo seja irrestrito (CCSDS, 2002, 2012, 2020).

Neste Modelo de Referência, há um enfoque específico na informação digital, seja esta a forma primária da informação mantida, ou informação de suporte para materiais preservados digital e fisicamente. Portanto, o modelo acomoda informação que é inerentemente não digital (por exemplo, uma amostra física), mas a modelagem e a preservação de tal informação não é abordada detalhadamente (CCSDS, 2002, 2012, 2020).

O *DLRM* tem sido entendido como uma estrutura conceptual que visa capturar as entidades significativas e suas relações no universo da biblioteca digital, com o intuito de desenvolver um modelo concreto. É o resultado do esforço dos grupos de investigação que integraram a *DELOS Network of Excellence on Digital Libraries*, com financiamento europeu e, posteriormente, de uma equipa de especialistas internacionais em Bibliotecas Digitais e do DL.org, projeto financiado pelo *Cultural Heritage and Technology Advanced Learning Unit* da *Information Society Directorate-General* da Comissão Europeia (Candela et al., 2011b, 2011a). A estrutura apresentada no Modelo de Referência foi concebida com o objetivo de coordenar abordagens, soluções e o desenvolvimento de sistemas na área da Biblioteca Digital, de forma que os sistemas da Biblioteca Digital pudessem ser descritos, classificados e avaliados de acordo com os elementos-chave deste modelo.

Este Modelo de Referência está fundamentado no *The Digital Library Manifesto*, que apresenta o cenário que rege toda a atividade e apresenta as principais noções que caracterizam o universo das Bibliotecas Digitais em termos abstratos. Começa por apresentar uma análise dos três tipos de ‘sistemas’ relevantes nesta área: Biblioteca Digital, Sistema de Biblioteca Digital e Sistema de Gestão de Biblioteca Digital. Também identifica os principais conceitos que caracterizam estes sistemas, classificando-os em vários domínios, cada um deles representativo de um aspeto particular do universo das Bibliotecas Digitais (Organização; Conteúdo; Utilizador; Funcionalidade; Política; Qualidade; Arquitetura), e apresenta os principais perfis que os atores podem desempenhar nas bibliotecas digitais (utilizadores finais da biblioteca digital, como criadores de conteúdo, consumidores de conteúdo e bibliotecários digitais; gestores de biblioteca digital, como designers de Biblioteca Digitais e administradores de sistema de biblioteca digital; desenvolvedores de aplicações de biblioteca digital). Por fim, apresenta um plano para traçar uma caracterização abrangente do universo das bibliotecas digitais, com base em diferentes artefactos que capturam o universo em diversos níveis de abstração, desde o muito abstrato

(o *Digital Library Reference Model*), até ao muito concreto (a implementação) (Candela et al., 2011b, 2011a).

O *DLRM* apresenta e descreve os principais conceitos e relações em conexão com cada um dos aspetos identificados no *Manifesto*, explicando a sua lógica, e apresentando exemplos da sua instanciação em cenários concretos. Conclui apresentando um conjunto de critérios para determinar a conformidade das bibliotecas digitais a este Modelo de Referência (Candela et al., 2011b).

Metodologia

Com recurso a uma análise comparativa desses instrumentos, propõe-se avaliar e comparar seus principais requisitos. A justificação desta proposta baseia-se numa combinação de fatores. Em primeiro lugar, os três modelos foram construídos com um propósito internacional; em segundo lugar, alguns desses modelos deram origem a normas ISO, tornando-os muito relevantes para o desenvolvimento de projetos na área de Biblioteconomia e Ciência da Informação; em terceiro lugar, são modelos reconhecidos pela indústria e pelos mercados para o desenvolvimento de sistemas dedicados a repositórios / arquivos digitais, sistemas de gestão de documentos eletrónicos e sistemas de gestão de bibliotecas digitais. O planeamento de sistemas tecnológicos com base nessas normas / requisitos traz garantias de um uso correto de classificações, tesouros, ontologias, entre outros tipos de SOC. Também promovem a sua configuração em sistemas tecnológicos e em relação aos processos de negócio nos quais esses sistemas tecnológicos são utilizados. Embora os modelos não tenham sido desenvolvidos primordialmente para SOC, podem servir de apoio no desenvolvimento de instrumentos / ferramentas / dispositivos tecnológicos que recorram à utilização de SOC.

Resultados

O Digital Library Reference Model

O Modelo de Referência de Biblioteca Digital descreve a Biblioteca Digital (DL) como uma organização potencialmente virtual, que coleta, gere e preserva de forma abrangente, a longo prazo, conteúdo digital e oferece às suas comunidades de utilizadores-alvo funcionalidades especializadas nesse conteúdo, de qualidade definida e de acordo com políticas codificadas abrangentes. Geralmente é o “sistema” final percebido pelos utilizadores finais como sendo a biblioteca digital. O Sistema de Biblioteca Digital (DLS) é o sistema de software implantado e em execução que implementa as funcionalidades do DL. O Sistema de Gestão da Biblioteca Digital (DLMS) é o sistema de software genérico que suporta a produção e administração de DLS e a integração de software adicional que oferece recursos mais refinados, especializados ou avançados.

Em seguida, define os conceitos que caracterizam e classificam esses sistemas. Para tanto, cada conceito também nomeia um domínio, um aspeto particular do universo da Biblioteca Digital. Estes incluem, por exemplo, a organização, o con-

teúdo, o utilizador, a funcionalidade, a política, a qualidade, e a arquitetura. Além disso, o domínio Recursos, considera o recurso como o conceito que captura as características de qualquer entidade DL. Cada recurso é identificado por um Identificador de Recurso, organizado de acordo com um Formato de Recurso, caracterizado por diversos Parâmetros de Qualidade, regulados por Políticas, expressos, descritos e comentados por um Objeto de Informação, agrupados em Conjuntos de Recursos.

O Reference Model for an Open Archival Information System (OAIS)

O Modelo de Referência OAIS considera que pode ser aplicável a qualquer arquivo e que os desenvolvedores de normas devem usar este Modelo de Referência como base para normalização adicional nesta área. Este modelo de referência não especifica um design ou uma implementação. As implementações reais podem agrupar ou separar funcionalidades de forma diferente. O modelo de referência considera que uma implementação de arquivo OAIS em conformidade deve suportar o modelo de informação apresentado, usando a *Unified Modeling Language (UML)* para a representação de objetos em diagrama de classes. Uma classe descreve as responsabilidades, o comportamento e o tipo de um conjunto partilhado de propriedades de objeto. Este modelo de referência explica que a informação é obtida dos dados da seguinte maneira: objetos de dados são interpretados usando a sua informação de representação, que retorna o objeto de informação.

O modelo de referência considera que um Arquivo OAIS conforme deve cumprir as seguintes responsabilidades: 1. Negociar e aceitar informação adequada dos Produtores de informação; 2. Obter controlo suficiente sobre a informação fornecida ao nível necessário para garantir a Preservação a Longo Prazo; 3. Determinar, isoladamente ou em conjunto com outras partes, quais comunidades se devem tornar a Comunidade Designada e, portanto, devem ser capazes de compreender a informação fornecida, definindo assim a sua Base de Conhecimento; 4. Garantir que a informação a preservar sejam de compreensão independente para a comunidade designada; 5. Cumprir com as políticas e procedimentos documentados que garantem que a informação seja preservada contra todas as contingências razoáveis, incluindo o desaparecimento do Arquivo, garantindo que nunca seja eliminado, a menos que tal seja permitido como parte de uma estratégia aprovada; 6. Disponibilizar a informação preservada para a comunidade designada e permitir que a informação seja disseminada como cópia dos, ou rastreável aos, objetos de dados originais enviados com evidências que comprovem sua autenticidade. Além disso, o modelo de referência afirma que um Arquivo OAIS em conformidade pode fornecer aos utilizadores serviços adicionais que vão além dos exigidos de um arquivo OAIS. Também pressupõe que os implementadores usarão este modelo de referência como um guia ao desenvolver uma implementação específica para fornecer serviços e conteúdo identificados. Além disso, ele não assume ou endossa qualquer plataforma de computação específica, ambiente de sistema, paradigma de design de sistema, metodologia de desenvolvimento de sistema, sistema de gestão de bases de dados, paradigma de design de bases de dados, linguagem de definição de dados, linguagem de comando, interface de sistema, interface de utilizador, tecnologia ou suporte

necessários para implementação. Por esse motivo, e para encontrar critérios relevantes para os arquivos OAIS, optamos por usar uma norma separada. Considerando que o Modelo de Referência OAIS continha um roteiro que incluía a necessidade de uma norma de certificação de Arquivos, em 2003, o Research Libraries Group (RLG) e a National Archives and Records Administration (NARA) criaram um grupo de trabalho para abordar especificamente a certificação de repositórios digitais. Esse grupo de trabalho publicou *Trustworthy Repositories Audit & Certification: Criteria and Checklist* (Center for Research Libraries & Online Computer Library Center, 2007), que foi a base para a norma a partir da qual se podem desenvolver processos de acreditação e certificação para qualquer tipo de repositórios digitais – o *Audit and Certification of Trustworthy Digital Repositories* (CCSDS, 2011), também disponível como ISO 16363: 2011. Esta norma (doravante TDR / ISO 16363) destina-se principalmente aos responsáveis pela auditoria de repositórios digitais e para aqueles que trabalham ou são responsáveis por repositórios digitais, que buscam uma medição objetiva da fiabilidade do seu repositório.

O Modular Requirements for Records Systems (MoReq 2010)

O MoReq 2010 considera que a informação de arquivo se refere às informações que têm um valor intrínseco que os torna importantes o suficiente para serem armazenados e conservados pelo seu valor probatório. Afirma a importância de obter uma compreensão do contexto de negócio para avaliar se uma parte da informação deve ser retida e gerida como informação de arquivo ou não, bem como a sua relevância e significado para a organização. Para tanto, o MoReq 2010 apresenta a norma de gestão documental ISO 15489 (ISO, 2001, 2016), que lista os processos de gestão documental e propõe que uma organização deve usar um sistema de informação de arquivo para implementar esses processos. A ISO 15489 também define um sistema de informação de arquivo como um sistema de informação que captura, gere e fornece acesso à informação de arquivo ao longo do tempo. O MoReq 2010 toma os processos de gestão documental listados na ISO 15489 e adiciona um nível muito maior de especificidade quanto à forma como esses processos devem ser realizados por um sistema de informação de arquivo. Isso permite atingir um maior grau de rigor ao que pode ser alcançado simplesmente seguindo a ISO 15489 em um planejamento de sistema de informação de arquivo. Tal define uma especificação para os sistemas de informação de arquivo conformes ao MoReq2010 (MCRS), que podem ser interoperáveis entre eles, porque eles não só entenderão suas próprias entidades (incluindo a informação) e seus próprios processos, mas também podem exportá-los para um formato normalizado e compreendido por outro MCRS.

O MoReq 2010 define tipos de entidade e como eles devem ser geridos por um MCRS, para torná-los o mais uniformes possível na forma como a sua metainformação é representada e o seu histórico de eventos é gerido, e nos seus controles de acesso e ciclo de vida da entidade. Ao contrário das entidades em outros sistemas de informação, as entidades num MCRS são destruídas, em vez de eliminadas, deixando uma entidade residual que permanece no MCRS. Entidades residuais são um conceito importante nos sistemas de informação de arquivo, pois indicam entidades

que já estiveram presentes no sistema: sem elas, não seria possível reconstruir o contexto completo de um registo histórico.

Dentro de um MCRS, entidades de diferentes tipos são nominalmente descritas como sendo geridas por diferentes serviços de acordo com uma “arquitetura baseada em serviços”. Estes são baseados nos processos listados na ISO 15489.

O MoReq2010 não requer que os fornecedores implementem os *Model Role Services* e *Model Metadata Services* exatamente como são especificados nos requisitos funcionais, que são apresentados apenas como padrão. Eles só precisam ser cumpridos como tal quando o fornecedor deseja oferecer suporte a módulos avançados, como o módulo de importação. Além disso, os módulos *Plug-in* representam conjuntos alternativos, mas igualmente válidos, de funcionalidades que atingem o mesmo objetivo, mas de maneiras diferentes. Os fornecedores podem escolher a abordagem a adotar. Cada MCRS deve fornecer suporte e ser certificado em pelo menos uma implementação da funcionalidade. Também pode oferecer suporte e ser certificado em relação a outros módulos de *plug-in* alternativos da mesma série.

Cada serviço MoReq2010 gere entidades pertencentes a um determinado número de tipos de entidades.

A maioria das entidades possui três conjuntos de informação associados a elas. A metainformação é a informação que descreve a entidade, contida nas definições dos elementos de metainformação, e dividida em metainformação do sistema (definida pelo MoReq2010) e metainformação contextual (definida pelo fornecedor e / ou utilizador). O histórico de eventos é um conjunto de eventos associados à entidade que armazena informação sobre as diferentes funções que foram executadas na entidade. A lista de controlo de acesso é uma lista de entradas de controlo de acesso que especifica que utilizadores e grupos podem executar funções na entidade, onde conjuntos específicos de funcionalidades são definidos coletivamente como funções.

Comparação entre Requisitos

No caso do DLRM, os domínios também definem os critérios. Apresenta um total de 94 critérios, dos quais 20 são orientados para o conteúdo, nove são orientados para o utilizador, 18 são orientados para a funcionalidade, 20 são orientados para as políticas, 16 são orientados para a qualidade e 11 são orientados para a arquitetura. Estes critérios são divididos em tipos de critérios obrigatórios, recomendados e opcionais. Os critérios obrigatórios englobam as características essenciais que o DLRM considera que uma biblioteca digital deve ter. Os critérios recomendados incluem os recursos que as “boas” bibliotecas digitais devem ter, como boa prática. Os critérios opcionais estão relacionados com as características que os DLRM consideram que tornam uma biblioteca digital mais adequada para um ou outro propósito ou que lhe agregam funcionalidade única.

Cada critério DLRM é apresentado com: uma declaração de uma única frase; uma explicação que descreve o critério e porque é significativo; um ou mais identificadores numéricos do(s) conceito(s) DLRM correspondente(s); um ou mais identificadores numéricos da(s) relação(ões) DLRM correspondente(s); um exemplo que mostra como o critério pode ser atendido; e sugestão de tipo de evidência (por exemplo, análise de materiais documentais, evidência observacional ou testemunho).

O DLRM também estabelece que o Formulário de Avaliação da Lista de Verificação da Conformidade do Modelo de Referência deve incluir as referidas informações, e a identificação do auditor / avaliador que preenche o formulário; a declaração de cumprimento de critérios relativos à biblioteca digital; o tipo de evidência que foi usada para confirmar se o critério é ou não satisfeito pela biblioteca digital; e a fonte da evidência que foi usada para determinar a declaração de cumprimento do critério.

O *Audit and Certification of Trustworthy Digital Repositories* (TDR / ISO 16363) fornece métricas / requisitos normativos através dos quais um repositório digital pode ser avaliado. Esses 109 requisitos são agrupados em três secções: Infraestrutura Organizacional (25 requisitos); Gestão de Objetos Digitais (60 requisitos); Gestão de Riscos de Infraestrutura e Segurança (24 requisitos). A primeira secção refere-se a: Governança e Viabilidade Organizacional; Estrutura Organizacional e o Pessoal; Política de *Accountability* e Enquadramento dos Procedimentos de Preservação; Sustentabilidade Financeira; Contratos, licenças e encargos. A segunda refere-se a: aquisição de conteúdo na Ingestão; Criação de pacotes de informação de arquivo na Ingestão; Planeamento de Preservação; Preservação dos pacotes de informação de arquivo; Gestão da Informação; Gestão de Acesso. A terceira secção inclui: Gestão de riscos de infraestrutura técnica; Gestão de Riscos de Segurança. Esse número de requisitos leva em consideração que algumas métricas / requisitos mais gerais são especificados em duas ou mais métricas / requisitos específicos.

Cada requisito é apresentado com: uma declaração da métrica numa única frase; um texto de apoio explicando por que a métrica é necessária; exemplos de maneiras como o repositório pode demonstrar que está a atender ao requisito; e uma discussão sobre o requisito e o seu contexto.

O MoReq2010 apresenta uma primeira parte com requisitos sobre os serviços essenciais. Estes são divididos em requisitos funcionais relativos a cada um dos serviços e requisitos de serviços do sistema que são comuns a todos os outros serviços principais. Os requisitos funcionais tendem a concentrar-se no comportamento específico exigido pelo sistema: Serviços do sistema; Serviço de utilizadores e grupos; Serviço de modelo de perfis; Serviço de classificação; Serviço de registo de informação de arquivo; Serviço de modelo de metainformação; Serviço de seleção e eliminação; Serviço de retenção; Serviço de pesquisa e relatórios; Serviço de exportação. Também apresenta requisitos não funcionais relacionados com as características ou qualidades desejáveis do MCRS, como o desempenho, a escalabilidade, a capacidade de gestão, a portabilidade, a segurança, a privacidade, a usabilidade, a acessibilidade, a disponibilidade, a confiabilidade, a capacidade de recuperação, a capacidade de manutenção, o suporte, a garantia e a conformidade. Estes especificam aspetos qualitativos do sistema de informação de arquivo que não são necessariamente explicitados apenas pelos requisitos funcionais. O MoReq2010 também apresenta uma segunda parte com requisitos relativos aos módulos *plug-in*, a interface (que inclui a Interface Gráfica de Utilizador – GUI e a Interface de Programação de Aplicativos – API), a classificação (relativa à Classificação Hierárquica) e Componentes (quanto aos Componentes Eletrónicos). Cada um desses tipos apresenta requisitos funcionais e não funcionais. O MoReq2010 apresenta um total de 344 requisitos, tendo em consideração que os requisitos não funcionais para escalabilidade saltam um número na contagem dos requisitos.

Cada requisito funcional é apresentado com uma declaração do requisito e uma discussão sobre o requisito e o seu contexto, descrevendo o requisito, a sua importância e a sua relação com outros requisitos, serviços e funções. Os requisitos não funcionais são apresentados com uma declaração do requisito e perguntas ou declarações, que devem ser respondidas ou verificadas ao avaliar o seu cumprimento.

O Quadro 1 fornece um resumo da análise cruzada entre os requisitos do MoReq 2010 e os requisitos dos outros documentos.

Quadro 1. Comparação dos requisitos do MoReq 2010 com os requisitos/critérios DLRM e TDR / ISO16363

MoReq 2010		DLRM	TDR/ISO 16363
Requisitos funcionais dos serviços essenciais	Serviços do sistema	Referem-se apenas às funções e não especificam atributos de metainformação a serem usados	
	Serviço de utilizadores e grupos	Apenas afirma que o ator deve ser identificado, descrito, ter um perfil e executar ações que aplicam funções e dizem respeito a recursos, e que ele pode pertencer a mais de um grupo	Afirma que o repositório deve ter mecanismos para verificar adequadamente a identidade do produtor de todos os materiais
	Serviço de modelo de perfis	Afirma apenas que cada ator deve ter pelo menos um perfil, deve haver perfis de Gestor da Biblioteca Digital, Desenvolvedor de Software da Biblioteca Digital e Utilizador Final; cada ator deve um perfil atribuído para aceder aos recursos; que devem existir políticas de acesso para regular o uso da Biblioteca Digital pelos atores	Considera que a equipa do repositório deve ter perfis, responsabilidades e autorizações delineadas relativas à implementação de mudanças no sistema; deve ter definido a sua comunidade designada e respetiva(s) base(s) de conhecimento; e deve ter essas definições apropriadamente acessíveis
	Serviço de classificação	Afirma apenas que todo	–
	Serviço de registo de informação de arquivo	Objeto de Informação deve pertencer a pelo menos uma Coleção, que deve ter um identificador único, pelo menos um elemento de Metainformação a ele associado, pode estar associado a várias Metainformações e Anotações, e deve ter uma intenção bem definida, ou seja, o conjunto de critérios que caracterizam os membros da coleção.	–

MoReq 2010	DLRM	TDR/ISO 16363
	<p>Deve ter uma extensão bem definida, ou seja, o conjunto de Objetos de Informação pertencentes à coleção</p>	
<p>Serviço de modelo de metainformação</p>	<p>Afirma apenas que cada objeto de informação e coleção deve ter pelo menos um ou mais elementos de metainformação associados a ele; que cada registo de metainformação deve estar em conformidade com um formato explícito e conhecido, e que as funções podem ser enriquecidas com metainformação e anotações</p>	<p>Considera que os repositórios devem especificar claramente a informação que precisa ser associada a cada informação de conteúdo em particular, no momento do seu depósito, ter acesso às ferramentas, métodos e recursos necessários para fornecer informação de representação fidedigna para todos os objetos digitais que contém, para determinar que informação de representação é necessária para tornar cada objeto de dados compreensível para a comunidade designada. Também afirma que os repositórios devem: ter acesso à Informação de Representação, ferramentas ou métodos necessários para garantir que a Informação de Representação necessária seja persistentemente associada aos Objetos de Dados relevantes; ter mecanismos para criar, identificar ou coletar qualquer Informação de Representação extra necessária; especificar os requisitos mínimos de informação para permitir que a comunidade designada descubra e identifique material de interesse, capture ou crie e mantenha ligações e associações bidirecionais entre cada pacote de informação arquivado e a sua informação descritiva ao longo do</p>

	MoReq 2010	DLRM	TDR/ISO 16363
			tempo.
	Serviço de seleção e eliminação	Nenhum dos serviços tem requisitos comparáveis nos outros documentos, uma vez que visam a preservação dos objetos de informação	
	Serviço de retenção		
	Serviço de pesquisa e relatórios	Afirma apenas que cada ator deve ser fornecido com perfis com funções para aceder e descobrir recursos, e que deve haver perfil com funções para consulta e pesquisa dos recursos	Considera que os repositórios devem especificar todos os aspetos apropriados ao acesso; seguir políticas e procedimentos que permitam a disseminação de objetos digitais rastreáveis aos originais, com evidências que comprovem a sua autenticidade; especificar requisitos mínimos de informação para permitir que a Comunidade Designada descubra e identifique o material de interesse; registar e analisar todas as falhas e anomalias de gestão de acesso
	Serviço de exportação	–	Afirma apenas a necessidade de planear e executar um plano de sucessão, que consideramos poder incluir os objetos de informação e exportação de dados
Requisitos não-funcionais dos serviços essenciais e dos módulos de <i>plug-in</i>	Desempenho	Apresenta requisitos gerais sobre os parâmetros de qualidade	Afirma a necessidade de mecanismos de revisão, atualização e desenvolvimento contínuo das suas Políticas de Preservação para que o repositório possa ter tecnologias de hardware e software adequadas aos serviços que presta
	Escalabilidade	Aparenta não apresentar requisitos	Apresenta requisitos gerais sobre monitorização, avaliação e substituição de mudanças de tecnologia
	Capacidade de Gestão	Apresenta requisitos gerais sobre os parâmetros de qualidade, que podem ser percecionados como	Apresenta requisitos gerais sobre monitorização da inadequação da informação de

	MoReq 2010	DLRM	TDR/ISO 16363
		equivalentes aos requisitos de auditoria MoReq2010	representação para a comunidade designada, e a necessidade de uma análise sistemática sobre segurança de riscos, atividades de preservação, mudanças de tecnologia e relatórios de falhas e anomalias
	Portabilidade	Afirma apenas que cada componente de arquitetura pode explorar um ou mais outros componentes de arquitetura não conflitantes	-
	Segurança	-	Apresenta requisitos gerais sobre controlos de riscos de segurança, definição de responsabilidades e planos para preparação e recuperação de desastres; afirma a necessidade de um programa de desenvolvimento profissional ativo que forneça à equipa oportunidades de desenvolvimento de competências e conhecimentos e para cumprir com as políticas de acesso e outras políticas e procedimentos que permitam a disseminação de objetos digitais
	Privacidade	-	-
	Usabilidade	-	
	Acessibilidade	-	Apresenta requisitos gerais sobre controlos de riscos de segurança, definição de responsabilidades e planos para preparação e recuperação de desastres; afirma a necessidade de um programa de desenvolvimento profissional ativo que forneça à equipa oportunidades de desenvolvimento de competências e para conhecimentos e para

MoReq 2010	DLRM	TDR/ISO 16363
		cumprir com as políticas de acesso e outras políticas e procedimentos que permitam a disseminação de objetos digitais
Disponibilidade	Nenhum dos dois parecem apresentar requisitos que	possam ser percebidos nestes termos
Confiabilidade		
Recuperabilidade	–	Apresenta requisitos que indicam a necessidade de: um plano de sucessão, planos de contingência e / ou acordos de custódia, e monitoração do ambiente organizacional, hardware / software e <i>backups</i> , além de registrar e agir sobre relatórios de erros nos dados ou nas respostas aos utilizadores, falhas de gestão e anomalias; e reagir também a atualizações de segurança
Capacidade de Manutenção	–	
Suporte	–	Refere-se aos requisitos relativos à identificação e gestão das operações de preservação e metas de riscos associados à infraestrutura do sistema, o número e localização de cópias de todos os objetos digitais e o compromisso com um cronograma regular de autoavaliação e certificação externa
Garantia	Nenhum dos dois parecem apresentar requisitos que possam ser percebidos nestes termos	
Conformidade		
	–	Refere-se aos requisitos relativos à identificação e gestão das operações de preservação e metas de riscos associados à infraestrutura do sistema, o número e localização de cópias de todos os objetos digitais e o compromisso com um cronograma regular de autoavaliação e certificação externa

MoReq 2010		DLRM	TDR/ISO 16363
	Módulos de Interface	Nenhum dos dois parece apresentar requisitos que possam ser especificamente percebidos como relacionados aos requisitos sobre <i>plug-ins</i> no MoReq 2010	
	Módulos de Classificação		
	Módulos de Componentes		

Referência: CCSDS (2011), Candela et al. (2011b, 2011a), DLM Forum (2011)

Em relação ao DLRM, além dos equivalentes percebidos já listados na tabela anterior, o Quadro 2 fornece um resumo para a análise cruzada entre os critérios do DLRM e os requisitos MoReq 2010 e TDR / ISO 16363.

Quadro 2. Comparação dos critérios DLRM com os requisitos MoReq 2010 e TDR / ISO16363

DLRM	MoReq 2010	TDR/ISO 16363
Critérios orientados para o conteúdo	Parecem relacionados com alguns requisitos dos serviços de modelo de metainformação e de registo de informação de arquivo	Parecem relacionado a alguns requisitos relativos aos conteúdos dos pacotes de informação de submissão e arquivo (objetos de dados e metainformação)
Critérios orientados para o utilizador	Podem ser percebidas relações com aspetos dos serviços de utilizadores e grupos, e de modelo de perfis	Parecem relacionados com alguns critérios referentes aos utilizadores (incluindo produtores, comunidade designada e equipa)
Critérios orientados para a funcionalidade	Podem ser percebidas relações com aspetos dos serviços do sistema e outros requisitos de serviços essenciais em geral	Parecem relacionados a alguns grupos da secção de gestão de objetos digitais, como a ingestão, planeamento de preservação, preservação de pacotes de informação de arquivo, gestão de informação e acesso
Critérios orientados para as políticas	Embora não haja nenhum requisito do MoReq2010 que pareça declarar diretamente os aspetos da política, eles podem ser inferidos das políticas gerais de gestão de informação de arquivo	Parecem relacionados aos critérios quanto à necessidade de Políticas de Coleta, Acesso, Divulgação e, principalmente, Preservação
Critérios orientados para a qualidade	Parecem estar amplamente relacionados com requisitos não funcionais de desempenho e capacidade de gestão, em termos de avaliação e auditoria	Parecem estar amplamente relacionados com critérios relativos a auditorias financeiras, autoavaliação, avaliação de risco-benefício e certificação externa
Critérios orientados para a arquitetura	Parecem ter alguma relação com o requisito não funcional de portabilidade e com os requisitos mais específicos do modelo de interface	Parecem ter alguma relação com os critérios relativos à gestão de risco de infraestrutura técnica

Referência: Candela et al. (2011b, 2011a), CCSDS (2011), DLM Forum (2011)

No que diz respeito à ISO16363, além dos equivalentes percebidos já listados nas tabelas já apresentadas, o Quadro 3 fornece um resumo para a análise cruzada entre os requisitos TDR / ISO 16363 e os requisitos / critérios MoReq 2010 e DLRM.

Quadro 3. Comparação dos requisitos TDR / ISO16363 com os requisitos/critérios DLRM e MoReq 2010 e

TDR/ISO 16363	DLRM	MoReq 2010
Critérios relativos à Infraestrutura Organizacional	Parecem não conseguir abordar um conjunto de critérios especificamente relativos ao ambiente organizacional	
Critérios relativos à Gestão de Objetos Digitais	Parecem ter uma ampla relação com os critérios orientados para a funcionalidade	Parecem ter uma ampla relação com os requisitos dos serviços essenciais
Critérios relativos à Gestão de Riscos de Infraestrutura e Segurança	Parecem estar amplamente relacionados com alguns critérios orientados para a qualidade e a arquitetura	Parecem estar amplamente relacionados a alguns requisitos funcionais para serviços de sistema e de modelo de perfil, e também com os requisitos não funcionais de desempenho, escalabilidade, capacidade de gestão, segurança, recuperabilidade e suporte

Referência: CCSDS (2011), Candela et al. (2011b, 2011a), DLM Forum (2011)

Discussão

O MoReq2010, tendo mais requisitos, parece ter uma preocupação maior no caso dos serviços de sistema, nas funcionalidades do serviço e especificações de metainformação para a inicialização, falhas, concretização, tipos de entidades, definição de funções, eventos dos serviços. Essa preocupação demonstra a importância de registrar qualquer evento ou ação, quando e por quem foi executado. A importância de registrar qualquer evento ou ação é mostrada para cada serviço essencial. Por exemplo, os serviços de utilizadores e grupos também mostram uma preocupação maior em especificar atributos de metainformação e processos de criação, eliminação, destruição e atualização de utilizadores e grupos, e registrar qualquer evento ou ação referente a esses serviços, quando e por quem foi executado. Os serviços de modelo de perfil, classificação e de registo de informação de arquivo apresentam as mesmas preocupações de metainformação e registo de eventos ou ações em relação, respetivamente, aos perfis e às listas de controlo de acesso, às classes e à informação de arquivo e agregações de informação de arquivo. Os serviços de seleção e eliminação e de retenção mostram a mesma preocupação em relação às ações de seleção, eliminação e de retenção. Os serviços de pesquisa e relatórios, e de exportação também apresentam a mesma preocupação em relação, respetivamente, às pesquisas e relatórios, e à exportação de entidades utilizando o esquema XML MoReq2010. Até mesmo o serviço de modelo de metainformação apresenta definições de elementos de metainformação e de registo de eventos ou ações sobre o sistema, metainformação contextual e as suas interfaces.

A ISO13636, estando no meio dos outros documentos relativos ao número de requisitos, dá atenção aos critérios de infraestrutura organizacional relacionados com a governança e viabilidade organizacional, estrutura e recursos humanos, para além da responsabilidade procedimental e enquadramento das políticas de preservação, sustentabilidade financeira e contratos, licenças, e encargos, além da gestão da

ingestão, preservação, informação e acessos dos objetos digitais, e da gestão de riscos de infraestrutura técnica e segurança.

Em relação ao DLRM, tendo o menor número de requisitos dos documentos analisados, os seus critérios parecem ter uma maior atenção para as entidades dos domínios e suas relações.

Percebe-se então que cada Modelo representa um universo SOC específico, e com enfoques diferentes. O MoReq2010 centra-se nos serviços/funções do sistema tecnológico (mesmo os requisitos não funcionais são declarados em relação ao serviço, portanto, a função), o TDR / ISO16363 cobre a infraestrutura organizacional, funções de gestão do objeto digital e gestão da infraestrutura técnica e riscos de segurança, e o DLRM centra-se nas entidades/domínios. Isso apresenta problemas ou desvantagens se esses modelos de referência forem usados do ponto de vista do universo SOC. Um dos problemas é a percepção de que, do ponto de vista do universo SOC, cada um parece carecer de aspetos que aparecem nos outros. Além disso, nenhum desses documentos apresenta explicitamente preocupações com relação à camada da Organização do Conhecimento onde eles seriam aplicados, mas apenas se relacionam com as camadas inferiores em relação a entidades, informação/objetos de dados digitais ou físicos, e informação registada.

A comparação destas ferramentas permite perceber que o desenvolvimento de sistemas de organização do conhecimento tecnológicos precisa considerar diferentes aspetos que são apresentados pelos instrumentos analisados. Do DLRM é importante reter a necessidade da lista de critérios / requisitos para identificar as entidades significativas e suas relações no universo SOC específico, em termos dos seus sistemas organizacionais e tecnológicos relevantes, e explicitar os principais conceitos que caracterizam esses sistemas, classificando-os em domínios que representam aspetos particulares desse universo SOC. O DLRM também apresenta a importância de definir os perfis dos diferentes atores dentro do SOC, como utilizadores, gestores e / ou desenvolvedores. Do OAIS e do TDR / ISO16363 é importante reter a necessidade de definir os conceitos, ambiente e relações internas e externas, informação, objetivos, enquadramento, responsabilidades obrigatórias e mecanismos de cumprimento de responsabilidades, no universo específico SOC. O OAIS e o TDR/ISO16363 também introduzem a importância de identificar métricas normativas como requisitos de avaliação para um universo SOC específico, incluindo aspetos como a infraestrutura organizacional, entidades funcionais, responsabilidades de gestão de informação / conhecimento, infraestrutura técnica e gestão de risco de segurança. Do MoReq2010 é importante reter a necessidade de obter uma compreensão do contexto de negócios para avaliar como a informação / conhecimento deve ser gerido, considerando a sua relevância e significado para a organização. Outros aspetos a reter do MoReq2010 é a necessidade de identificar os tipos de entidades e funções para a gestão de informação / conhecimento no universo SOC específico, e caracterizar os serviços que suportam essas funções, garantindo o registo completo, como metainformação, relativamente à gestão de entidades e operações dos serviços. O MoReq2010 apresenta também a importância da modularidade do SOC (especialmente no que se refere à interface, organização da informação / conhecimento e componentes para o universo SOC específico), da especificidade aprofundada da lista de critérios e da distinção entre serviços essenciais e de *plug-in*, e requisitos

funcionais e não funcionais (relativos aos aspetos qualitativos desejáveis não necessariamente explicitados pelos requisitos funcionais).

Esses elementos devem ser levados em consideração na análise, avaliação e criação de ferramentas para o desenvolvimento de sistemas de organização do conhecimento tecnológicos.

Conclusões

O estudo de modelos de referência para o desenvolvimento de SOC permitiu uma comparação entre o DLRM, o OAIS e ISO16363, e o MoReq2010. Cada um representa um universo SOC particular, contextualizado em bibliotecas digitais, repositórios de arquivos digitais e sistemas de gestão de informação de arquivo eletrónicos. Além de apresentar uma súmula da estrutura e modelos conceptuais de cada um, a comparação de requisitos / critérios mostrou que os critérios DLRM são centrados em entidades / domínios, as métricas MoReq2010 são centradas em serviços / funções do sistema tecnológico (mesmo os requisitos não funcionais são declarados em relação ao serviço, portanto a função), e os requisitos da ISO16363 cobrem a infraestrutura organizacional, gestão de objetos digitais e gestão da infraestrutura técnica e de risco de segurança.

Outra constatação importante é que o desenvolvimento de sistemas de organização do conhecimento tecnológicos deve considerar aspetos distintos que são apresentados pelas diferentes ferramentas analisadas. Esses aspetos devem ser considerados ao analisar, avaliar e criar ferramentas para ajudar no desenvolvimento, avaliação e certificação de SOC.

Parece faltar uma camada de Organização do Conhecimento, uma vez que esses documentos mostram a necessidade de camadas relacionadas com entidades, informação/objetos de dados digitais ou físicos e informação registada. Isso sugere que há espaço para o desenvolvimento de modelos de referência de SOC que aproveitem os aspetos destacados nesta análise e, igualmente importante, apresentem critérios que levem em consideração a camada superior da Organização do Conhecimento, como um guarda-sol que inclua as demais camadas.

Referências

- Bratková, E., & Kucerová, H. (2014). Knowledge Organization Systems and Their Typology. *Revue of Librarianship*, 25(2), 1-25.
- Candela, L., Athanaspoulos, G., Castelli, D., El Raheb, K., Innocenti, P., Ioannidis, Y., Katifori, A., Nika, A., Vullo, G., & Rosse, S. (2011a). *Digital Library Reference Model In a Nutshell*. DL.org Consortium. http://www.dlorg.eu/uploads/Booklets/booklet21x21_nutshell_web.pdf
- Candela, L., Athanaspoulos, G., Castelli, D., El Raheb, K., Innocenti, P., Ioannidis, Y., Katifori, A., Nika, A., Vullo, G., & Rosse, S. (2011b). *The Digital Library Reference Model. DL.org Project Deliverable*. DL.org Consortium. <https://www.coar-repositories.org/files/D3-2b-Digital-Library-Reference-Model.pdf>
- CCSDS. (2002). *Reference Model for an Open Archival Information System (OAIS)*. Consultative Committee for Space Data Systems. <https://public.ccsds.org/Pubs/650x0b1s.pdf>

- CCSDS. (2011). *Audit and Certification of Trustworthy Digital Repositories*. Consultative Committee for Space Data Systems.
- CCSDS. (2012). *Reference Model for an Open Archival Information System (OAIS)*. Consultative Committee for Space Data Systems. <https://public.ccsds.org/Pubs/650x0m2.pdf>
- CCSDS. (2020). *Reference Model for an Open Archival Information System (OAIS) Draft*. Consultative Committee for Space Data Systems. https://cwe.ccsds.org/moims/docs/MOIMS-DAI/Draft%20Documents/OAIS%20v3/650x0p21_final.doc?d=w4c3aae475b0f4ac5be15902f5c18ad0a
- Center for Research Libraries & Online Computer Library Center. (2007). *Trustworthy Repositories Audit & Certification: Criteria and Checklist (Version 1.0)*. CRL.
- DLM Forum. (2001). *MoReq: Model Requirements for the Management of Electronic Records–MoReq Specification*. DLM Forum. https://web.archive.org/web/20110720161237/http://dlmforum.eu/index.php?option=com_jotloader§ion=files&task=download&cid=55_53897762e6a9f201462ee81eda67e670&Itemid=100&lang=en
- DLM Forum. (2008). *MoReq 2: Model Requirements for the Management of Electronic Records–Update and Extension*. DLM Forum. https://web.archive.org/web/20110929085515/http://www.dlmforum.eu/index.php?option=com_jotloader§ion=files&task=download&cid=24_92cddeb6dd5f23e213b7b06db66b5fd7&Itemid=37&lang=en
- DLM Forum. (2011). *MoReq2010: Modular Requirements for Records Systems–Volume 1: Core Services & Plug-in Modules. Version 1.1*. DLM Forum.
- Hodge, G. (2000). *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files*. Council on Library and Information Resources. <http://www.clir.org/pubs/reports/pub91/contents.html>
- ISO. (2001). *ISO 15489-1:2001 Information and documentation–Records management–Part 1: General*. International Standard Organization.
- ISO. (2016). *ISO 15489-1:2016 Information and documentation–Records management–Part 1: General (2.^a ed.)*. International Standard Organization.
- Mazzocchi, F. (2018). Knowledge organization system (KOS). *Knowledge Organization*, 45(1), 54-78.

(Página deixada propositadamente em branco)

ORGANIZACIÓN Y DESCRIPCIÓN DE INFORMACIÓN EN WIKIS SEMÁNTICOS: WIKIBASE, SEMANTIC MEDIAWIKI Y CARGO

Jesús Tramullas¹, Ana I. Sánchez-Casabón², Piedad Garrido-Picazo³

¹Depto. Ciencias de la Documentación, Univ. de Zaragoza, tramullas@unizar.es,
ORCID iD 0000-0002-5374-9993

²Depto. Ciencias de la Documentación, Univ. de Zaragoza, asanchez@unizar.es,
ORCID iD 0000-0002-0908-1615

³Depto. Informática e Ingeniería de Sistemas, Univ. de Zaragoza, piedad@unizar.es,
ORCID iD 0000-0002-1750-7225

Resumen

Este trabajo revisa las características y prestaciones que deben ofrecer los wikis semánticos. Para ello, define un modelo teórico ideal, a partir del cual se desarrolla una parrilla de análisis de funcionalidades. Esta parrilla se aplica a tres aplicaciones ampliamente consideradas como wikis semánticos. Los resultados muestran que, aunque son capaces de implementar estructuras básicas de web semántico, hay diferencias en sus aproximaciones y prestaciones. Se establece como un desarrollo necesario mejorar el procesamiento de ontologías y la integración de datos enlazados que son capaces de llevar a cabo.

Palabras clave: web semántico, wiki, ontologías, Wikibase, Cargo, Semantic Mediawiki

Introducción

Los wikis, que definen tanto a una tecnología como a un modelo de trabajo, se han popularizado desde finales de la década de 1990 como las plataformas por antonomasia para la edición colaborativa. El éxito de *Wikipedia*, así como el auge de la gestión del conocimiento, durante la década de 2000, trajeron consigo una levada tasa de adopción de los wikis en las organizaciones, aplicados en diferentes contextos y con diferentes objetivos. Todo ello ha conducido a que se consideren una expresión clásica del trabajo en colaboración en entornos digitales, como herramientas que se adaptan con facilidad y elevadas prestaciones a entornos de trabajo. Ello es debido a su naturaleza dual, que permite procesar tanto información estructurada como no estructurada, pudiendo ser adaptados para prototipar estructuras de datos, ontologías o modelos de negocio. Las características técnicas y funcionalidades de los wikis han sido recogidas en Tramullas y Garrido Picazo (2009a).

Estas capacidades han facilitado su adopción como herramienta para el web semántico, y hacen posible su aplicación en diversos contextos, como portales de conocimiento, repositorios de información, o productos de documentación (Gil et alii, 2015). Desde la década de 2000, se han desarrollado y puesto en marcha diversas herramientas wiki con capacidad semántica, entendiendo como tal el uso y aplicación de modelos y tecnologías del web semántico. La funcionalidad de anotación que incorporan los wikis dota a estos de capacidades eficientes para el etiquetado semántico de documentos, y, en consecuencia, para la creación de recursos de información etiquetada y estructurada semánticamente. Ejemplos de ello son *Wikidata* o *DBpedia*, que van siendo progresivamente integradas en aplicaciones semánticas de diferentes tipos, incluyendo recursos como catálogos de bibliotecas, catálogos de archivos o bases de datos patrimoniales (Saorín y Pastor-Sánchez, 2018). *Wikidata*, el proyecto de anotación semántica colaborativa auspiciado por los proyectos Wikimedia, lanzado en 2012, se ha convertido en poco tiempo en el recurso semántico de referencia (Vrandečić y Krötzsch, 2014). Un número creciente de asociaciones y relaciones entre diferentes artículos de Wikipedia se genera automáticamente desde *Wikidata*. como por ejemplo las *infoboxes*, incorporando progresivamente relaciones semánticas entre los contenidos. (Sáez y Hogan 2018).

Las características de un wiki semántico han sido recogidas por Tramullas y Garrido Picazo (2009a: 109-110). Entre las principales deben mencionarse especialmente las características propias de los wikis, la autoría en colaboración, el uso de estándares web, la anotación de enlaces, el etiquetado en RDF, el uso de ontologías, la edición de metadatos, la navegación facetada, la interoperabilidad, y el uso de motores de búsqueda semántica. Desde mediados de la década de 2000 puede encontrarse en la bibliografía un buen número de trabajos que abordaban el wiki semántico, y presentaban diferentes herramientas y prototipos para tal fin (Bry et alii, 2012). Especial importancia tuvo en ese momento la integración y desarrollo de ontologías en los wikis semánticos, en cuanto se consideraba que la anotación de entidades, los enlaces creados entre ellas, y la generación de las tripletas RDF correspondientes, podían ser la base de la generación casi automática de ontologías de dominio (Tramullas y Garrido Picazo, 2009b).

Durante la década de 2010 se produjo en el mundo wiki un proceso de consolidación del mercado alrededor de la plataforma *MediaWiki* (Koren, 2020), así como el desarrollo y lanzamiento de *Wikidata*, sobre la plataforma de *Wikibase*. Otras herramientas diferentes que estaban disponibles para el prototipado o la implementación de wikis semánticos (como *IkeWiki*, *AceWiki*, *Platypus Wiki*...) han ido cesando en su desarrollo, lo que ha conducido a que en el momento de redactar este texto, las opciones disponibles estén basadas en extensiones o complementos, o conjunto de ellos, incorporadas a instalaciones de *MediaWiki*, como es el caso de *Semantic MediaWiki* y *Cargo*. *Ontowiki*, herramienta que se identifica también como wiki semántico, no ha sido objeto de análisis, ya que se orienta en particular a la construcción colaborativa de ontologías (Frischmuth et alii, 2015).

En contraste con este aprovechamiento creciente en productos finales, y con la bibliografía sobre prototipos y aplicaciones de wikis semánticos, son todavía escasos los trabajos que revisen comparativamente las capacidades de representación y organización del conocimiento en las herramientas para wikis semánticos (De Dauw, 2021).

Metodología

En contraste con este aprovechamiento creciente en productos finales, y con la bibliografía sobre prototipos y aplicaciones de wikis semánticos, no se identifican trabajos que revisen comparativamente las capacidades de representación y organización del conocimiento en las plataformas indicadas. La mayoría de ellas se centran o en aplicaciones específicas, o en estudios de caso, lo que dificulta la posible toma de decisiones para la implementación de proyectos sobre recursos de información semánticos. En este contexto, este trabajo tiene como objetivo principal llevar a cabo un análisis comparativo de las funcionalidades que ofrecen diferentes herramientas para implementar estructuras de conocimiento del web semántico en un wiki,

Para el desarrollo del trabajo se han seguido las siguientes fases:

i. Identificación de las características y de las funcionalidades que definen a un wiki como semántico (Tramullas y Garrido Picazo (2009a: 109-110).

ii. Desarrollo de una matriz de análisis de funcionalidades, con la cual confrontar las características y prestaciones semánticas de los wikis analizados. Estas funcionalidades, y los parámetros asociados, se han organizado en cuatro grandes categorías: a) funcionalidades wiki clásicas o estándares; b) modelado semántico y recuperación de información; c) presentación y exploración de información; y d) intercambio de información estructurada. Entre los parámetros analizados se encuentran, entre otros, los correspondientes a la creación de modelos, la importación de ontologías, la generación de RDF, y el acceso a datos enlazados. También han sido tenidas en consideración las funcionalidades referidas a la anotación en colaboración. Los criterios incluidos en la matriz de análisis se recogen en el apartado Resultados de este trabajo.

iii. Selección de las herramientas de software para wikis semánticos objeto de evaluación. Tras la revisión bibliográfica previa, y la comprobación del estado de desarrollo de las diferentes herramientas disponibles, se han seleccionado tres wikis semánticos: *Semantic MediaWiki (SMW)*, *Wikibase* y *Cargo* (tabla 1). Se trata de tres herramientas de software consideradas como semánticas, y utilizadas como plataformas de implementación y producción (WikiApiary, 2021), por lo cual, al menos en un principio, deberían cumplir con los requisitos establecidos en el modelo teórico.

iv. Ejecución del análisis comparativo de funcionalidades, mediante la contrastación contra la matriz de análisis previamente definida.

Tabla 1: wikis semánticos analizados.

Herramienta	Versión	URL de referencia
<i>Semantic MediaWiki</i>	3.2.3	https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki
<i>Wikibase</i>	1.35	https://wikiba.se/
<i>Cargo</i>	2.8	https://www.mediawiki.org/wiki/Extension:Cargo

Semantic MediaWiki (Krötzsch et alii, 2007) fue lanzado en 2005. Se trata de una extensión de *MediaWiki*, que permite añadir anotaciones semánticas a las páginas del wiki que se trate. Su progresivo desarrollo ha favorecido la creación e integración, a su vez, de extensiones propias, orientadas a la edición, anotación, recuperación y presentación usando criterios y estructuras semánticas. Es la herramienta más ampliamente utilizada para el desarrollo de wikis semánticos en el momento de escribir este trabajo.

Wikibase (Vrandečić y Krötzsch, 2014), presentada en 2012, está formada por un conjunto de extensiones para *MediaWiki*. La aplicación se organiza en dos componentes principales: por una parte está el repositorio de datos, y por otra el cliente. Emplea etiquetado en JSON en lugar de wikitexto. Estructura la información en entidades, que pueden ser elementos individuales, a los cuales se atribuyen propiedades, a través de la semántica. *Wikibase* es la plataforma sobre la cual se construye *Wikidata*, la base de conocimiento libre de los proyectos Wikimedia.

Cargo (Koren, 2015) fue lanzado en 2015, y también es una extensión para *MediaWiki*. Funcionalmente es bastante similar a *Semantic MediaWiki*, pero estructura la información utilizando una arquitectura relacional, gracias al uso de un analizador o *parser*, sin la integración directa de tecnología semántica.

Para el análisis de las funcionalidades se ha aplicado un doble protocolo. En primer lugar, se ha instalado un entorno de prueba basado en contenedores (*Docker*), en el cual se han ejecutado instancias limpias de las tres herramientas. En segundo lugar, se han identificado instancias en producción de las tres aplicaciones, que han sido exploradas para identificar y contrastar las funcionalidades operativas en un entorno real de producción.

Resultados

La matriz de análisis de prestaciones se ha estructurado en cuatro grandes categorías de actividad, correspondientes a:

i. Funcionalidades wiki estándares: se han considerados aquellas funcionalidades que caracterizan a un wiki clásico, en especial las referidas a los procesos de edición colaborativa.

ii. Modelado semántico y recuperación de información: en esta categoría se ha considerado como base la definición de modelos estructurados semánticamente, la anotación semántica, la integración de tripletas RDF, y la capacidad de recuperación de información mediante SPARQL.

iii. Presentación de información: uno de los objetivos de un wiki es el acceso directo a la información. En consecuencia, se han valorado funcionalidades de presentación visual de la información, como la combinación de entidades, o la generación automática de navegación basada en etiquetado semántico.

iv. Intercambio de información estructurada: en el desarrollo del web semántico resulta clave la capacidad para etiquetar datos en diferentes formatos y ponerlos a disposición de terceros, así como consultar de forma dinámica otros recursos semánticos, integrando la información obtenida mediante *Linked Open Data*.

Tabla 2. Matriz de análisis comparativo.

	<i>Semantic MediaWiki</i>	<i>Cargo</i>	<i>Wikibase</i>
<i>Funcionalidades wiki estándares</i>			
Edición de texto completo	SI	SI	NO
Historial de ediciones	SI	SI	SI
Edición colaborativa	SI	SI	SI
Enlaces (internos/externos)	SI	SI	SI
Ficheros adjuntos	SI	SI	NO
<i>Modelado semántico y recuperación de información</i>			
Creación de modelos	SI	SI	SI
Anotación en wikitexto	SI	NO	NO
Formularios para datos estructurados	SI	SI	SI
Tripletas RDF	SI	NO	SI
Búsqueda a texto completo	SI	SI	SI
Búsqueda estructurada	SI	SI	SI
Búsqueda SPARQL	SI (requiere complementos)	NO	SI (requiere complementos)
<i>Presentación de información</i>			
Combinación de entidades	SI	SI	NO
Campos calculados	SI	SI	NO
Presentación de formularios	SI (combinable con wikitexto)	SI	SI
Integración de gráficos	SI	SI	SI
Navegación semántica	SI	SI	SI
<i>Intercambio de información estructurada</i>			
Importación/exportación RDF	SI	NO	SI
Importación/exportación JSON	SI	NO	SI
Importación/exportación ontologías	SI	NO	SI
Importación/exportación vocabularios SKOS, DC...	SI (depende configuración específica)	NO	SI (depende configuración específica)
Integración de <i>Linked Open Data</i>	Parcial	NO	Parcial
	<i>Semantic MediaWiki</i>	<i>Cargo</i>	<i>Wikibase</i>

SMW y *Cargo* responden a todas las características deseables de edición y gestión de contenidos de un wiki clásico (tabla 2). *Wikibase*, dada su orientación específica para el tratamiento de datos estructurados, utiliza formularios (figura 1), y no permite generar páginas de wikitexto tradicionales, ni tampoco adjuntar ficheros, ya que para ello utiliza, si es necesario, los URL/URI de los mismos. No se trata de una limitación que permita afirmar que *Wikibase* no sea un wiki.

WIKIDATA

Main page
Community portal
Project chat
Create a new Item
Recent changes
Random Item
Query Service
Nearby
Help
Donate

Lexicographical data
Create a new Lexeme
Recent changes
Random Lexeme

Tools

What links here
Related changes
Special pages
Permanent link
Page information
Concept URI
Cite this page

Item Discussion

International Society for Knowledge Organization (Q1666905)

international professional association for scholars of knowledge organization ✎ edit

ISKO

[In more languages](#)
Configure

Language	Label	Description	Also known as
English	International Society for Knowledge Organization	international professional association for scholars of knowledge organization	ISKO
Spanish	No label defined	No description defined	
Catalan	No label defined	No description defined	
Galician	No label defined	No description defined	

All entered languages

Statements

instance of ✎ edit

▼ 0 references

+ add reference

professional society ✎ edit

▼ 0 references

+ add reference

+ add value

inception ✎ edit

▼ 0 references

+ add reference

+ add value

field of work ✎ edit

▼ 0 references

+ add reference

+ add value

founded by ✎ edit

▶ 1 reference

+ add value

Figura 1. ISKO en Wikidata (instancia de Wikibase. Fuente: <https://www.wikidata.org/wiki/Q1666905>).

El modelo de datos de las herramientas revisadas se fundamenta en propiedades y clases (ver tabla 2). El modelo de *SMW* y *Cargo* es de tipo flexible, ya que puede crearse y adecuarse a contextos diferentes, según las necesidades y requerimientos del proyecto. *Wikibase*, en cambio, usa un modelo más complejo, fijo, en el cual la selección y uso de las propiedades determina la descripción de la entidad. Debe tenerse en cuenta que, aunque trabajen con modelos de datos estructurados, ello no supone automáticamente que una aplicación sea semántica, ya que debe cumplir otros requerimientos. *SMW* y *Wikibase* introducen el modelo de tripletas RDF (*Resource Description Framework*). Sin embargo, lo hacen desde una perspectiva principalmente sintáctica. Esto quiere decir que están orientados a construir esquemas específicos para un contexto dado, lo que puede suponer dificultades para incorporar esquemas externos. Este problema ha sido abordado a través de la creación de extensiones complementarias que faciliten la integración. *Cargo*, en cambio, usa estructuras relacionales, usando las funcionalidades de SQL para crear la estructuras de conocimiento y los procesos de interrogación, sin utilizar tecnología semántica. En todas las herramientas el almacenamiento de información es gestionado en una estructura relacional, como *MySQL*, *MariaDB* o similar.

La interrogación y recuperación de información deben aprovechar el modelado y estructura semántica de las entidades. El lenguaje estándar para la consulta en el web semántico es SPARQL (*Protocol and RDF Query Language*). Cabe destacar que ninguna de las plataformas revisadas es capaz de responder directamente a consultas SPARQL. *SMW* necesita exportar la información a un almacenamiento externo RDF, para que resulte consultable a través de SPARQL, o añadir extensiones complementarias que ofrezcan una funcionalidad similar, como por ejemplo la API *Ask*. Para dar respuesta a esta limitación, *Wikibase* incluye en su instalación *ElasticSearch* y *SPARQL Query service*, por ejemplo. Finalmente, *Cargo* hace uso de un procesador sintáctico que transforma las consultas en sentencias clásicas de SQL, sin estar preparado para SPARQL.

La capa de presentación también muestra diferencias funcionales. *Wikibase* y *Cargo* sólo trabajan con formularios, mientras que *SMW* puede usar formularios estructurados al mismo tiempo que permite la anotación semántica directamente en la edición de contenidos. (figura 2). Para trabajar con formularios para datos estructurados, tanto *SMW* como *Cargo* requieren la instalación de la extensión *Page Forms*. Aunque la unidad de referencia para todas las herramientas es la página wiki, *SMW* y *Cargo* tienen la capacidad de presentar en la misma página diferentes entidades, mientras que *Wikibase* se ve limitada a una entidad por página. Además, *SMW* y *Cargo* son capaces de computar automáticamente el contenido de campos, lo que *Wikibase* no hace.

El uso de RDF hace posible que *SMW* y *Wikibase* puedan exportar sus contenidos a ficheros de texto con este etiquetado. Además, también son capaces de etiquetar las estructuras y los datos en JSON. *Cargo*, en cambio, al no usar estructuras realmente semánticas en sus modelos, sólo permite exportar los datos a CSV.

Otra de las características necesarias en una aplicación de web semántico es la capacidad para incorporar estructuras provenientes de fuentes externas, como otras bases de conocimiento y/o *Linked Open Data*. *SMW* y *Wikibase* incorporan entre sus funcionalidades la capacidad para importar ontologías, mientras que *Cargo* no ofrece

funcionalidades de este tipo. La documentación básica de *SMW* recoge que es capaz de importar entidades y vocabularios etiquetados en DC, RDF, SKOS, OWL y FOAF.

Los resultados obtenidos del análisis de las funcionalidades de las aplicaciones permiten afirmar que dos de las tres herramientas de software analizadas, *Semantic MediaWiki* y *Wikibase*, hacen posible implementar estructuras semánticas, cumpliendo con los requisitos y las características fundamentales para poder clasificar una herramienta de software como del web semántico. Las dos herramientas revisadas pertenecen a la que se ha dado en denominar segunda generación de wikis semánticos, que se basan en el uso de ontologías para organizar la información. Sin embargo, difieren en aspectos referidos a la creación de estructuras y de páginas (páginas y formularios frente a formularios exclusivamente), a la posibilidad de usar anotaciones semánticas, y a la recuperación de información (servicios SPARQL). Tampoco sus prestaciones son similares en el momento de capturar o aprovechar datos enlazados tomados de fuentes externas. *Cargo*, aunque se le considera como una herramienta semántica, en realidad es un conjunto de extensiones capaces de trabajar con datos estructurados, pero sin las características de web semántico necesarias para que pueda ser considerado un wiki semántico. Debe destacarse que estas herramientas requieren de la configuración de entornos que incorporen módulos o extensiones complementarias de software, para poder configurar productos y servicios de información que puedan aprovechar las ventajas inherentes al tratamiento semántico del contenido.

The screenshot shows a Semantic MediaWiki page for the inscription AK-1.1. The page includes a navigation sidebar on the left with sections like Contents, Navigation, Search, Maintenance, and Tools. The main content area is titled 'AK-1.1' and features a metadata table with the following entries:

Transliteration:	?[u]aleri[ri]enaise
Original script:	𐌱𐌰𐌶𐌰𐌱𐌰𐌶𐌰𐌶𐌰𐌶
Object:	AK:1 rock (stone) (Inscriptions: AK-1.1 , AK:1.2, AK:1.3, AK:1.4, AK:1.5, AK:1.6, AK:1.7, AK:1.8, AK:1.9, AK:1.10, AK:1.11, AK:1.12, AK:1.13, AK:1.14, AK:1.15, AK:1.16, AK:1.17, AK:1.18, AK:1.19, AK:1.20, AK:1.21)
Position:	bottom, left
Orientation:	90°
Script:	North Italic script (Magre alphabet)
Direction of writing:	sinistroverse
Letter height:	7 – 11.5 cm
Number of letters:	15
Number of lines:	1
Craftsmanship:	engraved
Current condition:	damaged
Date of inscription:	unknown
Date derived from:	
Type:	prob. votive
Language:	Raetic
Meaning:	unknown
Alternative sigla:	none

Below the metadata table, there is an 'Images' section with five thumbnails:

- Object AK-1 rock - picture area 1 with inscriptions **AK-1.1**, AK:1.2 and AK:1.3.
- Inscription **AK-1.1**.
- Tracing of inscriptions on AK-1 rock.
- Inscriptions **AK-1.1**, AK:1.2 and AK:1.3 - picture area 1.
- Inscription **AK-1.1**.

The 'Commentary' section contains the following text:

Not previously published. Examined by TIR in July 2014.
Image in MANDL 2011: Abb. 138 (photo).
Length min. 52 cm; the last character at about 80 cm above the ground. Running downward in a slight curve on a small protrusion on the very left edge of the rock wall, the leftmost one of a group of three (Bildstelle 1, with AK:1.2 and AK:1.3).
Traces of characters can be seen starting from the darker area of rock above the inscription, but no reading can be offered. The first usable group consists of the lower parts of two hastae, the second one possibly 'l' with the angle all but gone. 'ah' are fairly well legible. 'f' are unambiguous. Next, 'f' with the curve made up of three scratches, followed by 'l', is unproblematic. Close to 'l' a clearly visible hasta featuring a doubtful assortment of bars; a half-sized 's' appears to be attached to its upper end; this making no sense, it seems the most likely reading epigraphically, but is phonetically suboptimal. The next hasta, though slightly curved, is probably 'l', 'll' are unambiguous. Well legible 'ri' are the last letters which fit onto the face of the protrusion; the last two letters had to be engraved into the concave part below it and are accordingly hard to read. If the last letter is 'l', its two lower bars appear to be turned upward.

Figura 2. Inscripción del Thesaurus Inscriptionum Raeticarum (instancia de *Semantic MediaWiki*. Fuente: <https://www.univie.ac.at/raetica/wiki/AK-1.1>).

Conclusiones

Los wiki son herramientas que deben utilizarse en un entorno colaborativo, en el cual pueden desarrollar todo su potencial (Walk y Strohmaier. 2014). En su estado actual de desarrollo, ofrecen dos funciones básicas: además de trabajar con información y datos no estructurados, también procesan datos estructurados. Esto significa que, además de la información en texto continuo, hay datos en tablas y formularios. Pueden ser utilizados en una amplia variedad de escenarios, desde documentación a portales corporativos, pasando por intranets, herramientas de calidad y de gestión de proyectos, o elaboración y publicación de informes. El potencial de los wikis que tratan información estructurada, integrando tecnologías del web semántico, se ha aplicado en educación (Zablith, 2016), investigación biomédica (Lampa et alii, 2017), o entornos empresariales (Aveiro y Pinto, 2015)

Una evaluación de funcionalidades y prestaciones puede servir para determinar la capacidad teórica y funcional de una herramienta, pero no para establecer si su aplicación, en un contexto dado, será exitosa. Las herramientas de wiki semántico revisadas tienen capacidad para gestionar estructuras de datos estructurados, pero esos datos estructurados no se convierten automáticamente en web semántico; son necesarias anotaciones y etiquetado semánticos, y procesamientos posteriores para hacer posible el enlace de esos datos desde y a otros repositorios. Técnicamente, para asegurar el desarrollo futuro de los wikis semánticos será necesario mejorar y ampliar los mecanismos que permitan la integración automática de ontologías, así como el acceso a datos enlazados (*Linked Open Data*).

Referencias

- Aveiro, D., & Pinto, D. (2015). Universal Enterprise Adaptive Object Model: A Semantic Web-Based Implementation of Organizational Self-Awareness. *Intelligent Systems in Accounting, Finance and Management*, 22(1), 3-28. <https://doi.org/10.1002/isaf.1363>
- Bry, F., Schaffert, S., Vrandečić, D., & Weiland, K. (2012). Semantic Wikis: Approaches, applications, and perspectives. *Reasoning Web. Semantic Technologies for Advanced Query Answering*, LNCS 7487, 329-369. https://doi.org/10.1007/978-3-642-33158-9_9
- De Dauw, J. (2021). Semantic MediaWiki vs Wikibase vs Cargo. *Professional.Wiki*. <https://professional.wiki/en/articles/managing-data-in-mediawiki>
- Di Iorio, A., & Rossi, D. (2018). Capturing and managing knowledge using social software and semantic web technologies. *Information Sciences*, 432, 1-21. <https://doi.org/10.1016/j.ins.2017.12.009>
- Fischer, J., Gantner, Z., Rendle, S., Stritt, M., & Schmidt-Thieme, L. (2006). Ideas and improvements for semantic wikis. *European Semantic Web Conference ESWC 2006*, LNCS 4011, 650-663. https://doi.org/10.1007/11762256_47
- Frischmuth, P., Martin, M., Tramp, S., Riechert, T., & Auer, S. (2015). OntoWiki – An authoring, publication and visualization interface for the Data Web. *Semantic Web*, 6(3), 215–240. <https://doi.org/10.3233/SW-140145>
- Gil, Y., Kapoor, D., Markham, R., & Ratnakar, V. (2015). The Provenance Bee Wiki. *Proceedings of the 8th International Conference on Knowledge Capture*, ACM, 1-8. <https://doi.org/10.1145/2815833.2815847>

- Koren, Y. (2015). Cargo and the future of semantic mediawiki. *SMWCon Spring 2015*.
- Koren, Y. (2020). *Working with MediaWiki*. WikiWorks Press. <http://workingwithmediawiki.com/book/>
- Krötzsch, M., Vrandečić, D., Völkel, M., Haller, H., & Studer, R. (2007). Semantic Wikipedia. *Web Semantics*, 5(4), 251-261. <https://doi.org/10.1016/j.websem.2007.09.001>
- Lampa, S., Willighagen, E., Kohonen, P., King, A., Vrandecic, D., Grafstrom, R., & Spjuth, O. (2017). RDFIO: extending Semantic MediaWiki for interoperable biomedical data management. *Journal Of Biomedical Semantics*, 8, art. 35. <https://doi.org/10.1186/s13326-017-0136-y>
- Zablith, F. (2016). Towards a Linked and Reusable Conceptual Layer Around Higher Education Programs. En D. Mourontsev & M. d'Aquin (Eds.), *International Conference on Open Data for Education*, LNCS 9500, (pp. 86-102). https://doi.org/10.1007/978-3-319-30493-9_5
- Sáez, T., & Hogan, A. (2018). Automatically Generating Wikipedia Info-boxes from Wikidata. *Companion of the The Web Conference 2018 on The Web Conference 2018 – WWW '18*, ACM, 823–1830. <https://doi.org/10.1145/3184558.3191647>
- Saorín, T., & Pastor-Sánchez, J.-A. (2018). Wikidata y DBpedia: viaje al centro de la web de datos. *Anuario ThinkEPI*, 12, 207. <https://doi.org/10.3145/thinkepi.2018.31>
- Tramullas, J., & Garrido Picazo, P. (2009a). Wikis semánticos. En L. Codina, M. C. Marcos, & R. Pedraza (Eds.), *Web semántica y sistemas de información documental* (pp. 99-126). Trea.
- Tramullas, J., & Garrido Picazo, P. (2009b). Organización de información en herramientas wiki: aplicación de ontologías en wikis semánticos. *IX Congreso ISKO España, Nuevas perspectivas para la difusión y organización del conocimiento*, Univ. Politécnica de Valencia, 605-612.
- Vrandečić, D., & Krötzsch, M. (2014). Wikidata: A free collaborative knowledgebase. *Communications of the ACM*, 57(10), 78-85. <https://doi.org/10.1145/2629489>
- Walk, S., & Strohmaier, M. (2014). Characterizing and predicting activity in semantic mediawiki communities. En *SWCS 2014 Semantic Web Collaborative Spaces CEUR Workshop Proceedings*, 1275:21-34. http://ceur-ws.org/Vol-1275/swcs2014_submission_3.pdf
- Wikiapiary (2021). https://wikiapiary.com/wiki/Main_Page

O USO DA INTELIGÊNCIA ARTIFICIAL NA ORGANIZAÇÃO DO CONHECIMENTO: UMA REVISÃO SISTEMÁTICA

Roberta Jerônimo da Silva¹

¹Universidade do Porto, Faculdade de Engenharia, rojeronimo17@gmail.com,
ORCID iD 0000-0002-3004-5029

Resumo

Inteligência Artificial (IA) é um tema presente em muitos segmentos da sociedade. Termo concebido por John McCarthy, em 1956. O século XX teve como um dos maiores impactos da revolução informacional o aumento de dados, informações e conhecimentos produzidos e utilizados por empresas, instâncias governamentais, ou em determinados campos de atuação profissional. Como reflexo desse evento, evidenciou-se que o armazenamento, a organização e o manuseio desses dados, informações e conhecimentos, unicamente através de atividades humanas ou com tecnologias limitantes, havia se tornado algo crítico, resultando em prejuízos e comprometimento dos processos para a organização e o acesso dos dados, informações e conhecimentos, ou ainda de uma possível perda dos mesmos. As consequências disso poderiam resultar em falhas e prejuízos na tomada de decisão por parte das empresas, órgãos governamentais e outros campos. A IA tem como propósito criar mecanismos e/ou máquinas tecnológicas que tenham a competência de realizar atividades humanas: como as capacidades de raciocinar, perceber, compreender a linguagem, aprender, dentre outras atividades realizadas pelos humanos. Ela veio para facilitar e tornar menos complexas inúmeras atividades, deixando-as de rápida execução. Faz-se necessário verificar como o uso da IA pode, de fato, colaborar para que empresas, governos e outras instâncias da sociedade civil possam se utilizar, de maneira eficaz, do acúmulo de dados, informações e conhecimentos gerados na contemporaneidade. Neste estudo, busca-se identificar como tem ocorrido a aplicação da IA na Organização do Conhecimento (OC) e na Ciência da Informação (CI). Assim, o objetivo é identificar a evolução da IA na OC entre (2011-2020) e se realmente a IA tem a capacidade de ocupar o lugar do humano. Realizou-se uma revisão sistemática com 12 artigos de pesquisa coletados na Library, Information Science & Technology Abstracts (LISTA) e na Web of Science (WoS) utilizando o modelo PRISMA. Os resultados apontaram um número reduzido de aplicação/estudos, apesar de apresentar um aumento no último ano. Os assuntos identificados foram: indexação automática, relacionada a Recuperação da Informação (RI); ligação entre OC e Tecnologia da Informação (TI), no âmbito da teoria e prática; sistemas e serviços de OC e Recuperação do Conhecimento (RC). Tendo os temas, que prevaleceram referem-se a ontologia e semântica. Temáticas que ultimamente têm gerado um interesse do meio

acadêmico, uma vez que a ontologia é um recurso importante para a OC e principalmente no contexto da web semântica. Com as tecnologias, as ontologias desenvolvem-se como um novo padrão de representação da informação, principalmente nos ambientes digitais. Indica-se, no entanto, que em relação a conceituação, ainda permanece uma certa indefinição. Ademais, as técnicas de IA ainda são insuficientes para todo tipo de objeto informacional, uma vez que a tecnologia da IA não tem habilidade de trabalhar e lidar com problemas de conhecimento, problema denominado de lacuna semântica. O estudo, por fim, aponta para indícios de que a IA facilita e aperfeiçoa as atividades, mas necessita trabalhar em conjunto com as técnicas/ferramentas tradicionais, teorias e os métodos humanos da OC. Conclui-se que o uso exclusivo da IA ainda não é uma realidade.

Palavras-chave: Organização do Conhecimento, Inteligência Artificial, Revisão sistemática.

Introdução

A IA é um tema presente em muitos segmentos da sociedade. Acredita-se que de alguma forma ela está nas rotinas diárias das pessoas seja nos espaços da vida pessoal e/ou profissional.

Termo criado por John McCarthy em 1956, durante a conferência realizada na instituição estadunidense Dartmouth College. A conferência demonstrava no período que em um futuro próximo os computadores teriam a aptidão de copiar o pensamento humano. (Carr, 2011; Oliveira, 2020). No entanto, Tegmark (2019) aponta que alguns fundadores da área: John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon mostraram-se confiantes nas previsões onde acreditavam que dentro de poucos meses com uso de computadores arcaicos teriam a capacidade de resolver obstáculos em relação à linguagem relacionada a abstrações e conceitos enfrentados pelos seres humanos. O que de fato não ocorreu. Tegmark (2019) afirma que provavelmente uma IA superior a humana não ocorrerá neste século.

IA é uma área da Ciência da Computação e também encontra-se no limite de outras áreas científicas: Matemática, Psicologia, Filosofia, Biologia e Medicina (Oliveira, 2019). O objetivo dela é criar mecanismos e/ou máquinas tecnológicas que tenham a competência de realizar atividades humanas, como a capacidade de raciocinar, perceber, compreender a linguagem, aprender, ver, ouvir e efetuar atividades das mais diversas efetuadas pelos humanos. Ela veio para facilitar e realizar atividades complexas de forma mais rápida. “A inteligência artificial impactou muitas áreas do ser humano atividade, em parte devido à velocidade com que pode processar informações em um mundo sobrecarregado, economizando esforço humano e aparentemente oferecendo uma maneira mais objetiva de avaliar e responder a uma variedade de situações” (Broughton, 2019, p. 597, tradução da autora).

IA é o resultado da junção de muitos algoritmos e sua organização de modo inteligente, “Algoritmo é um conjunto de instruções passo a passo, tão explícitas que até algo com uma ‘mente’ tão literal como um computador pode segui-las” (Polson & Scott, 2020, p. 9-10).

É essencial salientar que muitas das ideias utilizadas pela IA tem mais de séculos passados. O motivo da IA ter um destaque somente nos últimos anos é decorrente de três facilitadores:

- a) crescimento exponencial nas últimas décadas da velocidade dos computadores (denominação de Lei de Moore);
- b) crescimento cada vez maior da quantidade de dados disponíveis. Reflete na necessidade de algoritmos mais inteligentes; e
- c) computação em nuvem é um facilitador na redução de custos para armazenamento e análise de dados em larga escala através de uma infraestrutura de uma empresa (Polson & Scott, p. 2020).

Com a junção destes três facilitadores e criação de ideias, ocasionam usar a IA para aplicação na resolução de problemas concretos.

Para as instituições ou setores, que possuem um número em crescimento de dados, informações e conhecimento, sendo eles muitas vezes organizados unicamente através de atividades humanas ou com tecnologias limitantes. Acaba por dificultar e comprometer o processo de organizar e acessar os dados, informações e conhecimentos ou ainda pode suceder a perda deles. Consequentemente pode prejudicar na tomada de decisão. Sendo deste modo verificar como o uso da IA pode de fato colaborar.

Buscando identificar como tem ocorrido a aplicação da IA na OC. Onde, de acordo com Dahlberg (2006), a OC é uma ciência que organiza de modo estruturado e sistematizado os conceitos. A OC, vai além das atividades como descrição de documentos, indexação e classificação. Ela é uma área interessada nos processos e nos sistemas da OC e que encontra-se em um campo maior, a Ciência da Informação (CI). Hjørland (2008).

A CI tem como característica fundamental a sua interdisciplinaridade. O que se pode confirmar através de Le Coadic (1996) que percebe a interdisciplinaridade como “[...] uma colaboração entre diversas disciplinas, que leva a interações, isto é, uma certa reciprocidade, de forma que haja, em suma, enriquecimento mútuo”. E também por Saracevic (1995; 1996), que sustenta que a CI é naturalmente interdisciplinar. Além de outras características gerais: ligação à tecnologia da informação; participante ativa no desenvolvimento da sociedade da informação, no âmbito social e humano e na área da tecnologia. (Saracevic 1995; 1996).

Assim, por este motivo e com sentido de recolher todos registros que apresentem a relação da IA com a OC. Efetuará a pesquisa usando os termos OC e CI em conjunto com IA. Deste modo será feito uma revisão sistemática, usando o modelo PRISMA, para averiguar como ocorre a aplicação da IA na OC no período de (2011-2020), compreender a evolução da IA na OC e identificar se a presença humana chegou ao ponto de ser eliminada nas atividades da OC, o que foi confirmado que não.

Metodologia

O estudo analisa os artigos com o intuito de responder às seguintes questões:

- a) Como tem ocorrido a evolução da aplicação da IA na OC no período (2011-2020)?
- b) Quais são os assuntos de aplicação da IA na OC?

c) O uso unicamente da IA é a solução para as atividades da OC?

Efetuiu uma revisão sistemática da literatura, “[...] meio de identificar, avaliar e interpretar todas as pesquisas disponíveis relevantes para uma questão de pesquisa específica, ou área de tópico ou fenômeno de interesse” (Kitchenham, 2004, p. 1, tradução da autora), dos artigos de pesquisa sobre IA na OC entre 2011-2020. E utilizou a abordagem PRISMA, detentora de várias etapas: identificação, triagem, elegibilidade e inclusão. Iniciou duas pesquisas em separado (IA agrupada com OC e posteriormente IA com CI) para um levantamento completo da aplicação da IA na OC.

Utilizou duas bases de dados Library, Information Science & Technology Abstracts (LISTA) e Web of Science (WoS). “WoS é a única base de dados bibliométrica que cobre um século de indicadores baseados em citações para todas as disciplinas, bem como, desde 1973, todos os autores e suas afiliações institucionais” (Moed, 2005 citado por Shu et al., 2019, p. 1141; tradução da autora) e além de ser multidisciplinar. A LISTA também é uma base de dados multidisciplinar, mas tem foco na CI e tecnologia.

Selecionou artigos de pesquisa analisadas por pares, período (2011-2020), idiomas (inglês, português e espanhol). Para a etapa da leitura foi determinado que o termo *knowledge organization* e os correspondentes em português e espanhol deveriam aparecer no título do artigo ou nos termos do assunto da base de dados, ou nas palavras-chave do autor ou no resumo. Caso contrário seria excluído.

Na pesquisa IA e OC, usou os termos “artificial intelligence” AND “knowledge organization”. Na LISTA selecionou a opção **todo texto**, pois na opção **assunto** não houve resultado. Retornaram 27 (23 em inglês, 2 em português, 2 em espanhol). Em seguida, conforme a verificação da presença do termo *knowledge organization* no texto, ficaram 8 na LISTA e após a leitura permaneceram 7. Na WoS escolheu a opção **topic** (título, resumo, palavras-chave do autor e Palavras-chave Plus), retornaram 12 (11 em inglês e 1 em português). No entanto, um que estava com título em inglês, o texto estava em espanhol. Efetuando o mesmo critério da LISTA, permaneceram os 11, sendo que existiam 7 em comum, e dentre estes 1 tinha sido excluído, ficaram 4 para a leitura e permaneceram para o estudo. Para a pesquisa IA e OC ficou um total de 11 artigos.

Na pesquisa IA e CI os termos foram “artificial intelligence” AND “information science”. Na LISTA, opção **assunto**, 10. Na WoS, opção **topic**, 36. Descartou 1 por ser artigo de opinião, outro por estar na LISTA e após análise da existência do termo *knowledge organization* no texto, permaneceram 3. Sendo 2 comuns a primeira pesquisa (IA e OC). Totalizando para revisão, após as duas pesquisas, 12 artigos. Todo o processo das duas pesquisas é mostrado nos diagramas de fluxo PRISMA (Figura 1 e Figura 2).

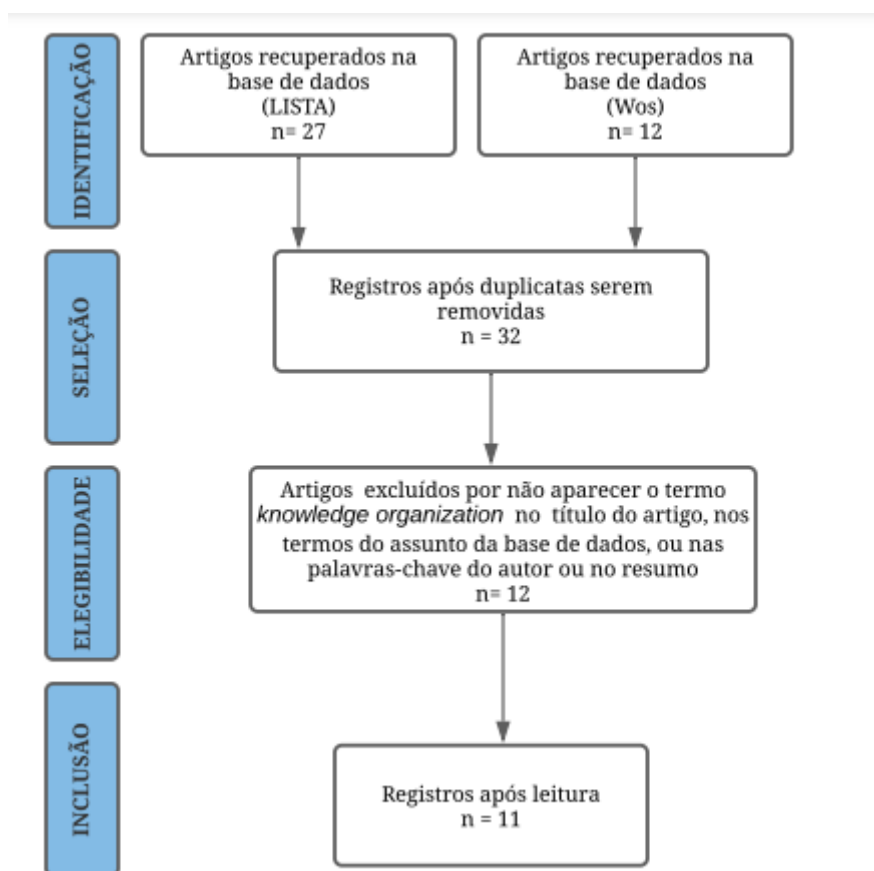


Figura 1. Diagrama de fluxo PRISMA do conjunto de pesquisa Inteligência Artificial e Organização do Conhecimento

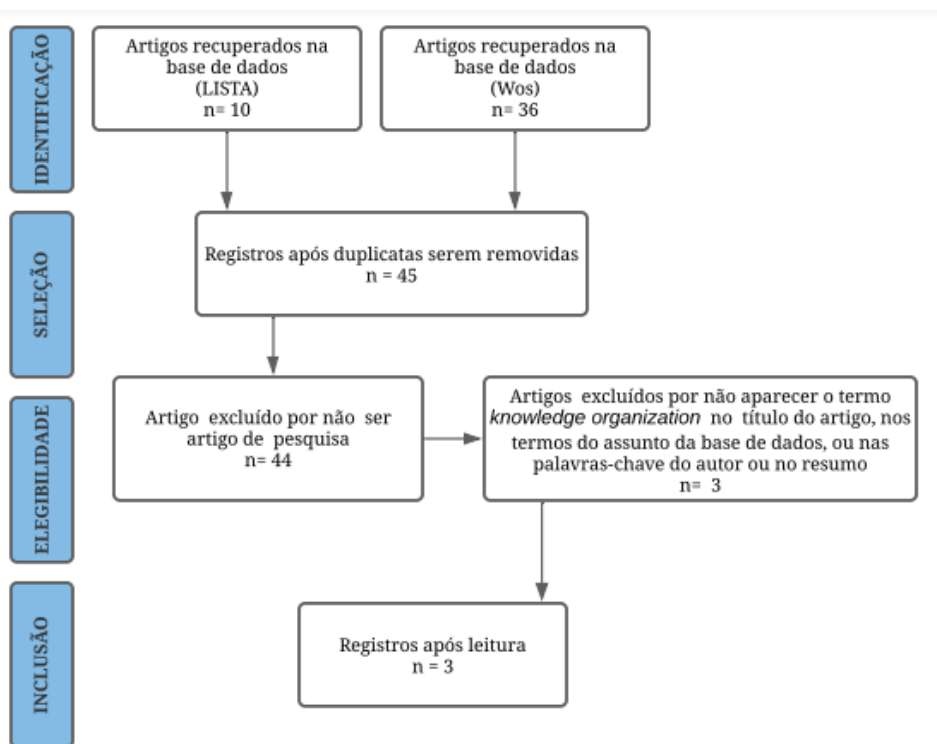


Figura 2. Diagrama de fluxo PRISMA do conjunto de pesquisa Inteligência Artificial e Ciência da Informação

Resultados e discussões

Para realização da análise dos resultados, inicialmente em relação à evolução da quantidade de publicações durante o período do estudo, classificou os 12 artigos por ano de publicação e 2 períodos de 5 anos. Em 2011-2015 (5) e 2016-2020 (7), o maior número de publicações ocorreu em 2020 (4), prevalecendo 1 documento por ano e outros sem nenhuma publicação (2012 e 2017). Observa-se um número um pouco maior de pesquisas no segundo período, mas constata que a aplicação/estudos da IA na OC permanece baixa, tendo uma média de 1,2 por ano. Destaca-se o ano de 2020, em que teve 4 estudos sobre o uso da IA.

A aplicação da IA foram em variados assuntos, prevalecendo a questão da ontologia e semântica, direta ou indiretamente (Arguello Casteleiro, 2020; Benson, 2011; Giunchiglia et al., 2014; Kåhre, 2013; MacFarlane, 2016; Qin, 2020; Xiu et al., 2020). Temáticas que ultimamente têm gerado um interesse do meio acadêmico, uma vez que a ontologia é um recurso importante para a OC e principalmente no contexto da web semântica. (Figueiredo & Almeida, 2017). E como foi descrito por Roa et al. (2014) baseado em (Uschold e Jasper, 1999) “Ontologias são usadas em diferentes campos da administração, ciência e engenharia para modelar fenômenos, situações ou domínios de interesse. São uma base para a transferência de conheci-

mento, desenvolvendo a compreensão de um domínio, ou estimulando o objeto que representam. O uso de ontologias foi generalizado para sistemas de computador e humanos. Por exemplo, ontologias podem facilitar a comunicação entre humanos, a interoperabilidade entre sistemas de computador, e também melhorar a especificação, confiabilidade e capacidade de reutilização de sistemas de computador”. (Roa et al., 2014, p. 1, tradução da autora).

Com as tecnologias, as ontologias desenvolvem-se como um novo padrão de representação da informação, principalmente nos ambientes digitais. Verificou-se em Benson (2011) que os sistemas de OC baseados em ontologia em conjunto com uso de IA nas práticas de catalogação de bibliotecas, permite resultados mais significativos na descrição de fotografias. E o outro fator é que sistemas de OC para fotografias possuem a aptidão de raciocinar a respeito de conceitos e relações, o que não ocorre nos sistemas que fazem uso exclusivamente de palavras. Assim, o autor descreve problemas nos modelos tradicionais que representam as fotografias em sistemas de informação de biblioteca. Portanto, trabalhando em conjunto com a ontologia de fotografia, teoria geral de conteúdo de fotografias e sistemas de base de conhecimento. Estas ferramentas proporcionam uma semântica de alta qualidade tanto para a leitura dos humanos quanto para o processamento das máquinas de modo eficaz (Benson, 2011).

No artigo de MacFarlane (2016), ele afirma que a recuperação de informações de texto por meio da indexação automática apresenta um bom resultado. Onde faz uso de palavras-chave e recursos estatísticos para classificar o documento conforme a importância para o utilizador ao contrário dos documentos multimídia. No entanto, funciona para alguns tipos de aplicações, mas para outros não, apresentando um sucesso limitado, problema este conhecido como “lacuna semântica”. Portanto, levanta questões se os computadores teriam a capacidade de indexar automaticamente e o que é preciso do conhecimento humano pelo *software* para fazer e substituir ou precisará de mais humanos.

A proposta é usar técnicas de IA, apesar de certo êxito, mas em relação a conceitualização é complexo. O autor aponta que é difícil todo o conhecimento que um indexador adquiriu ao longo do tempo sobre o assunto que trabalha ser transmitido para outras pessoas, como também para *softwares*. Para a solução de obstáculos complexos, como é o caso da RI de multimídia, os resultados são lentos e pouco significativos.

Outra dificuldade encontra-se relacionada aos problemas dos aspectos semânticos do conteúdo, não existindo até o momento tecnologias que identificam conceitos intrínsecos. Consequentemente é sugerido o uso de ontologias, sendo que para isto precisa do trabalho humano, distanciando portanto da indexação automatizada. Assim, o autor afirma que a lacuna semântica permanecerá por algum tempo. As técnicas de IA ainda não são suficientes, uma vez que a tecnologia da IA não tem habilidade de trabalhar e lidar com problemas de conhecimento.

Indexação automática, relacionada a RI, além de ser investigada por MacFarlane (2016), foi o conteúdo da pesquisa de Broughton (2019). Os outros temas presentes nos demais artigos foram: ligação entre OC e Tecnologia da Informação (TI) no âmbito da teoria e prática (Martínez-Ávila et al., 2014; Martínez-Ávila, 2015), sistemas e serviços de OC (Benson, 2011; Bo & Meifang, 2020), aplicação de Recuperação do Conhecimento e IA para melhorar a OC (Qin, 2020).

Qin (2020) argumenta sobre utilidades na aplicação de RC e IA para melhorar as funções de OC. Além de apontar que as ontologias e dados relacionados com padrões de OC e RC propicia base teórica e metodológica para inovar OC na era da IA. Destaca que as comunidades de OC e IA evoluíram de forma separadas, no entanto esta separação vem diminuindo e ambas têm se reunido para a evolução em tecnologias de web semântica e ciência de dados (*data science*). Apesar das técnicas antigas funcionarem e serem imprescindíveis, elas não têm como atender a quantidade exponencial de dados e informações digitais. Qin (2020) verifica objetivos variados dos paradigmas de OC e RC, no entanto existem semelhanças em métodos e funções. E afirma que com “Essa nova descoberta sugere que a comunidade OC pode examinar as semelhanças de metodologia e função com mais detalhes para identificar quais novas oportunidades podem existir para a comunidade OC causar um impacto na IA”. (Qin, 2020, p. 11, tradução da autora). Assim, infere que é notório o desenvolvimento da RC em IA disponibilizar extensa vantagens para a OC, apesar de ser preciso de mais pesquisas para identificar o caminho da OC unindo com a RC.

Na Tabela 1 é apresentado os assuntos e suas quantidades respectivas. Alguns artigos por tratar de mais de um tema foi inserido em mais de um, com isto o número ultrapassou o total dos 12 artigos.

Tabela 1. Assuntos dos artigos

ASSUNTOS	QUANTIDADES
Ontologia e semântica	07
Indexação automática	02
Teoria e prática (OC e TI)	02
Sistemas e serviços de OC	02
Recuperação do Conhecimento	01

E dentro destes temas, prevalecem direta ou indiretamente ligados na área da Biblioteconomia, campo pertencente a CI. Seja através de ser trabalhado o material no ambiente das bibliotecas e arquivos (Benson, 2011), e usando técnicas e/ou ferramentas da área (Benson, 2011; Broughton, 2019; Giunchiglia et al., 2014; Kähre, 2013; MacFarlane, 2016) ou por aversão ao termo, por causa da ligação da OC com a Biblioteconomia tradicional, a do papel. Em que muitos dos conceitos teóricos de origem da OC foram para TI, mas a TI não quer a OC como parâmetro (Martínez-Ávila et al., 2014; Martínez-Ávila, 2015) e novas possibilidades para OC, naturalmente para os profissionais da Biblioteconomia, como para CI como um todo. (Qin, 2020).

Muitas das novas tecnologias relacionadas na área da IA desprezam as teorias e os métodos humanos. “É importante ter mais do que mero conhecimento do assunto ou habilidades de TI para fazer o trabalho de OC. Existe todo um resultado de pesquisas e teorias em OC (muitas delas das ciências humanas e sociais) que podem ajudar a resolver muitos dos problemas atuais com marcação social e outras práticas colaborativas”. (Martínez-Ávila, 2015, p. 494, tradução da autora).

No artigo de Martínez-Ávila, San Segundo & Zurian apontam alguns pontos em relação aos desafios e oportunidades para os profissionais em OC na ligação com as TI, destaca-se no artigo, a atribuição de termos. Onde precisa de teoria tanto para os humanos quanto para as máquinas. Pois ambos podem apresentar resultados ineficientes por falta de teoria e assim “Uma formação correta das teorias e paradigmas na indexação e atribuição de termos poderia beneficiar tanto os processos de classificação e indexação humanos quanto o desenvolvimento de ferramentas de atribuição automática” (Martínez-Ávila, San Segundo & Zurian, 2014, p. 9, tradução da autora).

Observado também artigos no âmbito da saúde (Arguello Casteleiro et al., 2020; Xiu et al., 2020). Os tipos de materiais estudados foram literatura biomédica (Arguello Casteleiro et al., 2020) e registros médicos eletrônicos (Xiu et al., 2020). Nos demais artigos, outros objetos informacionais foram investigados, como por exemplo: bases de dados (Broughton, 2019), fotografias (Benson, 2011), documentos multimídia (MacFarlane, 2016), músicas (Souza & Café, 2018), e suprimentos portuários (Bo & Meifang, 2020).

Em relação ao objeto informacional música, trabalhado em Souza & Café (2018). Os autores apontam ser um assunto complexo no âmbito da OC. E para este tipo de objeto o uso de técnica favorável originária da IA e da linguística, Análise de Sentimento (AS), conhecido também como Mineração de Opinião ou IA emocional. Técnica esta que tem como finalidade identificar opiniões e emoções em textos, efetuando-as como positivas ou negativas através de processamento automático da linguagem natural (PLN). Identifica-se que apesar de ser uma novidade na OC, ela tem sido aplicada. Assim, Souza & Café (2018) em sua pesquisa descrevem a experiência com letras de música.

O objetivo é viabilizar aos pesquisadores uma valiosa fonte de dados com intuito de serem efetuados estudos que trabalhem com as questões da linguística e emocional da informação musical. Uma vez que, “Do ponto de vista linguístico, observar os aspectos morfológicos, sintáticos e semânticos das letras nos ajudam a identificar estas visões de mundo carregadas de emoção e beleza. Percebe-se, portanto, que o sentimento é um elemento fundamental para os estudos sobre o funcionamento linguístico das letras de música. Com isto em mente, iniciou-se esta pesquisa cujo objetivo foi tratar morfossintaticamente o léxico das letras de música e classificá-las segundo a dimensão emocional”. (Souza & Café, 2018, p. 277, tradução da autora). Ou seja, uma forma de OC.

Em Bo & Meifang (2020) relatam que com o uso do *big data* e IA, utiliza a tecnologia como modo de obter auxílio da teoria de sistemas de serviços de conhecimento para criação de modelo de serviço de conhecimento da cadeia de suprimentos portuários. Onde o armazenamento automático de conhecimento do sistema é também uma forma para OC inicial conforme as pesquisas e utilização da empresa. Concluem que com a união do conhecimento, processamento de linguagem natural e tecnologias de *big data* e IA nos serviços de conhecimento da cadeia de abastecimento portuário é de grande relevância para o desenvolvimento de empresas no ramo.

As pesquisas apontam que a utilização da IA na OC facilita, aprimora e inova na realização das atividades. Como também oportuniza novas opções de pesquisa em

OC (Souza & Café, 2018). No entanto, seu uso exclusivo ainda não é uma realidade. Técnicas tradicionais e teorias da OC são necessárias, como apontado por (Benson, 2011; Kahre, 2013; Martínez-Ávila et al., 2014; Martínez-Ávila, 2015). E mesmo no decorrer dos anos dos estudos, nota-se que o trabalho híbrido é uma realidade ainda atual. (Broughton, 2019).

Trabalho híbrido é a conclusão que chega Broughton (2019). Ela realiza uma análise sobre questões éticas na IA e na OC no contexto da indexação automática na RI em banco de dados. O estudo salienta que apesar das máquinas terem a capacidade de analisar, categorizar, indexar e classificar documentos e informações de modo automaticamente, não possui ainda a moral e ética inseridas. Portanto o modelo melhor a ser aplicado ainda é o híbrido. Onde deverá utilizar o vocabulário controlado para aperfeiçoar o funcionamento do sistema e a necessidade da intervenção do humano, mas levanta a expectativa de que no futuro as máquinas possuam comportamento ético e moral, não havendo nenhuma mediação dos humanos e assim não serem subjetivos como os humanos.

Conclusões

A IA oportuniza novas oportunidades em vários segmentos: pesquisa, aprendizado e profissional. Como também em uma diversificação de objetos informacionais da OC. A IA está para auxiliar o trabalho humano, onde ainda prevalece o trabalho em conjunto, máquina e homem.

Em alguns casos o uso da IA é limitada, como foi relatado por MacFarlane (2016) em relação ao problema da “lacuna semântica”, assim, “É nossa visão que os profissionais da organização do conhecimento, em geral, e tesouros, em particular, serão necessários para apoiar a recuperação de informação multimídia para o futuro previsível, isto é, indexadores humanos para classificar ou marcar objetos multimídia para que o utilizador pode encontrá-los na pesquisa”. (MacFarlane, 2016, p. 183, tradução da autora). E também constatado a importância do uso das técnicas/ferramentas tradicionais, teorias e os métodos humanos da OC na IA. Visto que auxilia tanto os processos humanos como também o avanço de ferramentas tecnológicas.

Conclui-se que apesar de toda a evolução na IA, o uso unicamente da IA nas atividades na OC, ainda não é uma realidade.

A principal limitação deste estudo diz respeito à quantidade de artigos recuperados. O que pode apresentar implicação em relação aos assuntos identificados nos documentos. Deste modo, recomenda-se uma investigação mais abrangente, utilizando outras bases de dados pertinentes.

Referências

Arguello Casteleiro, M., Demetriou, G., Read, W., Fernandez Prieto, M., Maroto, N., & Maseda Fernandez, D., Nenadic, G. Klein, J., Keane, J. & Stevens, R. (2020). Semantic Deep Learning: Prior Knowledge and a Type of Four-Term Embedding Analogy to

- Acquire Treatments for Well-Known Diseases. *Journal Of Biomedical Semantics*, 9 (1), e16948. <https://doi.org/10.1186/s13326-018-0181-1>
- Benson, A. (2011). OntoPhoto and the Role of Ontology in Organizing Knowledge. *Knowledge Organization*, 38 (2), 79-95. <https://doi.org/10.5771/0943-7444-2011-2-79>
- Bo, Y., & Meifang, Y. (2020). Construction of the knowledge service model of a port supply chain enterprise in a big data environment. *Neural Computing And Applications*, 1-12. <https://doi.org/10.1007/s00521-020-05044-w>
- Broughton, V. (2019). The Respective Roles of Intellectual Creativity and Automation in Representing Diversity: Human and Machine Generated Bias. *Knowledge Organization*, 46 (8), 596-606. <https://doi.org/10.5771/0943-7444-2019-8-596>
- Carr, N. (2011). A igreja da Google. In N. Carr, *A geração superficial: o que a Internet está fazendo com nossos cérebros* (pp. 162-191). Agir.
- Dahlberg, I. (2006). Knowledge Organization: A New Science?. *Knowledge Organization*, 33 (1), 11-19. https://www.ergon-verlag.de/isko_ko/downloads/ko3320061c.pdf
- Figueiredo, F. de C., & Almeida, F. G. (2017). Ontologias em ciência da informação: um estudo bibliométrico no Brasil. *Ciência da Informação*, 46 (1), 23-33. <http://revista.ibict.br/ciinf/article/view/4011>
- Giunchiglia, F., Dutta, B., & Maltese, V. (2014). From Knowledge Organization to Knowledge Representation. *Knowledge Organization*, 41 (1), 44-56. <https://doi.org/10.5771/0943-7444-2014-1-44>
- Hjørland, B. (2008). What is Knowledge Organization (KO)?. *Knowledge Organization*, 35(2-3), 86-101. <https://doi.org/10.5771/0943-7444-2008-2-3-86>
- Kåhre, P. (2013). Library and information science's ontological position in the networked society: using new technology to get back to an old practice. *Information Research-an International Electronic Journal*, 18(3). Retrieved 25 February 2021, from http://informationr.net/ir/18-3/colis/paperC13.html#_YDfyt-j7TIU
- Kitchenham, B. (2004). – Procedures for performing systematic reviews (Joint technical report No. Keele university technical report TR/Se-0401 // NICTA technical report 0400011T.1). <https://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf>
- Le Coadic, Y. (1996). *A Ciência da Informação*. Briquet de Lemos.
- MacFarlane, A. (2016). Knowledge Organisation and Its Role in Multimedia Information Retrieval. *Knowledge Organization*, 43 (3), 180-183. <https://doi.org/10.5771/0943-7444-2016-3-180>
- Martínez-Ávila, D. (2015). Knowledge Organization in the Intersection with Information Technologies. *Knowledge Organization*, 42 (7), 486-498. <https://doi.org/10.5771/0943-7444-2015-7-486>
- Martínez-Ávila, D., San Segundo, R., & Zurian, F. (2014). Retos y Oportunidades En Organización Del Conocimiento En La Intersección Con Las Tecnologías de La Información. *Revista Española De Documentación Científica*, 37 (3), e053. <https://doi.org/10.3989/redc.2014.3.1112>
- Oliveira, A. (2019). *Inteligência Artificial*. Fundação Francisco Manuel dos Santos.
- Polson, N. & Scott, J. (2020). *Inteligência Artificial*. Vogais.
- Qin, J. (2020). Knowledge Organization and Representation under the AI Lens. *Journal Of Data And Information Science*, 5 (1), 3-17. <https://doi.org/10.2478/jdis-2020-0002>
- Roa, H., Sadik, S., & Indulska, M. (2014). Ontology usefulness in human tasks: seeking evidence. In *Proceedings of the 25th Australasian Conference on Information Systems, 8th – 10th December, Auckland, New Zealand* (1-10). Auckland, New Zealand. <http://hdl.handle.net/10292/8178>

- Saracevic, T. (1996). Ciência da Informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, 1 (1), 41-62. <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/viewFile/235/22>
- Saracevic, T. (1995). Interdisciplinary nature of information science. *Ciência da Informação*, 24 (1), 1-9. https://brapci.inf.br/_repositorio/2011/05/pdf_aac5068b8b_0016893.pdf
- Shu, F., Julien, C. & Larivière, V. (2019). Does the Web of Science Accurately Represent Chinese Scientific Performance? *Journal of the Association for Information Science & Technology* 70 (10), 1138-1152. <https://doi.org/10.1002/asi.24184>
- Souza, R., & Café, L. (2018). Análise de sentimento aplicado ao estudo de letras de música. *Informação & Sociedade: Estudos*, 28 (3), 275-286. <https://doi.org/10.22478/ufpb.1809-4783.2018v28n3.34884>
- Tegmark, M. (2019). *Life 3.0: ser-se humano na era da Inteligência Artificial*. D.Quixote.
- Xiu, X., Qian, Q., & Wu, S. (2020). Construction of a Digestive System Tumor Knowledge Graph Based on Chinese Electronic Medical Records: Development and Usability Study. *JMIR Medical Informatics*, 8 (10), e18287. <https://doi.org/10.2196/18287>

O GÊNERO CINEMATOGRAFICO ENQUANTO FERRAMENTA DE AUXÍLIO A CATEGORIZAÇÃO DE FILMES DE TERROR EM SERVIÇOS DE *VIDEO ON DEMAND*

Rafael Rodrigo do Carmo Batista¹, Thiago Henrique Bragatto Barros²,
Rita do Carmo Ferreira Laipelt³

¹Universidade Federal do Rio Grande do Sul, rafael.bibufsm@gmail.com,
ORCID iD 0000-0002-1978-1818

²Universidade Federal do Rio Grande do Sul, bragato.barros@ufrgs.br,
ORCID iD 0000-0001-7439-5779

³Universidade Federal do Rio Grande do Sul, ritacarmo@yahoo.com.br,
ORCID iD 0000-0002-7429-8490

Resumo

Busca-se avaliar as categorias utilizadas para a descrição de filmes de terror no serviço Video on Demand Netflix. Apresenta o conceito de gênero cinematográfico e destaca um histórico dessa teoria ao longo do desenvolvimento dos estudos cinematográficos. Apresenta trabalhos relacionados à indexação e categorização audiovisual dentro da Ciência da Informação, destacando as abordagens dos pesquisadores sobre o tratamento da informação audiovisual. A pesquisa se constitui como exploratória, pois se concentra na investigação e coleta das categorias existentes dentro da netflix. A análise foi realizada a partir da observação da interface da Netflix, destacando os elementos representativos dentro da categoria “terror”. Foram coletados os termos descritivos de 20 filmes dentro da plataforma. Os termos coletados da netflix foram comparados com os gêneros dos filmes presentes no Internet Movie Database – IMDb, através da correspondência semântica. Foi possível observar que o serviço utiliza uma seleção de termos muito próxima aos termos encontrados nos gêneros utilizados no IMDb, realizando algumas adaptações para evitar a ambiguidade com outras categorias presentes na plataforma. Conclui indicando a relevância dos estudos de tratamento da informação audiovisual na atualidade e que os gêneros cinematográficos podem ser utilizados para compreender a dinâmica da informação audiovisual e suas formas de representação. Destaca a Netflix como um exemplo positivo nesse contexto.

Palavras-chave: Informação Audiovisual, Representação da Informação, Categorização, *Video on Demand*, Netflix

1 Introdução

A informação presente em filmes apresenta algumas particularidades em relação aos outros gêneros documentais. A primeira delas diz respeito a como tornar o enredo de um filme em algo aproveitável, em termos de representação. A síntese de seu conteúdo não revela uma uniformidade de discurso: um filme pode adquirir significados variados de acordo com a percepção do usuário.

Filmes que apresentem violência, por exemplo, podem ser associados a qualquer abordagem, indo desde fins cômicos como excessivamente dramáticos, herança do caráter polisêmico da imagem. Essa polissemia é resolvida em primeiro lugar através da criação dos gêneros cinematográficos e de seu papel na construção de uma classificação para cada tipo.

Os gêneros cinematográficos são a base para a constituição temática da informação em filmes. Em serviços de *Video on Demand*, como a *Netflix*, as etiquetas (ou tags) gerais de filmes são utilizadas para categorizar o conteúdo presente na plataforma, de modo a fornecer ao espectador experiências adequadas às suas expectativas. A pesquisa busca, nesse sentido, avaliar se essas etiquetas utilizadas pela Netflix são adequadas para representar filmes de um gênero específico – o terror – a partir da correspondência semântica entre elas e as etiquetas utilizadas para representar esses mesmos filmes no Internet Movie Database – IMDb. Avaliar esse processo pode fornecer subsídios para o trabalho de classificação e etiquetagem de recursos audiovisuais.

2 Gênero cinematográfico e a construção temática do filme

Do ponto de vista de obras como o Dicionário Teórico e Crítico de Cinema analisam o termo gênero através de duas perspectivas: a filosófica, onde a palavra gênero designa “a ideia geral de um grupo de seres e objetos que possuem caracteres comuns” (Aumont; Marie, 2003, p. 141) e o cinematográfico, empregado desde o século XVII, que consiste em uma “categoria de obras que têm caracteres em comum (de enredo, de estilo, etc)” (Aumont; Marie, 2003, p. 141). O dicionário continua o desenvolvimento do verbete indicando que os gêneros sempre foram presentes no desenvolvimento da arte, mas sua definição nunca foi fixa, sendo produto de “flutuações” (Aumont; Marie, 2003).

2.1 Histórico da teoria dos gêneros

Altman (2000) se debruça com mais profundidade sobre a evolução histórica da chamada “teoria clássica dos gêneros”. Sua análise parte da ideia de que os gêneros cinematográficos “em muitos aspectos, são um prolongamento dos estudos sobre gêneros literários” (Altman, 2000, p. 33, tradução nossa). Assim, segundo Altman, “os estudos de gênero cinematográfico têm se constituído, pelo menos nas últimas décadas, como um terreno a parte dos gêneros literários, desenvolvendo seus próprios postulados, *modus operandi*, e seus próprios métodos de estudo (Altman, 2000, p. 34).

A constituição desse gênero enquanto universo temático consiste basicamente em uma série de características compartilhadas. Por exemplo, um filme que apresenta seu enredo em cenários inóspitos do Oeste, com cowboys e saloons é associado ao gênero faroeste (*western*). Filmes com monstros, fantasmas e elementos fantásticos são geralmente associados ao horror.

Existem ainda elementos mais discretos como o uso da violência: filmes muito violentos são considerados como cinema extremo (próximo do horror), enquanto filmes de ação se apropriam da violência (em menor escala) como recurso estilístico. São esses elementos que trazem a riqueza da informação audiovisual, bem como a dificuldade de representá-la. De que forma a violência pode ser encarada como extrema ou não? Um filme cujo enredo se centra na relação entre um casal pode ser considerado como algo além de um romance? Essas questões impactam diretamente na forma com que esses filmes são classificados em sistemas de busca.

Se os gêneros cinematográficos surgiram a partir de consequências relacionadas à distribuição de filmes, os serviços de *Video on Demand* – VOD nascem a partir de um outro tipo de demanda e contexto. A distribuição desses filmes em diversos tipos de mídia dependia, inicialmente, das videolocadoras. Esses estabelecimentos ofereciam, através de aluguéis, um acervo rico de filmes e gêneros. Contudo, não havia sistemas de buscas robustos e muitas vezes as pesquisas dos espectadores dependiam do atendente desse local. Os encartes dos VHS, por exemplo, continham sinopses, mas essas sinopses não informavam com exatidão o conteúdo ou tipo de filme. A figura do atendente servia como um mediador do conteúdo fílmico, indicando com mais exatidão para um espectador qual filme seria de seu interesse.

3 O terror enquanto gênero cinematográfico: características e histórico

O cinema de terror surge nos primórdios da sétima arte, acompanhando as primeiras experiências de utilização do cinematógrafo. O ilusionista francês Georges Méliès foi o inventor do cinema fantástico (Primati, 2014, p. 11), utilizando a fantasia e a mágica para iludir e impressionar os espectadores. Cousins afirma que “seu papel nos primórdios do cinema não foi negligenciado” (2006, p. 3).

Nos anos posteriores, diretores europeus como Fritz Lang, F.W. Murnau e Robert Wiene, interessados em “aspectos reprimidos e primitivos dos seres humanos” (Cousins, 2006, p. 95) e se inspirando no movimento expressionista alemão criam filmes com uma estética única e marcante, abordando temas como a loucura, a ficção científica e o estranho. A partir de 1928, o cinema de terror entra em uma fase de grande prosperidade e popularidade. Surgem os monstros clássicos que, através da *Universal Studios*, fundam as características mais fundamentais do gênero.

A partir da década de 30, o gênero terror seria então caracterizado e explorado como um produto comercial. Tal processo segmentou o terror em subgêneros, com variadas especificidades. A divisão por subgêneros permite estabelecer critérios mais objetivos para um espectador. Por ser um gênero muito fluido, o terror é apreciado por pessoas que gostam de todos os outros gêneros (como o romance e a comédia) assim como pessoas que não gostam de violência extrema. Alguns subgêneros relacionados a monstros (como vampiros e criaturas pequenas) refletem filmes

que alcançaram sucesso comercial, como o subgênero *Aliens*, que surgiu na esteira de produções consagradas como *Alien* e *Predator*. Outros subgêneros como *Technology*, *Horror Romance*, *Gothic Horror* e *Comedy Horror* indicam a fluidez do terror entre outros gêneros populares no cinema.

4 Indexação e categorização de recursos audiovisuais: abordagens da ciência da informação

O processo de indexação de um filme é sensivelmente diferente de outros gêneros documentais, por características evidenciadas pela literatura da Ciência da Informação. Antes de partir para essas diferenças, se faz necessário apontar alguns percalços conceituais presentes dentro do contexto dos estudos sobre indexação no geral.

Bentes Pinto (2001, cf. Santos et al, 2018, p. 250) considera que para “um determinado vídeo seja inserido em um sistema de recuperação da informação, seja ele um catálogo online ou repositório digital, o indexador deve analisá-lo conceitualmente”. O uso da análise conceitual já aponta que a representação da informação audiovisual se diferencia por sua natureza temática em relação a outros tipos de informação.

Cordeiro (1996), ao estudar o processo de indexação aplicado aos recursos audiovisuais, apresenta a primeira complexidade relacionada: filmes podem possuir interpretações plurais, tanto por parte do autor, quanto de quem consome. Sendo o indexador um consumidor, sua metodologia pode apresentar um viés que afeta diretamente, por exemplo, o uso de um termo para representar um filme. Um indexador que não conheça o gênero *noir* pode simplesmente incluir filmes como *Blade Runner* apenas em uma categoria de ficção científica.

Se apropriando de outros princípios listados por Sorlin (1994), a autora concorda que “não se descreve informação sobre filmes não vistos, baseados em informes de imprensa” (Cordeiro, 1996, p. 5). A autora também questiona o método para se representar imagens, afirmando que “ocorrerá uma inaceitável perda de informação” (Cordeiro, 1996, p. 6).

5 Procedimentos metodológicos

A pesquisa se constituiu como exploratória, pois se concentra na investigação e coleta das categorias existentes dentro do serviço VOD estudado. A abordagem utilizada se aproxima da análise qualitativa pois aborda, em seus aspectos, o desenvolvimento das teorias pesquisadas (tanto a de gênero cinematográfico, quanto a da indexação audiovisual), além de se apropriar de um instrumento de análise próprio dessa abordagem, a análise temática.

O processo de categorização, proposto por Bardin, fornece os elementos necessários para auxiliar o processo de avaliação das categorias disponíveis nos serviços VOD. Categorias, que, de acordo com autora, “são rubricas ou classes, as quais reúnem um grupo de elementos sob um título genérico, agrupamento esse efetuado em razão das características comuns desses elementos.” (Bardin, 1977, p. 117). Bardin

define esse processo como “uma operação de classificação de elementos constitutivos de um conjunto, por diferenciação, e seguidamente, por reagrupamento segundo gênero (analogia), com critérios previamente definidos.” (Bardin, 1977 p. 117). Ela ainda situa o primeiro objetivo da categorização como “fornecer, por condensação, uma representação simplificada dos dados brutos” (Bardin, 1977, p. 119).

Os dados trabalhados foram as categorias retiradas da *Netflix*. Esses dados consistem em termos que a plataforma utiliza para categorizar filmes dentro da sua interface (exemplo: “filmes de terror assustadores” – associado a filmes como “Cemitério Maldito” na interface do serviço). Os dados de testagem foram retirados de 20 filmes incluídos na sessão “terror”. Os filmes foram selecionados na sessão “mais populares”, ou seja, os 20 primeiros filmes de terror mais populares entre os usuários da plataforma.

Os dados associados a cada filme foram retirados no campo “mais informações”, onde é possível verificar os termos descritores nos campos “gênero”. Por fim, os termos foram comparados, através da ideia de categorização proposta por Bardin, com os gêneros ou termos descritivos associados a cada filme no Internet Movie Database – IMDb. O IMDb se constitui como a base de dados mais relevante sobre cinema na atualidade, sendo utilizada como base referencial para a pesquisa. Os gêneros indexados no IMDb correspondem de maneira mais eficiente à divisão clássica presente na literatura. A comparação entre cada termo foi medida através da correspondência semântica (por exemplo, se um dos descritores da Netflix for “filme assustador” e um dos termos associados ao filme no IMDb for “susto”, a equivalência será considerada como eficiente).

6 Resultados obtidos

A natureza do processo de indexação em plataformas como a Netflix se baseia em facilitar a busca, por parte do espectador, do filme que deseja ver e de filmes semelhantes. De acordo com a página de ajuda da plataforma “se o que você está buscando não estiver disponível, a Netflix sugerirá títulos semelhantes que podem ser interessantes para você”. Indica ainda que “se os termos da busca forem muito vagos, a Netflix oferecerá sugestões de pessoas, títulos, gêneros ou diretores que possam ser relevantes”. Os tipos de busca elencados representam, de maneira geral, a principal dinâmica da plataforma. A Netflix trabalha com “cerca de 76.897 diferentes classificações para os seus filmes” (Teixeira, 2015, p. 18), apresentando-as em categorias gerais (presentes na plataforma) e códigos especiais.

Em outubro de 2020, são verificados 155 filmes do gênero “terror” na *Netflix no catálogo brasileiro*. A divisão desses filmes foi organizada, dentro da interface, da seguinte forma: **filmes baseados em livros, em alta (filmes mais vistos pelos espectadores), filmes sobre serial killers e de terror sangrento, Filmes com sobrenatural, terror e sobrevivência, filmes e séries europeus, filmes de suspense, ação e aventura, originais Netflix, populares na Netflix, comédia, filmes independentes, filmes com mulheres fortes, filmes de terror sobrenatural, terror adolescente, filmes psicológicos, lançamentos, filmes sinistros, filmes estrangeiros de suspense e sugestões baseadas em outros filmes que o espectador viu.**

Ao verificar as informações de cada filme, é possível clicar em uma janela com toda ficha técnica e informações sobre duração. As categorias são denominadas como “gêneros”, mas a Netflix ainda inclui metatags relacionadas a “cenas e momentos”, que tentam traduzir o teor das cenas através de palavras-chave:

Sobre It: A Coisa

Direção: **Andy Muschietti**

Elenco: **Jaeden Martell, Jeremy Ray Taylor, Sophia Lillis, Bill Skarsgård, Finn Wolfhard, Chosen Jacobs, Jack Dylan Grazer, Wyatt Oleff, Nicholas Hamilton**

Roteiro: **Chase Palmer, Cary Joji Fukunaga, Gary Dauberman**

Gêneros: **Filmes baseados em livros, Filmes de terror, Criaturas e monstros, Filmes de terror sobrenatural, Filmes de monstros**

Cenas e momentos: **Violentos, Assustadores**

Classificação etária: **16** **Violência, Linguagem imprópria**

Fonte: *Netflix*, 2021

Figura 1. Informações sobre o filme It

O campo “gêneros” indica os gêneros cinematográficos associados a cada obra. Esses gêneros não correspondem com exatidão aos gêneros consagrados pela literatura em cinema, mas a ajustes comerciais e temáticos que a plataforma organiza. Os gêneros são distribuídos internamente dentro da Netflix como seções, onde conjuntos de filmes com temáticas similares são agrupados. No IMDb é possível constatar um rigor maior na indexação desses gêneros, respeitando inclusive a possibilidade de associação entre filmes de terror e outros gêneros, como a ação.

Tabela 1. Correspondência semântica entre as categorias da Netflix e os gêneros do IMDb

Título do filme / Título Original	Termos coletados da Netflix	Termos coletados no IMDB	Correspondência semântica
O Ataque dos Vermes Malditos / Tremors	Filmes de ficção científica, Comédia, Filmes de terror	Comédia, Terror	Todos os gêneros do IMDb são contemplados na categorização da Netflix
Mãe! / Mother	Filmes de terror	Drama, Terror , Mistério	Apenas um gênero do IMDb corresponde a categorização utilizada na Netflix (Terror)
A Espreita do Mal / I See You	Mistério, Filmes de Suspense, Filmes de Terror	Crime, Drama, Terror, Mistério, Suspense	Apenas dois gêneros do IMDb não possuem correspondência com a categorização da Netflix (Crime e Drama)
Operação Overlord / Overlord	Filmes com temática militar, Ação e aventura, Filmes de terror	Ação, Aventura, Terror, Ficção Científica, Guerra	Um gênero do IMDb não corresponde a categorização da Netflix (Ficção Científica). O gênero “guerra” está

			próximo da categorização “filmes com temática militar”
A Visita / The Visit	Filmes de terror	Terror , Mistério, Suspense	Apenas um gênero do IMDb corresponde a categorização utilizada na <i>Netflix</i> (Terror)
O Orfanato / El Orfanato	Espanhol, Filmes de Terror	Drama, Terror , Mistério, Suspense	Apenas um gênero do IMDb corresponde a categorização utilizada na <i>Netflix</i> (Terror)
O Homem nas Trevas / Don't Breathe	Filmes de Terror	Crime, Terror , Suspense	Apenas um gênero do IMDb corresponde a categorização utilizada na <i>Netflix</i> (Terror)
Vozes e vultos / Things Heard & Seen	Filmes baseados em livros, Filmes de suspense, filmes de terror	Drama, Terror , Mistério, Suspense	Dois gêneros do IMDb não correspondem a categorização utilizada na <i>Netflix</i> (Drama e Mistério)
Fragmentado / Split	Filmes de suspense, filmes de terror	Terror, Suspense	Todos os gêneros do IMDb são contemplados na categorização da <i>Netflix</i>
A Bruxa / The VVitch: A New England Folktale	Filmes de Terror , Filmes Independentes	Drama, Terror , Mistério	Apenas um gênero do IMDb corresponde a categorização utilizada na <i>Netflix</i> (Terror)
Halloween	Filmes de Terror	Crime, Terror , Suspense	Apenas um gênero do IMDb corresponde a categorização utilizada na <i>Netflix</i> (Terror)
Resident Evil: a degeneração / Resident Evil: Degeneration	Japonês, Animes de Terror , Anime baseado em <i>games</i>	Animação, Ação, Terror , Ficção Científica, Suspense	Três gêneros do IMDb não correspondem a categorização utilizada na <i>Netflix</i> (Ficção Científica, Suspense e Ação)
Zumbilândia / Zombieland	Comédia, Filmes de terror , Filmes de terror com zumbis	Aventura, Comédia , Fantasia, Terror , Ficção Científica	Três gêneros do IMDb não correspondem a categorização utilizada na <i>Netflix</i> (Ficção Científica, Aventura e Fantasia)
O Garoto Sombrio / The Boy	Filmes baseados em livros, Filmes de Terror , Filmes Independentes	Drama, Terror , Suspense	Dois gêneros descritos no IMDb não possuem correspondência com a categorização da <i>Netflix</i> (Drama e Suspense)
Dragão Vermelho / Red Dragon	Mistério, Filmes baseados em livros, Filmes de suspense	Crime, Drama, Suspense	Apenas um gênero do IMDb corresponde a categorização utilizada na <i>Netflix</i> (Suspense)

Homunculus	Japonês, Mistério, Filmes de suspense , Filmes de terror	Drama, Fantasia, Mistério , Suspense	Dois gêneros do IMDb não possuem correspondência com a categorização da <i>Netflix</i> (Fantasia, Drama)
O Hospedeiro / Gwoemul	Coreano, Filmes de Ficção Científica , Ação e aventura , Filmes de terror , Criaturas e monstros	Ação , Drama, Terror , Ficção Científica	Apenas um gênero do IMDb não possui correspondência com a categorização da <i>Netflix</i> (Drama)
Vampiros de John Carpenter / Vampires	Filmes baseados em livros, Ação e aventura , Filmes de terror	Ação , Terror , Suspense	Apenas um gênero do IMDb não possui correspondência com a categorização da <i>Netflix</i> (Suspense)
O Bar / El Bar	Espanhol, Comédia , Filmes de Suspense	Comédia , Terror, Suspense	Apenas um gênero do IMDb não possui correspondência com a categorização da <i>Netflix</i> (Terror)
Lenda Urbana / Urban Legend	Filmes teen, Filmes de terror	Terror , Mistério, Suspense	Apenas um gênero do IMDb possui correspondência com a categorização da <i>Netflix</i> (Terror)
Guerra Mundial Z / World War Z	Filmes de ficção científica, Filmes baseados em livros, Ação e aventura	Ação, Aventura, Terror, Ficção Científica	Apenas um gênero do IMDb não possui correspondência com a categorização da <i>Netflix</i> (Terror)

Referência: elaborada pelo autores

5.1 Análise da categorização de filmes de terror na *netflix*

A Netflix categoriza os filmes de acordo com os gêneros mais populares, utilizando nomenclaturas próximas aos termos utilizados na figura 1. Entretanto, a organização segue uma lógica baseada em interpretações mais próximas de possíveis termos a serem utilizados pelos espectadores no momento da busca. A própria indicação dessas tags em “cenas e momentos” demonstra como a empresa utiliza as sensações na hora de representar esse tipo de conteúdo. Utilizando o exemplo da figura, o filme IT foi tagueado como “violento”, sendo que a violência do filme não é um elemento fundamental na construção de seu enredo ou conteúdo, sendo mais um elemento relacionado à classificação indicativa.

Os dados apresentados na tabela 1 indicam que, semanticamente, a lógica de etiquetagem da Netflix apresenta muita similaridade com os termos descritores de gênero presentes no IMDb. Contudo, a equivalência total entre termos registrados no IMDb e no serviço VOD foi verificada em apenas dois filmes. Esses termos indicam o próprio gênero principal (terror) e um gênero paralelo (suspense, comédia). A Netflix utiliza um sistema de etiquetagem baseada nos gêneros cinematográficos puros, sem indicar subgêneros relacionados. Exemplos como “filmes de terror

assustadores” podem ser encontrados no campo “cenas e momentos”, mas esses, como o próprio nome sugere, estão associados a cenas específicas. A tabela indica a ausência ou adição de alguns gêneros relacionados na Netflix em relação aos gêneros registrados no IMDb, o que pode indicar que a Netflix explora o potencial temático do seu acervo de maneira diversificada. Também poderia ser um indício de que incluir ou excluir um outro gênero pode afetar a localização de um filme dentro do catálogo, o que pode ser comercialmente desvantajoso para a plataforma. Um exemplo disso seria a ausência da etiqueta drama no filme “Garoto Sombrio” (mantendo-o dentro da sessão de terror) e na adição da etiqueta “ficção científica” no filme O Ataque dos Vermes Malditos.

A Netflix também utiliza muitas etiquetas relacionadas a outros aspectos dos filmes (como, por exemplo, filmes baseados em livros, filmes de determinado país e filmes independentes). Esses termos não se constituem como gêneros, sendo a sua atribuição outra preferência da Netflix para organizar o seu catálogo de acordo com elementos em comum entre os filmes.

Os dados apresentandos indicam que a equivalência entre os termos utilizados no IMDb e as etiquetas presentes na Netflix apresentam similaridades, ainda que as ressalvas anteriores revelem certas preferências dentro da plataforma. É mais importante redirecionar os filmes de interesse aos espectadores (e suas buscas realizadas) do que necessariamente obter todas as relações possíveis.

6 Considerações finais

Os estudos sobre representação da informação audiovisual na Ciência da Informação vêm alcançando relativo destaque em período recente na área. A indexação audiovisual, no campo do tratamento da informação, se constitui no campo de maior destaque e evidência. Entretanto, ainda não se nota de maneira efetiva a integração dos conceitos relacionados aos estudos audiovisuais com o desenvolvimento de novos processos ou práticas. A análise fílmica surge como um dos poucos marcos de integração, mas ainda não é o suficiente para o estabelecimento de novos marcos.

Os gêneros cinematográficos, nesse contexto, podem se constituir como elementos de integração na medida em que eles estabelecem fronteiras e características próprias dos filmes. No gênero terror, naturalmente multifacetado, a delimitação dos tipos de filmes presentes na literatura indica possibilidades de termos a serem atribuídos.

A categorização presente em plataformas *Video on Demand* oferece um panorama de organização da informação audiovisual que pode oferecer novas perspectivas aos pesquisadores. Por exemplo, pode indicar que termos são utilizados como etiquetas, como se dá a relação entre filmes de gêneros opostos e outros tópicos. Soma-se a outras possíveis contribuições para a pesquisa um aprimoramento da utilização de *tags* para a representação de filmes, no sentido de atuar como possível balizador dessas *tags*, equacionando uma terminologia mais próxima do usuário com um efetivo termo de representação afim de evitar ambiguidades. A *Netflix* oferece um exemplo a ser mais bem observado nesse contexto.

Referências

- Altman, R. (1999). *Film/genre*. British Film Institute.
- Aumont, J; Marie, M. (2003). *Dicionário teórico e crítico de cinema*. Papirus.
- Bardin, L. (1977). *Análise de conteúdo*. Edições 70
- Bentes Pinto, V. (2001). Indexação documentária: uma forma de representação do conhecimento registrado. *Perspectivas em Ciência da Informação*, 6(2), 223-234. <http://portal.deperiodicos.eci.ufmg.br/index.php/pci/article/view/423/23>
- Cordeiro, R. (1996). Informação Cinematográfica e Textual: da geração à interpretação e representação de imagem e texto. *Ciência da Informação*, 25(3), 461-465. <http://revista.ibict.br/ciinf/article/view/648>.
- Cousins, Mark. (2013). *História do cinema*. Martins Fontes.
- Horror Genres and Sub-genres*. (2014). Retrieved January 14, 2021, from <https://popcornhorror.com/genres/>.
- Netflix*. Retrieved January 20, 2021, from <https://www.netflix.com/br/>.
- Nogueira, L. (2010). *Manuais de cinema II: Gênero Cinematográfico*. Labcom Books. http://labcom.ubi.pt/ficheiros/nogueira-manual_II_generos_cinematograficos.pdf
- Primati, C. (2014) Prefácio. In: ALMEIDA, C. (ed.) *Cemitério Perdido dos Filmes B: Redux*. (pp. 11-13). Editora Estranho.
- Santos, F. E. P., Farias, M. G. G, Feitosa, L. T., Sobrinho, H. C. (2018). Definição de metadados e critérios de indexação para documentário em repositório audiovisual. *Revista Brasileira de Biblioteconomia e Documentação*, 14 (3), 237-261. <https://rbbd.febab.org.br/rbbd/article/view/1092>
- Sorlin, Pierre.(1994). Indispensáveis e enganosas, as imagens, testemunhas da história. *Estudos Históricos*, 7(13), 87.
- Teixeira, F. d. S. (2015). O impacto da netflix na produção e consumo de conteúdo audiovisual [Undergraduate Thesis, Universidade Federal do Rio de Janeiro]. <https://pantheon.ufrj.br/handle/11422/5252>

UMA PROPOSTA DE MODELO CONCEITUAL PARA ESTRUTURAR UMA BASE DE DADOS DE LIÇÕES APRENDIDAS EM PROJETOS

Mauricio Augusto Cabral Ramos Junior¹

¹Universidade Federal Fluminense, macrjunior@id.uff.br, ORCID iD 0000-0002-3567-1678

Resumo

A Gestão de Projetos é uma disciplina fortemente apoiada no conhecimento, muito utilizada para executar as estratégias organizacionais e, sendo assim, o aprimoramento do conhecimento organizacional se torna um fator crítico de sucesso em projetos. Uma forma para as organizações que os executam aprimorarem seu conhecimento é registrar lições aprendidas, que representam conhecimentos relevantes obtidos a partir de experiências vivenciadas. Neste cenário, mediante uma pesquisa de abordagem qualitativa, natureza aplicada, procedimento bibliográfico e inflexão ao método hipotético-dedutivo, esta comunicação apresenta um modelo conceitual básico para incorporar lições aprendidas em projetos ao conjunto de conhecimentos organizacionais, um problema recorrente na gestão de projetos diante, principalmente, da ausência de métodos, procedimentos e formatos estruturados para fazê-lo. O modelo foi concebido a partir de um conjunto básico de metadados estatisticamente representativo de registros desta natureza, permitindo uma modelagem conceitual que produziu um Diagrama Entidade-Relacionamento, um meio de representação deste modelo e que mostra as entidades e respectivas inter relações de uma estrutura básica de dados para registrar lições aprendidas em projetos.

Palavras-chave: *Lição aprendida, Metadados, Modelo conceitual, Organização do Conhecimento, Projeto.*

Introdução

A Gestão de Projetos é uma disciplina baseada no aprimoramento do conhecimento através da experimentação, nos conhecimentos tácito e explícito e no compartilhamento formal de conhecimentos (Jugdev e Wishart, 2014). Organizações dedicadas a esta disciplina, como APM (2012), AXELOS (2017) e PMI (2017), incluíram em suas bases de conhecimento processos e práticas para que equipes de projeto possam gerenciar de maneira efetiva o conhecimento obtido durante os projetos, que oferecem oportunidades únicas para aprimorá-lo (Anbari, Carayannis e Voetsch, 2008). Segundo estas organizações, uma forma de se realizar isto é através do registro de lições aprendidas.

Lições aprendidas são ensinamentos relevantes obtidos através de experiências e que podem contribuir para a aprendizagem organizacional (Taniguchi e Onosato, 2018). Elas são um dos mais relevantes valores agregados que se pode conseguir nos projetos (Walker, 2008), constituindo-se em uma das quatro variáveis diretamente ligadas ao sucesso na gestão destes (Besteiro, 2012), o que demanda sua incorporação ao conjunto de conhecimentos organizacionais e não apenas sua difusão entre as equipes de projetos (NATO, 2011) (Mmatsele e Carl, 2012) (Almeida e Soares, 2014) (Bell, van Waveren e Steyn, 2016) (Michell e McKenzie, 2017). Isto se justifica na percepção de que só existe sentido em se analisar a informação (ou o conhecimento que pode representar) diante de um conjunto de conhecimentos preexistentes compartilhados com um indivíduo ou grupo de indivíduos, ou seja, lidar com a informação é (re)contextualizar o conhecimento que dela se pode obter para aplicação prática em uma necessidade concreta (Capurro, 2003).

Para Kotnour (1999), as lições aprendidas funcionam como um processo para identificar e registrar ações a serem evitadas ou disseminadas – e as respectivas medidas a serem implementadas – e instrumento para compartilhar o conhecimento que representam. Segundo ABPMP (2013), trabalhadores do conhecimento operam diante de *casos*, situações únicas, não previsíveis e que podem ser conduzidas através de um processo – uma reunião de atividades e condutas realizados por indivíduos ou máquinas para se obter um ou mais resultados – onde uma lista de tarefas (e não exatamente um fluxo de trabalho) é realizada sob uma forma e ordem que somente podem ser definidas após se compreender com o que se depara. Desta maneira, as lições aprendidas podem ser vistas como registros originados em situações imprevistas e que representam um conhecimento relevante para uma organização obtido através da execução racional de atividades correlacionadas.

Considerando as lições aprendidas enquanto processo, o conhecimento pode assumir diferentes instâncias ou estágios ao longo do tempo. Segundo a visão de AXELOS (2017), um novo conhecimento identificado deve inicialmente ser denominado *lição*, adquirindo a denominação *lição aprendida* apenas se provocar uma mudança, o que remete à visão de Scharmer (2001) sobre o conhecimento tácito, que o dividiu em conhecimento auto transcendente (uma força situada entre os conhecimentos explícito e tácito e que provoca neste diferentes estágios) e conhecimento tácito incorporado (o conhecimento em uso). Na visão de Oliveira (2005), por sua vez, embora possuam uma estreita relação entre si, informação e conhecimento diferem quando se considera seu emprego: ao contrário do conhecimento propriamente dito, informação é um conhecimento cuja utilização ou aplicação ainda não foi desencadeada por uma experiência ou prática.

Transportando estas visões para a Gestão de Processos (ABPMP, 2013, p. 50), pode-se dizer que, no processo de registro de lições aprendidas, o conhecimento sofre “transições de estado ... os vários estágios do desenvolvimento dos produtos de trabalho à medida que avançam e são transformados através do processo”. Um estado representa uma condição única que um objeto pode estar ou ter. As transições de estados representam “um comportamento dinâmico que move um item de um estado para outro” e são acionadas “por atividades completas, eventos ou outros estímulos”, sendo que “um evento apenas pode causar uma transição se o objeto for afetado pelo evento no seu estado atual” (IIBA, 2011, p. 216).

Porém, a julgar pelo conteúdo dos guias de boas práticas de organizações de destaque dedicadas à disciplina (AAPM, 2009) (APM, 2012) (IPMA, 2015) (PMI, 2017), o registro de lições aprendidas em projetos é apresentado mais como uma única atividade (um registro propriamente dito) que um processo, comprometendo sua eficácia como instrumento de compartilhamento de conhecimentos e, em última análise, aprendizagem. Uma possível consequência desta percepção é a constatação de que as lições aprendidas não costumam ser adequadamente aplicadas para aprimorar as práticas organizacionais (Walker, 2008) (Norang e Nooshin, 2016) (Matthies, 2017) por causa, principalmente, i) do excesso de informações, ii) da ausência de procedimentos e métodos para sua avaliação e implementação (Matthies, 2017) e da ausência de um formato estruturado para registrá-las (Michell e McKenzie, 2017) (Abdellatif, Capretz e Ho, 2018).

Neste contexto e considerando que tecnologias de informação e comunicação têm, em geral, o objetivo de organizar, formalizar e difundir o conhecimento (Rautenberg, 2009), esta comunicação assumiu como seu objetivo buscar responder a seguinte questão: *como incorporar lições aprendidas ao conjunto de conhecimentos organizacionais?*, sugerindo um modelo conceitual básico para estruturar uma base de dados de registros de lições aprendidas em projetos. O modelo foi elaborado a partir de um conjunto básico de metadados estatisticamente representativo de registros desta natureza (Ramos Junior e Marcondes, 2020) e que foi obtido mediante a análise de cinquenta e quatro formulários de registro de lições aprendidas em projetos de organizações de onze países.

O método utilizado para a criação deste modelo consistiu em uma inflexão ao método hipotético-dedutivo de Karl Popper (Popper, 1975), que possibilitou inicialmente i) analisar o conjunto de metadados citado para apurar entidades que dele pudessem se derivar ou ser incluídas para aprimorar a representatividade do modelo e ii) estabelecer os relacionamentos entre as entidades, produzindo um Diagrama Entidade-Relacionamento (Chen, 1976). Compartilhando da visão de AXELOS (2017), os resultados obtidos nesta análise sugeriram que, de uma maneira geral, quando as organizações registram lições aprendidas em projetos, o fazem sem indicar as transições de estado entre a identificação de uma informação relevante que represente um conhecimento potencial (lição) e sua transformação em conhecimento propriamente dito (lição aprendida). Em um segundo momento, esta ausência foi então contemplada no modelo conceitual elaborado, para aprimorá-lo.

A fim de atingir seu objetivo, esta comunicação discorre através da seguinte estrutura, subdividida em seções: além desta introdução, apresenta em seguida a metodologia utilizada, os resultados obtidos com sua aplicação e as conclusões, onde sugere a realização de trabalhos futuros.

Metodologia

A comunicação se fez através de uma pesquisa de abordagem qualitativa (por não se preocupar em quantificar valores, mas sugerir o que se convém fazer a partir da interpretação de seu objeto), natureza aplicada (por buscar produzir conhecimentos para aplicação prática em problemas específicos), objetivo exploratório e procedimento bibliográfico (Gerhardt e Silveira, 2009).

Ao se propor a oferecer uma solução experimental e provisória (Popper, 1975) para o problema de como incorporar lições aprendidas ao conjunto de conhecimentos organizacionais, espera-se contribuir para satisfazer a necessidade de um método contínuo de apuração destas lições, possibilitando um registro com uma riqueza de detalhes a mais próxima possível de representar qualquer experiência significativa vivenciada (Chirumalla, 2013) (Veronese, 2014), além de uma solução viável para o problema do excesso de informações (Matthies, 2017) e da ausência de um formato estruturado (Michell e McKenzie, 2017) (Abdellatif, Capretz & Ho, 2018) ao registrá-las.

Uma vez que a Organização do Conhecimento estuda leis, princípios e procedimentos empregados para se estruturar um conhecimento especializado e com a finalidade de representar e recuperar de forma eficiente a informação contida em documentos (Barité, 2015), espera-se ainda contribuir para o debate em torno de um tema que lhe é fundamental: a Recuperação da Informação, termo bastante relacionado à Ciência da Computação e Ciência da Informação, cujo objeto de estudo são as relações entre termos, documentos e coleções de documentos e que se constitui em um conjunto de abordagens úteis para que usuários de sistemas de organização do conhecimento possam identificar mais facilmente documentos relevantes (Hørland, 2021). Diante da importância da estrutura e do formato que sistemas computacionais de registro de lições aprendidas devem ter para favorecer a recuperação de seus dados (Fosshage, 2013), há dificuldades para se classificar, recuperar e utilizar as lições aprendidas em projetos organizacionais (Duffield e Whitty, 2012).

O método utilizado foi uma inflexão ao método hipotético-dedutivo, que comumente transcorre pela i) definição do problema (reconhecimento de fatos e descoberta e enunciação do problema), ii) elaboração de uma solução teórica (e dedução de suas possíveis consequências, sob a forma de suposições e hipóteses) e iii) realização de testes de falseamento (tentativas de refutar as hipóteses através de observações e experimentações) (Popper, 1975) (Marconi e Lakatos, 2003). Por esta comunicação possuir um objetivo exploratório, a metodologia não se propôs a realizar estes testes, uma atividade que, porém, foi recomendada nas conclusões como uma sugestão de trabalho futuro.

A escolha deste método se justifica perante o próprio método científico, que se debruça sobre problemas considerados relevantes e critica as frequentes tentativas experimentais e provisórias para solucioná-los (Popper, 1975). Desta forma, a pesquisa buscou construir um modelo conceitual, uma representação parcial de um domínio capacitada a responder a uma questão (Mayr e Thalheim, 2021) e que decorre de um processo de abstração de uma conjuntura a partir da perspectiva de indivíduos nela interessados e que decorre a partir da definição de uma origem, dos conceitos envolvidos, da forma de representação e da compreensão de seus usuários ou partes interessadas envolvidas (Thalheim, 2011).

O tipo de modelo conceitual escolhido foi o Modelo Entidade-Relacionamento, que representa a estrutura informacional do domínio de um problema em termos de entidades e relacionamentos (e suas cardinalidades) graficamente configurados em um Diagrama Entidade-Relacionamento, um dos mais conhecidos modelos utilizados em análise estruturada, modelagem conceitual e projetos de bancos de dados (Song e Chen, 2009) (Amran, Mohamed e Bahry, 2018). Sua estruturação se deu

através da identificação das entidades principais e definição de suas inter relações (Martin, 1976) (Amran, Mohamed e Bahry, 2018).

O caráter semântico cada vez mais presente na busca e recuperação de informações revela a importância dos metadados para o estabelecimento do conhecimento e de evidências (considerando sua capacidade de representar as características dos documentos e garantir sua autenticidade) (Feinberg, 2018) e a representação das maneiras como as entidades de um sistema se comportam e relacionam, e não apenas suas propriedades (Mayernik, 2020). Foram estes os motivos que justificaram a estruturação do modelo conceitual a partir da utilização do conjunto básico de metadados criado por Ramos Junior e Marcondes (2020) para registrar lições aprendidas em projetos.

O conjunto básico de metadados foi obtido a partir da análise de cinquenta e quatro formulários de registro de lições aprendidas em projetos, elaborados por organizações públicas ou privadas de onze países da África, América do Norte, América do Sul, Ásia, Europa e Oceania. Trata-se de um conjunto estatisticamente representativo, que permite sua generalização e ressalta sua validade, cujos atributos e respectivas descrições se encontram listados na Tabela 1:

Tabela 1 – Conjunto básico de metadados para registrar lições aprendidas em projetos

#	NOME E DESCRIÇÃO DO ATRIBUTO	METADADO (padrão <i>Dublin Core™</i>)
1	Código do projeto (projeto onde foi identificada a lição aprendida)	<i>DC.Identifier.Id_projeto</i>
2	Código do cliente (cliente do projeto onde foi identificada a lição aprendida)	<i>DC.Identifier.Id_cliente</i>
3	Código do documento de lições aprendidas (documento associado à lição aprendida)	<i>DC.Identifier.Id_documento</i>
4	Código da lição aprendida (identificador da lição aprendida)	<i>DC.Identifier.Id_lição</i>
5	Código da categoria da lição aprendida (categoria associada à lição aprendida)	<i>DC.Subject.Categoria</i>
6	Elemento de análise (elemento cuja análise originou a lição aprendida)	<i>DC.Subject.Elemento</i>
7	Item de elemento de análise (item que compõem o elemento de análise ou a ele se refere)	<i>DC.Subject.Item_elemento</i>
8	Título da lição aprendida (título da lição aprendida)	<i>DC.Title.Titulo</i>
9	Sumário da lição aprendida (breve descrição da lição aprendida e seu contexto)	<i>DC.Description.Sumario</i>
10	Fase do ciclo de vida do projeto (fase do ciclo de vida do projeto onde a lição aprendida foi identificada)	<i>DC.Coverage.Fase</i>
11	Fato ocorrido (descrição do fato que originou a lição aprendida)	<i>DC.Description.Fato</i>
12	Causas do fato ocorrido (causas que provocaram o fato que originou a lição aprendida)	<i>DC.Description.Causa</i>
13	Causas comuns ao fato ocorrido (causas referentes a fatos que originaram outras lições aprendidas, suscitando a existência de um possível padrão)	<i>DC.Description.Causa_comum</i>

14	Consequências do fato ocorrido (descrição das consequências do fato que originou a lição aprendida)	<i>DC.Description.Consequencia</i>
15	Período que originou a lição aprendida (período do ciclo de vida do projeto onde a lição aprendida foi identificada)	<i>DC.Coverage.Periodo</i>
16	Identificador de risco (grau de risco associado à não adoção de medidas que eliminem ou mitiguem um risco)	<i>DC.Coverage.Id_risco</i>
17	Termo associado à lição aprendida (palavra-chave associada à lição aprendida)	<i>DC.Coverage.Termo_associado</i>
18	Nome do <i>stakeholder</i> que a identificou (denominação da parte interessada no projeto que identificou a lição aprendida)	<i>DC.Contributor.Stakeholder</i>
19	Descrição da lição aprendida (descrição das características da lição aprendida)	<i>DC.Description.Descricao</i>
20	<i>Link</i> para associar lições aprendidas (associação entre uma ou mais lições aprendidas)	<i>DC.Relation.Link</i>
21	Grau de impacto da lição aprendida (grau de impacto decorrente da incorporação da lição aprendida às melhores práticas organizacionais)	<i>DC.Coverage.Impacto</i>
22	Aplicação da lição aprendida (possível aplicação da lição aprendida)	<i>DC.Description.Aplicacao</i>
23	Ação recomendada (ação que deve ser executada para incorporar a lição aprendida às melhores práticas organizacionais)	<i>DC.Description.Acao</i>
24	Resultado da revisão da lição aprendida (resultado da análise da lição aprendida, com vistas à sua aprovação ou rejeição como tal)	<i>DC.Description.Resultado</i>
25	Histórico de revisões – data (data de uma revisão do conteúdo da lição aprendida)	<i>DC.Date.Historico_data</i>
26	Histórico de revisões – versão (última versão do registro da lição aprendida, gerada após sua última revisão)	<i>DC.Identifier.Historico_versao</i>
27	Histórico de revisões – descrição (descrição da revisão do conteúdo da lição aprendida)	<i>DC.Description.Historico_descricao</i>
28	Histórico de revisões – revisor (responsável pela revisão do conteúdo da lição aprendida)	<i>DC.Contributor.Historico_revisor</i>
29	Histórico de revisões – aprovado por (responsável pela aprovação da revisão do conteúdo da lição aprendida)	<i>DC.Contributor.Historico_aprovado_por</i>

Referência: adaptado de Ramos Junior e Marcondes (2020).

Resultados

A aplicação da metodologia apresentada transcorreu através das seguintes atividades:

i) Identificação das entidades principais, bem como suas descrições, apresentadas na Tabela 2:

Inicialmente, buscou-se identificar as entidades principais a partir do conjunto básico de metadados. Catorze entidades principais foram identificadas neste primeiro momento, com suas respectivas descrições. Em seguida, fez-se uma análise mais aprofundada do conjunto básico de metadados para verificar se, diante do referencial teórico aqui apresentado, haveria a necessidade de se incluir novos metadados (e suas respectivas entidades, se fosse o caso).

Considerando as lições aprendidas como um produto a ser desenvolvido e entregue por um respectivo processo de registro e as visões de Scharmer (2001), Oliveira (2005) e AXELOS (2017) sobre o conhecimento, a análise do conjunto básico de metadados apresentado na Tabela 1 mostrou que, de um modo geral, as organizações que executam estes processos não contemplam as transições de estados. Isto significa que, do momento em que uma informação é identificada como lição (um conhecimento – auto transcendente – potencialmente capaz de contribuir para a melhoria das práticas organizacionais) até ser registrada como lição aprendida (conhecimento tácito incorporado), não há neste conjunto nenhum metadado que indique o estado momentâneo do produto neste processo.

Diante desta constatação e observando o mencionado na seção Introdução sobre as ocorrências de transições de estado em processos, uma proposta para solucionar esta ausência é a criação de um metadado que seja associado ao histórico de revisões do conteúdo do registro de uma lição aprendida – o metadado denominado *DC.Identifier.Id_status*, cuja incorporação ao conjunto básico de metadados possibilitaria, ao se revisar um determinado registro, indicar uma transição em seu estado, se fosse o caso. Outro efeito prático da adoção desta proposta diz respeito ao caráter temporal dos documentos.

A Organização do Conhecimento direciona sua atuação para os documentos, que se referem a registros textuais ou gráficos criados em meio eletrônico ou não, de sentido comprobatório ou instrutivo (Buckland, 2018), instrumentos de preservação e compartilhamento da cultura que se constituem de um suporte, conteúdo e formato textual (Almeida e Torres, 2013) e possuem um caráter temporal que motiva sua classificação e destinação (Arquivo Nacional, 2019). Uma vez que, por motivos diversos, nem toda lição se torna uma lição aprendida (o que suscita seu descarte por não ter sido considerada relevante a ponto de justificar sua incorporação ao conjunto de conhecimentos organizacionais), o metadado proposto pode contribuir para o controle da destinação das lições.

Tabela 2 – Entidades principais do modelo conceitual

#	NOME E DESCRIÇÃO DA ENTIDADE	METADADOS RELACIONADOS
1	PROJETO (projetos onde lições aprendidas foram identificadas)	<i>DC.Identifier.Id_projeto</i>
2	CLIENTE (clientes dos projetos)	<i>DC.Identifier.Id_cliente</i>
3	DOCUMENTO (documentos associados a lições aprendidas)	<i>DC.Identifier.Id_documento</i>
4	LIÇÃO (lições aprendidas nos projetos)	<i>DC.Identifier.Id_lição</i>
		<i>DC.Title.Titulo</i>
		<i>DC.Description.Descricao</i>
		<i>DC.Description.Sumario</i>
		<i>DC.Description.Fato</i>
		<i>DC.Coverage.Periodo</i>
		<i>DC.Description.Causa</i>
		<i>DC.Description.Causa_comum</i>
		<i>DC.Description.Consequencia</i>
	<i>DC.Description.Acao</i>	
	<i>DC.Description.Aplicacao</i>	
5	FASE (fases do ciclo de vida dos projetos onde as lições aprendidas foram identificadas)	<i>DC.Coverage.Fase</i>
6	CATEGORIA (categorias das lições aprendidas)	<i>DC.Subject.Categoria</i>
7	ELEMENTO (elementos analisados nas lições aprendidas)	<i>DC.Subject.Elemento</i>
8	ITEM (itens referentes aos elementos analisados nas lições aprendidas)	<i>DC.Subject.Item_elemento</i>
9	RISCO (riscos associados às lições aprendidas)	<i>DC.Coverage.Id_risco</i>
10	PALAVRAS-CHAVE (palavras-chave associadas às lições aprendidas)	<i>DC.Coverage.Termo_associado</i>
11	STAKEHOLDER (partes interessadas nos projetos e que foram responsáveis por identificar as lições aprendidas)	<i>DC.Contributor.Stakeholder</i>
12	CORRELAÇÃO (correlações entre lições aprendidas semelhantes em suas características e propósitos)	<i>DC.Relation.Link</i>
13	IMPACTO (graus de impacto das lições aprendidas)	<i>DC.Coverage.Impacto</i>
14	REVISÃO (histórico de revisões realizadas nos conteúdos das lições aprendidas)	<i>DC.Description.Resultado</i>
		<i>DC.Date.Historico_data</i>
		<i>DC.Identifier.Historico_versao</i>
		<i>DC.Description.Historico_descricao</i>
		<i>DC.Contributor.Historico_revisor</i>
	<i>DC.Contributor.Historico_aprovado_por</i>	
15	STATUS (<i>status</i> de registro de lições aprendidas)	<i>DC.Identifier.Id_status</i>

ii) Definição das relações (e respectivas cardinalidades) entre as entidades principais da Tabela 2, apresentadas nas Figuras 1 e 2 sob a forma de um Diagrama Entidade-Relacionamento, onde os retângulos arredondados representam as entidades principais e as linhas contínuas, suas inter relações e respectivas cardinalidades:

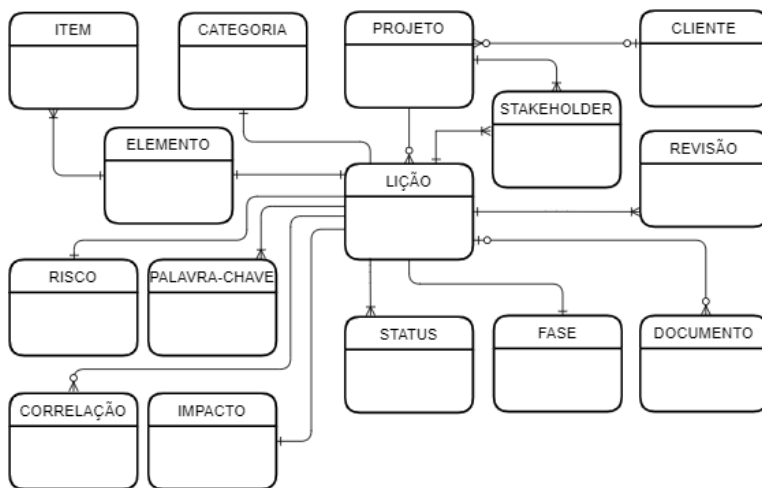
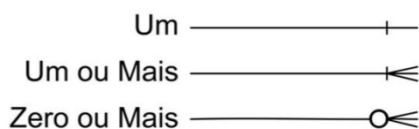


Figura 1. Diagrama Entidade-Relacionamento



Referência: Martim (1976)

Figura 2. Legenda das cardinalidades das relações entre as entidades principais

Conclusões

A partir de um conjunto básico de metadados para registrar lições aprendidas em projetos (Ramos Junior e Marcondes, 2020), esta comunicação apresentou um modelo conceitual (representado por um Diagrama Entidade-Relacionamento) para estruturar uma base de dados referentes a estas lições. A inferência estatística obtida neste conjunto de metadados garantiu sua representatividade, prenunciando assim que o modelo oferece uma estrutura de dados básica para a construção de sistemas computacionais para registrar lições aprendidas em projetos. Portanto, ao oferecer uma solução ao problema que orientou a pesquisa, este modelo sintetiza uma resposta ao excesso de informações referentes a lições aprendidas em projetos, à ausência de procedimentos e métodos para avaliá-las e implementá-las e à ausência de um formato estruturado para registrá-las, apresentando as entidades básicas (e suas interações) envolvidas no registro de lições que, se incorporadas ao conjunto de conhecimentos organizacionais, se transformarão de fato em lições aprendidas.

Ao buscar especificamente contribuir para a representação e recuperação de informações sobre lições aprendidas em projetos, a análise deste conjunto de metadados revelou a ausência de um metadado que permitisse indicar uma transição de

estado (algo inerente ao objeto de qualquer processo) em um registro. Propôs-se então a inclusão do metadado (no padrão *Dublin Core™*) *DC.Identifier.Id_status*, que também foi contemplado pelo modelo conceitual. Em última análise, isto observa o transcorrer do conhecimento por diferentes instâncias ou estágios de um processo de registro de lições aprendidas em projetos. Neste sentido, cita-se a diferença conceitual estabelecida entre lição e lição aprendida por AXELOS (2017) no âmbito da Gestão de Projetos, que destaca o caráter processual deste registro (remetendo-o à Gestão de Processos) e, indiretamente, o documental, ao sugerir que uma lição pode não se transformar em lição aprendida, o que remete à temporalidade dos documentos e sua inerente necessidade de classificação e destinação.

Portanto, a fim de se oferecer as melhores condições possíveis para uma efetiva organização e representação do conhecimento que encerram, é aconselhável que uma abordagem sobre o registro de lições aprendidas em projetos seja realizada de maneira interdisciplinar, combinando perspectivas oriundas da Gestão de Projetos, Gestão de Processos, Gestão de Documentos e Gestão da Informação. Em última análise, acredita-se ser este o caminho mais adequado para uma efetiva busca e recuperação da informação e disseminação do conhecimento organizacional útil à gestão de projetos.

A título de recomendação para a realização de pesquisas futuras, por se tratar de um modelo conceitual básico, acredita-se que o mesmo possa ser empregado com sucesso em diferentes organizações que executam projetos, sugerindo-se então a prévia realização de testes de falseamento a partir desta hipótese. Por outro lado, sua natureza básica sugere eventuais modificações em sua estrutura visando sua adequação a diferentes necessidades informacionais presentes nas organizações.

Referências

- Almeida, M. V., Soares, A. L. (2014). Knowledge sharing in project-based organizations: Overcoming the informational limbo. *International Journal of Information Management*, 34(6), 770-779.
- Abdellatif, T. M.; Capretz, L. F.; Ho, D. (2018). Searching for relevant lessons learned using hybrid information retrieval classifiers: A case study in software engineering. In *International Workshop on Professional Search* (ACM SIGIR – ProfS 2018), Ann Arbor, MI, United States.
- Almeida, M. B., Torres, S. (2013). O conceito de documento na ciência da informação e arquivologia. In *Encontro Nacional de Pesquisa em Ciência da Informação* (ENANCIB 2013), Florianópolis, SC, Brasil.
- Amran, N., Mohamed, H., & Bahry, F. D. S. (2018). Developing Human Resource Training Management (HRTM) Conceptual Model Using Entity Relationship Diagram (ERD). *International Journal of Academic Research in Business and Social Sciences*, 8(12), 1444-1459.
- Anbari, F. T., Carayannis, E. G., Voetsch, R. J. (2008). Post-project reviews as a key project management competence. *Technovation*, 28(10), 633-643.
- ARQUIVO NACIONAL. (2019). *Conselho Nacional de Arquivos*. Rio de Janeiro, RJ, Brasil. Recuperado em: <<http://conarq.arquivonacional.gov.br/o-conselho.html>>.
- American Academy of Project Management (AAPM). (2009). *IPMC Project management handbook for managers & executives*. Colorado Spring: AAPM.

- Association of Business Process Management Professionals (ABPMP). (2013). *Guia para o gerenciamento de processos de negócios – Corpo comum de conhecimento (BPM CBOK)* (1a ed.). Brasília: ABPMP.
- Association of Project Management (APM). (2012). *APM body of knowledge*. (6a ed.). Buckinghamshire: APM.
- Axelos Limited. (2017). *Managing successful projects with PRINCE2®*. (6a ed). London: Axelos Limited.
- Barité, M. (2015). *Diccionario de organización del conocimiento: Clasificación, indización, terminología*. (6a ed). Montevideo: CSIC.
- Bell, L., Van Waveren, C., Steyn, H. (2016). Knowledge-sharing within the project-based organization: A knowledge-pull framework. *The South African Journal of Industrial Engineering*, 27(4), 18-33.
- Besteiro, E. N. C. (2012). *Escala de mensuração dos fatores críticos de sucesso no gerenciamento de projetos* (Tese de doutorado). Universidade Estadual de Campinas, Campinas, SP, Brasil.
- Buckland, M. K. (2018). Document theory. *Knowledge organization*, 45(5), 425-436.
- Capurro, R. (2003). Epistemologia e ciência da informação. In: *Encontro Nacional de Pesquisa em Ciência da Informação* (ENANCIB 2003), Belo Horizonte, MG, Brasil.
- Chen, P. P. (1976). The entity-relationship model – Toward a unified view of data. *ACM Transactions on Database Systems*, 1(1), 9-36.
- Chirumalla, K. (2013). *Development of a methodology for lessons learned practice* (Tese de Doutorado). Department of Business Administration, Technology and Social Sciences, Luleå University of Technology, Luleå, Sweden.
- Duffield, S., Whitty, S. J. (2012). A systemic lessons learned and captured knowledge (SLLCK) model for project organizations. In PMglobal (Ed.), *9th Project Management Australia Conference* (PMOz 2012). Melbourne: Eventcorp PtyLtd.
- Feinberg, M. (2018). Factotem: What is Information Access for? *Cataloging & Classification Quarterly* 56(8), 665-682.
- Fosshage, E. (2013). *Considerations for implementing and organizational lessons learned process*. Albuquerque: Sandia National Laboratories.
- Gerhardt, T. E., Silveira, D. T. (Orgs). (2009). *Métodos de pesquisa*. (1a ed.). Porto Alegre: Editora da Universidade Federal do Rio Grande do Sul.
- Hørland, B. (2021). Information retrieval and knowledge organization: A perspective from the philosophy of science. *Information*, 12(3):article 135.
- International Institute of Business Analysis (IIBA). (2011). *Um guia para o corpo de conhecimento de análise de negócios (Guia BABOK)* (2a ed.). São Paulo: IIBA.
- International Project Management Association (IPMA). (2015). *Individual competence baseline for project, programme & portfolio management*. (4a ed.). Zurich: IPMA.
- Kotnour, T. (1999). A learning framework for project management. *Project Management Journal*, 30(2), 32-38.
- Jugdev, K., Wishart, P. (2014). Mutual caring – resolving habituation through awareness: supporting Meaningful Learning From Projects. *Project Management Journal*, 45(2), 66–82.
- Marconi, M. de A., Lakatos, E. M. (2003). *Fundamentos de metodologia científica*. (5a ed.). São Paulo: Atlas.
- Martin, J. (1976). *Principles of Data-Base Management*. Englewood Cliffs: Prentice-Hall.
- Matthies, B. A. (2017). Text mining approach for extracting lessons learned from project documentation: An illustrative case study. *International Journal of Management, Knowledge and Learning*, 6(2), 153-174.
- Mayernik, S. M. (2020). Metadata. *Knowledge Organization*, 47(8), 696-713.

- Mayr, H. C.; Thalheim, B. (2021). The triptych of conceptual modeling: A framework for a better understanding of conceptual modeling. *Software and Systems Modeling*, 20, 7-24.
- Michell, V., McKenzie, J. (2017). Lessons learned: structuring knowledge codification and abstraction to provide meaningful information for learning. *VINE Journal of Information and Knowledge Management Systems*, 47(3), 411-428.
- Mmatseleng, M., Carl, M. (2012). Project reviews: the vehicle for learning in organisations. *African Journal of Business Management*, 6(44), 10853-10861.
- NATO. *The NATO Lessons Learned Handbook*. (2016). (3a ed). Monsanto: NATO Joint Analysis and Lessons Learned Centre.
- Norang, A., Nooshin, S. M. (2016). Designing a lessons learned model to improve the success of new product development in project based organizations. *Management Science Letters*, 6, 759-766.
- Oliveira, V. P. (2005). Uma informação tácita. *DataGramaZero*, 6(3).
- Popper, K. S. (1975). *A lógica da pesquisa científica* (2a ed.). São Paulo: Cultrix.
- Project Management Institute (PMI). (2017). *Um guia do conhecimento em gerenciamento de projetos (Guia PMBOK®)*. (6a ed.). Newtown Square: Project Management Institute, Inc.
- Ramos Junior, M. A. C., Marcondes, C. H. (2020). Uma proposta de metadados para o registro de lições aprendidas em projetos. *Informação & Informação*, 25(3), 212-238.
- Rautenberg, S. (2009). *Modelo de conhecimento para mapeamento de instrumentos da gestão do conhecimento e de agentes computacionais da engenharia do conhecimento baseado em ontologias* (Tese de doutorado). Universidade Federal de Santa Catarina, Florianópolis, SC, Brasil.
- Scharmer, C. O. (2001). Self-transcending knowledge: sensing and organization around emerging opportunities. *Journal of Knowledge Management*, 5(2), 137-150.
- Song, I., Chen, P. P. (2009). Entity Relationship Model. In: LIU L., ÖZSU M.T. (Eds.) *Encyclopedia of Database Systems*. Boston: Springer.
- Taniguchi, A., Onosato, M. (2018). Effective Use of Lessons Learned to Conduct the Project Review for ERP Implementation. *International Journal of Information Technology and Computer Science*, 10, 1-15.
- Thalheim, B. (2011). The theory of conceptual models, the theory of conceptual modeling and foundations of conceptual modeling. In Embley, D. W., Thalheim, B. (Eds.). *Handbook of Conceptual Modeling* (1. ed., Cap. 17, 1-37). Berlin Heidelberg: Springer-Verlag.
- Veronese, G. (2014). Métodos para captura de lições aprendidas: em direção à melhoria contínua na gestão de projetos. *Revista de Gestão e Projetos*. São Paulo, 5(1), 71-83.
- Walker, L. W. (2008). Learning lessons on lessons learned. In *PMI Global Congress 2008*, North America, Denver, CO. Project Management Institute, Inc, Newtown Square, Philadelphia, United States.

IDENTIFICACIÓN Y VALIDACIÓN DE FORMATOS DE FICHEROS ORIENTADA A LA PRESERVACIÓN DIGITAL DE DOCUMENTOS. DE TIK A DROID

María-José Baños-Moreno, María González Balanza,
Javier Lurquí López, Francisco-José Valentín-Ruiz

¹ Odilo, mbm41963@um.es, ORCiD iD 0000-0001-9137-1330

² Odilo, mgonzalez@odilotid.es, ORCiD iD 0000-0002-9577-8290

³ Odilo, jlurqui@odilo.us, ORCiD iD 0000-0001-5656-0750

⁴ Odilo, fjvalentin@odilotid.es, ORCiD iD 0000-0001-5343-920X

Resumen

Se expone la problemática derivada de la identificación y verificación de formatos de ficheros como elemento clave de todo sistema de preservación digital, ya que la cantidad de ficheros que preservar en las organizaciones no deja de crecer, pero también la variedad de los formatos de estos ficheros.

Para este propósito, las herramientas de archivo y preservación de Odilo, A3W-AE y Preserver, han empleado el API de TIK A, pero en los últimos meses se ha revelado como insuficiente para atender las necesidades cada vez más acuciantes de las organizaciones con que trabaja Odilo. En este sentido, el objetivo de este trabajo es evolucionar hacia otra herramienta que permita una identificación y verificación de formatos, versiones de formatos y extensiones de ficheros de forma más exhaustiva con respecto a TIK A.

Para ello, se realiza una revisión bibliográfica y sendos análisis de tipo funcional y técnico sobre las dos herramientas más destacadas en la bibliografía para la identificación y verificación de formatos: JHOVE y DROID.

Las pruebas realizadas evidencian que DROID cumple adecuadamente con el propósito establecido, por lo que las aplicaciones A3W-AE y Preserver comenzaron a trabajar con el API de DROID para esa identificación y verificación de formatos. Finalmente, este desarrollo se incorporó a la versión en producción de dichas herramientas y se instalaron en cliente.

El cambio de TIK A a DROID ha permitido mejorar la identificación y verificación de los formatos de ficheros de clientes, proporcionando una respuesta adecuada a las necesidades de las organizaciones. Este API permitirá, además, mejorar otros servicios proporcionados por Odilo, relacionados con la preservación digital.

Palabras clave: preservación digital, A3W-AE, API TIK A, API DROID

Introducción

La preservación digital es el conjunto de actividades que pretende hacer frente a los retos que se les plantean a personas y organizaciones en el uso a largo plazo de sus objetos digitales (Baños-Moreno, 2018). Como ya apuntaba Álvarez Wong (2017), a medida que pasa el tiempo estos retos son más patentes, puesto que la cantidad de recursos generados en formas digitales, sin un respaldo físico, no deja de crecer. Así ha ocurrido en Odilo desde que comenzamos a trabajar con esta cuestión. El número de clientes de los servicios de archivo y preservación ha crecido y, con él, la complejidad, variedad y volumen de los paquetes que preservar.

Así, hoy se pueden encontrar proyectos de digitalización con distinta trayectoria. Muchos fracasaron, se abandonaron o se dieron por finalizados una vez que no hubo material que digitalizar. No se pensó en la preservación de los documentos digitales a los que dieron lugar. No se valoraron otros factores como los responsables de la preservación, los planes de preservación, posibilidades de migración de formatos, la necesidad de versiones estables de los ficheros, herramientas de visualización, alojamiento en sistemas fiables, copias de seguridad redundantes, metadatos de descripción, de metadatos de trazabilidad, metadatos técnicos, cuestiones de interoperabilidad etc. Hablamos de una serie de conceptos ya no tan desconocidos, pero con grandes implicaciones. También para las empresas que suministran servicios de preservación digital. Quizá incluso más, pues se percibe con mayor frecuencia cómo las dificultades crecen y varían para todas las organizaciones con que trabajan.

El problema de la variedad de formatos en la preservación digital

Uno de los problemas identificados es la variedad de formatos, versiones y extensiones de los ficheros con que se trabaja. De entrada, cualquier entidad querrá preservarlo todo, con un volumen de documentación cada vez mayor, y de forma automática. Surge, además, una segunda cuestión, cómo gestionar esa variedad de forma ágil, que permita conocer las características de dichos archivos, así como los riesgos asociados a estos y gestionarlos de forma eficaz. De ahí que contemos con herramientas de análisis y verificación de las características de los ficheros. Estas validaciones forman parte de una gestión mayor de los paquetes de cualquier organización. En el siguiente gráfico, se muestra el proceso completo por el que pasa un expediente desde que ingresa en la aplicación A3W-AE de Odilo hasta que se pone a disposición de un usuario final, pasando por su almacenamiento en un repositorio de preservación digital. La identificación de formatos de los ficheros asociados a ese expediente es una de las acciones de validación a que se someten durante el ingreso.

La identificación y verificación de los formatos están (o deben estar) muy presentes en cualquier plan de preservación de una organización. En realidad, se trata de diversos aspectos, los formatos varían y proliferan y, además, se deben revisar periódicamente y de forma iterativa, porque la tecnología cambia (Giusti, 2016). También se debe valorar la multiplicidad de extensiones vinculadas a esos formatos (Moro Cabero, 2018), la obsolescencia asociada a los tipos de ficheros, la evolución

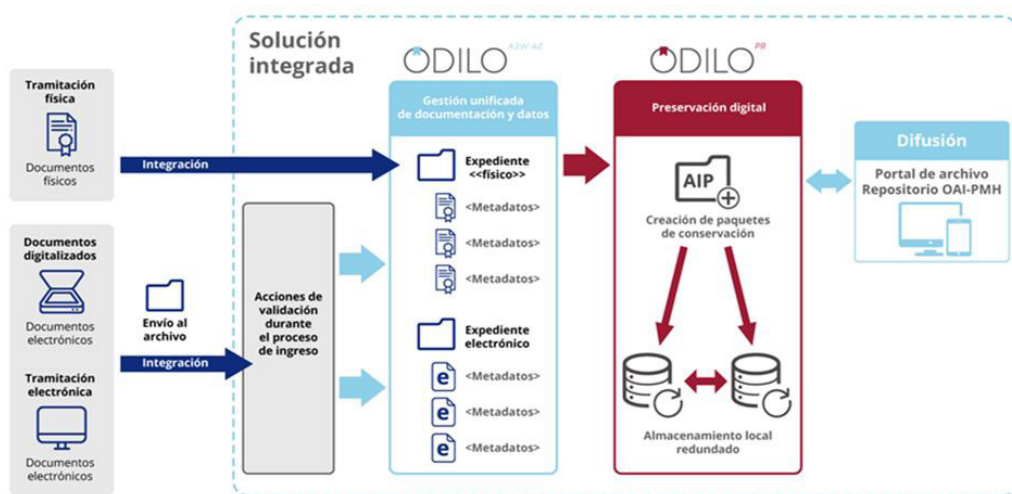


Figura 1. Proceso de ingreso, validación, almacenamiento y difusión de un expediente con la solución de archivo y preservación de Odilo

de los formatos, la dependencia tecnológica con respecto al software utilizado para su gestión, así como la necesidad de hacer frente a un volumen cada vez mayor de documentación electrónica (Candás-Romero, 2006). Barbedo (2019, p. 6) resume todo como la necesidad de mantener los documentos electrónicos a lo largo del tiempo.

Esto implica tomar decisiones en distintas partes del ciclo documental:

1. Aplicando *metodologías de preservación digital integradas con la gestión documental corporativa* (Valentín-Ruiz, 2020), en la fase administrativa del documento, sea en papel (si son sometidos a un proceso de digitalización) o electrónico. Desde ese momento, *comienza el problema de la preservación y el acceso del mismo modo* (Giusti, 2016).

2. Empleando repositorios de preservación confiables, que aseguren la aplicación de *metodologías y herramientas para evaluar el grado de cumplimiento de estándares o buenas prácticas aceptadas en un sistema en particular* (Térmens y Leija, 2017), una vez que pasan a su fase archivística.

Por esto, es importante conocer y evaluar aplicaciones para la identificación y validación, de los formatos de ficheros, tanto en la fase administrativa como en la archivística: la identificación permite conocer la estructura del objeto de datos. Esto es cuál es el formato del fichero; la validación constata la consistencia entre el fichero y sus metadatos. Es decir, el formato identificado para el fichero, ¿es realmente su formato? (Moro Cabero, 2018; Abrams, 2005). La correcta identificación de los formatos de los ficheros es una de las primeras cuestiones técnicas que abordar en materia de preservación digital, pues, *sólo conociendo perfectamente los objetos de preservación se les podrán aplicar los procedimientos técnicos que*

aseguren su conservación y validez a largo plazo (Térmens y Ribera, 2009). El objetivo último es permitir la accesibilidad al documento desde el momento en que se crea hasta que se expurga o PARA SIEMPRE. Dos palabras que evidencian el gran reto de la preservación digital.

La necesidad de evolucionar de herramienta: de TIKa a DROID

Como parte de la gestión de la preservación digital a través de las aplicaciones A3W-AE y Preserver (que soportan los servicios de archivo y preservación de Odilo), hace unos años, se analizaron algunos de los sistemas de identificación y validación de formatos de ficheros existentes por entonces. Más allá de si la decisión fue errónea o no, la herramienta seleccionada fue TIKa, orientada a la detección y extracción de metadatos de ficheros. La API de TIKa:

1) Identifica el formato y extensión de una lista limitada de formatos básicos. Este control comprueba cuál el formato y la extensión del fichero adjunto. Para el formato PDF/A, además se realizan comprobaciones desarrolladas ad hoc. Para este proceso, se emplean distintos parsers, como AdobeFontMetricParser, AudioParser, HtmlParser, ImageParser o XMLParser.

2) Verifica si el formato y la extensión de los ficheros identificados inicialmente son reales y tienen un comportamiento asociado o no, pero sólo para algunos formatos de tipo MIME, no se tienen en cuenta otros tan extendidos como el mp3. Mantener la verificación MIME, implica el análisis uno a uno de cada formato y su posterior inclusión en el proceso, si es posible. El problema es que, de continuar por este camino, se seguiría sin dar respuesta ficheros que no son de tipo MIME. De ahí que fuera necesario contar con desarrollos propios que realicen este tipo de validación para esos formatos ajenos a TIKa. Estos desarrollos se aplicaron sobre los siguientes formatos: mp3, png, pdf; mp4, docx, doc, odt, ods, odp y odg.

Con los desarrollos ad hoc aumentando en cada versión de las aplicaciones de Odilo, era evidente que TIKa resultaba insuficiente para un análisis pormenorizado de los formatos de los ficheros de las organizaciones. Era necesario evolucionar hacia otro sistema más completo, que permitiera identificar y verificar una mayor variedad de ficheros, sin necesidad de recurrir a desarrollos propios. A través de una revisión bibliográfica se determinó que dos de las aplicaciones más extendidas y asentadas para este propósito, la identificación y verificación de formatos, son JHOVE y DROID.

JHOVE (JSTOR/Harvard Object Validation Environment), creado por la Harvard University, es capaz de identificar formatos, pero sólo 17 oficialmente. Para cualquier otro formato sería necesario un desarrollo ad hoc. Fue descartado, por no incluir gran cantidad de formatos que analizar, por mucho que los que emplea están muy vinculados a la conservación (Moro Cabero, 2018).

DROID (Digital Record Object Identification) *fue creado por los archivos nacionales del Reino Unido en Java, como una herramienta para facilitar la exploración directa del registro Pronom* (Térmens y Ribera, 2009). Se trata de un recurso con

información sobre formatos de archivos, software y otros componentes técnicos relacionados.

Mediante un análisis funcional y otro de tipo técnico, desempeñados por el servicio de preservación y el equipo de desarrollo de Odilo, respectivamente, se evidenció que DROID proporcionaba resultados óptimos en cuanto a la identificación de todos los formatos y versiones empleados. JHOVE no fue capaz de detectar algunos formatos particulares, como algunos de Autodesk, por ejemplo.

Como resultado, se cambió todo el sistema de identificación y comprobación de formatos, versión y extensión de ficheros de los servicios de archivo y preservación de Odilo, pasando de TIKa a DROID. Esto implicó una importante simplificación del código de la aplicación, pues se eliminaron todos los desarrollos que se habían implementado ad hoc para la validación de todos aquellos elementos no contemplados por TIKa. Además, para facilitar el trabajo realizado por el servicio de preservación digital (descrito con detalle en Baños-Moreno, 2018), la lista de extensiones, asociadas a versiones de formatos y a los formatos en sí de ficheros, fueron preintroducidos en la base de datos como no admitidos. En el momento en que un formato es admitido, basta con cambiar su estado para aceptar ingestas de este tipo de fichero. De esta forma, se evitan posibles errores de transcripción y se facilita la incorporación de nuevos formatos al plan de preservación de la organización, si así se estima necesario.

La evolución de TIKa a DROID ha permitido mejorar e incrementar las posibilidades de identificación de extensiones, versiones de formato y formato de los ficheros que preservar de las distintas organizaciones con que trabaja Odilo, principalmente archivos de la administración pública, aunque también contamos con archivos de otro tipo de entidades. Se trata de un aspecto muy positivo, desde el punto de vista de la preservación digital, pues profundiza en el análisis de las características de los ficheros de dichas organizaciones, y facilitar la recomendación de migración de formatos, basadas en la información proporcionada en Pronom, entre otros. En el futuro, se pretende automatizar algunos de los procesos que actualmente son realizados exclusivamente por el equipo humano del servicio de preservación, como implementar alertas de recomendaciones de migración de formatos o bien proporcionar información técnica sobre los ficheros de forma legible para el usuario final, entre otras cuestiones.

Metodología

Como se ha evidenciado, era necesario cambiar de herramienta de identificación y verificación de los formatos de ficheros empleada en las aplicaciones A3W-AE y Preserver, pues es necesario atender a una cada vez mayor variedad de formatos. Es importante tener en cuenta que existen formatos específicos para cada tipo de objeto digital: gráficos, audio, videos, documentos, etc. Algunos formatos están asociados con una o más extensiones de archivo (Castilla Solano y Umaña Alpizar, 2019) y, además, para cada formato puede haber o no versiones del formato. La preservación digital es un proceso permanente que requiere tener en cuenta estas cuestiones (Adair et al., 2019).

Para ese cambio de API, en primer lugar, se analizaron varios de los registros de formatos en la actualidad. Un repositorio de este tipo recoge metadatos descriptivos, técnicos y administrativos sobre los formatos digitales e incluyen la definición de las características sintácticas y semánticas de los formatos registrados (Abrams, 2005). Este primer análisis permite profundizar en el conocimiento de algunos de los problemas más frecuentes, asociados a los formatos. Por ejemplo, es destacable la dificultad que entraña el uso de formatos asociados a Autocad, como DWG o DXF, frente a PDF en cuanto a cuestiones de preservación, pero también de difusión (Green et al. 2016), pues están mucho más ligados a un software propietario que evoluciona periódicamente y provoca la rápida obsolescencia de sus ficheros. En Narasaiah et al. (2021), por otro lado, la propuesta de preservación, en este caso de literatura gris, tenía que ver con formatos más conocidos, como TIFF, JPG y PDF/A. Estos formatos también son empleados con frecuencia entre las organizaciones que trabajan con las aplicaciones de Odilo. Se recoge los registros más destacados en el apartado Resultados.

A partir del conocimiento adquirido sobre características de formatos, versiones y extensiones se analizaron las principales herramientas de identificación y verificación de formatos más extendidas en la actualidad mediante una revisión bibliográfica.

Para los análisis posteriores, se trabajó con una muestra de veinte ficheros de distintos formatos y/o versiones de formatos más empleados entre esas instituciones que trabajan con los servicios de archivo y preservación de Odilo. Así, primero se aplicó un análisis de tipo funcional, desempeñado por el servicio de preservación digital a partir de las aplicaciones de usuario de las dos herramientas más destacadas, JHOVE y DROID. El propósito fue analizar la capacidad de estas herramientas para la identificación y verificación de los formatos de la muestra de ficheros. A continuación, se revisó la documentación de un software y del otro, desde un punto de vista técnico, y se realizaron catas de uso de las correspondientes APIs ya de forma integrada con A3W-AE y Preserver. En ambos casos, como se muestra en el apartado Resultados, DROID resultó más efectiva para el propósito de esta comunicación, por lo que se implementó como nueva herramienta de identificación y verificación de formatos, versiones y extensiones de formatos. Este desarrollo se instaló en versiones de pruebas de las aplicaciones A3W-AE y Preserver de Odilo y fue sometido a pruebas de funcionamiento y usabilidad por el equipo de Q&A (Quality & Assurance) de Odilo.

Resultados

En este apartado se recogen los resultados más relevantes tras aplicar la metodología descrita. En primer lugar, se indican los registros de formatos orientados a la preservación digital que pueden ser utilizados para la adquisición de conocimiento sobre formatos. A continuación, se especifican los resultados obtenidos con los análisis realizados sobre las herramientas de identificación y verificación de ficheros analizadas.

Registros de formatos de ficheros

Los repositorios o registros de formatos son una fuente esencial de consulta de información relacionada con los formatos de ficheros, ya que permiten conocer con detalle características de los ficheros y orientar la estrategia de preservación de una organización en un sentido o en otro. A lo largo de los últimos años han existido distintas propuestas de registros de formatos: existía la iniciativa *Unified Digital Format Registry* (UDFR) (<http://www.udfr.org/>), que pretendía unificar los registros del Global Digital Format Registry (GDFR), de las Bibliotecas de la Universidad de Harvard y Pronom (Giusti, 2016). Sin embargo, parece que tanto el GDFR como la propia iniciativa UDFR fueron descontinuados hace varios años; IANA Media Types (<http://www.iana.org/assignments/media-types/media-types.xhtml>), destacado por Abrams (2005), actualizado por última vez en 2012; la web de formatos de la Biblioteca del Congreso (<https://www.loc.gov/preservation/digital/formats/intro/intro.shtml>), que, como indican Térmens y Ribera (2009) está más orientada a dar consejos sobre los formatos más aceptables en un marco de preservación a largo plazo, resulta útil para complementar la información proporcionada por el siguiente repositorio; Pronom (<http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>), un recurso con información sobre formatos de archivos, software y otros componentes técnicos relacionados, creado por los archivos nacionales de Reino Unido, sobre la que se sustenta DROID. Este último es, probablemente, el repositorio que se actualiza con mayor frecuencia, precisamente por esa relación con DROID.

Herramientas para la identificación y validación de formatos, versiones de formatos y extensiones

Para conocer estas herramientas, principalmente, se valoraron los trabajos de Moro Cabero (2018), Térmens y Ribera (2009) y Abrams (2005). Otras publicaciones relacionadas con el objeto de estudio de esta comunicación no profundizaban en APIs concretas o reseñaban las mismas que las indicadas por los autores anteriores. Las dos herramientas más destacadas en la bibliografía resultaron ser JHOVE y DROID aunque se valoraron otras herramientas, como: XENA (*XML Electronic Normalising for Archives*) de los archivos australianos, que fue descontinuada, tal como se especifica en su página web (<https://sourceforge.net/projects/xena/>) y TrID, creado por Marco Pontello para la identificación de formatos pero que no permite su uso comercial, además, no da respuestas tan seguras como los programas analizados (Térmens y Ribera, 2009).

JHOVE (JSTOR/Harvard Object Validation Environment) (<http://jhove.openpreservation.org/>) es una herramienta de código abierto de la Open Preservation Foundation para la identificación, validación y caracterización de formatos (Depoortere, et al., 2020). No sólo identifica formatos y extensiones, sino que, establece un mecanismo por el cual se realiza un tratamiento específico para cada formato, considerando cada uno en un módulo individual, para poder validarlo y acceder a la información particular de cada formato. El principal inconveniente es

que existen 17 módulos oficiales: AIFF, ASCII, ByteStream, GIF, GZIP, HTML, JPEG, JPEG200, PDF, TIFF, UTF-8, WARC, WAVE, XML, EPUB, MP3 y ZIP. Para cualquier otro formato sería necesario su desarrollo ad hoc.

En cuanto a DROID (*Digital Record Object IDentification*) (<https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>), es un programa bastante completo, de código abierto y en software libre, con interfaz de usuario, ejecución desde línea de comando y un funcionamiento complejo para establecer perfiles de identificación (esto es, formatos preferentes y aceptables) y uso de patrones de búsqueda para archivos y carpetas.

Esta herramienta permite el reconocimiento de formatos desconocidos y trabaja en colaboración con la herramienta Pronom, informando tanto de la versión como de la extensión del fichero, incluso si ésta es incorrecta o se encuentra ausente (Moro Cabero, 2018). Para esa identificación automatizada de lotes de formatos (Depoortere et al., 2020), se emplea la última definición de su archivo *Signatures* (en versión 96 y actualizado a enero de 2020 en el momento en que se realizaba este análisis). Este fichero define los patrones de búsqueda dentro de los ficheros para lograr su identificación, así como los formatos posibles asociados estos patrones. Los patrones de búsqueda son secuencias de bytes, asociadas a posiciones dentro del archivo, que son usadas para las identificaciones. Podemos usar las actualizaciones que se facilitan desde su página, modificarlo para añadir las nuestras propias, o quitar las que no usemos, lo que nos permite adaptar el proceso a nuestras necesidades. Por otro lado, igual que con TIKa, para el formato PDF/A, además se realizan comprobaciones desarrolladas ad hoc que profundizan en su identificación.

En la siguiente tabla, se muestra parte de la lista de ficheros con que se han analizado JHOVE y DROID, tanto desde un punto de vista funcional (realizadas por el equipo de Q&A (Quality & Assurance) de Odilo), con sus respectivas aplicaciones de usuario, como desde el punto de vista técnico, mediante la revisión de la documentación y el uso del API de ambas herramientas. La muestra total estaba compuesta de 100 ficheros, aquí se recogen los 20 más frecuentes.

Tabla 1. Muestra de ficheros analizados en JHOVE

Fichero	JHOVE formato	DROID formato	JHOVE versión	DROID versión
fichero01.bmp	Bytestream	Windows Bitmap	–	3.0
fichero02.doc	Bytestream	Microsoft Word Document	–	97-2003
fichero03.docx	Bytestream	Microsoft Word for Windows	–	2007 onwards
fichero04.gif	PNG	Portable Network Graphics	–	1.0
fichero05.jpeg	JPEG	JPEG File Interchange Format	1.01	1.01
fichero06.jpg	JPEG	JPEG File Interchange Format	1.01	1.01
fichero07.odt	Bytestream	OpenDocument Text	–	1.2
fichero08.pdf	PDF	Acrobat PDF 1.7 – Portable Document Format	1.7	1.7
fichero10.tif	TIFF	Tagged Image File Format	6.0	–
fichero11.tif	TIFF	Tagged Image File Format	6.0	–
fichero12.txt	ASCII	Plain Text File	–	–
fichero13.xls	Bytestream	Microsoft Excel 97 Workbook (xls)	–	8
fichero14.xlsx	Bytestream	Microsoft Excel for Windows	–	2007 onwards

fichero15.pptx	Bytestream	Microsoft Powerpoint for Windows	–	2007 onwards
fichero16.avi	Bytestream	Audio/Video Interleaved Format	–	–
fichero17.mov	Bytestream	Quicktime	–	–
fichero18.m4v	Bytestream	MPEG-4 Media File	–	–
fichero19.mp4	Bytestream	MPEG-4 Media File	–	–
fichero20.wmv	Bytestream	Windows Media Video	–	–
fichero21.dwg	Bytestream	AutoCAD Drawing	–	2013/2014
fichero22.dxf	ASCII	Drawing Interchange File Format (ASCII)	–	R14

En ambos casos se pudo identificar el formato de buena parte de los ficheros empleados, aunque DROID mostró una identificación más clara. JHOVE no fue capaz de detectar algunos formatos particulares, como los de Autodesk. Además, varios de los ficheros tenían una extensión que no era acorde con la extensión real del fichero. En el proceso de verificación, DROID detectó este error, no así JHOVE. Por ejemplo, el fichero.04.gif tenía como extensión real .png. Por otro lado, DROID proporcionó la versión del formato de fichero en mayor cantidad que lo hizo JHOVE.

Por tanto, es evidente que DROID proporcionaba resultados óptimos en cuanto a la identificación de todos los formatos y versiones empleados con respecto a JHOVE, que incluso mostraba errores en el proceso de verificación de algunos de los formatos. Ya lo adelantaba Moro Cabero (2018), JHOVE no incluye gran cantidad de formatos que analizar, aunque los que emplea están muy vinculados a la conservación. En este sentido, el empleo de JHOVE, implicaría mensajes de error al intentar preservar determinados paquetes y no se superarían las limitaciones de TIKa que, en base, venían a ser similares.

Así, se decidió implementar DROID como nueva herramienta de identificación y verificación de formatos, por lo que se llevó a cabo el desarrollo previsto y se instaló en versiones de pruebas de las aplicaciones A3W-AE y Preserver.

Acción	Descripción	Estado	Fecha
Iniciar Ingreso	Empezando proceso de ingreso	Proceso de Ingreso	29/12/2020 10:50:31
Validaciones	Iniciando validaciones	Validación en Proceso	29/12/2020 10:50:31
Validación Hash	Validando hash recibido	Validación en Proceso	29/12/2020 10:50:32
Validación Hash	El hash del fichero es válido	Validación Correcta	29/12/2020 10:50:32
Validación SIP	Validando estructura y contenido del SIP	Validación en Proceso	29/12/2020 10:50:32
Validación SIP	La version 1.3 del formato dng del documento NP_060768.dng no está dado de alta en Gestión de formatos. No se puede preservar si no existe la versión en la aplicación.	Error en Ingreso	29/12/2020 10:50:33

Figura 2. Ejemplo de fichero con extensión DNG y versión 1.3 no dado de alta en la aplicación por el servicio de preservación

Los resultados fueron similares a los proporcionados por la aplicación de usuario de DROID aunque, desde las aplicaciones de Odilo se proporcionan mensajes que intentan clarificar el por qué se puede o no se puede preservar determinados ficheros, de acuerdo con los requerimientos de los formatos recogidos en el plan de preservación de la organización de que se trate. El propósito, identificar, con el mayor detalle posible el mayor volumen de formatos, versiones y extensiones y, a partir de dicha identificación, orientar a la organización hacia formatos asentados y vinculados a la preservación digital, como TIFF, JPEG o PDF se atendía conforme a lo esperado. Finalmente, DROID se instaló de forma definitiva en versiones de producción de los clientes, incluyendo otras mejoras asociadas al cambio y que se detallan en el siguiente apartado.

Conclusiones

La preservación digital representa uno de los desafíos impostergables que enfrenta la sociedad de la información (Castilla Solano y Umaña Alpizar, 2019). El volumen de ficheros que preservar cada vez va a ser mayor (Azorín et al. 2020) y también aumentarán las necesidades de control posterior, si queremos asegurar su pervivencia y legibilidad en cualquier momento (Valentín-Ruiz, 2020). La generalización de la administración electrónica hará que la preservación sea más compleja (Térmens y Ribera, 2009) y sólo ahonda en esta problemática. Por otro lado, la abundancia de formatos es uno de los mayores desafíos a los que se enfrentan los gestores documentales y los encargados de la preservación (Moro Caballero, 2018). El empleo de sistemas de preservación digital no puede llevarnos a pensar que cuando un fichero ha sido preservado, podemos olvidarnos de él. Todo lo contrario, es necesario continuar con el plan de preservación, tener en mente los riesgos asociados al mismo porque no van a desaparecer, al menos en un futuro cercano (Voutssás Márquez, 2014, p. 260), optar por algún método de preservación, como la retrocompatibilidad o la migración de información (Barbedo, 2019, p. 19-20), etc. Es decir, aplicar los principios de la preservación digital es necesario desde ya y para toda la vida. El cambio de TIKa a DROID es parte del recorrido previsto para mejorar la respuesta a los requerimientos del modelo OAIS y de la norma ISO 14721, a través de los servicios de archivo y preservación ofertados por Odilo.

La evolución de TIKa a DROID ha permitido mejorar e incrementar las posibilidades de identificación de extensiones, versiones de formato y formato de los ficheros que preservar de las distintas organizaciones con que trabaja Odilo, principalmente archivos de la administración pública, aunque también contamos con archivos de otro tipo de entidades. Se trata de un aspecto muy positivo, desde el punto de vista de la preservación digital, pues profundiza en el análisis de las características de los ficheros de dichas organizaciones y permite aplicar recomendaciones de migración de formatos basadas en la información proporcionada en Pronom. Esto no significa que DROID vaya a ser empleada para siempre, podemos encontrarnos con formatos que no sean identificados, ni por esta herramienta ni por JHOVE, como indica Johnston (2018).

En el futuro, se pretende automatizar algunos de los procesos que actualmente son realizados bajo una estricta revisión del servicio de preservación, como es proporcionar información técnica sobre los ficheros de forma legible para el usuario final. Esta función de caracterización, que identificaba Moro Cabero (2018) completaría la actividad realizada en la actualidad por el equipo humano del servicio de preservación de Odilo. Así, además de ofrecer recomendaciones de software de visualización y edición de ficheros, de acuerdo con su formato, versión y extensión, se podría mostrar, por ejemplo, el tamaño del fichero o la última fecha de modificación.

Por otro lado, atendiendo a la función de normalización, se pretende implementar alertas automatizadas de recomendaciones de migración de versión de formatos o de formatos. Se pretende también incidir en las recomendaciones de uso de formatos no propietarios o, si lo son, de tipo abierto (Giusti, 2016; Azorín et al. 2020). Esto facilitaría el cumplimiento de los niveles 3 y 4 de la NDSA en cuanto a formatos de ficheros (NDSA, 2019). En definitiva, se trata de conocer la viabilidad de los formato y posibles riesgos asociados a estos, facilitando así la valoración y selección de formatos (Moro Cabero, 2018; Abrams, 2005) a lo largo del tiempo, sometiendo al plan de preservación a revisiones periódicas pertinentes.

Referencias bibliográficas

- Abrams, S. (2005). Establishing a Global Digital Format Registry. *Library Trends*, 54(1), 125-143. [10.1353/lib.2006.0001](https://doi.org/10.1353/lib.2006.0001)
- Adair, A., Esteva, M. y Chang, B. (2019). Early exists strategies in digital preservation. 16th International Conference on Digital Preservation. Amsterdam, The Netherlands, 2019. https://ipres2019.org/static/pdf/iPres2019_paper_111.pdf
- Álvarez Wong, B. (2017). Los repositorios digitales para la conservación: Un acercamiento a la preservación digital a largo plazo. *Ciencias de la Información*, 48(2), 15-22. <https://www.redalyc.org/pdf/1814/181454540003.pdf>
- Azorín, C., Barrueco Cruz J. M. y Bernal, I. (2020). *Guía para la evaluación de los procesos de preservación en repositorios institucionales de investigación*. <http://hdl.handle.net/20.500.11967/634>
- Baños Moreno, M. J. (2018). El servicio de preservación digital como garantía de control de los riesgos asociados a la documentación de archivo: El caso de Odilo Preserver. *VII Seminario Hispano-Brasileño de Investigación en Información, Documentación y Sociedad*. Madrid, noviembre de 2018. <http://seminariohispano-brasileiro.org.es/ocs/index.php/viishb/viishbucm/paper/viewFile/351/35>
- Barbedo, F. (2019). *Recomendações para a produção de planos de preservação digital*. http://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2019/08/Recomendacoes_PPD_v2.pdf
- Candás Romero, J. (2006). El papel de los metadatos en la preservación digital. *El profesional de la información*, 15(2), 126-136. <https://rdu-demo.unc.edu.ar/xmlui/handle/123456789/875>
- Castilla Solano, M. G. y Umaña Alpízar, R. (2019) Modelo para la preservación de documentos digitales. *Revista del archivo nacional (RAN)*, 1983, 1-12, 129-182.
- Depoortere, R., Geeraert, F., Soyez, S. y Vandepontseele, S. (2020). A roadmap for establishing a Belgian web archive at the federal level. *In Monte Artium*, 13, 25-42. <https://doi.org/10.1484/J.IMA.5.122153>

- Giusti, M. R. (2016). Las dificultades de la preservación digital: problemas, desafíos y propuestas para los repositorios. *Conferencia Internacional BIREDIAL-ISTEC. San Luis Potosí, México, octubre de 2016*. <http://sedici.unlp.edu.ar/handle/10915/56288>
- Green, K., Niven, K. y Field, G. (2016). Migrating 2 and 3D Datasets: Preserving AutoCAD at the Archaeology Data Service. *ISPRS International Journal of Geo-Information*. 5(4). <https://doi.org/10.3390/ijgi5040044>
- Johnston, L. (2018). Creating a Holdings Format Profile and Format Risk and Digital Preservation Prioritization Matrix at the National Archives and Records Administration. *Proceedings of the 15th iPres International Conference on Digital Preservation, Boston, MA, USA*. <https://osf.io/ctw3g/>
- Moro Cabero, M. (2018). Identificación, caracterización y selección de formatos para la preservación del recurso digital. *MEI (Métodos de información)*, 9(16), 49-90. <https://doi.org/10.5557/IIMEI9-N16-049090>
- Narasaiah, I., Tripathi, M. y Awasthi, S. (2021). Digital Archiving of Grey Literature: a case study of Jawaharlal Nehru University, New Delhi. *Library Philosophy and Practice*. <https://digitalcommons.unl.edu/libphilprac/4820>
- NSDA (2019). *Levels of Digital Preservation*. <https://ndsa.org/publications/levels-of-digital-preservation/>
- Térmens, M. y Leija, D. (2017). Auditoría de preservación digital con NDS-A Levels. *Interinformación, Actas de XI Jornadas Españolas de Documentación. Zaragoza, mayo de 2009*. 139-146. <http://www.fesabid.org/zaragoza2009/www.fesabid.org/zaragoza2009/actas-fesabid-2009/139-146.pdf>
- Térmens, M. y Ribera, M. (2009). El control de los formatos en preservación digital. *FESABID*, 26(3), 447-456. <https://doi.org/10.3145/epi.2017.may.11>
- UNE-ISO 14721:2015, Sistemas de transferencia de datos e información espaciales. Sistema abierto de información de archivo (OAIS) Modelo de referencia. AENOR (Asociación Española de Normalización y Certificación), 2015.
- Valentín Ruiz, F. J. (2020). Retos tecnológicos de los archivos en tiempos de pandemia o cómo lograr la máxima eficiencia en la era de la virtualidad. *Tábula*, 23, 229-245. <https://doi.org/10.51598/tab.850>
- Voutssás Márquez, J. (2014). Metadatos para preservación digital. En: A. A. Rodríguez García. *La revolución de los datos bibliográficos, científicos y culturales*, 247-264. http://ru.iibi.unam.mx/jspui/handle/IIBI_UNAM/28

THE BRAZILIAN CURRENT RESEARCH INFORMATION SYSTEM: BRCRIS

Adilson Luiz Pinto¹, Washington Luís Ribeiro de Segundo²,
Luc Quoniam³, Thiago Magela Rodrigues Dias⁴,

¹PGCIN/UFSC, adilson.pinto@ufsc.br, ORCID iD 0000-0002-4142-2061

²IBICT, washingtonsegundo@ibict.br, ORCID iD 0000-0003-3635-9384

³UFMS, mail@quoniam.info, ORCID iD 0000-0002-6333-6594

⁴CEFET/MG, thiagomagela@gmail.com, ORCID iD 0000-0001-5057-9936

Resumo

Objetivos: A proposta visa recuperar e tratar um conjunto de dados para criação de sistemas de recomendações em dados abertos, utilizando os dados das seguintes bases: OasisBr/IBICT, BDTD/IBICT, Plataforma Lattes, Espacenet, WikiData, Diadorim/IBICT, Latindex, DOAJ, UlrichsWeb, ROR, Dimensions, Altmetric, Plataforma Sucupira, OpenAIRE, DBpedia, OpenCitations, SherpaRomeo, Miguelin e Google Scholar Metrics.

Métodos: Uso de Machine Learning e Ciência de dados para a construção de sistemas de recomendações para identificar revistas chave de áreas científicas; pesquisadores chave de temáticas específicas; pesquisadores chave a partir de suas participações em orientações/bancas de mestrado/doutorado, e; especialistas em inovação, a partir de grande participação em determinadas temáticas de patentes. Para podermos identificar este cenário foi estruturado um ecossistema de sistemas de informação (científica e tecnológica), focado em identificar Pesquisadores, Infraestrutura de Laboratório e Equipamentos, Projetos de Pesquisa, Financiamento, Instituições de Pesquisa e Outputs (conjunto de dados, patentes, artigos científicos, teses/dissertações, softwares). Para reunir e congregiar todos estes dados foi criado um grande banco de dados a partir do La Referencia, com representações e Vivo Ontology e Elastic Search.

Resultados: Construção de um sistema aberto para identificar a informação científica e tecnológica de qualidade do Brasil. Até o momento foi criada a estrutura para dois dos quatros sistemas de recomendações (especialistas em produção científica e especialistas em teses de dissertações), os outros dois sistemas de recomendações (especialistas em patentes e sistema de especialidade para editores científicos) estão em fase final, com proposta de métricas e visualização já definidas. Os formatos e dados utilizados foram em XML e JSON, no qual foi realizada uma organização deste conhecimento em semântica e classificação. O propósito final do estudo foi: Representar toda informação da Plataforma Lattes, Sistemas CRIS, Bancos de patentes e outras fontes de ecossistemas em C&T; Aplicar de forma integradora e representativa ao público todo o processamento dos dados, identificação de entidades e seus relacionamentos; reestruturação dos

dados, em poucos casos, visto que todo o sistema já está em concordância com os dados da La Referencia, matriz da normalização dos dados; aplicação integradora da deduplicação dos dados, para entidades, pesquisadores, e; consolidação do Dashboard de busca.

Conclusão: No início da proposta um dos pontos era criar sistemas de recomendações com visualizações e métricas que pudessem identificar os atores mais relevantes, porém o que se conseguiu até o momento (agosto/2021) foram algumas visualizações e algumas métricas de frequência, porém dentro do cronograma do grupo a proposta de estruturação de todo o processo está dentro dos prazos previstos, no qual se pretende que em dezembro/2021 já tenhamos o sistema para consulta pública e com todas as métricas informativas aplicadas. Por outro lado, a base estrutural dos sistemas já está construída e já é uma realidade para sistemas de recomendações de dados brasileiros, e podemos afirmar que este sistema é único no Brasil e na América Latina.

Palavras-chave: BrCRIS, Sistemas de Recomendações, Bases de dados e Repositórios, Fontes de Informação, Ciência de Dados.

Introdução

Com limitações de alocações de recursos públicos na área de pesquisa, é cada vez mais importante analisar os resultados alcançados (análises retrospectivas) bem como tentar otimizar o uso dos recursos no futuro (análises prospectivas). Qualquer análise desejada, precisará dos dados organizados e acessíveis. O projeto BrCRIS pretende o levantamento e estruturação dos dados, sobre o ecossistema da pesquisa brasileira, para propiciar essas análises em nível do Brasil.

O *Brazilian Current Research Information System* (BrCRIS) é a idealização de um ecossistema de sistemas, e tem a interoperabilidade como seu ponto central. Neste contexto, o grande problema do BrCRIS é reunir os dados de Projetos de pesquisa, financiamento, pesquisadores, infraestrutura de pesquisa, instituições de pesquisa e seus *outputs* em C&T.

A ideia original do BrCRIS nasceu em 2014, e foi proposta pelo Instituto de Brasileiro de Informação em Ciência e Tecnologia (IBICT), tendo como inspiração o Projeto PTCRIS, da mesma natureza, mas executado no contexto de Portugal pela agência portuguesa de financiamento, Fundação para Ciência e Tecnologia (FCT), que por sua vez teve seu surgimento alinhando ao sistema “*Directory of Research Information Systems*” que é uma iniciativa da euroCRIS para cartografar a informação científica disponível na Europa (Pereira, 2015).

O problema central do BrCRIS é o de que estes dados, no Brasil, estão dispersos em várias plataformas/repositórios e outros sistemas, com linguagem e indexações diferentes, e de difícil acesso. A **proposta** do BrCRIS é justamente a reunião de todos os dados/metadados de pesquisa em C&T brasileira em um único espaço, com relações e sistematização dos seus dados para uma cartografia de forma consistente.

Os métodos para esta proposta estão baseados em módulos criados em Python e na utilização da Plataforma LA Referencia, servindo como uma orquestração dos dados de inserção e de exportação. Os processos estão organizados da seguinte forma:

– A Plataforma LA Referencia é utilizada para armazenar e exportar os dados coletados, sendo também uma interface de dados de visualização e de busca infor-

macional e ao mesmo tempo servindo para a interoperabilidade de dados com outras plataformas e outros conjuntos de dados. A forma com que isso está sendo executado é via API's, permitindo um contexto semântico do universo de dados. A LA Referencia da visibilidade para a produção científica do ensino superior e instituições de pesquisas da América Latina, promovendo acesso livre e gratuito ao texto integral, privilegiando os trabalhos financiados pelo setor público e em acesso aberto.

– Um segundo aspecto conta com o mapeamento de um conjunto de plataformas e bases de dados, como: OasisBr/IBICT, BDTD/IBICT, Plataforma Lattes, Espace-net, WikiData, Diadorim/IBICT, Latindex, DOAJ, UlrichsWeb, ROR, Dimensions, Altmetric, Plataforma Sucupira, OpenAIRE, DBpedia, OpenCitations, SherpaRomeo e Google Scholar Metrics. Neste conjunto, algumas plataformas servem para a coleta, outras para homogeneização dos dados e, enfim, outras para certificar a existência dos dados.

– O terceiro passo é um modelo de organização das entidades deste conjunto de dados, visando identificar quais são: os projetos, serviços, programas de pós-graduação, cursos, organizações e pesquisadores no intuito de ter um arcabouço de dados de patentes, publicações, datasets e poder verificar como estas entidades se relacionam. Esta ordenação será executada pelas ferramentas Elasticsearch e Kibana.

O objetivo central desta proposta, a partir da reunião de todos os dados em C&T, é apresentar sistemas de recomendações em níveis de identificação de especialistas na pesquisa científica, na pesquisa de patentes, na tutoria de pesquisas (teses e dissertações), na avaliação de pesquisas (bancas de defesas) e no reconhecimento de categorias de revistas chave para áreas/campos de conhecimento.

Os principais resultados almejados são: (1) sistemas de busca com *dashboards* que tenha a dimensão de poder vislumbrar a organização e representação de dados científicos e tecnológico brasileiros; (2) que sejam sistemas de recomendações para identificação de especialidades, como revistas, pesquisadores, inventores, orientadores/tutores. Para a representação de todo este universo temos a junção de ser visto por (a) métricas e (b) visualização da gráfica da informação.

Desenvolvimento

Este estudo se caracteriza dentro da organização e representação da informação e do conhecimento, em especial porque contempla vários elementos de análise de domínio que Hjørland (2004, p. 17) propõe, como: produção de classificações especiais (no caso está sendo feita uma readequação de sistemas classificatórios de diferentes bases utilizadas); sistema de indexação e recuperação da informação; estudos bibliométricos; estudos terminológicos e linguagens para propósitos específicos; estudos sobre estruturas e instituições na comunicação científica, e; análise de domínio em cognição profissional e de inteligência artificial.

O estudo também aplica estruturação de análise em várias frente, como sistemas de recomendações de cunho científico/tecnológico, tendo como base: (1) revistas acadêmicas de uma determinada área/campo de conhecimento; (2) especialistas em determinadas áreas/campos pela sua publicação científica; (3) especialistas em

determinadas áreas/campos pela sua atividade tecnológica – produção de patentes, e; (4) especialistas em determinadas áreas/campos pelas suas orientações de teses/dissertações e pela participação assídua em bancas de defesas de teses/dissertações (Pinto et al, 2021).

A proposta de sistemas de recomendações está baseada na iniciativa *Current Research Information System*¹ (CRIS) que procura armazenar, gerenciar e trocar metadados contextuais para a atividade de pesquisa financiada por órgãos governamentais ou conduzida em uma organização executora de pesquisa. Várias iniciativas para a criação deste tipo de sistema estão sendo desenvolvidas atualmente (Bevan; Harrington, 2011; Jetten et al, 2019), com especial destaque para a EuroCRIS.

A continuação temos o modelo de ecossistema apresentado pelo BrCRIS.

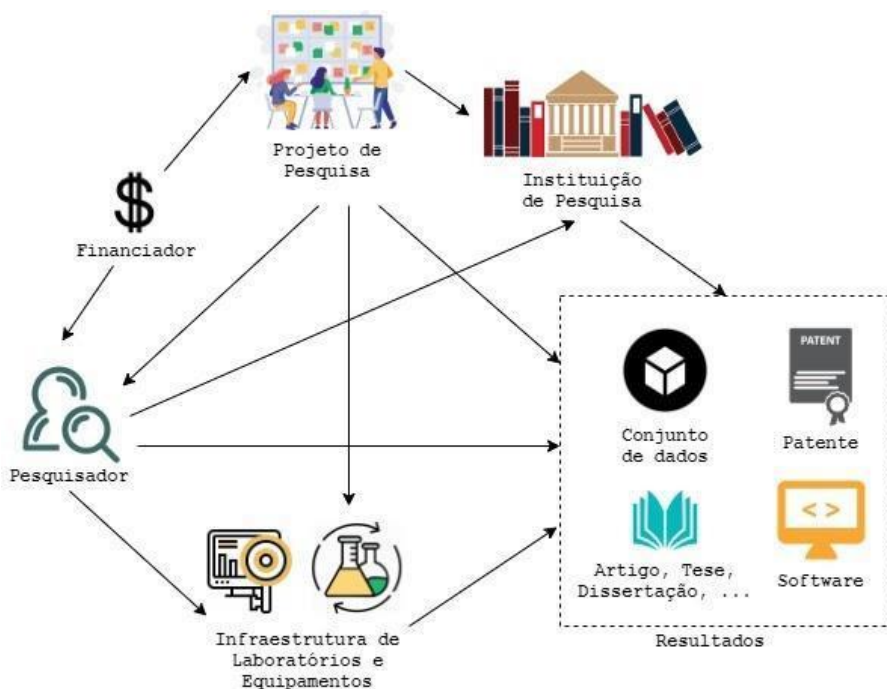


Figura 1. Ecossistema da pesquisa científica

O projeto BrCRIS é um ecossistema de sistemas, com foco em executar a interoperabilidade de dados científicos e tecnológicos. Este sistema integra a troca de dados de vários atores, como agências de financiamento, bases de dados científicas e tecnológicas, repositórios institucionais, plataformas governamentais, infraestrutura de pesquisa, bases de teses e dissertações entre outros atores.

Como ponto de vislumbrar este ecossistema, no proposta contorna todo o universo de dados priorizando a sua representação, neste caso em sistemas de recomendações.

¹ <https://www.eurocris.org/why-does-one-need-cris>

Resultados

Para este apartado iremos apresentar e discutir toda a cadeia informativa do BrCRIS, no qual a coleta e o enriquecimento de dados de repositórios e bases de dados foi realizada pela Plataforma LA Referencia, a partir de protocolos OAI-PMH e múltiplos formatos de dados em XML e JSON, tendo como foco:

- Adicionar fontes de informação (Plataforma Lattes, Sistemas CRIS, Bancos de patentes e outras fontes de ecossistemas em C&T);
- Processar dados, identificação de entidades e seus relacionamentos (exemplo: pessoas/publicações para relacionamento de autoria);
- Armazenar entidades e suas relações de forma flexível e adaptável aos ecossistemas de cada país, visando não duplicar esforços;
- Implementar mecanismos para identificar entidades repetidas em diferentes fontes e unificá-las (deduplicação);

Fornecer serviços de valor agregado para outros atores do sistema, como busca de produção de pesquisadores das plataformas e repositórios agregados;

- Alimentar um serviço próprio, com estatísticas, buscadores e um *dashboard*.

Em um segundo passo, temos as etapas de obtenção e transformação dos dados, no qual para os dados da Plataforma Lattes foi utilizado o protocolo SOAP do Extrator Lattes, conseguindo a extração de todos os Curriculum Vitae desta plataforma e ao mesmo tempo não sobrecarregando o sistema. Para as informações de patentes foi utilizado parte da Plataforma Lattes e conteúdos do Patent2Net (Ferraz et al, 2016) de recuperação da base de dados Espacenet. Em relação aos conteúdos de revistas, recuperamos informações das principais bases e diretórios de revistas científicas, no intento de ter a maior gama possível de dados destes objetos. Outro aspecto desta etapa foi gerar entidades, para sua normalização, e os campos de cada base utilizada na criação do BrCRIS.

Para que tudo fosse consolidado, foi realizada uma transformação de dados em formato único de XML, dentro da biblioteca de dados, onde foi trabalhado como estes dados poderiam ser inseridos no modelo de dados estabelecido pela Plataforma LA Referencia. Criou-se as entidades e suas instâncias com a agregação de múltiplos identificadores persistentes externos e internos ao modelo BrCRIS. Gerou-se assim os relacionamentos entre as instâncias de entidades, na construção de um *grafo de conhecimento* que é exportado para interface de busca e para o *dashboard* de visualização.

A terceira parte do BrCRIS foi gerar a infraestrutura, trabalhando *on premises* com os servidores do IBICT em GNU/Linux, de forma virtual e uso de software livre (*scripts* em Python). O foco da infraestrutura é apresentar dois universos (o mundo exterior e o mundo interior).

Em relação ao mundo exterior, o sistema é apresentado pelo Kibana para visualizar o *dashboard*. Também serão fornecidas API's, que permitem o consumo dos dados agregados e geração, outras buscas e visualizações. Já no mundo interior temos o Elasticsearch que facilita a organização e busca de grande quantidade de dados, e reflete toda a orquestração de organização dos dados, serviço de mensageria, e trabalha recebendo os dados diretamente com o LA Referencia, que por sua vez utiliza uma base PostgreSQL.

Como *dashboard* desta proposta, temos dois pontos chave: (1) a geração de indicadores pelo universo de dados coletados e exportados, (2) a visualização deste conteúdo.

Pensando em algo prático, o BrCRIS se apresenta com repertórios de especialidades, como já mencionado nos nossos objetivos, visando mostrar os esforços do projeto em: (1) identificar as especialidades de pesquisadores em determinado tema e/ou áreas de conhecimento; (2) identificar especialistas em orientações e/ou participação de bancas de teses/dissertações de áreas/campos de conhecimento; (3) identificar revistas específicas em determinadas áreas e/ou temas de conhecimento, e; (4) identificar especialistas (inventores ou detentores de tecnologia) em patentes.

Para os dados de produção científica, foi utilizado como base as informações recuperadas e tratadas do Diadorim, Latindex, Portal ISSN, DOAJ, UlrichsWeb, Plataforma Lattes (Doutores e Bolsistas de Produtividade), Google Scholar Metrics. Foram priorizados indicadores voltados a: (a) tempo de resposta da submissão até a publicação; (b) h-5 média; (c) taxa de publicação; (d) comitê da revista; (e) endogeneia, e; (f) idiomas de publicação. A visualização destes dados estiveram preocupados em identificar as áreas de conhecimento, a relação das áreas de conhecimento, a junção de terminologias nas áreas, a relevância, o crescimento/decréscimo de alguma tipologia de publicação, a clusterização de áreas/pesquisadores, os picos evolutivo das áreas, e a evolução das tipologias de publicação.

Para visualização, exploração e análise das informações coletadas no âmbito do projeto BrCRIS e carregadas na plataforma LA Referencia, foram desenvolvidas várias rotinas de exportação e indexação dos dados. Uma destas rotinas exporta os dados para o formato RDF, seguindo o modelo semântico, baseado na ontologia VIVO, adotado para o BrCRIS. O grafo resultante é importado na ferramenta VIVO, uma plataforma de visualização de dados do domínio de pesquisa acadêmica que permite não só a exibição de todos os atributos de uma determinada entidade (Pessoa, Organização, Publicação, etc.), como também a navegação dos dados, explorando-se os relacionamentos entre as diferentes entidades (por exemplo, as pessoas afiliadas a uma organização, as publicações de autoria de uma pessoa, o periódico contendo uma publicação, e assim por diante). A plataforma VIVO também oferece alguns gráficos, como a rede de coautores de um pesquisador (Figura 2), além de agregar métricas para publicações, como Altmetric e PlumX (Figura 3).

Em relação às teses e dissertações foram utilizadas fontes de dados da Biblioteca Digital Brasileira de Teses e Dissertações e conteúdos da Plataforma Lattes. Os tipos de indicadores que nortearam este sistema de recomendações foram: (a) Genealogia por até 11 níveis; (b) orientações por regiões, (c) país, (d) especialistas em temáticas em orientações ou por participação de bancas. A visualização representada esteve voltado em averiguar a genealogia científica, quais pesquisadores orientaram um conjunto de outros pesquisadores nos níveis de formação – graduação, mestrado e doutorado, a origem das defesas em nível de regionalização, as defesas fora do Brasil, as relações institucionais e as temporalidades dos pesquisadores que mais orientaram por áreas e/ou temáticas.

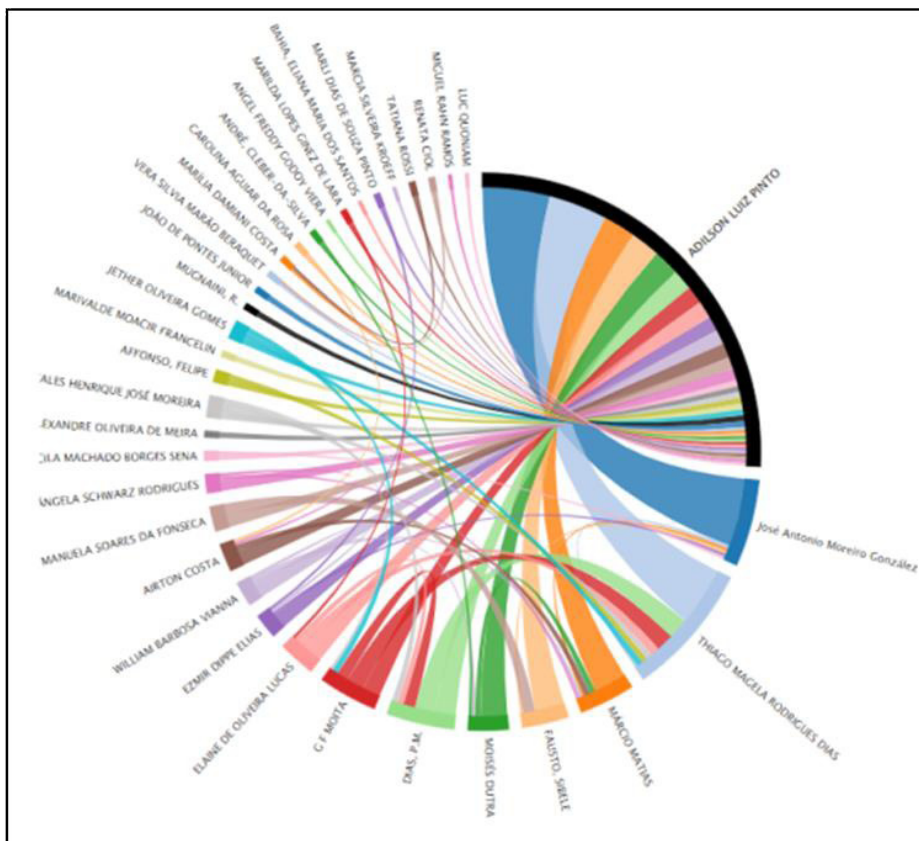





Figura 2. Rede de coautores

Science for Education network: The Brazilian proposal | Document    24

[10.1787/562a8659-en](https://doi.org/10.1787/562a8659-en)
<https://doi.org/10.1787/562a8659-en>

Overview Identity Additional Document Info Other View All

type
 book part

authors
[JESÚS PASCUAL MENA CHALCO](#)

Figura 3. Detalhe de uma publicação, com as métricas associadas no canto superior direito

Outra rotina importante realizada foi a indexação dos dados para a plataforma Elasticsearch. Uma vez indexados, os dados são visualizados na ferramenta Kibana, onde é possível criar dashboards para análise dos dados através de gráficos e totalizações envolvendo correlações entre diferentes variáveis. O dashboard apresentado na Figura 4 reúne análises sobre as teses defendidas no Brasil. O histograma mostra o total de teses defendidas por ano, enquanto as tabelas quantificam os totais de teses por instituição, por área de conhecimento, por orientador, etc.

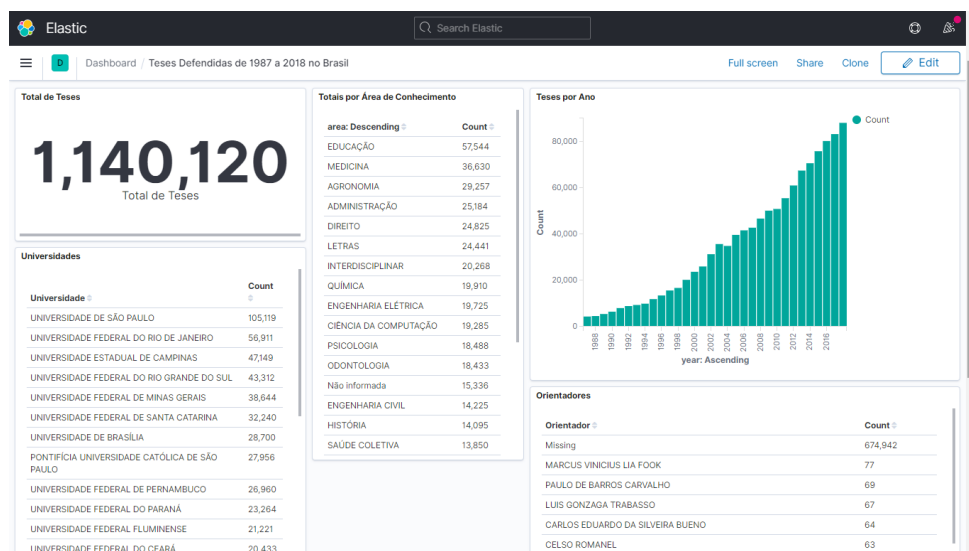


Figura 4. Dashboard para teses defendidas no Brasil

Em se tratando de conteúdos de revistas científicas, foi criado o sistema de recomendações voltado para editores científicos conseguirem identificar as especialidades de determinadas publicações e também tem o viés de identificação de possíveis pares científicos. Esta proposta foi estruturada utilizando as seguintes bases de dados: Latindex, Portal ISSN, DOAJ, UlrichsWeb, Plataforma Lattes (Bolsistas de Produtividade, Doutores/Mestres), Google Scholar Metrics. O foco deste sistema está em dados métricos, como: (a) h-5 média, (b) h-5 mediana, (c) produção de bolsistas, (d) produção de mestres.

A sua representação se dá por um dashboard (Figura 5) que concentra as estatísticas sobre periódicos, com a distribuição por idioma, a nuvem de palavras dada pelas palavras-chave associadas, bem como os totais de periódicos por área de conhecimento. Análises similares, bem como outros tipos de visualizações como gráficos de linha, de barras, de pizza, etc., podem ser construídas para qualquer uma das entidades coletadas e indexadas a partir da Plataforma LA Referencia, gerando indicadores a partir de qualquer variável de interesse.

Atendendo as necessidades dos dados de patentes, foi estruturada junção de dados a base Espacenet e da Plataforma Lattes para criar um sistema de recomendações para a produção de patentes. Os índices gerados trabalham com: (a) citações recebidas, (b) citações realizadas, (c) concessão/prorrogação, (d) família de patentes,

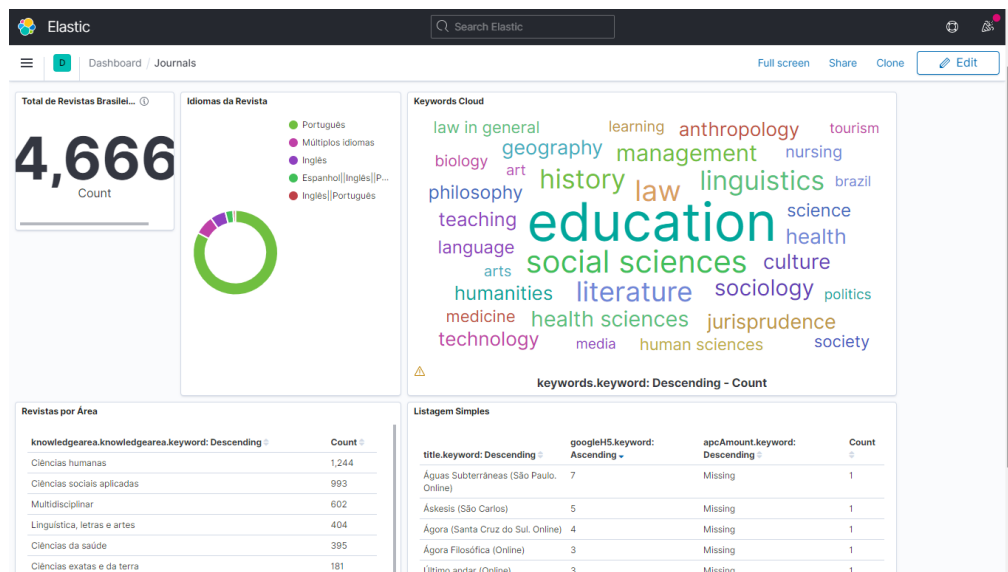


Figura 5. Dashboard para periódicos

(e) triádica, (f) classificação. As visualizações estimulam identificar quais são as citações recebidas/realizadas, saber se as patentes estão sendo protegidas em outro país e na triádica mundial, identificar se existe concessão/prorrogação das patentes, identificar a relação de inventores/depositantes, identificar valor econômico direto (tamanho da família), valor econômico indireto (anuidade) e valor tecnológico (citações recebidas).

Como resultado principal, busca-se a certificação de declarações de patentes nos Currículos dos pesquisadores, sendo o código informado, exato ao existente no Portal da Espacenet, ou aproximado, cujo casamento é obtido pela busca utilizando-se as estratégias já descritas.

Para além da certificação, tem-se as opções de visualização da informação agregada, nas classificações de grandes áreas e áreas dos pesquisadores descritos como inventores nos registros de patentes. Visualizações do tipo nuvem de tags são adequadas à verificação de frequência das áreas e palavras-chave envolvidas. A evolução temporal, quantidade por ano, de patentes e a distribuição espacial dos pesquisadores/instituições, requerentes das patentes são também alvo para a construção de dashboards de visualização. Uma atenção particular deve ser dada a análise das classificações patentárias IPC e CPC, por serem obrigatórias em qualquer patente, por ser independente da língua (se adequa para patente de qualquer país de origem e enfim por OMPI, fornecendo uma API reindexando automaticamente qualquer texto com essas classificações, fornecendo assim uma verdadeira passarela potencial entre ciência (publicações, teses) e tecnologia (áreas patentárias).

Ainda em relação às possíveis análises sobre a massa de dados agregada, há que se estabelecer índices de valor de patentes, observando-se, diferentes fatores: (1) se uma dada patente é classificada como triádica, o que significa que possui registros

nos escritórios de patentes do Estados Unidos, Europa e Japão; (2) se pertence a uma estratégia territorial (regional tipo mercosul, África, Ásia ou estritamente local, tipo Brasil); (3) se possui citações realizadas por outras patentes; (4) se ela possui um forte embasamento teórico devido a extensas referências patentária e bibliográfica; (5) se faz parte de uma família de patentes que caracteriza uma invenção.

Por fim, tem-se que a base de dados estabelecida deve ser utilizada também para a pesquisa sobre patentes que não possuem proteção no Brasil, ou que proteção não está vigente por várias razões (uma sendo a falta de pagamento das anuidades, todas razões que tornam o conhecimento da patente de domínio público e que portanto poderiam ser exploradas tecnologicamente ou comercialmente sem obrigação de pagamento de royalties). Há ainda a pesquisa sobre patentes que possam ser aplicadas à solução de problemas técnicos, transferência de tecnologia, com baixo custo na realização ou por finalidade. Estes casos caracterizam o que se denomina por Inovação Frugal (Bhatti, 2012). As patentes podem constituir uma verdadeira base de respostas técnicas.

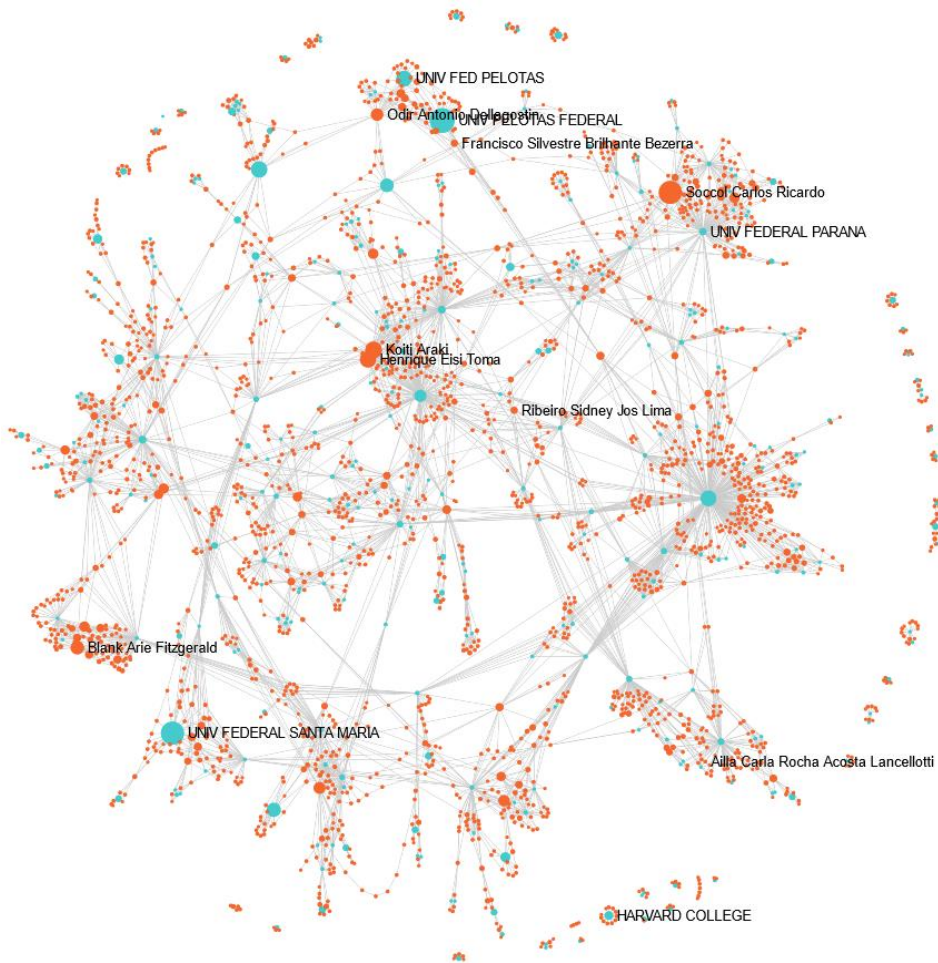


Figura 6. Representação por grafo de inventores e instituições produtoras de patentes

Conclusões

Os sistemas de recomendações criados são de quatro naturezas: (1) saber quem são os especialistas pelas publicações; (2) saber quem são os principais orientadores no país, por áreas de conhecimento; (3) as revistas e quais são direcionadas a determinadas temáticas e inclusive o impacto de suas publicações em termos de citações, e; (4) saber quais são as especialidades e especialistas em desenvolvimento de tecnologia no país, mais exatamente em termos de patentes.

A execução de tal projeto tem um grupo de pesquisadores voltados a Big Data, Data Science, Machine Learning, Ontologias e Tratamento e Disseminação da Informação. As limitações estão em conseguir reunir todo o universo de dados para a consolidação de um sistema integrador. Existem dados diversos com formatos diversos, portanto o grande detalhe aqui é re-indexar esta informação em um formato legível pelos sistemas criados.

Para questões mais teóricas, é um modelo único no Brasil e tende a surtir pelo menos 20 estudos da aplicabilidade, tratamento e consolidação dos dados. Também está sendo criado um banco de informação para termos um modelo teórico de sistemas para ser replicado por outros setores do IBICT.

Consideramos que este sistema irá trazer ótimos serviços públicos para a sociedade científica e tecnológica, tendo estes dados abertos para consultas e até mesmo para futuras colaborações de outros órgãos no país.

Os primeiros resultados já se fazem presente no módulo interno do projeto BrCRIS, porém teremos uma versão pública dos quatro sistemas e recomendações em outubro de 2021, já com modelos de representação para os bolsistas de produtividade em nível 1A no Brasil; das patentes da Universidade Federal do Mato Grosso do Sul; das informações de saúde, em especial de SARS-CoV-2, com instituições brasileiras de infectologia; de agregação dos dados de revistas científicas pelo Miguelin, plataforma de recomendações à editores científicos, e; integração dos dados da Base Digital de Teses e Dissertações do IBICT para identificação de pais científicos, especialidade por bancas de defesas e genealogia científica.

Referências

- Bevan, Simon J.; Harrington, John (2011). Managing research publications: lessons learned from the implementation of a current research information system. **Serials: The journal for the Serials Community**, 24(1): 26-30. Disponível em: <https://doi.org/10.1629/2426>. Acesso em: 20 ago. 2021.
- Bhatti, Yasser Ahmad. What Is Frugal, What Is Innovation? Towards a Theory of Frugal Innovation. **SSRN Electronic Journal**, fev. 2012. Disponível em: <https://dx.doi.org/10.2139/ssrn.2005910>. Acesso em: 10 ago. 2021.
- Ferraz, Renato Ribeiro Nogueira et al. (2016). Example of open-source OPS (Open Patent Services) for patent education and information using the computational tool Patent2Net. **World Patent Information**, (46): 21-31. Disponível em: <https://doi.org/10.1016/j.wpi.2016.05.002>. Acesso em: 20 ago. 2021.
- Hjørland, Birger (2004). Domain analysis: a socio-cognitive orientation for Information Science research. **Bulletin of the American Society for Information Science and**

- Technology**, 30(3): 17-21. Disponível em: <https://doi.org/10.1002/bult.312>. Acesso em: 18 abr. 2020.
- Jetten, Mijke; Simonsb, Ed; Rijndersc, Jan (2019). The Role of CRIS's in the research life cycle: a case study on implementing a FAIR RDM policy at Radboud University, the Netherlands. **Procedia Computer Science**, 146: 156-165. Disponível em: <https://doi.org/10.1016/j.procs.2019.01.090>. Acesso em: 12 ago.2021.
- LA Referencia, Red de repositorios de acceso abierto a la ciencia** (2021). <http://www.lareferencia.info/en/>. Acesso em 28/02/2021
- Pereira, Maria de Nazaré Freitas (2015). Developments in research information and CRIS in Brazil. In: **euroCRIS Strategic Membership Meeting Autumn 2015**, Barcelona. Disponível em: <http://hdl.handle.net/11366/418>. Acesso em: 20 ago. 2021.
- Pinto, Adilson Luiz et al (2021). Identificación de especialistas basado en orientaciones y participación en tribunales de tesis. In: **III LATmétrics, Medellín, Colombia**, Universidad de Antioquia, 2021.

ORGANIZAÇÃO DO CONHECIMENTO NO BRASIL: UMA ANÁLISE DO DOMÍNIO PELAS TESES DE DOUTORADO EM CIÊNCIA DA INFORMAÇÃO

Adilson Luiz Pinto¹, José Augusto Chaves Guimarães²,

¹PGCIN/UFSC, adilson.pinto@ufsc.br, ORCID iD 0000-0002-4142-2061

²PPGCI/UNESP, chaves.guimaraes@unesp.br, ORCID iD 0000-0002-0310-2331

Resumo

Objetivos: Considerando a organização do conhecimento como área nuclear e estratégica no contexto da Ciência da Informação, visto permitir que o conhecimento socialmente produzido seja acessado, apropriado e possa gerar novo conhecimento, realiza-se uma análise de domínio, sob a abordagem bibliométrica, das teses de doutorado defendidas nos treze programas de pós-graduação em Ciência da Informação do Brasil ao longo do tempo, a partir de três aspectos: os pais científicos, os número de orientações dos filhos científicos e a participação em bancas de defesa.

Métodos: Como base para o estudo utilizamos estratégia de busca dentro da Plataforma Lattes (Currículos do pesquisadores) pelos termos “Organização do Conhecimento e Representação do Conhecimento”, recuperando 227 especialistas no tema. Para identificar a depuração dos dados criamos alguns mecanismos em duas frentes: (1) redução dos especialistas em participação das teses (pais científico, orientador ou membro da banca); (2) aplicação de métricas para chegarmos de fato aos especialistas em Organização e Representação do Conhecimento (ORC), com uma análise minuciosa para não contabilizarmos outros temas como parte da ORC, como Gestão do Conhecimento, Sistemas de Conhecimento, Organizações empresariais, Política de informação, Estudos métricos. Para as especificações métricas utilizamos para identificar os orientadores e membros de bancas a Lei de Lotka, e para identificar os pais científicos utilizamos o modelo eletista de Price. Outro detalhe foi identificar os autores pelo nome de suas instituições, visando validar que alguns são de fato especialistas no tema.

Resultados: Os resultados evidenciam a consonância entre esses três universos o que permitiu identificar uma comunidade experiente na área, aspecto que se reforça pela presença predominante de bolsistas de Produtividade em Pesquisa do CNPq, o que atesta o capital social e intelectual desse grupo. Como destaque tivemos: (1) pais científicos tivemos como destaque os seguintes pesquisadores – Unesp_2, IBICT_4, Unesp_1, Unesp_3, UFF_2 e USP_9; (2) orientadores tivemos como destaque os seguintes pesquisadores – Unesp_2, IBICT_4, Unesp_3, IBICT_2, UFF_2, Unesp_4, UFMG_7, USP_9 e UFMG_6, e; (3) membros de bancas tivemos – Unesp_2, seguido por Unesp_3, USP_3, UFMG_8, UFMG_7, UFF_1, Unirio_2, USP_4, Unesp_1, IBICT_4 e UnB_1.

Conclusões: O estudo chegou a conclusões que no Brasil existem grupos temáticos no tema, com destaque a algumas instituições, como Unesp, UFMG, USP e IBICT, todas instituições da região sudeste do país. Outro detalhe que os destaques no cenário na ORC são bolsistas de produtividade na área de concentração da Ciência da Informação. Os destaques nas três categorias de análises (pais científicos, orientadores e membros de banca) também são maioritariamente das instituições de destaque, como pouca inserção de instituições da região nordeste e sul do país. Finalmente, na comparação geral de todas análises e todos os especialistas chegamos a uma métrica de match entre dois especialistas: Unesp_2 e IBICT_4.

Palavras-chave: Organização do conhecimento, Comunidades epistêmicas, pais científicos, filhos científicos.

Introdução

A Organização do Conhecimento remonta ao início da civilização, acompanhando o transcorrer da história humana desde os tabletes de argila na Mesopotâmia, passando pelas categorias aristotélicas, a árvore de Porfírio, o Trivium e o Quadrivium, as glosas dos monges copistas, as concordâncias bíblicas e a taxonomia de Lineu, entre outros (Witty, 1973; San Segundo, 1993). Mas é notadamente a partir do século XIX que a organização do conhecimento assume uma nova dimensão, com as contribuições de classificacionistas como Dewey, Otlet, Harris, Brown Bliss e Ranganaghan, indo além da classificação dos saberes rumo a uma classificação de documentos (Pombo, 1998).

Nesse contexto, destaca-se a existência de um conhecimento gerado pela sociedade que, uma vez socializado por conta de sua materialização ou registro, constitui-se naquilo que Michael Buckland (1991) denomina “informação como coisa”. Esse conceito de materialidade decorre, como destaca Hansson (2013), de um contexto de produção social de documentos, constituindo uma base material para o próprio conceito de epistemologia, na medida em que esta se torna tão somente discernível, para a sociedade, a partir da produção e disseminação de documentos. No entanto, para que esse conhecimento socialmente produzido possa ter um uso social, necessário se torna um processo mediador, de organização, em que se estabelecem “substitutos do conhecimento” (surrogates of knowledge) (Olson, 2002), de modo a que os contextos de produção e de uso possam ser colocados em diálogo.

Tem-se, assim, um processo helicoidal (e não cíclico, como por muito tempo se acreditou), em que um conhecimento A é produzido, recebe uma organização B, tem um uso C e, desse uso e apropriação, gera-se um conhecimento D, que recebe uma organização E, para um uso F, etc. (Guimarães, 2008a). E é exatamente nesse contexto que têm lugar os denominados Sistemas de Organização do Conhecimento – SOC que, por sua vez, apresentam-se em uma multitude de configurações, tais como glossários, dicionários, diretórios, taxonomias, esquemas de classificação, listas de cabeçalhos de assunto, tesouros, mapas conceituais ontologias, entre outras (Zeng, 2008). Desse modo, e em virtude de uma crescente complexidade da realidade social, novas perspectivas se abriram aos SOC, que passaram a assumir funções

mais amplas, pressupondo abordagens híbridas, com novos designs e novas denominações (Vickery, 2008; Lauruhn; Groth, 2016).

No âmbito da Ciência da Informação foi a partir das obras de Henry Evelyn Bliss “The organization of knowledge and the system of sciences” (Bliss, 1929) e “The organization of knowledge in libraries” (Bliss, 1933) que a expressão organização do conhecimento teve lugar. No entanto, a configuração teórica da organização do conhecimento como campo de estudos deu-se notadamente a partir dos estudos de Dagobert Soergel, em 1971, na obra *Organization of Knowledge and Documentation*, e de Ingetraut Dahlberg, em 1973, na tese de doutorado intitulada *Foundations of Universal Organization of Knowledge* (Dahlberg, 1995, 2006).

Tendo seu marco institucional e investigativo mais importante na International Society for Knowledge Organization – ISKO, criada em 1989, na Alemanha (Frankfurt), por Ingetraut Dahlberg e Dagobert Soergel, a organização do conhecimento pode ser definida como uma disciplina:

[...] dedicada ao estudo e desenvolvimento dos fundamentos e técnicas do planejamento, construção, gestão, uso e avaliação de sistemas de descrição, catalogação, ordenação, classificação, armazenamento, comunicação e recuperação dos documentos criados pelo homem para testemunhar, conservar e transmitir seu saber e seus atos, a partir de seu conteúdo, com o fim de garantir sua conversão em informação capaz de gerar novo conhecimento. (Esteban Navarro; Garcia Marco, 1995).

Tem-se, pois, um universo de estudo cujo conteúdo se constrói a partir de processos (análise, síntese, representação, condensação etc), instrumentos (sistemas de organização do conhecimento) e produtos (catálogos, índices, tags, etc) e que, por sua vez, assume uma dimensão epistemológica (teoria, metodologias, interdisciplinaridades etc), aplicada (ferramentas e o aporte tecnológico) e cultural (visando a promover a compatibilidade entre as necessidades de comunicação global e o respeito às especificidades locais) (Guimarães; Dodebei, 2015).

No caso brasileiro, a pesquisa e a formação de recursos humanos em organização do conhecimento assumiu especial ênfase a partir da década de 1970, com a criação dos cursos de pós-graduação *stricto sensu* na área de Ciência da Informação. Nesse sentido, verificou-se um forte investimento público do país em qualificar pesquisadores no exterior para que pudessem dar conta da formação de docentes de ensino superior. Como resultado desse investimento, notadamente voltado para instituições de excelência da Europa e da América do Norte (em especial dos Estados Unidos, do Reino Unido, da França e da Espanha) três correntes teóricas da organização do conhecimento emergiram nos cursos de pós-graduação brasileiros em Ciência da Informação: a norte-americana (subject cataloguing), mais voltada para os produtos da organização do conhecimento em bibliotecas, a inglesa, centrada nos instrumentos de organização do conhecimento que pudessem fazer frente a realidade documentais especializadas (indexing) e a francesa (analyse documentaire) cujo foco reside nos processos lógico-linguísticos envolvidos no fazer da organização – e na representação – do conhecimento registrado (Guimarães, 2008b).

A vista de tal cenário, e considerando o lastro científico advindo da atuação, desde 2009, do capítulo brasileiro da International Society for Knowledge Organization que, desde 2011 vem realizando congressos bienais sobre tal temática, e, considerando ainda a existência, na atualidade, de treze programas de doutorado em Ciência da Informação no Brasil (Universidade de Brasília – UnB, Universidade de São Paulo – USP, Universidade Estadual de Londrina – UEL, Universidade Estadual Paulista – Unesp, Universidade Federal da Bahia – UFBA, Universidade Federal da Paraíba – UFPB, Universidade Federal de Santa Catarina – UFSC, Universidade Federal do Rio de Janeiro e Instituto Brasileiro de Informação de Ciência e Tecnologia UFRJ/IBICT, Universidade Federal Fluminense UFF, Fundação Mineira de Educação e Cultura FUMEC e Universidade Federal de Minas Gerais UFMG com dois programas), coloca-se a questão: como vem se constituindo a comunidade epistêmica de organização do conhecimento no Brasil a partir da atuação dos programas de pós-graduação, a partir dos pais científicos (orientadores) e dos filhos científicos (orientações e bancas de doutorado) nessa temática?

Para tanto, considera-se como comunidade epistêmica uma rede de especialistas com reconhecida expertise e autoridade científica em um dado domínio de conhecimento (Haas, 1992) evidenciando tanto um capital científico quanto um capital social (Bourdieu, 2004).

Metodologia

A presente pesquisa insere-se metodologicamente no âmbito da análise de domínio, mais especificamente nas abordagens bibliométrica e epistemológica propostas por Hjørland (2002), em que a abordagem bibliométrica visa à identificação dos fenômenos, neste caso o capital social de orientadores (pais científicos) e da combinação de pesquisadores em participação de bancas de teses e orientações (filhos científicos) como subsídios para a caracterização da comunidade epistêmica brasileira na área. A abordagem epistemológica, por sua vez, emerge da análise dos dados, tornando-se possível identificar núcleos temáticos e, ainda, influências teóricas predominantes.

Para tanto, procedeu-se à extração de dados da Plataforma Lattes, do Conselho Nacional de Desenvolvimento em Ciência e Tecnologia – CNPq, que reúne mais de 7 milhões de currículos de pesquisadores – dados de janeiro de 2021. Os dados foram extraídos a partir de busca no campo Assunto, pelos termos “Organização do Conhecimento” e “Representação do Conhecimento”, buscando-se ainda identificar, entre eles, os pesquisadores que atuam como bolsistas de Produtividade em Pesquisa do CNPq na área de Ciência da Informação (aqui entendida como evidência de uma “elite investigativa”), o que resultou em um conjunto de 227 currículos Lattes, todos com titulação mínima de doutorado.

Os dados recuperados foram analisados a partir da lei de Lotka (1926) 20/80 para determinar os especialistas nas temáticas relacionadas ao campo de pesquisa, de modo a identificar os “pais científicos” dentro do conjunto dos $n=227$ currículos Lattes que eram os orientadores deste universo de pesquisadores. A formulação dessa aplicação

se dá pela: $Y(n) =$ quantidade de autores que produzem n documentos; $C =$ constante para cada temática, e; $n^2 =$ quadrado da frequência dos autores.

$$Y(n) = \frac{C}{n^2}$$

Em relação aos “filhos científicos”, aplicou-se o Modelo de Elitismo de Price (1963), considerando (E) 50% da totalidade de cada conjunto de dados, e; (\sqrt{n}) elevando a raiz quadrada, obtendo-se, assim, um número médio de especialistas. Para a representação da elite foram consideradas tanto orientações quanto participações em bancas de teses de doutorado realizadas ao longo da carreira de cada pesquisador.

$$E = \sqrt{n}$$

Relativamente às bancas de orientação dos “filhos científicos”, considerou-se um universo de 82 orientadores (com um total de 654 orientações o que, aplicando-se a lei dos 50%, levou a 327 de cuja raiz quadrada chegou-se aos 18 primeiros orientadores) e 113 para bancas (com um total de 1755 participações o que, aplicando-se a lei dos 50%, levou a 877,5 cuja raiz quadrada se representou por 29 participantes de bancas de doutorado, entretanto arredondamos para 31 participantes justamente porque os pesquisadores 29 ao 31 tinham a mesma frequência de participação em bancas de doutorado).

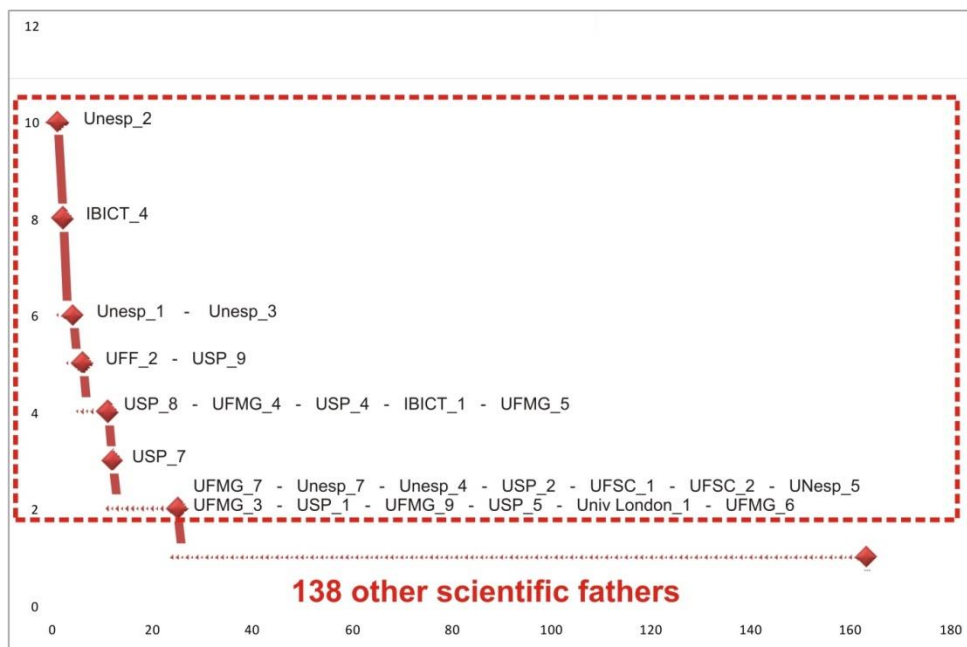
Resultados

Os resultados estão divididos em três blocos: lei de Lotka dos orientadores – os “pais científicos”; elite dos orientadores – os “filhos científicos”, e; elite das bancas de teses – também dos “filhos científicos”.

Pais científicos: são entendidos como orientadores e formadores de doutores, com papel relevante em dividir o seu conhecimento e atuando como importante fonte de informação e considerado um especialista sênior em uma determinada área e/ou temática de conhecimento. Na Organização do Conhecimento, a partir de informações autodeclaráveis da Plataforma Lattes, analisou-se a relação 20/80 dos pais científicos, com base $n=163$ pais científicos, sendo que os $n=25$ pais científicos mais relevantes perfazem 15,33% do universo analisado. Esse universo de pais científicos foi responsável por $n=89$ orientações, correspondendo a 39,20% do total de orientações recuperadas nos currículos analisados.

Vale destacar que 8 dos pais científicos são atualmente bolsistas de produtividade do CNPq (sendo, hierarquicamente e em escalara decrescente, 2 bolsistas nível 1A, 2 bolsistas nível 1B, 1 bolsista nível 1D e três bolsistas nível 2) e 7 são ex-bolsistas. Todo esse universo de bolsas e de ex-bolsistas é oriundo da área de Ciência da Informação. Nesse sentido, vale destacar que a classificação dos bolsistas de Produtividade Científica do CNPq tem por base os pesquisadores nível 2, ascendendo gradualmente aos níveis 1D, 1C, 1B e, finalmente, 1A.

Figura 1: pais científicos



Dados da Pesquisa, 2021

Outro detalhe de análise foi identificar as ações diretas para o tema Organização e Representação do Conhecimento, tema central desta proposta, no qual fica evidente algumas centralidades, como a de Unesp_2, IBICT_4, Unesp_1, Unesp_3, UFF_2 e USP_9.

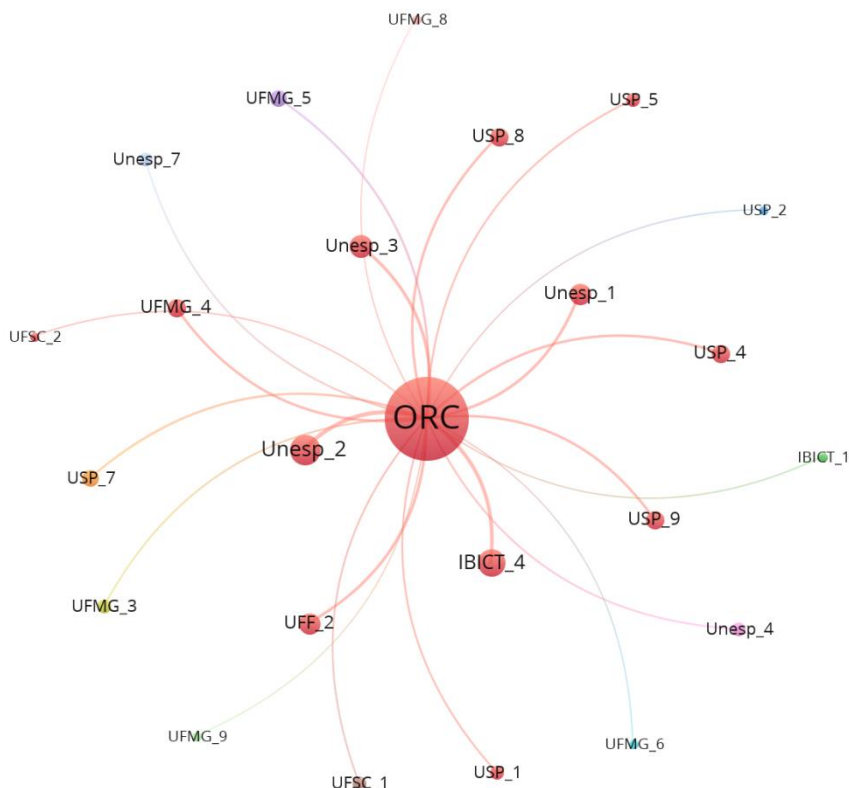
Os dados representados na figura 2 são vitais para analisar a coerência das duas primeiras figuras, com a validação dos pais científicos envolvidos com a Organização e Representação do Conhecimento.

Outro detalhe é que, como os dados da Plataforma Lattes são autodeclaráveis, uma simples indexação com parte do termo da pesquisa faz com que sejam destacados, como no caso de pesquisadores (pais científicos) que têm maior aderência de suas orientações e pesquisas em segmentações da Ciência da Informação como Gestão do Conhecimento, Sistemas de Conhecimento, Organizações empresariais, Política de informação e Estudos métricos, entre outros.

Relevante é ainda mencionar que esses temas podem ter relação com Organização e Representação do Conhecimento mas isso, entretanto, pouco foi identificado na depuração dos dados.

Número de orientações dos filhos científicos: Em relação às orientações de doutorado autodeclaradas na Plataforma Lattes sobre o tema “Organização do Conhecimento” tem-se um universo de n=227 pesquisadores; entretanto, realizou-se um refinamento para os orientadores dentro dos programas em Ciência da Informação com pelo menos uma orientação no período, chegando-se ao total de 84 orientações, com a somatória de 692 orientações.

Figura 2: Pais científicos com enfoque exclusivo na Organização e Representação do conhecimento



Fonte: Dados da Pesquisa, 2021.

A elite de autoridade ficou, assim, em 18 orientadores mais representativos, todos professores de cursos da Ciência da Informação, porém com formações diversas, como: Linguística, História, Educação, Comunicação e Cultura, Comunicação & Semiótica, Comunicação, Ciências da Comunicação, Ciência da Informação e Administração de Empresas.

A relação desse universo com os “pais científicos” revelou um número significativo de nomes coincidentes, tais como: IBICT_1, Unesp_2, UFMG_6, UFF_2, Unesp_3, USP_9, Unesp_4, UFMG_9 e IBICT_4. A elite foi responsável por 46,67% das orientações. Outro dado importante é que nesse universo registram-se quatro bolsistas de produtividade do CNPq 1A, um bolsista 1B, um bolsista 1C, 3 bolsistas nível 2 e 4 ex-bolsistas de produtividade, todos em Ciência da Informação, representando 72% do total.

A especificação de orientações exclusivas em Organização e Representação do Conhecimento também foi evidenciada em uma minuciosa investigação no teor das orientações, ilustrada na figura 4.

Figura 3: Os pesquisadores que mais orientaram

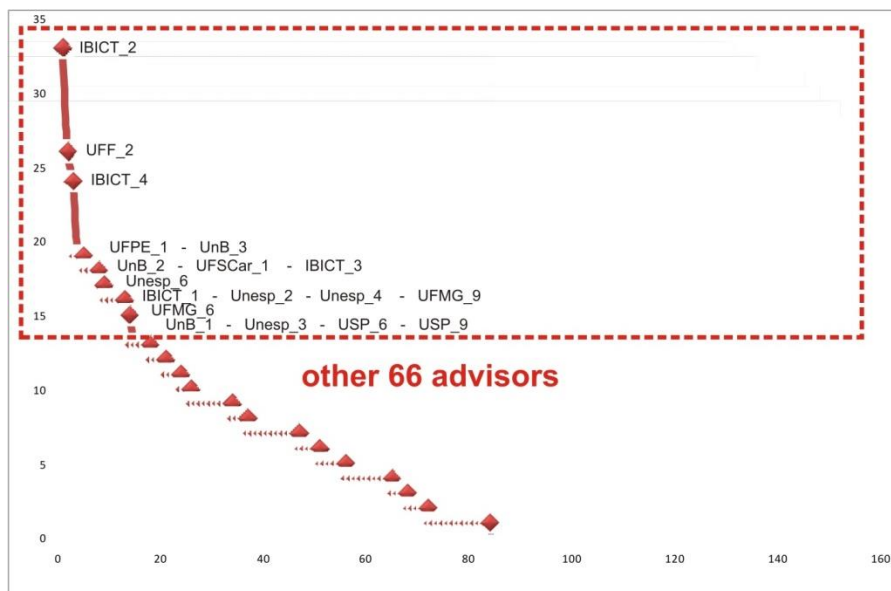
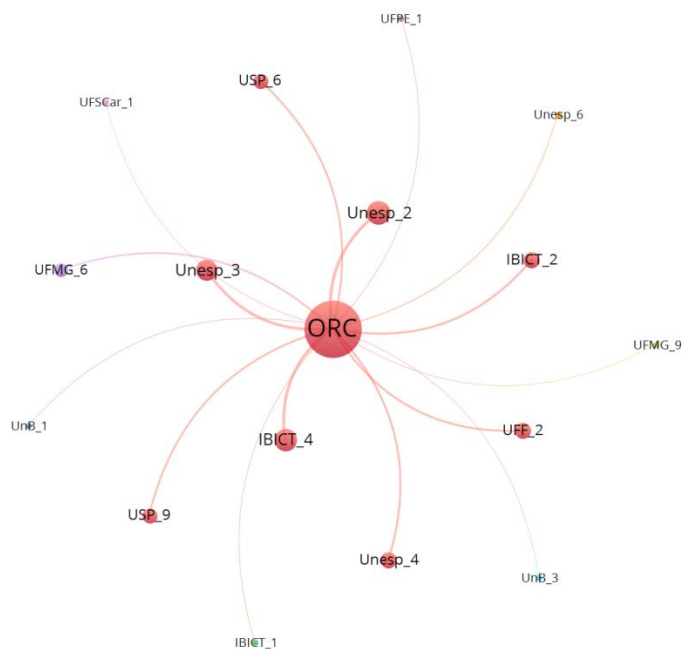


Figura 4: Orientadores com enfoque exclusivo na Organização e Representação do conhecimento



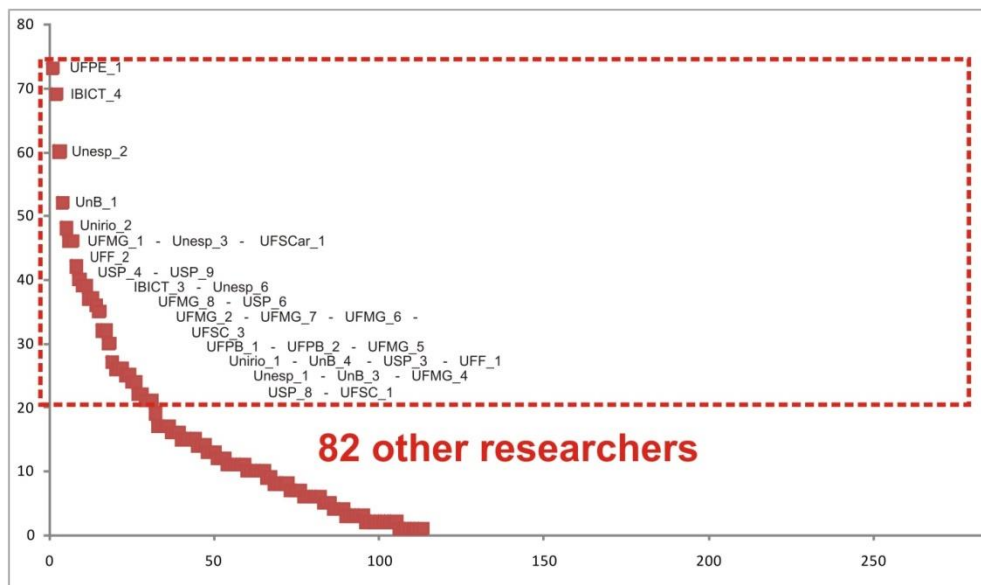
Desse universo de análise temos que ressaltar que dois destaques na figura 3 não aparecem na rede (UnB_2 e IBICT_3), em especial por terem algum elemento de publicação e de orientação que interaja com a Organização e Representação do Conhecimento, mas que de fato não tratam especificamente do tema estudado. Algo semelhante acontece com outros orientadores que se destacam com uma quantidade alta de orientações, porém mais articulados a temas outros, como Gestão do Conhecimento, Conhecimento empresarial e Estudos Métricos da Informação. Assim, os destaques da rede foram: Unesp_2, seguida por IBICT_4, Unesp_3, IBICT_2, UFF_2, Unesp_4, UFMG_7, USP_9 e UFMG_6. Vale lembrar que esses orientadores fazem parte de programas que possuem linha ou grupo de pesquisa destinado a trabalhar com o tema da Organização e Representação do Conhecimento.

Número de banca dos “filhos científicos”: Mais uma vez utilizou-se a mesma base de análise, os n=227 pesquisadores autodeclaráveis em trabalhar com o tema Organização do Conhecimento tendo-se, também, realizado um refinamento que, por sua vez, conduziu a 113 pesquisadores que participaram de pelo menos uma banca de doutorado.

Como já ressaltado na Metodologia, utilizaram-se os 31 primeiros da listagem com mais participações em bancas representando 63,41% de todas as bancas (base de cálculo 1113 bancas da elite ÷ 1755 que é o total de todas as bancas).

Em termos de formação acadêmica, esse grupo tem, em sua maioria, origem na Ciência da Informação, mas se verificam formações outras tais como: Análise do discurso, Ciências da Comunicação, Ciências Sociais, Comunicação, Comunicação & Semiótica, Comunicação e Cultura, Educação, Engenharia de Produção, Estudos Literários, História Econômica e História Social.

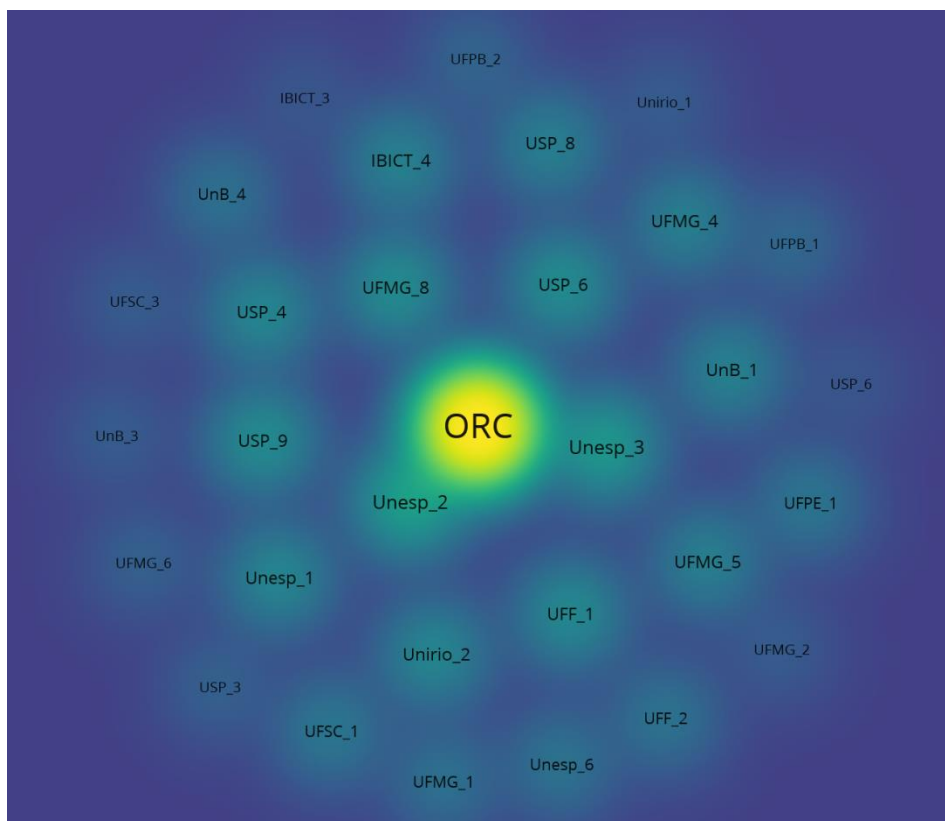
Figura 5: Os pesquisadores que mais participaram de banca de tese



Vale ainda mencionar, nesse grupo, a presença de três bolsistas 1A, um bolsista 1B, um bolsista 1C, três bolsistas 1D e dez bolsistas de nível 2 do CNPq, além de oito ex-bolsistas, o que representa 90% da elite identificada.

Para a especificidade de bancas de teses voltadas a Organização e Representação do Conhecimento observaram-se algumas particularidades, em que mais uma vez alguns pesquisadores se destacaram por serem autoridades da temática em todos os sentidos (produção e orientação) e acabam ilustrados na rede.

Figura 6: Bancas com enfoque exclusivo na Organização e Representação do conhecimento



Dados da Pesquisa, 2021

Os destaques dentro do tema Organização e Representação do Conhecimento são: Unesp_2, seguido por Unesp_3, USP_3, UFMG_8, UFMG_7, UFF_1, Unirio_2, USP_4, Unesp_1, IBICT_4 e UnB_1, todos participantes ativos de grupos de pesquisa e linhas que tratam o tema a fundo no Brasil.

Uma análise da vinculação institucional de todos os pesquisadores que se situam nas cinco primeiras posições em alguma das categorias de análise – pais científicos, orientações acadêmicas, participação em bancas e/ou enfoque exclusivo em ORC – evidenciou quatro programas de pós-graduação, de instituições distintas, todas

pertencentes à região sudeste do Brasil, nomeadamente nos estados de São Paulo, Rio de Janeiro e Minas Gerais, a saber: UNESP – Programa de Pós-Graduação em Ciência da Informação (Linha de pesquisa Produção e Organização da Informação), IBICT/UFRRJ – Programa de Pós-Graduação em Ciência da Informação (Linha de pesquisa: Comunicação, Organização e Gestão da Informação e do Conhecimento), USP – Programa de Pós-Graduação em Ciência da Informação (Linha de pesquisa: Organização da Informação e do Conhecimento) e UFMG – Programa de Pós-Graduação em Gestão e Organização do Conhecimento (Linha de pesquisa: Organização da Informação e do Conhecimento). A vista disso, pode-se dizer que nessas quatro instituições reside, primordialmente, o núcleo epistêmico da pesquisa nessa temática, no Brasil, visto aí se encontrarem seus mais proeminentes atores.

Por sua vez, uma comparação geral dos pesquisadores classificados nas cinco primeiras posições em todas as quatro categorias de análise – pais científicos, orientações acadêmicas, participação em bancas e enforque exclusivo em ORC – levou a dois pesquisadores, respectivamente Unesp_2 e IBICT_4.

Conclusões

A pesquisa revelou que o campo da Organização do Conhecimento ocupa uma posição altamente relevante no contexto da Ciência da Informação brasileira visto que se tem um conjunto de 227 teses de doutorado cujos autores autodeclararam tratar dessa temática. Essa relevância pode ser explicada por alguns fatos, tais como: a existência do Grupo Temático Organização e Representação do Conhecimento no âmbito da Associação Nacional de Pesquisa em Ciência da Informação – ANCIB, com intensa atuação, o capítulo brasileiro da ISKO, existente desde 2007 e promotor de eventos bienais e com forte articulação internacional, e a forte presença de pesquisadores da área de ORC no grupo de bolsistas de Produtividade em Pesquisa do CNPq

Essa relevância se reflete no fato de as três elites estudadas (“pais científicos”, orientadores e participantes de bancas) contarem com a presença de bolsistas de Produtividade em Pesquisa do CNPq, os quais podem ser institucionalmente considerados como pertencentes a uma elite científica nesse campo.

Se, por um lado, identificou-se, na elite dos “pais científicos”, uma relação umbilical com a Ciência da Informação e com o campo da Organização do Conhecimento, tal aspecto não se repete nas elites dos orientadores e integrantes de bancas, aspecto que pode se explicar pelo fato de muitos doutores haverem abordado temas de organização do conhecimento em programas de doutorado de outras áreas (muitas vezes por dificuldade de deslocamento para programas com foco mais específico nesse tema) e tal aspecto, por sua vez, se reflete na composição de bancas de defesa. Outra razão para tal decorre ainda pelo fato de, em dissertações e teses com temáticas de interface com outras áreas, a composição das bancas pressupõe, via de regra, contemplar tal interdisciplinaridade.

Infere-se, ainda, que o termo “organização do conhecimento”, como foi objeto de palavras-chave, nem sempre é utilizado com o devido rigor conceitual e terminológico inerente a esse campo do saber e muitas vezes acaba se confundindo com estudos de gestão do conhecimento, por exemplo.

Referências

- Bliss, H. E. (1929). *The Organization of Knowledge and the System of the Sciences*. New York: Henry Holt.
- Bliss, H.E. (1933). *The organization of knowledge in libraries and the subject-approach to books*. New York: The H. W. Wilson.
- Bourdieu, P. (2004). *Os usos sociais da ciência: por uma sociologia crítica do campo científico*. São Paulo: Unesp, 2004.
- Buckland, M. (1991). Information as thing. *Journal of the American Society of Information Science*, 42(5):351-360, June 1991.
- Dahlberg, I. (1995). Current trends in knowledge organization. In: Garcia Marco, F. J. (Ed.). *Organización del conocimiento en sistemas de información y documentación*. Zaragoza: Librería General, p. 7-26.
- Dahlberg, I. (2006). Knowledge organization: a new science? *Knowledge Organization*, 33(1):11-19.
- Esteban Navarro, M. A.; García Marco, F. G. (1995). Las “primeras jornadas sobre organización del conocimiento: organización del conocimiento e información científica”. *Scire*, Zaragoza, 1(1): 149-157, Ene./Jun.
- Guimarães, J.A.C. (2008a). *Ciência da Informação, Arquivologia e Biblioteconomia: em busca do necessário diálogo entre o universo teórico e os fazeres profissionais*. In: Fujita, M. S. L.; Guimaraes, J. A. C. (Org.). *Ensino e Pesquisa em Biblioteconomia no Brasil: a emergência de um novo olhar*. São Paulo: Cultura Acadêmica, p. 33-44.
- Guimarães, J.A.C. (2008b). A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). *Revista Ibero-americana de Ciência da Informação (RICI)*, 1(1):77-99, jan./jun.
- Guimarães, J.A.C.; Dodebei, V. (Orgs.). (2015). *Organização do conhecimento e diversidade cultural*. Marília: ISKO-Brasil; FUNDEPE.
- Haas, P. M. (1992). Epistemic communities and international policy coordination. *International Organization* 46(1): 1-35.
- Hansson, (2013)
- Hjørland, B. (2002) Domain analysis in information science: eleven approaches – traditional as well as innovative. *Journal of Documentation*, 58(4): 422-462.
- Lauruhn; Groth (2016)
- Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12); 317–323, 1926.
- Olson, H. A. (2002). *The power to name: locating the limits of subject representation in libraries*. Dordrecht: Kluwer Academic.
- Pombo, O. (1998). Da classificação dos seres à classificação dos saberes. *Leituras: Revista da Biblioteca Nacional de Lisboa*, v. 2, p. 19-33, 1998.
- Price, J. D. de S. (1963). *Little science, big science*. New York: Columbia University Press.
- San Segundo Manuel, R. (1996). *Sistemas de organización del conocimiento: la organización del conocimiento en las bibliotecas españolas*. Madrid: Universidad Carlos III de Madrid; Boletín Oficial de Estado.
- Vickery (2008) On ‘knowledge organisation’. Available at: <http://www.lucis.me.uk/knowlorg.htm#start>
- Witty, F.J. (1973). *The Beginnings of Indexing and Abstracting: Some Notes towards a History of Indexing and Abstracting in Antiquity and the Middle Ages*. *The indexer*, 8(4):193-198.
- Zeng (2008). Knowledge Organization Systems (KOS). *Knowledge Organization*, 35(2/3): 160-182.

O *SOFTWARE* LIVRE E A DESCRIÇÃO ARQUIVÍSTICA NO MEIO DIGITAL: O USO DO ATO M EM PORTUGAL E NO BRASIL

Susana Sofia Cunha¹, Maria Cristina Vieira de Freitas²

¹Doutoranda em Ciência da Informação, Faculdade de Letras, Universidade de Coimbra, susana.cunha@student.fl.uc.pt, ORCID iD 0000-0003-4444-4319

²Centro de Estudos Interdisciplinares do Século XX, Faculdade de Letras, Universidade de Coimbra, cristina.freitas@fl.uc.pt, ORCID iD 0000-0002-8849-8792

Resumo

A organização e o acesso à informação em arquivos dependem de processos e programas digitais. A opção por sistemas *open source* ganha terreno e assume-se como alternativa a *softwares* proprietários. O objetivo geral deste estudo é analisar aspetos da utilização do AtoM® em Portugal e no Brasil, percecionando o seu grau de aplicação e aceitação. Para tanto, delinearam-se objetivos específicos: recolha de trabalhos científicos em bases de dados e repositórios *online* nacionais e internacionais, comparação de instituições abordadas nos trabalhos e argumentos para a utilização do programa. A metodologia baseia-se em pesquisa bibliográfica de caráter seletivo e na análise de conteúdo de textos recolhidos. Os resultados obtidos nos 42 textos identificados e analisados revelam que existe um grande interesse por parte de investigadores/as e técnicos/as de instituições arquivísticas na utilização do programa em Portugal e no Brasil. A sua utilização pode ser vantajosa por trazer soluções interessantes e por incorporar mais-valias significativas a instituições com orçamentos e meios técnicos reduzidos e, principalmente, por permitir que os diferentes arquivos alcancem uma rede de colaboração e partilha de experiências técnicas e de uso da aplicação, o que, segundo se conclui, será um fator que poderá potenciar a sua difusão e o seu uso generalizado, no futuro.

Palavras-chave: AtoM, programas de código aberto, descrição arquivística, arquivos digitais, acesso à informação.

Introdução

A representação arquivística no meio digital é, hoje, incontornável. O uso de *softwares* pelas instituições culturais é uma realidade desde pelo menos há quatro décadas e, ao longo deste período, estas ferramentas têm dado provas suficientes da sua importância para a comunicação e, conseqüentemente, para o incremento do uso da informação. As ferramentas de código aberto ampliam as possibilidades de inovação, por meio do livre desenvolvimento e customização e, com isso, potenciam

não apenas a competitividade, mas também, e sobretudo, a cooperação interinstitucional, e abrem um leque de possibilidades para que possam atingir, com poucos recursos e algum esforço conjunto, a tão expetável quanto desejável transição tecnológica. Organizar, representar, dar acesso à informação e, conseqüentemente, ao conhecimento, são metas a atingir e a sustentar pelos serviços de arquivo que, até há algumas décadas atrás, pelo serviço que prestavam, julgavam-se imunes a estas inovações (Freitas, 2009).

Nesta era digital, e, sobretudo, nos tempos mais recentes, nos quais assistimos a um confinamento massivo da população, em face das ameaças provocadas pela pandemia COVID-19, os arquivos enfrentam desafios relacionados com o incremento das exigências dos/as utilizadores/as, no que tange às necessidades de um acesso, cada vez mais generalizado, à informação que mantêm sob a sua custódia. Num tal contexto, a apresentação dos documentos em suportes convencionais é substituída pelos meios e digitais e as instituições recorrem a plataformas desenvolvidas para o efeito. Nesta linha, encontram-se os *softwares* livres destinados à comunicação dos instrumentos de descrição arquivística. Assim, um programa de descrição arquivística normalizado, com navegação multinível e com a possibilidade de disponibilização *online* de conteúdos é indispensável para cumprir a legislação e, ao mesmo tempo, romper alguns paradigmas (Flores, 2016). No entanto, como referem António e Silva (2011), o recurso a ferramentas informáticas nem sempre é possível às instituições, seja por falta de soluções normalizadas ou por limitações orçamentais. Neste sentido, recorrem-se aos *softwares* livres, adaptando-os às necessidades e objetivos institucionais.

Dada a sua forte representação em bases de dados referenciais, acreditamos que um dos *softwares* possivelmente mais utilizados seja o AtoM®. O projeto ICA-Atom® surgiu em 2003, quando o Comité de Informação Tecnológica da *International Council on Archives* (ICA) publicou os requisitos funcionais para o *Open Source Archival Resource Information System* (OSARIS). A partir de 2005, e com o apoio da UNESCO, lançaram-se as bases para a construção de uma aplicação livre, *online*, de descrição arquivística para utilização geral (Bushey, 2012). Inicialmente denominada ICA-AtoM®¹, a partir de 2014 a manutenção e gestão do *software* ficam sob a alçada da empresa *Artefactual Systems* e este passa a denominar-se apenas AtoM (Moraes et al., 2019).

Partindo do pressuposto de que o AtoM®, sendo um *software* livre de código aberto, auxilia as instituições arquivísticas a desenvolver os seus projetos de descrição arquivística, potenciando o seu raio de ação, face aos interesses e necessidades informacionais dos/as utilizadores/as, o principal objetivo deste estudo é analisar aspetos da utilização do AtoM® em duas comunidades arquivísticas de uso distintas, mas muito próximas – Portugal e Brasil – e perceber o seu grau de aplicação e, conseqüentemente, de aceitação, em ambos os contextos. Para tal e conforme se verá, nos próximos pontos, foram realizadas recolhas de dados em publicações científicas e efetuadas análises comparadas, para responder à seguinte pergunta de

¹ AtoM como acrónimo de “*Access to Memory*” e o prefixo usado para assinalar que o produto era desenvolvido sob orientação ou parceria do ICA.

partida: quais são os argumentos invocados para a utilização do AtoM®, bem como as opções adotadas e os constrangimentos encontrados?

Metodologia

1) Objetivos

Este estudo pretende analisar aspetos da utilização do AtoM® como plataforma de descrição arquivística em Portugal e no Brasil. Para a sua prossecução será necessário identificar a produção científica sobre o tema, mapear e comparar dados de utilização do AtoM® em Portugal e no Brasil, para compreender as opções por trás da adoção desta plataforma que é usada para a descrição arquivística.

2) Abordagem metodológica

A abordagem privilegiou aspetos empíricos do problema. Numa primeira fase, procedeu-se a uma consulta da literatura científica nacional e internacional, visível em repositórios e bases de dados *online* como a *Web of Science (Information Science & Library Science)*, a *B-on (Biblioteca do Conhecimento Online)*, agregadora de repositórios como a *EBSCO (Library & Information Science Source)*, *RCAAP (Repositório Científico de Acesso Aberto em Portugal)* e *OASIS (Portal brasileiro de publicações científicas em acesso aberto)*, e a *BRAPCI (Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação)*. A pesquisa foi realizada nos meses de novembro e dezembro de 2020, por títulos e assuntos, artigos em acesso aberto e texto integral. Os termos utilizados para delimitar a pesquisa, em inglês, foram “ICA-AtoM” e “AtoM” conjugado com “*archiv**”. Aplicou-se como restrição temporal os anos de 2006 (data de lançamento da primeira versão do ICA-AtoM) e 2020. Selecionaram-se os resultados que se enquadravam no domínio da Biblioteconomia e Ciência da Informação (cf. Tabela 1). Descontando-se as redundâncias, restaram 42 trabalhos que constituem o *corpus* deste estudo, isto é, o “conjunto dos documentos tidos em conta para serem submetidos aos procedimentos analíticos” (Bardin, 2008, p. 122).

Tabela 1. Artigos recuperados nos diferentes repositórios/ bases de dados

BASE DE DADOS/ REPOSITÓRIO	ATOM		ATOM + ARCHIV*		ICA-ATOM	
	RECUPERADOS	RELEVANTES	RECUPERADOS	RELEVANTES	RECUPERADOS	RELEVANTES
WEB OF SCIENCE	14233	5	38	4	6	4
B-ON	484	0	2	1	2	1
EBSCO/LIIS	111	1	2	1	0	0
RCAAP	84	10	10	10	6	6
OASIS	1030	6	10	10	7	7
BRAPCI	23	14	8	3	32	18
TOTAIS	15965	36	70	29	53	36

Fonte: elaboração nossa, com base na consulta das fontes mencionadas.

Assim, após a identificação e a seleção dos trabalhos científicos a analisar, foi elaborada uma primeira tabela com as características bibliométricas a considerar para cada trabalho, nomeadamente, autoria, tipologia, instituição abordada, tipologia da instituição e país. Posteriormente, foi elaborada uma segunda tabela para registar as características destacadas em cada texto, nomeadamente, vantagens e desvantagens ou dificuldades da utilização do AtoM®. Estas categorias² respeitam as regras propostas por Bardin (2008) para que uma análise seja válida: são homogêneas (no sentido em que não “misturam” assuntos), exclusivas (já que um mesmo elemento do conteúdo não pode ser classificado em duas categorias diferentes) e adequadas ou pertinentes (estão adaptadas ao material de análise e refletem as intenções da investigação). A categorização das vantagens e desvantagens do AtoM® fez-se segundo o processo “por caixa”, pelo qual define-se um sistema de categorias e repartem-se da melhor maneira possível os elementos, consoante vão sendo encontrados.

Em suma, os dados foram tratados quantitativamente através de análises estatísticas simples (ocorrências e percentagens) e qualitativamente através da sua interpretação à luz dos objetivos propostos. Estes recursos analíticos, consistentes com o método de análise de conteúdo, levaram à categorização dos textos e à realização de inferências, com base na interpretação dos conteúdos, em cada uma das categorias fixadas.

É de salientar que não se pretende uma mera recolha de dados estatísticos, mas sim uma análise interpretativa que foi feita *a posteriori*, já que no início do trabalho era difícil prever que características os autores iriam destacar.

Resultados e discussão

Partindo dos dados recolhidos nas bases de dados e nos repositórios, foi possível identificar 42 trabalhos relacionados com a utilização do AtoM® em Portugal e no Brasil. Para esta primeira análise foram considerados todos os trabalhos recuperados, mesmo aqueles cujo autor e/ou tema se repete (desde que publicados em data e revista diferente e o texto não seja o mesmo). Não foram considerados trabalhos recuperados, simultaneamente, no mesmo idioma e publicados na mesma revista (caso de revistas bilingues). Pese embora a pesquisa tenha tido como referência o inicial o ano de 2006, o primeiro trabalho recuperado data de 2011, coincidente com o lançamento da versão 1.2 do ICA-AtoM®, melhorada graças ao contributo a nível financeiro e de codificação proporcionado pelos investidores, e data a partir da qual a *Artefactual Systems* dá continuidade ao projeto (Anjos et al., 2015). Os trabalhos mais recentes datam de 2020: um versa sobre a utilização do AtoM® para a descrição arquivística na Universidade Federal da Paraíba enquanto que o outro trata da promoção do acesso ao património documental postal brasileiro (cf. Apêndice I).

Dos trabalhos recuperados, 64% (27 trabalhos) são de autores/as brasileiros/as e 36% (15 trabalhos) de autores/as portugueses/as, o que reflete a maior incidência de

² De acordo com a definição de Bardin (2008, p. 145) “categorias são rubricas ou classes, as quais reúnem um grupo de elementos (...) sob um título genérico, agrupamento esse efectuado em razão das características comuns destes elementos”.

utilização e problematização do *software* no Brasil, sendo este resultado exepetável e parcialmente explicável pela dimensão continental do país e, naturalmente, pela existência de uma comunidade de uso/discussão maior do *software* em causa, o que, inclusive, facilita a sua difusão. Quanto à tipologia dos trabalhos, 43% (18) são artigos publicados em revistas da área da Ciência da Informação, Arquivística e Biblioteconomia, 24% (10) são publicações em atas e 5% (2) são comunicações apresentadas em encontros e seminários da área da Ciência da Informação. Ainda, 31% (13) desses trabalhos recuperados são dissertações de mestrado, o que parece indicar uma preocupação de pendor acadêmico em estudar e aprofundar a utilização de *softwares open source* na descrição e disponibilização de acervos documentais. É igualmente importante salientar que 64% (27) dos trabalhos recuperados consistem em apresentações de estudos de caso e aplicação efetiva do AtoM® em instituições públicas ou privadas, enquanto 36% (15) destes trabalhos são reflexões sobre o programa ou orientações teóricas e técnicas para a sua aplicação.

Passemos, então, à análise das instituições visadas nos trabalhos recuperados. O número mais significativo, com 14 ocorrências (33%), corresponde a projetos ligados à descrição de arquivos de instituições de ensino superior como institutos, universidades ou centros de investigação universitários. Mais uma vez, é no Brasil que encontramos um maior número de projetos (10; 24%), sendo que a maioria (6; 14%) aborda a utilização do AtoM® para a organização e descrição de arquivos da própria universidade ou de faculdades³. Este resultado é consistente com o de Saraiva e Nogueira (2014), que afirmam que este *software* é utilizado pela maioria das Instituições Federais de Ensino Superior brasileiras para descrever eletronicamente os seus acervos arquivísticos. Ainda, dentro do âmbito universitário brasileiro, destacam-se também projetos de descrição de acervos museológicos com o AtoM®⁴. De salientar que, por vezes, e sobretudo em dissertações de mestrado, a instituição objeto de estudo não é a instituição que concede o grau acadêmico⁵. São, ainda, cinco (12%) os trabalhos portugueses que tratam de aplicação do AtoM® a projetos ligados a universidades⁶. A análise dos trabalhos permite verificar, também, que o AtoM® é escolhido por instituições públicas municipais para descrição e divulgação dos seus arquivos (5; 12%)⁷. Os restantes trabalhos (4; 10%) focam a

³ Sendo estes: Instituto Federal de Rio Grande do Sul (arquivo da Reitoria), Universidade Federal de Ciências da Saúde de Porto Alegre, Universidade Federal de Paraíba (Arquivo Central), Universidade Federal de Santa Maria (Departamento de Arquivo Geral), Universidade Federal do Rio Grande – FURG (Conselho Universitário e Faculdade de Direito Clóvis Beviláqua) e Universidade Federal de Rio Grande do Sul – UFRGS.

⁴ Desenvolvidos pela Universidade Federal de Minas Gerais e pela Universidade Estadual Paulista.

⁵ Como exemplo, citam-se os trabalhos desenvolvidos na Universidade Federal de Santa Maria relacionados com arquivos de outras universidades, como a FURG, UFRGS ou o IFRGS.

⁶ Sendo estas: Universidade do Porto (um estudo para implementação no arquivo da Reitoria e duas dissertações de mestrado que abordam arquivos de arquitetura – 3; 7%); Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa (base de dados sobre arquivos de família – 1; 2%); Universidade Católica do Porto (plataforma de arquivos pessoais e de instituições religiosas – 1; 2%).

⁷ Sendo estes trabalhos acerca: do Arquivo Histórico Municipal de Santa Maria, do Arquivo

utilização do AtoM® em instituições de cariz diverso (políticas, militares ou fundações)⁸.

Quanto às características da aplicação mais destacadas e valorizadas pelos/as autores/as, bem como desvantagens ou dificuldades sentidas durante a utilização, incidimos a análise sobre 37 dos 42 trabalhos recolhidos; a diferença prende-se com o facto de cinco dos trabalhos não abordarem qualquer vantagem ou desvantagem na utilização do *software* e por isso não puderam ser contemplados nesta análise (cf. Apêndice II).

O facto de ser uma aplicação gratuita e de código aberto é desde logo apontado por 78% (29 textos recuperados) dos textos como vantagem em relação a *softwares* proprietários. A opção por esta aplicação pode significar poupanças consideráveis no orçamento de uma instituição, que se liberta assim da compra e do pagamento de taxas de licenciamento, atualização e apoio técnico. É de salientar que, sendo uma aplicação de código aberto, a instalação, alterações, resolução de problemas, etc., são por conta da entidade que o instala (ao contrário de um *software* proprietário), exigindo conhecimentos de informática para a instalação e parametrização, migração de dados e gestão do programa (C. G. da Silva, 2016). O *software* foi concebido para funcionar em todos os sistemas operativos, vantagem mencionada por 5% (2) dos textos. Apesar da simplicidade da instalação, 19% (7) dos textos referem alguma dificuldade nesta tarefa e necessidade de recurso ao exterior para acompanhamento dos serviços de informática da instituição (Caldeira, 2016). Como refere Vasques (2016), o apoio técnico é fundamental para garantir uma implementação e manutenção bem-sucedida de qualquer sistema de informação. Isto remete-nos para outra das vantagens apontadas nos textos analisados: a existência de uma comunidade ativa de utilizadores/as e técnicos/as, pertencentes à comunidade e à empresa *Artefactual Systems*, que através da partilha de experiências contribuem para resolver os problemas e encontrar soluções, sobretudo para instituições com orçamentos e meios técnicos reduzidos (H. Silva, 2016) (cf. Apêndice II).

Outra das características apresentadas como vantagem na adoção deste programa é ser executado em ambiente *Web*, não necessitando, por isso, de instalação local. Observa-se que 70% (26) dos textos destacam esta funcionalidade: Anjos et al. (2015) e Caldeira (2016) referem que a possibilidade de utilização do *software*, a partir de qualquer equipamento com ligação à Internet, facilita a autonomia do trabalho e da equipa, aspeto fundamental quando se trata de equipas grandes e/ou projetos com várias entidades, como, por exemplo, o da CGTP-IN, em Portugal, que possui estruturas espalhadas pelo país, ou do projeto PAPIR, que engloba descrições de oito entidades diferentes (Rocha, 2019). Pacheco (2013), inclusive, acentua que a utilização de um *software* baseado na *Web* pode ser mais fácil e permite que as informações sejam publicadas assim que são inseridas. Efetivamente, esta possibili-

Público de Pernambuco e da Secretaria de Cultura do Município de Londrina (neste caso, o AtoM é utilizado para efetuar o registo do Património Arquitetónico e Histórico da cidade de Londrina), no Brasil; e dos arquivos municipais de Albergaria-a-Velha e Torres Vedras, em Portugal.

⁸ Sendo estes: Arquivo da Social Democracia – PSD e Central Sindical CGTP-IN, Arquivo fotográfico Teófilo Braga, todos em Portugal, e no Arquivo da Marinha do Brasil.

dade é referida por 51% (19) dos textos já que, embora possa ser utilizado apenas para descrição, a divulgação *online* providencia o acesso rápido e fácil à documentação, além de minimizar a utilização do suporte analógico (Lima & Flores, 2016). C. G. da Silva (2016) relembra que a maior parte dos/as investigadores/as faz pesquisa na Internet e que a adoção de um *software* com essa funcionalidade é uma mais-valia para a instituição, na divulgação de conteúdos (cf. Apêndice II).

O AtoM® é multilíngue, quer em termos de produção de dados quer na possibilidade de navegação, vantagem apontada por 60% (22) dos textos analisados. Esta característica permite a criação de um sistema de informação arquivística internacional e acessível em qualquer lado (António & Silva, 2011). Outro fator relacionado com a possibilidade de difusão e divulgação dos conteúdos é a capacidade de o AtoM® permitir a interoperabilidade entre sistemas, habilitando a instituição a aderir a agregadores como o Portal Português de Arquivos ou outros de âmbito internacional (Pereira et al., 2016; H. Silva, 2016). Esta funcionalidade é apontada por 24% (9) dos textos, embora tenham sido relatadas dificuldades sentidas na sua implementação em 5% (2) dos trabalhos analisados (cf. Apêndice II).

Ainda, no que respeita a aspetos de instalação e configuração do *software*, é notória a valorização da possibilidade de parametrização e personalização mediante as necessidades da instituição (apontada em 43% (16) dos textos), mais concretamente a possibilidade de instalar *plug ins* para diferentes funcionalidades (referidas por 24% – 9), a possibilidade de alterar ou adicionar níveis de descrição (referidas por 11% – 4) e a criação de páginas estáticas (referidas por 5% – 2), com texto, imagens ou ligações para fornecimento de informações ou outros conteúdos. Apenas um texto (3%) menciona a possibilidade de georreferenciação (Flores, 2016) (cf. Apêndice II).

Ainda, verifica-se que 57% dos trabalhos analisados (21) valorizam o facto de o AtoM® poder ser utilizado por uma única instituição ou servir de plataforma multi-repositório ou multi-institucional. Esta funcionalidade pode ser aproveitada para gerir diferentes serviços dentro de uma autarquia, como refere H. Silva (2016), ou para reunir numa só plataforma o contributo de entidades arquivísticas diferentes, rentabilizando custos e recursos (Rocha, 2019). Como referem H. Silva (2016) e Silveira (2019) parece não existir limite para o número de entrada de registos no programa (vantagem referida em 5% (2) dos textos), o que reforça a utilização como multi-repositório. A possibilidade de criar várias contas de utilizador, com diferentes tipos de permissões, mencionada em 41% dos textos (15), torna-se muito útil quando se trata de uma equipa grande e a definição de perfis com diferentes graus de acesso, permissões e responsabilidade é importante para uma correta gestão da base de dados (Barradas, 2016; Pereira et al., 2016). No que respeita à descrição arquivística propriamente dita, destaca-se, desde logo, o facto de ter sido concebido com o apoio do ICA, promover a descrição multinível e respeitar as normas de descrição arquivística usadas em Portugal e no Brasil (84% dos textos (31) destacam este aspeto). Como acrescento, 46% dos textos (17) apontam como vantagem a compatibilidade com outras normas de descrição como a *Canadian Rules for Archival Description (RAD)*, o *Dublin Core* ou a *Encoded Archival Description (EAD)* (Vasques, 2016), bem como a possibilidade de mudar de norma de descrição durante a utilização (referida por 5%; 2 textos). Em cada campo existem notas explicativas que remetem para a norma em causa, facilitando a descrição (mencionado em 5% (2) dos textos) e

14% dos textos (5) apontam como vantajoso o programa alertar para a falta de preenchimento de campos de descrição obrigatórios segundo a ISAD(G) (código de referência, título, suporte e dimensão e data), bem como para a discordância de datas entre níveis de descrição hierarquizados superior ou inferiormente (Batista et al., 2016). Ainda assim, em 3% (1) dos trabalhos apontou-se como desvantagem o facto de o AtoM® permitir a repetição de códigos de referência. No AtoM®, existe a possibilidade de normalização e controlo descritivo dos produtores, evitando duplicação de entradas, bem como a uniformização das descrições de instituições e o controlo da informação relativa a entidades que lhes possam estar associadas (António & Silva, 2011). Esta mais-valia é apontada por 16% (6) dos textos analisados. Ainda, em 8% (3) dos textos valoriza-se a possibilidade de associar uma família ou pessoa (singular ou coletiva) a uma unidade de descrição, através dos registos de autoridade e de produtor, permitindo um maior controlo descritivo. Como refere H. Silva (2016), o AtoM® faz herdar automaticamente ligações que existam num nível superior, seguindo o princípio de não repetição da informação e contribuindo para a melhoria do desempenho do sistema. Esta funcionalidade é bastante útil para o preenchimento automático do campo “Produtor” (Batista et al., 2016) e é mencionada por 8% (3) dos textos. Igualmente, é mencionada a possibilidade de copiar registos (por 8% – 3) e de mover registos entre os vários níveis de descrição (por 11% – 4). Uma das mais-valias da aplicação é a possibilidade de importação e exportação de registos em vários formatos (EAD, EAC-CPF, CSV e SKOS (Rocha, 2019). Esta característica, destacada em 54% (20) dos textos, revela-se de extrema importância sobretudo para instituições que tenham os dados registados em folhas de cálculo ou outras bases de dados e que desta forma não perdem o trabalho realizado nem necessitam de o refazer. Três dos exemplos encontrados são o Arquivo Municipal de Albergaria-a-Velha (H. Silva, 2016), o Arquivo de Torres Vedras, que migrou dados da aplicação ArqBase (C. G. da Silva, 2016) e a Universidade do Porto, que possuía dados no GISA e Dspace (Pereira et al., 2016). Apesar da vantagem que isto representa, em 8% dos textos (3) revelaram-se alguma dificuldade na importação e migração de dados de outras plataformas para o AtoM®, situações que, no entanto, acabaram por ser resolvidas graças à comunidade de ajuda do AtoM® ou do recurso a serviços técnicos exteriores à instituição (Anjos et al., 2015; Pereira et al., 2016). A possibilidade de agregar objetos digitais à descrição, viabilizando e facilitando a descrição de outros tipos de recursos ou suportes, como a fotografia ou o vídeo, é também uma das mais-valias desta plataforma, mencionada em 41% dos trabalhos (15). Uma das condicionantes acerca desta funcionalidade refere-se à limitação de um objeto digital para cada registo, o que dificulta a inserção de documentos com mais do que uma página (5% (2) mencionam esta desvantagem); esta limitação pode ser ultrapassada com a utilização de documentos em PDF, como refere H. Silva (2016). A necessidade de preservar a autoria e a propriedade dos objetos digitais, nomeadamente, textos ou fotografias, pode ser mantida através de um *plug in* para colocação automática de marca de água, quando se faz o *upload* de objetos digitais (5% (2) mencionam esta mais-valia). No entanto, Barradas (2016) e Pereira *et al.* (2016) referiram a inexistência de uma funcionalidade para colocação de marca de água personalizada em objetos digitais de instituições arquivísticas diferentes (quando multi-repositório, por exemplo) (cf. Apêndice II).

Com o AtoM® podem gerir-se incorporações, doações, compras e outras entradas, bem com as respetivas fontes de aquisição e direitos patrimoniais, vantagem apontada em 5% dos trabalhos. Desta forma, dados dispersos por várias bases de dados, folhas de cálculo e outras aplicações passam a ser centralizados numa única plataforma (H. Silva, 2016). É, também, possível associar os documentos ao depósito fixo, facilitando a organização interna da instituição (vantagem mencionada em 14% (5) dos trabalhos); no entanto, esta informação não é passível de ser ocultada, quando da divulgação *online* do registo, o que constitui uma desvantagem, mencionada em um dos textos (3%). A plataforma permite associar taxonomias (vocabulários controlados), traduzindo os pontos de acesso em categorias e facilitando a gestão dos descritores, funcionalidade que é considerada vantajosa para a recuperação da informação em 46% (17) dos textos, nomeadamente, através de categorias como “Assunto” e “Locais” (embora seja possível definir mais categorias, como por exemplo o campo “nomes”, que dá origem a um registo de autoridade segundo a norma ISAAR(CPF) (Barradas, 2016) (cf. Apêndice II).

Uma das funcionalidades mais interessantes e úteis, porque preserva a consistência das informações inseridas na base de dados, é que qualquer alteração ou atualização de um termo das taxonomias é automaticamente efetuada em todos os documentos associados ao descritor em causa (mencionado por 8% (3) dos textos). A possibilidade de escolher o estado de publicação nível a nível e descrição a descrição (estado preliminar ou publicado) também se assume como uma vantagem em 27% (10) dos textos (cf. Apêndice II).

Do ponto de vista do *front office*, isto é, do/a investigador/a, salienta-se a utilização amigável e boa visibilidade dos menus (apontado por 30% – 11), a qualidade na visualização dos objetos digitais (5% – 2) e possibilidade de criar ou imprimir relatórios e listagens de registos (14% – 5). O AtoM® possibilita várias hipóteses de pesquisa e todos os campos das descrições arquivísticas são pesquisáveis: “pesquisa simples”, visível em todas as páginas da aplicação e onde se pode aplicar um termo ou palavra-chave simples, e “pesquisa avançada”, onde se podem utilizar filtros, característica apontada por 24% (9) dos textos. As desvantagens apontadas prendem-se com a ocultação de alguma informação na página principal (mencionada por 8% – 3) e a linguagem e estrutura demasiado técnica para um/a utilizador/a que não seja da área, resultando em dificuldades na pesquisa (3% (1) mencionaram esta desvantagem). (cf. Apêndice II). Esta desvantagem não é, contudo, uma particularidade do AtoM®, sendo devida à rigidez formal das normas de descrição arquivística, consideradas pouco amigáveis e pouco focalizadas nos interesses dos/as utilizadores/as

Pese embora as vantagens elencadas, importa referir algumas desvantagens e dificuldades encontradas e mencionadas pelos/as autores/as dos textos selecionados e que não foram já mencionadas. É ponto assente que os problemas que envolvem a preservação digital constituem um fator de preocupação, sobretudo quando se pensa na ameaça da degradação do próprio suporte digital, como referem Hedlund e Flores (2014). O AtoM® não garante preservação digital e segundo Barradas (2016) deve ser a entidade detentora da documentação a prever um plano de preservação digital e *backups* do *software*. Esta é uma preocupação manifestada em 32% (12) dos trabalhos analisados. Da mesma forma, o risco de perda de dados por problemas no servidor é apontado por um dos textos (3%) (cf. Apêndice II). As desvantagens

apontadas, no entanto, podem colmatar-se através da integração do AtoM® com algum *software* vocacionado para a preservação digital, tomando-se como exemplo o Archivematica®.

Ainda, 8% (3) dos textos lamentam o facto de não ser possível aceder a estatísticas de visualizações (apenas é possível ver os registos mais consultados recentemente) e ressalvam a importância da quantificação destas informações para, por exemplo, se ter conhecimento dos documentos mais procurados e consultados. Batista et al. (2016) apontam, também, como fundamental saber quais os termos de pesquisa utilizados e que não obtiveram resultado, de forma a poder corrigir essas lacunas e a corresponder às expectativas dos/as utilizadores/as. Outra lacuna apontada por um dos textos (3%) prende-se com o facto de não ser possível gerar relatórios individuais de desempenho, embora seja possível aferir e quantificar os registos criados, editados ou eliminados ao longo do tempo⁹.

Conclusão

A análise efetuada sobre a utilização do AtoM® em Portugal e no Brasil vem demonstrar o interesse que investigadores/as e técnicos/as de instituições arquivísticas têm no programa. Igualmente, demonstra que as características mais destacadas da aplicação são, de facto, as mais conhecidas e difundidas entre as suas comunidades de uso (como, por exemplo, *open source*, parametrizável, multi-repositório, multilingue, com permissão de importação e exportação de dados, com base assente nas normas internacionais de arquivo e na *Web*, com autonomia e independência tecnológica na gestão). Alguma surpresa, no entanto, reserva-se ao facto de a interoperabilidade não ter sido largamente recordada como aspeto vantajoso.

A maior incidência de trabalhos de autores/as brasileiros/as vai ao encontro do que é indicado no site do AtoM®, que conta com 53 instituições brasileiras registadas, ao lado de apenas 11 em Portugal. Importa, no entanto, referir que existem instituições que utilizam o AtoM® para descrição arquivística, mas optam por não fazer divulgação dos seus acervos *online*, pelo que este número é apenas um indicador, por baixo, da verdadeira dimensão da sua utilização.

Nota-se, ademais, que os trabalhos que incidiram sobre o AtoM® e que foram recuperados para compor este estudo são, maioritariamente, estudos de caso publicados no Brasil em forma de artigos em revistas científicas, vinculados a projetos desenvolvidos por Instituições Superiores de Ensino ou por arquivos municipais. Destes trabalhos, a maioria aborda aplicações efetivas do programa, demonstrando a sua capacidade de utilização para a descrição de acervos tão variáveis como arquivos históricos, municipais e de instituições de ensino, arquivos fotográficos, arquivos de família, arquivos de instituições religiosas, arquivos de partidos políticos, arquivos de arquitetura, arquivos de património histórico edificado e até estudos

⁹ Estas desvantagens foram identificadas nos textos analisados e refletem as experiências com as versões do programa utilizadas, à data, em cada um dos casos de estudo e têm vindo a ser corrigidas em versões mais recentes do programa. Refira-se, como exemplo, a possibilidade recente de aceder às estatísticas de visualização através de uma ligação ao *Google Analytics*.

para descrição de acervos museológicos. Isto deve-se, em parte, à capacidade de parametrização e personalização do programa, sendo possível adaptá-lo às necessidades das diferentes instituições, mas também ao facto de ter sido desenvolvido com base nas normas de descrição emanadas do ICA e de ser compatível com outras normas de descrição como Dublin Core, EAD ou RAD.

É notório, embora não suficientemente generalizado, o esforço e investimento das universidades portuguesas e brasileiras, bem como dos seus centros de investigação, no apoio a projetos de descrição de arquivos das próprias universidades e projetos de outras instituições, como arquivos públicos. Neste estudo, verificamos que 13 dos trabalhos analisados são dissertações de Mestrado, o que parece indicar a preocupação da academia em estudar e investir na aplicação do programa.

Um dos grandes trunfos do AtoM® é a possibilidade de agregar objetos digitais às descrições (imagem e vídeo). Esta capacidade é especialmente importante em arquivos de imagem, como arquivos fotográficos ou de arquitetura, para referir apenas alguns dos abordados nos trabalhos, mas também em arquivos históricos, acrescentando a mais-valia da digitalização do documento original, contribuindo para a sua preservação e para o acesso *online*, de qualquer parte, aos seus conteúdos informacionais.

Os textos analisados salientam que os/as investigadores/as do século XXI fazem as suas pesquisas *online*, em qualquer lugar com acesso à Internet, e a possibilidade de utilizar um programa que lhes permita essa facilidade traduz-se em mais-valias para si e para a instituição, que ganha utilizadores/as.

Os principais fatores que concorrem para a escolha deste *software* podem converter-se em vantajosos para a sua difusão no futuro. No entanto, e para isto, em projetos de instalação da aplicação, será preciso equacionar o problema da preservação digital. Esta dificuldade pode ser ultrapassada articulando-o, por exemplo, com o Archivematica, também de código aberto e desenvolvido pela *Artefactual System* (Jorente et al., 2016), a mesma empresa que dá assistência ao AtoM®. Ressalve-se, ainda, que as desvantagens apontadas nos trabalhos analisados se referem às versões do programa utilizadas à data e que algumas dessas lacunas têm vindo a ser discutidas e corrigidas nas versões mais recentes.

Finalmente, o mérito deste estudo está no facto de colocar num mesmo cenário analítico o uso do AtoM® em duas realidades simultâneas e com resultados qualitativamente coincidentes: Portugal e Brasil, delineando um “estado da arte” em duas comunidades arquivísticas de língua portuguesa, tendo como objeto de estudo um sistema aplicacional que se encontra há mais de uma década e meia em uso e experiência em ambos os contextos.

Como limitações, o facto de a amostra seletiva de textos analisados – apenas 37 – revelar um “instantâneo” que representa uma realidade contextual e não generalizável, para além de situar-se num terreno cambiante, uma vez que os desenvolvimentos aplicacionais a que o AtoM® se encontra sujeito prosseguem no dia-a-dia do seu ciclo de vida e de todo um investimento que é feito, com vista à melhoria contínua das suas funcionalidades, o que torna qualquer tipo de análise sob a perspetiva das vantagens/desvantagens/aceitação na comunidade rapidamente desatualizada.

Referências bibliográficas

- Anjos, J., Caixas, M., Vaquinhas, N., Marques, N., & Negrão, S. (2015). *AtoM Work in progress. Manual de utilizador da aplicação AtoM (Access to Memory)*. Rede de Arquivos do Algarve. https://rededearquivosdoalgarve.files.wordpress.com/2015/02/atom-work-in-progress{_}manual-de-utilizador.pdf
- António, R., & Silva, A. C. (2011). Arquivos definitivos na Web: Que futuro? *Actas do Encontro Nacional de Arquivos Municipais, 10*. <https://www.bad.pt/publicacoes/index.php/arquivosmunicipais/article/view/1>
- Bardin, L. (2008). *Análise de conteúdo* (4.^a ed.). Edições 70.
- Barradas, G. (2016). A difusão de arquivos fotograficos através do AtoM (Access Memory) O exemplo da fotografia de arquitetura. *Revelar, 1*, 80-96.
- Batista, P., Barradas, G., & Forjaz, M. (2016). A experiência de utilização do AtoM no Arquivo da Social-Democracia. *AtoM: Work in Progress*, 16-18. <https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf>
- Bushey, J. (2012). ICA-AtoM: Open-source software for archival description. *Archivi & Computer, 22*(1), 10-25.
- Caldeira, F. (2016). ICA-AtoM na CGTP-IN: uma ferramenta para a descrição documental. *AtoM: Work in Progress*, 7-14. <https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf>
- Flores, D. (2016). A adoção do AtoM (ICA-AtoM) para descrição, difusão e acesso de documentos arquivísticos e as perspectivas de preservação digital e autenticidade com sua interconexão aos repositórios arquivísticos digitais confiáveis—RDC-Arq. *AtoM: Work in Progress*, 28-37. <https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf>
- Freitas, M. C. V. de. (2009). *A arquivística sob o signo da mudança: Cenários arquivísticos (re) desenhados pelo documento eletrônico* [Tese de Doutoramento, Universidade de Salamanca]. <https://gredos.usal.es/handle/10366/76594>
- Hedlund, D. C., & Flores, D. (2014). Análise e aplicação do software livre ICA-AtoM como ferramenta para descrição e acesso às informações do patrimônio documental e histórico do município de Santa Maria-RS. *Informação Arquivística, 3*(1), 24-41. <https://doi.org/10.5433/1981-8920.2014v19n3p86>
- Jorente, M. J. V., Nakano, N., da Silva, T. C., & Batista, L. da S. (2016). O Marco Civil da Internet e a Ciência da Informação: Uma discussão sobre os softwares livres AtoM e Archivematica | The Internet Civil Rights Framework and Information Science: A discussion of AtoM and Archivematica software. *Liinc em Revista, 12*(1), 90-106. <https://doi.org/10.18617/liinc.v12i1.858>
- Lima, E. D. S., & Flores, D. (2016). A utilização do Ica-Atom como plataforma de acesso, difusão e descrição dos documentos arquivísticos de instituições públicas. *Informação & Informação, 21*(3), 207. <https://doi.org/10.5433/1981-8920.2016v21n3p207>
- Moraes, H. A. R., Zafalon, Z. R., & Barroso, T. de B. (2019). Descrição arquivística, records in context (RiC) e Access to Memory (AtoM): Análise exploratória da literatura científica. *Revista Digital de Biblioteconomia e Ciencia da Informacao, 17*(1), 1-25. <https://doi.org/10.20396/rdbci.v17i0.8652807>
- Pacheco, A. F. D. (2013). *O arquivo de Manuel Marques no Centro de Documentação de Arquitectura: Aplicação da plataforma ICA-AtoM para a organização do acervo e o acesso à informação online (Dissertação de Mestrado)* [PhD Thesis, Universidade do Porto]. <https://repositorio-aberto.up.pt/handle/10216/68973>

- Pereira, J., Eugénia, M., & Gerós, M. I. (2016). AtoM – Resultados de estudo para a sua adoção na U.Porto. *AtoM: Work in Progress*, 38-46.
- Rocha, J. A. (2019). O uso do AtoM em Portugal. O panorama geral e a plataforma PAPIR como estudo de caso. *IX Seminário Internacional de Saberes Arquivísticos: «Arquivo(s) e Identidade(s)»*, 1-10. <http://casacomum.org/>
- Saraiva, N. L., & Nogueira, R. F. (2014). Software Livre: Um Diferencial Competitivo para a Arquivística. *Informação Arquivística*, 3(2), 136-142. <https://doi.org/10.18377/2316-7300/informacaoarquivistica.v3n2p136-142>
- Silva, C. G. da. (2016). AtoM-Access to Memory: A memória partilhada de Torres Vedras. *AtoM: Work in Progress*, 19-27. <https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf>
- Silva, H. (2016). O Acesso à Memória (AtoM) no Município de Albergaria-a-Velha. *12º Encontro de Arquivos Municipais*.
- Silveira, M. M. G. da. (2019). Uso do Software-Livre “Atom” na gestão e na difusão de acervos: Um projeto-piloto do Centro de Referência da Música de Minas – Museu Clube da Esquina para a Rede de Museus e Espaços de Ciências e Cultura da Universidade Federal de Minas Gerais (UFMG). *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia*, 14(1). <https://doi.org/10.22478/ufpb.1981-0695.2019v14n1.44905>
- Vasques, E. (2016). Considerações informáticas acerca do AtoM. *AtoM: Work in Progress*, 47-53. <https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf>

APÊNDICE I – LISTA BIBLIOGRÁFICA DE DOCUMENTOS SELECIONADOS PARA COMPOR A AMOSTRA DE TEXTOS USADOS NO ESTUDO (N=37)

Referências bibliográficas	
1	Anjos, J., Negrão, S., & Marques, N. (2016). Projeto de Acesso à Memória do Algarve. In <i>AtoM: Work in Progress</i> (pp. 59-64). Albufeira: Rede de Arquivos do Algarve. Retrieved from https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf
2	António, R. (2016). Atom – Acesso to Memory: Passado, presente e futuro. In <i>AtoM: Work in Progress</i> (pp. 54–57). Albufeira: Rede de Arquivos do Algarve. Retrieved from https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf
3	António, R., & Silva, A. C. (2011). Arquivos definitivos na Web: que futuro? <i>Actas Do Encontro Nacional de Arquivos Municipais</i> , (10). Retrieved from https://www.bad.pt/publicacoes/index.php/arquivosmunicipais/article/view/1
4	Barradas, G. (2016). A difusão de arquivos fotograficos através do AtoM (Access Memory) O exemplo da fotografia de arquitetura. <i>Revelar</i> , 1, 80–96.
5	Batista, L. da S. (2018). <i>O redesign do sistema Access to Memory (AtoM) para a curadoria digital de acervos museológicos heterogêneos (Dissertação de Mestrado)</i> . Universidade Estadual Paulista —Júlio de Mesquita Filho (Unesp). Retrieved from https://repositorio.unesp.br/handle/11449/154326
6	Batista, P., Barradas, G., & Forjaz, M. (2016). A experiência de utilização do AtoM no Arquivo da Social-Democracia. In <i>AtoM: Work in Progress</i> (pp. 16–18). Albufeira: Rede de Arquivos do Algarve. Retrieved from https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf
7	Bezerra, E. de S., Bandeira, P. M., & Silva, J. T. e. (2020). O ica-atom como ferramenta para descrição de documentos arquivísticos da universidade federal da Paraíba. <i>Revista Conhecimento Em Ação</i> , 5(2), 17–49. https://doi.org/10.47681/rca.v5i2.36059
8	Caldeira, F. (2016). ICA-AtoM na CGTP-IN: uma ferramenta para a descrição documental. In <i>AtoM: Work in Progress</i> (pp. 7–14). Albufeira: Rede de Arquivos do Algarve. Retrieved from https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf
9	Cé, G., & Flores, D. (2016a). Implantação do software ica-atom como ferramenta de difusão de acervos arquivísticos em ifes: o estudo de caso da ufcspa. <i>Informação Arquivística</i> , 5(2), 22–44. https://doi.org/10.22478/ufpb.1981-0695.2018v13n2.43019
10	Cé, G., & Flores, D. (2016b). O conceito e a aplicabilidade do ica-atom (atom) como ferramenta de descrição, difusão e acesso do patrimônio documental arquivístico. <i>Informatio</i> , 21(2), 88–100.
11	Chaves, E. M. L. (2018). <i>Descrição arquivística de documentos fotográficos em sistemas informatizados (Dissertação de Mestrado)</i> . Universidade de São Paulo.
12	Conrado, F. H. (2014). <i>Arranjo, descrição e difusão do patrimônio documental arquivístico da Universidade Federal de Rio Grande do Sul (Dissertação de Mestrado)</i> . Universidade Federal de Santa Maria. Retrieved from https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/trabalhoConclusao/viewTrabalhoConclusao.jsf?popUp=true&id_trabalho=1551222
13	Fantinel, E. G., Flores, D., Hedlund, D. C., Santos, A. G. dos, & Rodrigues, M. D. M. (2017). O acesso, via web, ao patrimônio documental da universidade federal do rio grande – FURG: a aplicabilidade da digitalização e do software atom (ica-atom). <i>Informação Arquivística</i> , 6(1), 318–333. https://doi.org/10.22478/ufpb.1981-0695.2018v13n2.43046
14	Flores, D. (2016). A adoção do AtoM (ICA-AtoM) para descrição, difusão e acesso de documentos arquivísticos e as perspectivas de preservação digital e autenticidade com sua interconexão aos repositórios arquivísticos digitais confiáveis – RDC-Arq. In <i>AtoM: Work in Progress</i> (pp. 28–37). Albufeira: Rede de Arquivos do Algarve. Retrieved from https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf
15	Francisco, L. C. (2019). <i>Análise da presença dos requisitos de classificação em softwares livres: DSpace, Alfresco e ICA-Atom (Dissertação de Mestrado)</i> . Universidade Estadual de Londrina
16	Gago, A. B. (2017). A importância dos registos de autoridade arquivística nos arquivos de família: o Arquivo Almada e Lencastre Bastos-um caso prático. <i>RevCEDOUA</i> , 30, 451–493. https://doi.org/10.14195/2182-7974_30_7
17	Hedlund, D. C. (2014). <i>O Patrimônio Fotográfico De Santa Maria Em Ambiente Digital (Dissertação de Mestrado)</i> . Universidade Federal de Santa Maria.
18	Hedlund, D. C., & Flores, D. (2014). Análise e aplicação do software livre ICA-AtoM como ferramenta para descrição e acesso às informações do patrimônio documental e histórico do município de Santa Maria-RS. <i>Informação Arquivística</i> , 3(1), 24–41. https://doi.org/10.5433/1981-8920.2014v19n3p86

19	Jorente, M. J. V., Nakano, N., Silva, T. C. da, & Batista, L. da S. (2016, May). O Marco Civil da Internet e a Ciência da Informação: uma discussão sobre os softwares livres AtoM e Archivemática The Internet Civil Rights Framework and Information Science: a discussion of AtoM and Archivemática software. <i>Liinc Em Revista</i> , 12(1), 90–106. https://doi.org/10.18617/liinc.v12i1.858
20	Jorente, M. J. V., Silva, T. C. da, Mesa, A. M., & Silva, M. H. da. (2016). AtoM como recurso descritivo Web de representação de arquivos públicos: Brasil e Colômbia. <i>Brazilian Journal of Information Science</i> , 10(3), 94–99. https://doi.org/10.36311/1981-1640.2016.v10n3.11.p94
21	Lima, E. D. S., & Flores, D. (2016). A utilização do Ica-Atom como plataforma de acesso, difusão e descrição dos documentos arquivísticos de instituições públicas. <i>Informação & Informação</i> , 21(3), 207. https://doi.org/10.5433/1981-8920.2016v21n3p207
22	Moraes, H. A. R. (2018). <i>Records in contexts – a conceptual model for archival description (RIC-CM): análise da proposta de um padrão internacional integrado de descrição arquivística (Dissertação de Mestrado)</i> . Universidade Federal de São Carlos.
23	Morais, C. S. P. de. (2013). <i>Estudo Orgânico-Funcional e Tratamento Arquivístico dos Concursos para Obtenção do Diploma de Arquiteto realizados na Escola de Belas Artes do Porto : utilização do software ICA-AtoM (Dissertação de Mestrado)</i> . Universidade do Porto.
24	Pacheco, A. F. D. (2013). <i>O arquivo de Manuel Marques no Centro de Documentação de Arquitectura: Aplicação da plataforma ICA-AtoM para a organização do acervo e o acesso à informação online (Dissertação de Mestrado)</i> . Universidade do Porto. Retrieved from https://repositorio-aberto.up.pt/handle/10216/68973
25	Pereira, J., Eugénia, M., & Gerós, M. I. (2016). AtoM – Resultados de estudo para a sua adoção na U.Porto. In <i>AtoM: Work in Progress</i> (pp. 38–46). Albufeira: Rede de Arquivos do Algarve.
26	Rocha, J. A. (2019). O uso do AtoM em Portugal. O panorama geral e a plataforma PAPIR como estudo de caso. In <i>IX Seminário Internacional de Saberes Arquivísticos: “Arquivo(s) e Identidade(s)”</i> (pp. 1–10). Coimbra: Faculdade de Letras da Universidade de Coimbra. Retrieved from http://casacomum.org/
27	Salcedo, D. (2020). Memória postal brasileira no Arquivo Público de Pernambuco. <i>InCID: Revista de Ciência Da Informação e Documentação</i> , 10(2), 158–174. https://doi.org/10.11606/issn.2178-2075.v10i2p158-174
28	Santos, A. G. dos. (2012). <i>Descrivendo o património documental da FURG: faculdade de Direito Clovis Bevilacqua (1959-1972) (Dissertação de Mestrado)</i> . Universidade Federal de Santa Maria.
29	Santos, H. M. dos. (2014). Aplicação do ICA-AtoM no acervo fotográfico do departamento de arquivo geral da UFSM. <i>Biblos: Revista Do Instituto de Ciências Humanas e Da Informação</i> , 28(3), 5–14. Retrieved from https://periodicos.furg.br/biblos/article/view/5357
30	Saraiva, N. L., & Nogueira, R. F. (2014). Software Livre: Um Diferencial Competitivo para a Arquivística. <i>Informação Arquivística</i> , 3(2), 136–142. https://doi.org/10.18377/2316-7300/informacaoarquivistica.v3n2p136-142
31	Seiffert, C. C. (2019). <i>Acesso e transparência ativa de informações e documentos do IFRS através da plataforma arquivística AtoM (Dissertação de Mestrado)</i> . Universidade Federal de Santa Maria – Brasil.
32	Silva, C. G. da. (2016). AtoM-Access to Memory: a memória partilhada de Torres Vedras. In <i>AtoM: Work in Progress</i> (pp. 19–27). Albufeira: Rede de Arquivos do Algarve. Retrieved from https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf
33	Silva, T. C. da. (2017). <i>AtoM, normas e padrões para a descrição e representação de informações e objetos digitais em ambientes digitais web (Dissertação de Mestrado)</i> . Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp).
34	Silva, H. (2016). O Acesso à Memória (AtoM) no Município de Albergaria-a-Velha. In <i>12º Encontro de Arquivos Municipais</i> . Castelo Branco.
35	Silveira, M. M. G. da. (2019). Uso do Software-Livre “Atom” na gestão e na difusão de acervos: um projeto-piloto do Centro de Referência da Música de Minas – Museu Clube da Esquina para a Rede de Museus e Espaços de Ciências e Cultura da Universidade Federal de Minas Gerais (UFMG). <i>Múltiplos Olhars Em Ciência Da Informação</i> , 8(1). https://doi.org/10.22478/ufpb.1981-0695.2019v14n1.44905
36	Vasques, E. (2016). Considerações informáticas acerca do AtoM. In <i>AtoM: Work in Progress</i> (pp. 47–53). Albufeira: Rede de Arquivos do Algarve. Retrieved from https://rededearquivosdoalgarve.files.wordpress.com/2016/11/livro-de-actas.pdf
37	Xavier, J. F. (2014). <i>O uso de metadados para a preservação digital no Arquivo da Marinha: a implementação do software ICA-AtoM (Dissertação de Mestrado)</i> . Universidade Federal do Rio de Janeiro.

**APÊNDICE II – VANTAGENS E DESVANTAGENS NO USO DO ATOM:
DISTRIBUIÇÃO PELAS CATEGORIAS E SUBCATEGORIAS DE ANÁLISE (N=37)**

Vantagens		Desvantagens	
Característica	n %	Característica	n %
1. Instalação e configuração			
Aplicação livre e <i>open source</i>	29 78,4	Dificuldades na instalação/ recurso ao exterior (não há apoio técnico)	7 18,9
Simplicidade de instalação	2 5,4	Dificuldades na agregação a outros Interfaces (Portal Português de Arquivos)	2 5,4
Comunidade de ajuda	13 35,1	Lentidão no carregamento de informação	1 2,7
Ambiente <i>Web</i>	26 70,3	Não apresenta estatísticas de pesquisa	3 8,1
Disponibilização <i>online</i> das descrições	19 51,4	Versão portuguesa incompleta (1.3)	2 5,4
Criação de páginas estáticas	2 5,4	Não permite relatórios de desempenho	1 2,7
Possibilidade de Parametrização/ personalização	16 43,2	Não assegura a preservação digital (ligação a <i>Archivematica</i> necessária)	12 32,4
Gestão de diferentes utilizadores /vários perfis	15 40,5		
Interoperabilidade com outros interfaces	9 24,3		
Interface de tradução multilíngue	22 59,5		
Possibilidade de georreferenciamento	1 2,7		
<i>Plug ins</i> adaptáveis	9 24,3		
Multiplataforma (windows, Linux...)	2 5,4		
Permite armazenar grande volume de informação	2 5,4		
2. Descrição arquivística			
Utilização individual ou multi-institucional/ Múltiplos repositórios	21 56,8	Impossibilidade de ocultar alguns campos de descrição (localização física)	1 2,7
Respeita normas de descrição arquivística do ICA / Descrição multinível	31 83,8	Dificuldade na importação e exportação de registos	3 8,1
Compatível com outras normas de descrição	17 45,9	Permite repetição de códigos de referência	1 2,7
Opção de mudar de norma de descrição	2 5,4	Só permite um objeto digital por registo	2 5,4
Alerta para o não preenchimento de campos obrigatórios (ISAD(G))	5 13,5	Marca de água comum a todas as instituições arquivísticas	2 5,4
Notas explicativas de preenchimento	2 5,4	<i>Thumbnails</i> gerados não são uniformes	1 2,7
Possibilidade de copiar registos	3 8,1	Perda de dados	1 2,7
Possibilidade de mover registos entre descrições	4 10,8		
Preenchimento automático do campo “Produtor” (herdado)	3 8,1		
Possibilidade de alterar ou adicionar níveis de descrição	4 10,8		
Diferentes estados da descrição (Preliminar/ Publicado)	10 27		
Criação de Registos de autoridade (não duplica a informação)	6 16,2		
Alerta para não concordância em relação a níveis superiores (data)	1 2,7		
Gestão de incorporações, doações, etc	2 5,4		

Ligação ao depósito físico	5	13,5
Possibilidade de importação e Exportação de registos	20	54
Taxonomias/ Pontos de acesso divididos em 3 categorias (vocabulário controlado)	17	45,9
Atualização automática das taxonomias em todas as descrições	3	8,1
Possibilidade de importação de objetos digitais	7	18,9
Agregação de objetos digitais (imagem, vídeo, som)	15	40,5
Colocação de Marca de Água	2	5,4
Criação de relações autores/ documentos	3	8,1

3. Front office

Utilização amigável	11	29,7	Ocultação de alguma informação na página principal (não aparecem todos os fundos)	3	8,1
Pesquisa simples ou avançada (com filtros)	9	24,3	Linguagem demasiado técnica para quem não é da área (dificuldades na Pesquisa)	1	2,7
Sugestão de termos de pesquisa	1	2,7			
Qualidade de visualização dos objetos digitais	2	5,4			
Possibilidade de impressão de listagem de registos (criação de relatórios/ instrumentos de pesquisa)	5	13,5			

Fonte: elaboração nossa com base na amostra de textos consultada (n=37).

(Página deixada propositadamente em branco)

UM MODELO SEMÂNTICO BASEADO EM ONTOLOGIA PARA O CRIS BRASILEIRO

Vivian S. Silva¹, Tales Moreira², Thiago M. Dias³, Josir Gomes⁴,
Washington C. Segundo⁵

¹IBICT, vivianss@gmail.com, ORCID iD 0000-0001-6797-2400

²IBICT, tales.info@gmail.com

³CEFET-MG, thiagomagela@cefetmg.br, ORCID iD 0000-0001-5057-9936

⁴IBICT, josireg@gmail.com, ORCID iD 0000-0001-7629-1404

⁵IBICT, washingtonsegundo@ibict.br, ORCID iD 0000-0003-3635-9384

Resumo

Objetivos: Elaboração de uma representação semântica dos dados coletados no âmbito do projeto BrCris, que propõe-se a coletar, integrar e disponibilizar informações relativas ao universo de pesquisa científica no Brasil. Esta representação semântica visa a ser compatível com os padrões internacionais e permitir a disponibilização dos dados para consulta pública de forma acessível. **Método:** Análise comparativa de padrões internacionais para seleção do modelo a ser adotado como base. Mapeamento sistemático entre os elementos do modelo relacional do BrCris e as classes e propriedades da ontologia VIVO. A base para a criação do modelo semântico do BrCris foi o modelo de entidades e relacionamentos, representação interna composta de entidades, atributos e relacionamentos. Cada um destes elementos foi mapeado para um elemento da ontologia VIVO, através de uma busca sistemática para identificar os elementos equivalentes. Quando uma classe ou propriedade equivalente não foi encontrada na ontologia, um novo elemento foi criado, dando origem a uma extensão local da ontologia. A criação de novos elementos foi reduzida ao menor número possível, sempre com o objetivo de maximizar o reuso e restringir a customização do modelo ao mínimo necessário. Para reforçar a integração e consistência entre elementos novos e existentes, sempre que possível, classes e propriedades foram criadas como subclasses e subpropriedades de classes e propriedades da ontologia VIVO. **Resultados:** Um modelo semântico composto por um subconjunto da ontologia VIVO, acrescido de uma extensão local que visa cobrir informações específicas do ambiente de pesquisa científica brasileiro. Este modelo é composto por 18 classes, sendo que 16 delas foram reutilizadas da ontologia VIVO, e apenas duas criadas na extensão local: as classes “brcris:GraduateProgram” e “brcris:RefereeRole”, definidas como subclasses das classes “vivo:Program”, da sub-ontologia VIVO Core, e “obo:BFO_0000023” (Role), da sub-ontologia OBO Foundry. Completam o modelo 78 propriedades de objeto e 133 propriedades de dados. Todas as propriedades de objeto possuem uma propriedade inversa, permitindo que a navegação dos dados representados pelo modelo possa ser feita em qualquer direção. **Conclusão:** A escolha da ontologia VIVO como base para o modelo semântico do BrCris privilegia o

reuso e a aderência aos padrões internacionais, facilitando o consumo dos dados por agentes externos. O modelo semântico final, composto por um subconjunto da ontologia VIVO e uma extensão local destinada a acomodar informações específicas coletadas no âmbito do projeto, cobre os principais conceitos da área, como Pessoas, Organizações, Publicações, Projetos, Patentes, etc., bem como os relacionamentos mais importantes entre eles. A escolha da ontologia VIVO como base e a maximização do reuso como premissa resultaram em um modelo semântico formal, expressivo e compatível com representações já utilizadas por sistemas similares em diversos países, garantindo acessibilidade, utilidade e interoperabilidade aos dados do projeto BrCris.

Palavras-chave: BrCris, Modelo Semântico, Ontologia VIVO

Introdução

O objetivo do projeto BrCris é coletar, integrar e disponibilizar informações relativas ao universo de pesquisa científica no Brasil, catalogando e traçando relacionamentos entre pesquisadores, as organizações às quais pertencem, os projetos dos quais participam e como são financiados, e todos os produtos por eles gerados, como publicações, patentes e software. O Projeto BrCris foi idealizado em 2014 no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) inspirado no modelo de um CRIS nacional proposto por Portugal, o PTCRIS¹. Desde então, o IBICT iniciou uma sequência de estudos e parcerias interinstitucionais para a execução do projeto, e em 2020 houve a implementação formal de um Projeto de Pesquisa para a construção do BrCris. O objetivo final do projeto é disponibilizar dados consolidados da produção científica nacional para que a comunidade acadêmica brasileira possa realizar estudos utilizando ferramentas computacionais livres e de fácil acesso sem a necessidade de utilizar sistemas proprietários. Uma das estratégias utilizadas para que estes dados sejam acessíveis e reutilizáveis foi a criação de um modelo semântico que atenda aos padrões internacionais, sendo compatível com as representações já utilizadas por sistemas similares em outros países.

Os modelos mais utilizados por *Research Information Systems* (RISs) são o metamodelo do OpenAIRE *Research Graph* (Manghi, et al., 2019), o *Common European Research Information Format* (CERIF) (Jörg, 2010), e a ontologia VIVO (Corson-Rikert, et al., 2010). O metamodelo adotado pelo OpenAIRE é relativamente simples, e, por ser apenas um *XML Schema*, tem um número bastante limitado de entidades e atributos, o que demandaria muita customização para que toda a informação do BrCris fosse acomodada. O CERIF, por outro lado, é um modelo bastante abrangente, porém bastante complexo e voltado para a implementação física de bancos de dados relacionais. Apesar de disponibilizar um modelo semântico, este é apenas uma abstração do modelo real, e não reflete a maneira como os dados são de fato representados quando armazenados. A ontologia VIVO é um meio-termo entre estes dois modelos, representando todos os elementos importantes do domínio (Pessoas, Organizações, Publicações, Projetos, Financiamento, Patentes, etc.), bem como seus principais atributos e relacionamentos, no nível semântico,

¹ <https://ptcris.pt>

além de refletir diretamente a forma como os dados serão armazenados, seja em OWL ou simplesmente em RDF.

Além de apresentar o melhor equilíbrio entre expressividade e complexidade, a ontologia VIVO vem sendo largamente utilizada em diversos projetos, como a disponibilização de *Linked Open Data* na Universidade de Québec (Héon, Dickner, Jerabek, & Belkouch, 2020), a representação de informações artísticas e históricas da coleção *Albrecht Haupt* (Walther & Rubach, 2020), o sistema de mapeamento de especialistas do consórcio de universidades italianas CINECA (Rollandi, 2020), e principalmente RISs, como o da Universidade Técnica de Chemnitz (Nguyen Hai Vu, Langer, & Gaedke, 2020) e o da Universidade de Osnabrück (Feldschnieders, 2020), para citar apenas algumas dos mais recentes. Um estudo comparativo (Lezcano, Jörg, & Sicilia, 2012) mostrou que a cobertura da ontologia VIVO é equivalente à do CERIF, reforçando sua posição como padrão na área.

O modelo semântico do BrCris foi então desenvolvido através do mapeamento entre os elementos do modelo de entidades e relacionamento no qual os dados são armazenados internamente, na plataforma LA Referencia, e as classes e propriedades da ontologia VIVO.

Modelagem Conceitual através de Ontologias

A ampla utilização da ontologia VIVO para representação de dados do domínio de pesquisa acadêmica é um reflexo da crescente importância de ontologias como recurso de modelagem conceitual. Segundo Gruber (2007), no contexto de sistema de banco de dados, ontologias podem ser vistas como um nível de abstração de modelos de dados, análogo a modelos relacionais ou hierárquicos mas com o intuito de modelar o conhecimento a respeito de indivíduos, seus atributos e seus relacionamentos com outros indivíduos. Por serem especificadas em linguagens que não se prendem a estruturas de dados ou estratégias de implementação, ontologias se posicionam no nível semântico, enquanto esquemas de bancos de dados se restringem aos níveis lógico ou físico (Gruber, 2007). Dois pontos importantes a serem destacados são o fato de ontologias serem um recurso formal, isto é, expresso de forma a ser interpretável por máquina; e proverem uma visão compartilhada de um domínio, e não uma visão individual (Guarino, Oberle & Staab, 2009). Estes aspectos garantem que uma ontologia seja não só um meio de representação de dados no nível semântico, mas também uma forma de permitir que estes dados sejam utilizados por diferentes agentes, que desta forma se comunicam utilizando um vocabulário comum.

Wand, Storey, & Weber (1999) argumentam que, dado que modelos conceituais se destinam a capturar o conhecimento sobre um domínio do mundo real, o significado dos elementos do modelo devem se basear em modelos da realidade. Neste contexto, ontologias, que têm primariamente o foco em prover modelos da realidade, permitem que o significado de cada elemento do modelo conceitual seja definido de forma rigorosa, levando a uma representação mais efetiva. Os sólidos princípios e o formalismo envolvidos na criação de ontologias, que levam em consideração teorias sobre a estrutura e o comportamento do mundo real, resultam em melhores modelos conceituais, garantindo que eles de fato sejam representações fiéis do domínio de

interesse (Shanks, Tansley, & Weber, 2003). Um estudo empírico realizado por Verdonck et al. (2019) confirma que, de fato, mesmo novatos produzem modelos de alta qualidade quando utilizam técnicas de modelagem conceitual orientadas a ontologias, reforçando as vantagens deste recurso na criação de representações semânticas expressivas, consistentes e precisas.

Ontologia VIVO

A ontologia VIVO é um modelo que descreve o domínio de pesquisa acadêmica, e é composta por classes e propriedades para representar uma rede de pesquisadores, as instituições e projetos com as quais estão envolvidos e as publicações e demais produtos resultantes de suas pesquisas. Apesar de ter sido desenvolvida como um componente da plataforma VIVO, uma ferramenta de navegação e visualização de informações relacionadas a atividades acadêmicas, a ontologia pode ser utilizada de forma independente em diferentes sistemas.

A principal vantagem da ontologia VIVO é reutilizar diversas outras ontologias já bastante consolidadas. Entre elas se destaca a *Basic Formal Ontology* (BFO), uma ontologia de fundamentação que fornece uma sólida base conceitual às classes e propriedades do modelo. Além da BFO e da VIVO Core, elementos das seguintes ontologias são integrados em uma estrutura semântica unificada: Bibliographic Ontology (BIBO), Event Ontology (EO), Friend of a Friend (FOAF), Geopolitical Ontology (GEO), Software Ontology (SWO), Simple Knowledge Organization System (SKOS) e vCard, entre outras.

Modelo de Entidades e Relacionamentos BrCris

O modelo de entidades e relacionamentos é um formato interno de representação da informação e foi criado de modo a servir como base para o esquema do banco de dados relacional onde os dados são armazenados, na plataforma LA Referencia, representando os principais objetos coletados e a forma como se inter-relacionam.

O BrCris reúne dados de diversas origens, sendo que, inicialmente, foram elencados os principais repositórios de dados que seriam objeto de interesse para o projeto. Diversos critérios foram adotados para a seleção dos repositórios a serem utilizados, entre eles: a consistência e atualização dos dados, o acesso aberto aos conjuntos de interesse, a amplitude dos repositórios e o reconhecimento dos dados pela comunidade científica brasileira. Como resultado, são agregados diversos repositórios, em diferentes formatos e com características distintas. As fontes de dados coletadas, e as entidades extraídas de cada uma delas, são:

- Plataforma Lattes²: Dados curriculares, sendo considerada a principal fonte de dados do BrCris, trazendo dados pessoais dos indivíduos, local de trabalho, publicações, eventos, cursos e revistas. Estes dados foram transformados nas seguintes entidades do modelo: *Person*, *OrgUnit*, *Publication*, *Event*, *Program*, *Course* e *Service*;

² <http://lattes.cnpq.br>

- Oasisbr³: Repositório mantido pelo IBICT que fornece dados confiáveis sobre publicações científicas em acesso aberto. Os dados foram mapeados para as entidades *Publication*, *Service* e *Person*;
- BDTD⁴: Assim como o Oasisbr, a BDTD também é mantida pelo IBICT. Fornece dados confiáveis sobre teses e dissertações brasileiras. Os dados foram mapeados para as entidades *OrgUnit*, *Publication*, *Course* e *Person*;
- PICTI⁵: Plataforma de instituições em Ciência, Tecnologia e Inovação também desenvolvida pelo IBICT, no contexto do Projeto BrCris. Os dados foram mapeados para a entidade *OrgUnit*;
- Plataforma Sucupira⁶: Dados de Programas de Pós-Graduação que fornecem uma base confiável de programas brasileiros. Os dados foram mapeados para as entidades *Program*, *Publication*, *Course* e *OrgUnit*;
- Instituições do INEP⁷: Assim como os programas da Plataforma Sucupira, o INEP fornece uma base confiável, mas de instituições, sendo mapeadas para a entidade *OrgUnit*.
- Dados da API e dados Abertos da CAPES⁸: Dados confiáveis fornecidos pela CAPES, mapeados para diversas entidades, como *Person*, *OrgUnit*, *Publication*, *Program* e *Course*.
- Revistas: O processamento de dados do conjuntos de revistas fornece um conjunto também confiável, sendo mapeadas para a entidade *Service*. Estes dados são provenientes das plataformas Diadorim, Latindex, DOAJ e UlrichsWeb.

As diversas fontes de dados mapeadas se complementam, possibilitando a criação de um conjunto padronizado e consistente, validado através de dados provenientes de diversas instituições brasileiras. Agregando-se todos estes repositórios, é possível a adoção de técnicas que permitem a vinculação de conjuntos inicialmente isolados, possibilitando a construção de um grande conjunto de dados interligados, facilitando a criação de consultas que inicialmente não seriam possíveis. O modelo de entidades e relacionamentos é, portanto, o responsável por organizar e integrar todas estas informações para que sejam utilizadas como entrada para os objetivos do projeto.

O modelo de entidades e relacionamentos foi elaborado com base nas características dos dados coletados, visando atender às demandas de armazenamento interno do projeto. Porém, neste formato, estes dados seriam de difícil acesso, interpretação e integração por agentes externos. A criação de um modelo semântico padronizado propõe-se a superar este problema. A ontologia VIVO atende aos requisitos do

³ <https://oasisbr.ibict.br>

⁴ <https://bdttd.ibict.br>

⁵ <https://brcris.ibict.br/picti>

⁶ <https://sucupira.capes.gov.br/sucupira>

⁷ <http://inep.gov.br>

⁸ <https://dadosabertos.capes.gov.br/dataset>

projeto BrCris não apenas em relação à aderência a padrões internacionais, mas também no que se refere à sua finalidade central de geração de métricas, já que o objetivo desta ontologia é permitir a avaliação do impacto de pessoas e organizações, bem como a identificação e reuso dos produtos de seus trabalhos.

Metodologia

A base para a criação do modelo semântico do BrCris foi o modelo de entidades e relacionamento, representação interna composta de entidades, atributos e relacionamentos. Cada um destes elementos foi mapeado para um elemento da ontologia VIVO, através de uma busca sistemática para identificar os elementos equivalentes. Quando uma classe ou propriedade equivalente não foi encontrada na ontologia, ou em outras ontologias existentes, um novo elemento foi criado, dando origem a uma extensão local da ontologia. A criação de novos elementos foi reduzida ao menor número possível, sempre com o objetivo de maximizar o reuso e restringir a customização do modelo ao mínimo necessário.

O mapeamento foi realizado da seguinte forma:

1. Cada entidade do modelo original foi mapeada para uma classe da ontologia;
2. Cada relacionamento foi mapeado para uma propriedade de objeto, ou uma cadeia de propriedades, caso a ontologia represente a relação por meio de uma classe intermediária (por exemplo, a relação entre “Person” e “Organization” é mediada pela classe “MemberRole”);
3. Cada atributo foi mapeado para uma propriedade de dados, uma propriedade de objeto ou uma cadeia de propriedades (por exemplo, dados pessoais, como nome, email, telefone, etc., são atributos no modelo original, mas representados como uma cadeia de propriedades na ontologia VIVO, que utiliza classes e propriedades da ontologia vCard para representar este tipo de dado);
4. Na ausência de um elemento equivalente na ontologia VIVO, antes de se recorrer à criação de um novo elemento, foi realizada uma busca externa na seguinte ordem:
 - a. Em sub-ontologias da ontologia VIVO, que, apesar de incluir várias outras ontologias, nem sempre as incorpora de forma completa, importando apenas um subconjunto de suas classes e propriedades. Por exemplo, a ontologia Dublin Core Terms (namespace “dcterms”) faz parte da ontologia VIVO, que, entretanto, originalmente não incluía a propriedade de dados “dcterms:provenance”. Esta propriedade foi identificada no mapeamento como equivalente ao atributo “provenance” da entidade “Journal” do modelo de entidades e relacionamentos, sendo portanto incorporada à ontologia VIVO, no namespace de origem, ou seja, “dcterms”.
 - b. Em outras ontologias do domínio. Por exemplo, a entidade “Bibliographic Reference” não tem nenhum equivalente na ontologia VIVO, nem em nenhum subconjunto externo (não incluído na ontologia VIVO) de suas sub-

-ontologias. Porém, a Bibliographic Reference Ontology (BiRO)⁹, que descreve registros, referências e coleções bibliográficas, contém a classe “*biro:BibliographicReference*”, que foi então reutilizada dentro da ontologia VIVO, evitando a criação desnecessária de uma nova classe na extensão local.

5. Na criação de novos elementos, seu tipo (propriedade de objeto ou de dado) foi definido a partir de elementos similares na ontologia, adotando-se o tipo do elemento mais próximo semanticamente (por exemplo, “*brcris:cnpqIdentifier*” foi definida como uma propriedade de dados por ser similar a “*vivo:identifier*”, do mesmo tipo). Sempre que possível, as propriedades foram criadas como subpropriedades de propriedades já existentes. O mesmo princípio foi adotado na criação de novas classes, sempre definidas como subclasses de classes já existentes na ontologia VIVO.

Resultados

O modelo semântico final é composto por um subconjunto da ontologia VIVO (com inclusões ocasionais de elementos adicionais de suas sub-ontologias) acrescido de uma extensão local que engloba as classes e propriedades destinadas a representar informações específicas do projeto BrCris. Como a ontologia VIVO é bastante extensa, foi adotada a abordagem MIREOT (*Minimum Information to Reference an External Ontology Term*) (Courtot, et al., 2011) para evitar a inclusão de elementos desnecessários. A Tabela 1 mostra o mapeamento entre as entidades do modelo original e as classes da ontologia VIVO, que serviu como ponto inicial e referência para o mapeamento do restante do modelo. As entidades em itálico são elementos não existentes ou representados como relacionamentos ou atributos no modelo original, mas que também correspondem a classes na ontologia. O *namespace* identifica a ontologia de origem, sendo que o *namespace* “*brcris*” denota as classes criadas na extensão local. As classes “*brcris:GraduateProgram*” e “*brcris:RefereeRole*” foram definidas como subclasses das classes “*vivo:Program*” e “*obo:BFO_0000023*” (*Role*), respectivamente, reforçando a integração e consistência entre elementos novos e existentes.

Tabela 1. Mapeamento entre entidades e classes

Entidade Original	Classe Mapeada	Entidade Original	Classe Mapeada
Person	foaf:Person	Funding	vivo:Grant
OrgUnit	foaf:Organization	Event	event:Event
Service	obo:ERO_0000005 (Service)	<i>BibliographicReference</i>	biro:BibliographicReference
<i>Journal</i>	bibo:Journal	<i>ResearchGroupMembership</i> <i>ProjectMembership</i>	vivo:MemberRole
Publication	bibo:Document	<i>Authorship</i>	vivo:Authorship
Course	vivo:Course	<i>Advisoring</i>	vivo:AdvisingRelationship
GraduateProgram	brcris:GraduateProgram	<i>Refereeing</i>	brcris:RefereeRole
Patent	bibo:Patent	<i>Editoring</i>	vivo:Editorship
Project	vivo:Project	<i>ResearchArea</i>	skos:Concept

⁹ <https://sparontologies.github.io/biro/current/biro.html>

A Tabela 2 apresenta as estatísticas do modelo final. O número relativamente grande de novas propriedades de dados (88 ao todo) deve-se ao fato de algumas entidades coletadas apresentarem um grande nível de detalhe, não previsto na ontologia VIVO. Apesar de aumentar o grau de customização, a inclusão destas propriedades garante que nenhum dado seja perdido. Como estão agrupadas no novo *namespace* (“brcris”), podem ser facilmente filtradas caso não sejam do interesse do usuário final.

Tabela 2. Estatísticas do modelo semântico do BrCris

Elemento	Ontologia VIVO	Extensão BrCris	Total
Classe	16	2	18
Propriedade de objeto	50	28	78
Propriedade de dados	45	88	133

A Figura 1 mostra um recorte do modelo final, ilustrando a classe “bibo: Document” (equivalente à entidade “Publication”) e alguns de seus principais relacionamentos. É possível notar que toda propriedade de objeto na ontologia VIVO (em azul na figura) possui uma propriedade inversa, e este mesmo padrão foi adotado para as novas propriedades (em verde).

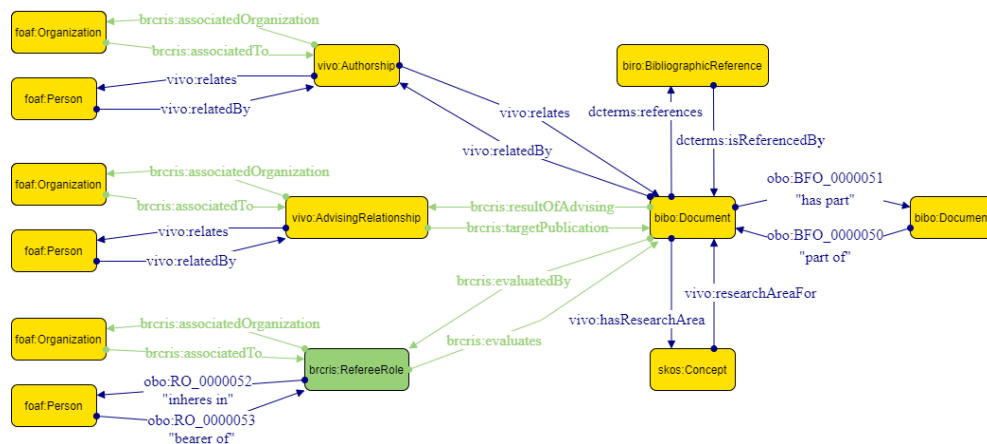


Figura 1. Recorte do modelo semântico para publicações

A Figura 2 apresenta outro recorte, mostrando como a nova classe “brcris:Graduate Program” se integra ao restante da ontologia, relacionando-se com as classes “foaf:Organization”, “vivo:Course”, “bibo:Document” e “skos:Concept” (que representa áreas de pesquisa ou de conhecimento). Alguns destes relacionamentos são também customizados, como a propriedade “brcris:hasGraduate Program” entre “foaf:Organization” e “brcris:GraduateProgram”, mas também foi possível reutilizar propriedades, como “vivo:offers”, entre “brcris:hasGraduate Program” e “vivo:Course”.

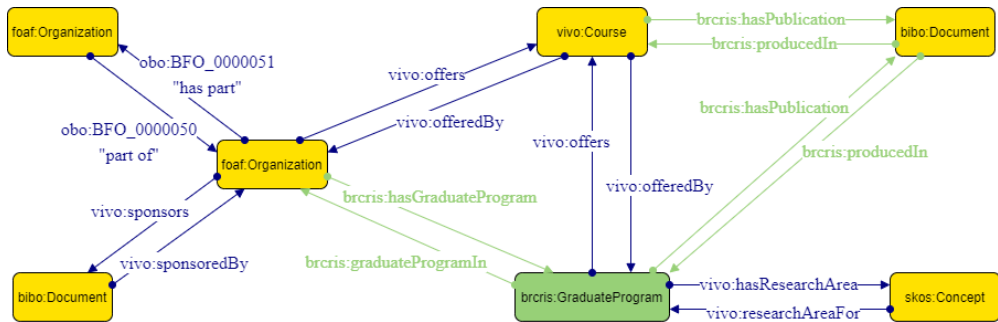


Figura 2. Recorte do modelo semântico para organizações

Os resultados apresentados se referem à primeira versão do modelo semântico, que, assim como qualquer ontologia, está em constante evolução para melhor representar seu domínio. Neste caso particular, esta evolução também se destina a atender às demandas do projeto BrCris, cujo universo de informações também tende a se expandir com a inclusão de novas fontes de dados.

Conclusões

O projeto BrCris reúne, integra e inter-relaciona informações relativas ao universo de pesquisa científica no Brasil, e uma representação semântica adequada é crucial para que estas informações sejam acessíveis, reutilizáveis e interoperáveis. O desenvolvimento do modelo semântico do BrCris teve como princípio central a aderência a padrões internacionais para representação de dados na área, resultando na escolha da ontologia VIVO como base para o modelo. O modelo semântico final, composto por um subconjunto da ontologia VIVO e uma extensão local destinada a acomodar informações específicas coletadas no âmbito do projeto, cobre os principais conceitos da área, como Pessoas, Organizações, Publicações, Projetos, Patentes, etc., bem como os relacionamentos mais importantes entre eles.

A escolha da ontologia VIVO como base e a maximização do reuso como premissa resultaram em um modelo semântico formal, expressivo e compatível com representações já utilizadas por sistemas similares em diversos países, garantindo acessibilidade, utilidade e interoperabilidade aos dados do projeto BrCris.

Referências

- Corson-Rikert, J., Mitchell, S., Lowe, B., Rejack, N., Ding, Y., & Guo, C. (2010). The VIVO ontology. *Synthesis lectures on semantic web: theory and technology*.
- Courtot, M., Gibson, F., Lister, A. L., Malone, J., Schober, D., Brinkman, R. R., & Ruttenberg, A. (2011). MIREOT: The minimum information to reference an external ontology term. *Applied Ontology*.
- Feldschnieders, D. (2020). VIVO as a standalone CRIS at University Osnabrück. *Proceeding of the 11th International VIVO Conference*. Technische Informationsbibliothek (TIB).

- Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology? *Handbook on ontologies*. Springer, Berlin, Heidelberg.
- Gruber, T. (2009) Ontology. *Encyclopedia of database systems*.
- Héon, M., Dickner, N., Jerabek, A., & Belkouch, R. (2020). VIVO as a Linked Open Data Enabler for the Université du Québec Network. *Proceedings of the 11th International VIVO Conference*. Technische Informationsbibliothek (TIB).
- Jörg, B. (2010). CERIF: The common European research information format model. *Data Science Journal*.
- Lezcano, L., Jörg, B., & Sicilia, M.-A. (2012). Modeling the context of scientific information: Mapping VIVO and CERIF. *Proceedings of the International Conference on Advanced Information Systems Engineering*. Springer.
- Manghi, P., Bardi, A., Atzori, C., Baglioni, M., Manola, N., Schirrwagen, J., & Principe, P. (2019). *The OpenAIRE Research Graph Data Model*. Zenodo.
- Nguyen Hai Vu, D., Langer, A., & Gaedke, M. (2020). TUCfis: Applying VIVO as the new RIS of the Technical University of Chemnitz. *Proceedings of the 11th International VIVO Conference*. Technische Informationsbibliothek (TIB).
- Rollandi, L. (2020). The CINECA “find an expert” solution based on VIVO. *Proceedings of the 11th International VIVO Conference*. Technische Informationsbibliothek (TIB).
- Shanks, G., Tansley, E., & Weber, R. (2003). Using ontology to validate conceptual models. *Communications of the ACM*.
- Verdonck, M., Gailly, F., Pergl, R., Guizzardi, G., Martins, B., & Pastor, O. (2019). Comparing traditional conceptual modeling with ontology-driven conceptual modeling: An empirical study. *Information Systems*.
- Walther, T., & Rubach, B. (2020). Representation of Art-Historical Information with Vitro. *Proceedings of the 11th International VIVO Conference*. Technische Informationsbibliothek (TIB).
- Wand, Y., Storey, V. C., & Weber, R. (1999). An ontological analysis of the relationship construct in conceptual modeling. *ACM Transactions on Database Systems (TODS)*.

AS REDES SOCIAIS DE COLABORAÇÃO INTERNACIONAL DE AUTORES IMPACTO NAS REVISTAS E NA DIFUSÃO DO CONHECIMENTO EM CIÊNCIA DA INFORMAÇÃO

Ana Gouveia Coelho¹

¹Universidade de Coimbra, gouveiacoelho@fl.uc.pt,
ORCID-ID 0000-0001-8520-2522

Resumo

O domínio da ciência da informação (CI) inclui uma disciplina designada de bibliometria que antecede a constituição da CI como campo autónomo do saber. Esta socorre-se de indicadores que visam avaliar qualitativamente a produção científica de determinado autor ou de uma revista. O JIF ou CiteScore (SCOPUS) são exemplo de indicadores bibliométricos com base no número de citações durante determinado período de tempo. Existem diversas variáveis que influenciam o número de citações para além da qualidade (subjéctiva) do artigo. Uma delas é a inserção dos autores em redes sociais de colaboração internacional que se relacionam com um efeito multiplicador da difusão do conhecimento. Objetivou-se perceber se, no domínio da CI, as revistas com maior CiteScore plasmavam redes internacionais de autores mais intensas e também, em termos globais, quais os países que mais influenciam atualmente, a construção do pensamento científico em CI. Utilizou-se um método de tipo exploratório e comparativo. Para o efeito recolheram-se na base de dados SCOPUS os cinquenta artigos mais citados no período 2016-2020 e compararam-se com as redes sociais de colaboração internacional de autores plasmadas nos cinquenta artigos mais citados de cada uma das primeiras seis revistas da CI, dos 1.º e 2.º quartis. Verificou-se que as revistas do 1.º quartil detinham redes sociais internacionais mais complexas dos que as revistas do 2.º quartil. Também, em termos globais, os dados indiciam o crescimento da China nas redes sociais de colaboração internacional. Perspetivam-se mudanças no protagonismo da difusão do conhecimento também em ciência da informação

Palavras-chave: *Redes sociais internacionais de autores, CiteScore, difusão do conhecimento*

Introdução

O domínio da ciência da informação inclui uma área disciplinar designada bibliometria, cujo conceito, antecede a constituição da ciência de informação como campo autônomo do saber. Este foi utilizado a primeira vez por Otlet, em 1934, como uma medida de quantidade aplicada aos livros. No entanto, atribui-se a Pichard (1969) a designação tal como a conhecemos hoje, isto é, utilizada para descrever “todos os estudos que procuram quantificar os processos de comunicação escrita” (Pinheiro, 1983). A bibliometria utiliza como ferramentas uma série de indicadores, de índole estatística, que visam caracterizar, de forma quantitativa e também qualitativa, a produção científica de determinado investigador ou a qualidade de determinada revista.

O JIF (Journal Impact Factor) foi introduzido no início da década de 60, por Eugene Garfield (1925-2017). Constitui um índice utilizado para a ordenação das revistas científicas de acordo com a sua qualidade, apesar de se basear num cálculo matemático (Bordons, Fernandez & Gomez, 2002). É atualizado anualmente pelo Journal Citations Reports-Clarivate associado à Thompson Reuters. O JIF é calculado com base no número médio de citações de artigos publicados nos dois anos precedentes. Existem outros indicadores, cujo cálculo tem como referência o número de citações dum artigo num período de tempo, como por exemplo, o fornecido pela base de dados SCOPUS (calculado para quatro anos invés de dois), designado por CiteScore (2016). Como outros indicadores bibliométricos é ainda de destacar, como exemplo, o índice h (Hirsch, 2005) ou índice g (Egghe, 2006), que surgiram como uma tentativa de contrabalançar as reconhecidas insuficiências e limitações do JIF. Contudo, graças à simplicidade aritmética dessa tipologia de indicadores (JIF, CiteScore), que ajuda a sua compreensão e fácil acesso, continuam a ser largamente utilizados (Glanzel & Moed, 2002). Saliente-se que apesar das críticas que lhe são imputadas, apresenta também robustez, visível na própria percepção dos leitores acerca da qualidade das revistas. É disso exemplo um estudo realizado por Saha, Saint & Christakis (2003), incidindo em revistas da área da saúde e que demonstrou uma correlação ($r=0,82$) entre o JIF e a percepção pessoal dos leitores da qualidade das revistas em estudo. Ou seja, mesmo pesando sobre este indicador, ou outros semelhantes, uma série de críticas, a sua capacidade em caracterizar qualitativamente determinado periódico é válida. O que não exclui o uso conjunto de outros indicadores bibliométricos, tais como o índice h ou índice g (já referidos) ou a lei de Bradford, ou ainda outros altimétricos (Glanzel & Moed, 2002), visando tornar mais robusta uma análise que, apesar de tudo, encerra sempre alguma dose de subjetividade, inerente a uma caracterização, que não é puramente matemática e também limitada a determinado período temporal e espacial. Este último fator (espaço) é especialmente verdadeiro, se considerarmos que a localização geográfica dos investigadores constitui uma variável que influi no seu impacto e performance e logo nas citações que amealham. As comunidades científicas localizadas em países com maiores restrições orçamentais, por exemplo, na aquisição de assinaturas de bases de dados que possibilitem o acesso a informação contida em revistas de internacionais, ficam em desvantagem na participação no processo global, pelo isolamento que as restrições económicas tendem a impor, ainda que possam ter uma produção científica bastante rica, a sua capacidade em difundir esse conhecimento

fica limitada. Dado que que o sistema tradicional de métricas ainda é o mais relevante no que toca à avaliação da performance ou do impacto dos investigadores e das revistas, diversos estudos têm identificado variáveis, que nada têm a ver com a qualidade intrínseca ou originalidade de um artigo e que influenciam tudo o resto. Dessas variáveis, destaca-se a visibilidade dos investigadores decorrente da sua participação em redes sociais de co-autoria, o que é influenciado, não apenas pela vontade dos investigadores, mas também pela possibilidade de acesso a essas redes. (Tahamtan, Ashefar e Ahamdzadeh, 2016); (D'Angelo et al., 2011).

Face ao exposto é de salientar que a importância da variável cooperação dos autores na elaboração de artigos com maiores citações, encontra-se demonstrada por diversos estudos quantitativos e que consubstanciam a teoria na qual, se alicerçam mesmo, os vários algoritmos de análise bibliométrica, construídos para prever o impacto de determinado artigo. Muitos conceitos chave da ciência das redes (network science), tem as suas raízes na teoria dos grafos (Barabasi, 2016). Os grafos de forma natural, representam fluxos de interdependência entre objetos relacionados. De acordo com Higgins e Ribeiro (2018), a “linguagem dos grafos é a ferramenta fundamental para descrever a morfologia das redes sociais”. Foi König (1990), matemático húngaro, quem primeiro sistematizou a “teoria dos grafos”. Mais recentemente, estudos desenvolvidos por Newman (2001), demonstraram a total aplicabilidade da teoria dos grafos para a visualização de redes sociais e por tal, a sua aplicabilidade no seu estudo. Neste caso, os vértices denotam entidades e as ligações, a comunicação entre as entidades, num misto de modelização e representação gráfica (Chen, 2006); (Ahmed & Chen, 2016). Como exemplo empírico, pode-se ainda referir que os artigos publicados em coautoria tendem a ser mais apresentados em seminários, conferências internacionais, o que os torna, desde logo, mais visíveis e por isso mais citados. Ou seja, um artigo pode receber mais citações apenas por ser mais “barulhento”. (Bosquet & Combes, 2013). Tal consubstancia o princípio da difusão do conhecimento em redes sociais, que se espalha para outros domínios, aparentemente até, não relacionados com o domínio de origem. Como exemplo, podemos citar os estudos desenvolvidos por Klavans e Boyack (2007) no âmbito das redes de co-citações na descoberta de campos de estudo multidisciplinares que deixa bem patente o poder das representações visuais, ou ainda o estudo, desenvolvido pelos mesmo autores, também com base na rede de co-citações, acerca das relações entre o trabalho desenvolvido por Albert Einstein e por Pablo Picasso, de forma a responder à pergunta de investigação, se seria uma coincidência ou não, o desenvolvimento do cubismo e da teoria da relatividade no início do século XX.

A medição prática do nível da difusão do conhecimento, pode ser feita com recurso à representação gráfica de redes de informação que espelham indicadores, tais como, a densidade, centralidade, diâmetro e a conectividade das redes e que refletem o potencial da mesma, em difundir a informação (Newman, 2003). Ou seja, quanto mais complexas foram as redes mais potencial de difusão de conhecimento detêm e no nosso caso, os artigos assim produzidos, em co-autoria, também detêm maior potencial de colher mais citações que por sua vez, vão contribuir para aumentar o JIF ou CiteScore das revistas que os publicam.

Uma rede social intensa tende a multiplicar e ampliar a transferência de informação, o que foi demonstrado por modelos designados de “capital social”, que enfati-

zam a importância da coesão e da amplitude dessas redes no processo (Chang & Chuang, 2011). O capital social constitui um fator que influencia a competitividade das organizações (Reagns & McEvily, 2003) e logicamente, também das revistas. Consequentemente, podemos afirmar que a extensão da rede profissional/social na qual os investigadores se inserem, pode ser usada como indicador para prever a performance futura dos investigadores e das organizações que os filiam. Contudo, para que tal se verifique a rede deve ser ampla, não só em número de investigadores que dela fazem parte, mas também, em termos espaciais. Ou seja, não basta apenas existir co-autoria. Se a rede for próxima, há uma grande possibilidade de se receber informação redundante e que não contribui para mais conhecimento, o que se refletirá em menor performance dos seus membros (Abbasi, Altmann & Hossain, 2011). Demonstrou-se que o efeito multiplicador da difusão do conhecimento em redes sociais é bastante ampliado se se tratar duma rede internacional. (Bathelt et al., 2004). Os artigos mais citados, são, duma forma geral, aqueles que resultam do trabalho em equipe de investigadores de diferentes países, num padrão complexo, em que o conjunto constituído pela qualidade e dinâmica, imprimida pela visibilidade, contribuem para a contabilidade mais elevada de citações (Aksnes, 2003); (Glänzel & Shoepflin 1995). Por outro lado, segundo Tahamtan, Ashefar e Ahamdzadeh (2016), a colaboração internacional com alguns países como por exemplo, os Estados Unidos, aumenta, à partida, o número provável de citações. Tal pode implicar um efeito perverso, firmado na perpetuação de determinados padrões e tornando muito difícil de almejar o mesmo reconhecimento de autores e revistas, pertencentes a países ultraperiféricos não conectados com a rede principal.

Logicamente, se o número de citações de artigos elaborados por autores que participam em redes internacionais de difusão de conhecimento é mais elevado, também o JIF ou o Cite Score das revistas que publicam os artigos desses autores, tenderá a aumentar. É lícito pensar que nas revistas com maior JIF ou CiteScore, transparecem redes internacionais de colaboração mais complexas (maior densidade, diâmetro, número de nodos e ligações), ou seja, integram mais artigos de autores inseridos nessas redes internacionais. tal deverá ser verdadeiro para a Ciência da Informação, à semelhança de outros domínios, sujeita aos mesmos fenómenos e idiossincrasias no processo de difusão de conhecimento, difusão essa, que se torna visível em grafos de redes sociais e que se interligam de forma mais eficaz quando conectam diversos países numa influência recíproca

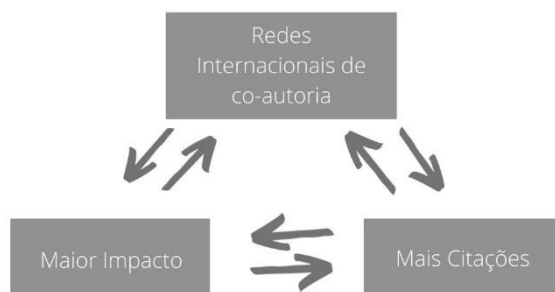


Fig 1. As revistas com maior impacto JIF ou CiteScore, publicam artigos mais citados de autores inseridos em redes sociais internacionais de co-autoria.

Foi essa hipótese que se procurou testar e consubstanciou a problemática deste estudo e a pergunta de investigação que se procurou responder:

“Será que as revistas do domínio da ciência da informação com maior CiteScore, integram redes autorais de colaboração internacional mais complexas?”

A procura da resposta para a interrogação acima implicou o delineamento de uma estratégia, baseada nos recursos que permitissem a sua operacionalização. Para o efeito, foram firmados os seguintes objetivos:

Objetivos

Objetivo Geral:

Verificar se as revistas da Ciência da informação (CI) com maior impacto de acordo com o CiteScore (SCOPUS) refletem redes autorais de colaboração internacional mais complexas

Objetivos específicos:

– Proceder à comparação da complexidade das redes autorais de colaboração internacional entre as revistas dos 1.º e 2.º quartis de acordo com os dados fornecidos pela base de dados SCOPUS.

– Perceber entre quais países se verificam as redes de colaboração internacional de autores mais intensas e daí deduzir quais os países que se destacam como protagonistas no processo de difusão do conhecimento em ciência da informação.

Metodologia

Foi utilizado um método qualitativo de tipo descritivo e comparativo. Para a operacionalização, procedeu-se da seguinte forma:

1 – Na base de dados Scopus selecionar as revistas classificadas com sendo da ciência da informação (library and information sciences).

2 – Selecionar as seis primeiras revistas dos 1.º e 2.º quartis e, concomitantemente, selecionar os cinquenta artigos mais citados de cada revista, durante o período 2016-2020.

3 – Os dados bibliométricos, nomeadamente, dos cinquenta artigos mais citados para cada revista no período considerado, foram operacionalizados com o software R, pacote Bibliometrix, interface biblioshiny.

4 – Para cada uma das revistas foi calculado, através do software, as redes de colaboração entre países, apresentadas sobre a forma de um grafo para melhor visualização.

Resultados

Procedeu-se à ordenação das primeiras seis revistas classificadas como sendo da ciência da informação (library and information science) dos 1.º e 2.º quartis, por ordem decrescente do CiteScore. Concomitantemente foi consultada a base de dados Scopus para a recolha do valor atribuído ao CiteScore, bem como o número total de artigos indexados em cada revista e o número total de citações atribuídas aos cinquenta artigos mais citados, durante o período de tempo (2016-2020)

Tabela 1 – As primeiras seis revistas da Ciência da Informação

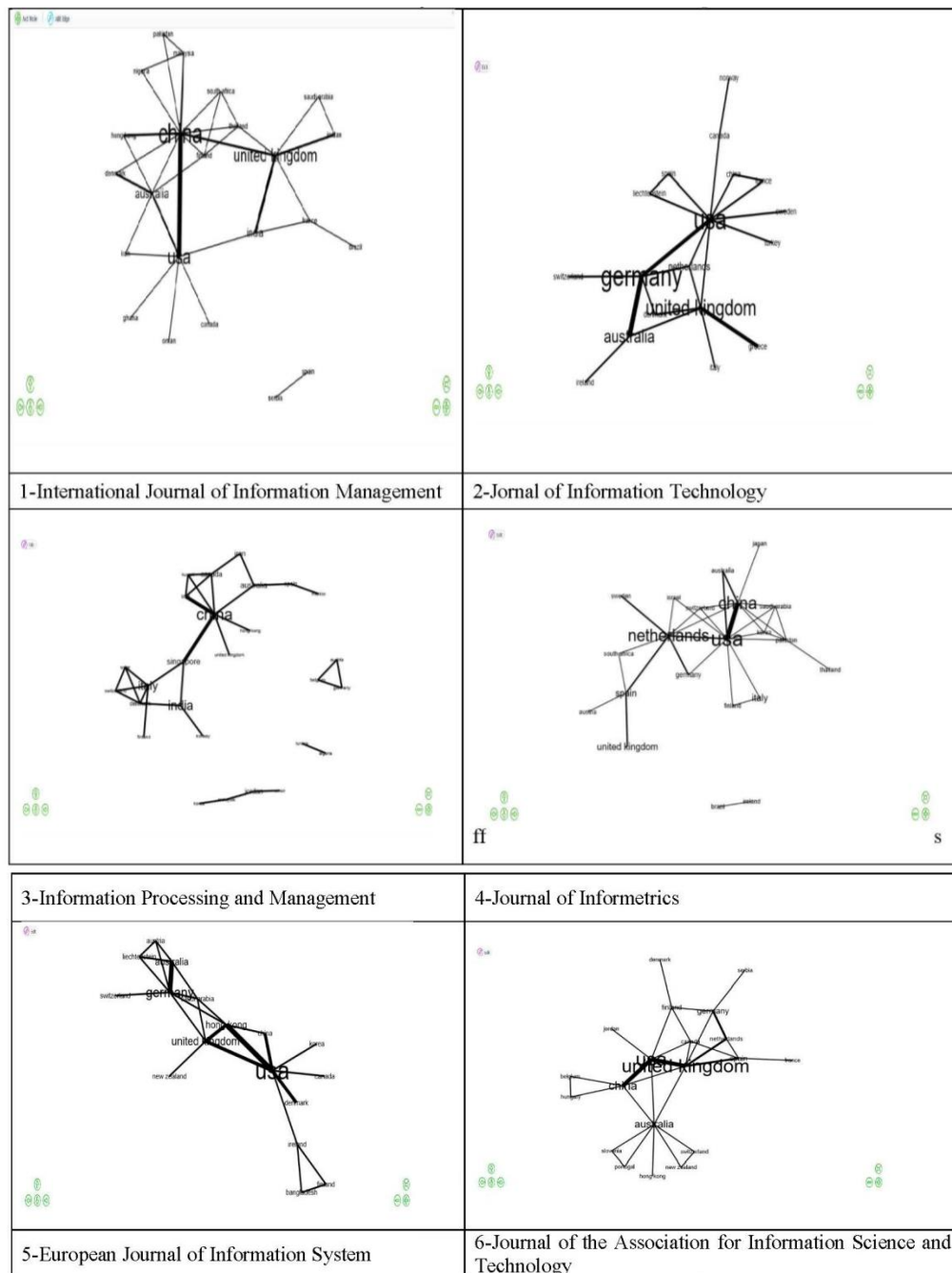
Título da Revista	CiteScore	Highest percentile	N.º documentos indexados 2016-2020	N.º de citações colhidas nos 50 artigos mais citados (2016-2021)
1º Quartil				
International Journal of Information Management	14,1	99.0%1/227/Library and Information Sciences	659	7048
Journal of Information Technology	9,8	98.0%4/227/Library and Information Sciences	93	1640
Information Processing and Management	8,6	98.0%5/227/Library and Information Sciences	614	3174
Journal of Informetrics	8,4	97.0%7/227/Library and Information Sciences	394	2863
European Journal of Information Systems	8,3	96.0%8/227/Library and Information Sciences	170	1579
Journal of the Association for Information Science and Technol	7,9	95.0%10/227/Library and Information Sciences	740	3659
2º Quartil				
Computers in the Schools	1,9	73.0%61/227/Library and Information Sciences	86	349
Electronic Library	1,9	73.0%60/227/Library and Information Sciences	318	687
New Review of Academic Librarianship	1,9	74.0%58/227/Library and Information Sciences	113	296
Reference Services Review	1,9	74.0%59/227/Library and Information Sciences	173	327
Bottom Line	1,8	72.0%64/227/Library and Information Sciences	92	348
Informing Science	1,8	72.0%63/227/Library and Information Sciences	65	185

Fonte: SCOPUS (adaptado)

Verifica-se que no conjunto das s revistas do 1.º Quartil existe uma revista que se demarca das restantes, a International Journal of Information Management. Já o segundo lugar, ocupado pela revista Journal of Information Technology, revela um CiteScore cerca de 5 pontos abaixo do primeiro lugar e espelha a maior diferença de todo o conjunto. As restantes posições, quer pertençam ao primeiro, quer pertençam ao segundo quadril, seguem uma linha decrescente num contínuo sem grandes discrepâncias. O maior CiteScore, que se destaca em relação aos outros, atribuído à primeira revista do 1.º quartil, indicia que esta será um veículo privilegiado da difusão do conhecimento em CI, o que se reflete de forma evidente no maior número de citações atribuídas aos 50 artigos mais citados, que se distancia grandemente das outras fontes de informação.

A visualização das redes de cooperação social internacional associadas às revistas, através dos 50 artigos mais citados em cada uma delas, permite a comparação e a verificação, se existe ou não evidência de uma maior complexidade das redes associadas às revistas do 1.º quartil, face às revistas do 2.º quartil e também, quais os países protagonistas na construção do conhecimento em ciência da informação, durante o período considerado

Tabela 2 – Redes internacionais de colaboração associadas às revistas do 1.º Quartil da CI

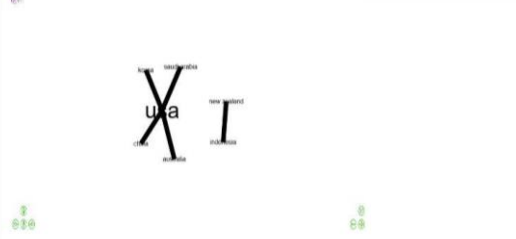
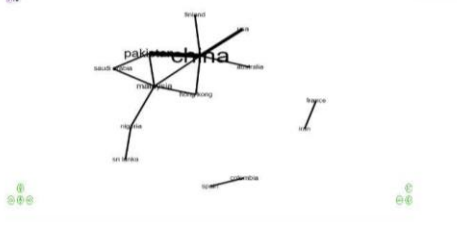

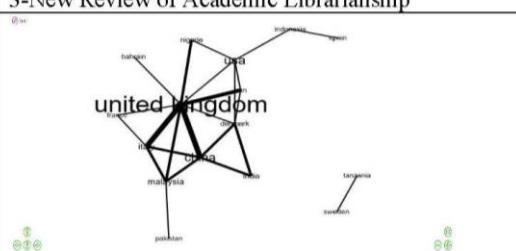
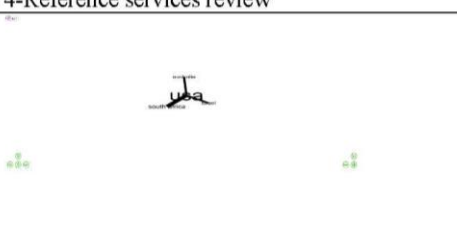


Nota: Os grafos foram calculados tendo apenas por base os 50 artigos mais citados em cada uma das revistas

Fonte: SCOPUS

Output: R Biblioshiny

Tabela 3 – Redes internacionais de colaboração entre autores associadas às revistas do 2.º Quartil

	
<p>1-Computers in the Schools</p> 	<p>2-Electronic Library</p> <p>Não foi indexado nenhum artigo produzido em coautoria internacional</p>
<p>3-New Review of Academic Librarianship</p> 	<p>4-Reference services review</p> 
<p>5-Botton line</p>	<p>6-Informing Science</p>

Nota: Os grafos foram calculados tendo apenas por base os 50 artigos mais citados em cada uma das revistas

Fonte: SCOPUS

Output: R Biblioshiny

É evidente a diferença existente entre a maior complexidade das redes sociais internacionais dos autores associados às revistas do 1.º quartil em relação às mesmas redes associadas às revistas do 2.º quartil.

Esta complexidade refere-se ao número de nodos e ligações, bem como o diâmetro das ligações que aumenta em conformidade com o número de artigos produzidos, no geral, a uma maior conectividade no caso das revistas do 1.º quartil, que preconiza uma maior difusão do conhecimento. A comparação dos grafos ilustra que o número de artigos produzidos em coautoria internacional é sempre superior nas revistas do primeiro quartil relativamente às do segundo quartil, pois para ambos os casos se recolheram a mesma quantidade de artigos, isto é, os cinquenta artigos mais citados. Salienta-se ainda, que todas as revistas do primeiro quartil refletem a presença dos EUA e/ou China, mas também de outros países numa abrangência interna-

cional muito maior, do que a verificada na cobertura internacional das redes do 2.º quartil, que também não excluem a presença dos EUA e/ou China. A presença praticamente constante destes dois países, quer no 1.º quartil, quer no 2.º quartil, permite induzir do protagonismo desses países, no processo de construção do conhecimento em CI. Por via das citações são predominantemente colhidas informações, difundidas daquelas regiões do globo e que se espalham, influenciado a construção do conhecimento numa forma global.

Conclusões

No domínio da CI e no período considerado, os dados recolhidos e a respetiva operacionalização, corroboram a correlação entre as redes internacionais de autores e o impacto das revistas que publicam esses artigos, fruto da colaboração internacional. É ponto assente que a construção do conhecimento é resultado de um processo social e colaborativo, em que as informações ao serem partilhadas, se conjugam cognitivamente, num crescendo, cujo resultado final é mais do que a mera soma aritmética das partes, tal como um sistema. Faz sentido que os autores, que carregam consigo experiências muito diferentes, fruto da envolvimento cultural e de aprendizagem diversas, quando se cruzam numa rede colaborativa, os resultados dessa partilha, tendem a se potenciar de forma mais eficaz, resultando em artigos mais apelativos e que colhem mais citações. Essas maiores citações, contribuem para aumentar o impacto das publicações científicas, que os acolhem. Ou seja, as revistas que conseguem atrair ou escolher, artigos publicados no âmbito dessas redes internacionais de autores, tendem a ter mais sucesso. Tal traduz-se num ciclo de retroação positiva, em que as revistas com maior impacto também são mais atraentes para esses autores ou consórcios de autores internacionais, provavelmente com maiores financiamentos, que tenderão a produzir, através dos diferentes contributos, artigos que colhem maiores citações, sendo que a inversa também se verifica, ou seja, os autores colhem as maiores citações, porque publicam em revistas de maior reputação.

No âmbito das diferentes redes internacionais de colaboração verifica-se que todas englobam os EUA e/ou a China, mesmo as do 2.º quartil. Por sua vez, verifica-se que as revistas com maior reputação demonstram redes mais complexas e extensas e que incluem também importantes fluxos entre os EUA e a China, mas também de outros países que no seu conjunto espelham uma conectividade espalhada por todos os continentes. Este é um ponto de distinção em ter em conta, a inclusão de contributos oriundos de países periféricos, mas que imprimem uma maior riqueza de informação, outras perspetivas, que ditam o maior sucesso dos artigos.

Da observação e face ao exposto, considerando ainda que o modus mais utilizado pelos investigadores na pesquisa de informação dá-se através de bases de dados internacionais, como por exemplo a SCOPUS a qual se utilizou para a realização deste estudo, deduz-se também, que o conhecimento em CI, se encontra tendencialmente consubstanciado em construções teóricas emanadas dos EUA e da China.

Salienta-se ainda que considerando a revista com maior CiteScore, a *International Journal of Information Management*, verifica-se que a China constitui o nodo com maior densidade e centralidade e conectividade, nomeadamente face ao nodo

correspondente aos EUA. Apesar dos grafos terem sido cálculos tendo por base os cinquenta artigos mais citados, tal poderá constituir um indício que o protagonismo na construção do conhecimento em ciência da informação, atribuída tradicionalmente aos EUA, está, atualmente, a deslocar-se também para aquele país.

Limitações ao estudo: As principais limitações têm a ver com o número de revistas utilizadas para comparação, bem como, a única base de dados analisada foi a Scopus e o CiteScore.

Recomendações futuras: Apesar da análise bibliográfica ser amplamente utilizada, nomeadamente para o estudo de campos emergentes de investigação científica, a influência internacional, contextualizada pelo ambiente social e humano dos países que mais emanam investigação científica, patente em bases de dados de acesso global e a respetiva influência cultural na construção do pensamento científico, constitui ainda, um campo pouco explorado.

Referências

- Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, 5(4), 594-607. <https://doi.org/10.1016/j.joi.2011.05.007>
- Ahmed, N. M., & Chen, L. (2016). An efficient algorithm for link prediction in temporal uncertain social networks. *Information Sciences*, 331, 120-136. <https://doi.org/10.1016/j.ins.2015.10.036>
- Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159-170. <https://doi.org/10.3152/147154403781776645>
- Barabási, A.L. (2016). *Network Science*. Cambridge University Pres. <http://networksciencebook.com/>
- Bathelt, H., MalMBERG, A., & Maskell, P. (2004). Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28(1), 31-56. <https://doi.org/10.1191/0309132504ph4690a>
- Bordons, M., Fernandez, M. T., & Gomez, I. (2002). Advantages and limitations in the use of impact factor measures for the assessment of research performance in a peripheral country. *Scientometrics*, 53(2), 195-206. <https://doi.org/10.1023/A:1014800407876>
- Bosquet, C, Combes, P. (2013). Are academics who publish more also more cited? Individual determinants of publication and citation records? *Scientometrics*, 97, pp. 831-857
- Chang, H. & Chuang, S. (2011). Social capital and individual motivations on knowledge sharing: Participant involvement as a moderator. *Information & Management*, 48(1), 9-18. <https://doi.org/10.1016/j.im.2010.11.001>
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359-377. <https://doi.org/10.1002/asi.20317>
- D'Angelo, C. A., Giuffrida, C., & Abramo, G. (2011). A Heuristic Approach to Author Name Disambiguation in Bibliometrics Databases for Large-Scale Research Assessments. *Journal of the American Society for Information Science and Technology*, 62(2), 257-269. <https://doi.org/10.1002/asi.21460>
- Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69(1), 131-152. <https://doi.org/10.1007/s11192-006-0144-7>

- Glanzel, W., & Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics*, 53(2), 171-193. <https://doi.org/10.1023/A:1014848323806>
- Glanzel, W., & Schoepflin, U. (1995). A Bibliometric Study on Aging and Reception Processes of Scientific Literature. *Journal of Information Science*, 21(1), 37-53. <https://doi.org/10.1177/016555159502100104>
- Higgins, S. S., & Ribeiro, A. C. (2018). *Análise de Redes em Ciências Sociais*. ENAP. https://repositorio.enap.gov.br/bitstream/1/3337/1/Livro_Analise%20de%20Redes%20em%20Ci%C3%A2ncias%20Sociais.pdf
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569-16572. <https://doi.org/10.1073/pnas.0507655102>
- Klavans, R., & Boyack, K. W. (2007). Is there a convergent structure of science? A comparison of maps using the ISI and scopus databases. *Proceedings of ISSI 2007 – 11th International Conference of the International Society for Scientometrics and Informetrics*. 437-448.
- Newman, M. (2001). The structure of scientific collaboration networks. *PNAS*, 98(2). <https://www.pnas.org/content/98/2/404>
- Newman, M. (2003). The Structure and Function of Complex Networks. *Computer Physics Communications*. 147 (40). [https://doi.org/10.1016/S0010-4655\(02\)00201-1](https://doi.org/10.1016/S0010-4655(02)00201-1)
- Pinheiro, L. (1983). Lei de Bradford: uma reformulação conceitual. *Ciência da Informação* 12(2) p. 59-80
- Reagans, R., & McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly*, 48(2), 240-267. <https://doi.org/10.2307/3556658>
- Saha, S., Saint, S., & Christakis, D. A. (2003). Impact factor: A valid measure of journal quality? *Journal of the Medical Library Association*, 91(1), 42-46.
- Tahamtan, I., Afshar, A. S., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics*, 107(3), 1195-1225. <https://doi.org/10.1007/s11192-016-1889-2>

(Página deixada propositadamente em branco)

DIVULGAÇÃO *ONLINE* DE COLEÇÕES DE FOTOGRAFIA: A UTILIZAÇÃO DO ATOM EM PORTUGAL

Susana Sofia Cunha

Doutoranda em Ciência da Informação, Faculdade de Letras, Universidade de Coimbra,
susana.cunha@student.fl.uc.pt ORCID iD 0000-0003-4444-4319

Resumo

A fotografia ganhou lugar nos arquivos e despertou para a necessidade de tratamento arquivístico que satisfaça os utilizadores. Não existe um modelo único convencionado para descrição de fotografia, adaptando-se normas gerais de descrição arquivística ou modelos desenvolvidos por cada instituição. O AtoM, *software* livre de código aberto, a funcionar em ambiente *Web*, permite que instituições disponibilizem acervos *online*, cumprindo as normas do ICA. O objetivo geral é analisar a utilização do AtoM para divulgação *online* de coleções de fotografia em instituições portuguesas, no sentido de avaliar a adequação da plataforma para a disponibilização desta tipologia documental. A metodologia adotada consistiu numa pesquisa bibliográfica de carácter seletivo e na análise de conteúdo resultante do levantamento de instituições utilizadoras do AtoM. A análise das 12 instituições que disponibilizam *online* acervos fotográficos permitiu concluir que a utilização do AtoM é adequada à descrição e disponibilização destes documentos e pode trazer vantagens às instituições que o utilizam.

Palavras-chave: AtoM; arquivo fotográfico; descrição de fotografia; arquivo *online*

Introdução

O documento fotográfico ganhou lugar nos arquivos e despertou para a necessidade de tratamento arquivístico adequado às necessidades dos utilizadores. Pavão (1997) é pioneiro, em Portugal, no que respeita às orientações precisas para o tratamento de coleções fotográficas. O seu modelo de organização cumpre as orientações da ISAD(G) e baseia-se numa descrição multinível, partindo do geral para o particular, com a introdução de campos específicos para a descrição de fotografia, nomeadamente “imagem” (o que vemos representado, registado apenas uma vez, independentemente do suporte em que se encontra) e “espécie” (a espécie fotográfica propriamente dita, já que a mesma imagem pode ter diferentes suportes: negativo, prova, entre outros). Sánchez Vigil (2006) propõe que a análise documental da imagem fotográfica se realize sob três pontos de vista: identificação geral (dados imprescindíveis e nem sempre disponíveis, como a identificação do autor, título, data), características técnicas (suporte, processo, formato) e descrição de conteúdo

(o que é representado). A descrição deverá, sempre que possível, ser completada com indexação e, segundo Shatford Layne (1994), os termos retirados de uma imagem poderão ser classificados em quatro categorias, através das proposições “Quem” (identificação do objeto), “Onde” (localização no espaço), “Quando” (localização no tempo cronológico ou no momento do registo da imagem) e “O Quê /Como” (descrição de ações, atividades ou eventos). Estes elementos podem ter como fonte de informação a própria fotografia ou ser resultado de pesquisas feitas pelos responsáveis pela descrição.

Não existe um modelo único ou normas universais convencionadas para a organização e descrição intelectual de fotografia, em parte devido ao carácter específico desta documentação, nomeadamente a sua polissemia, subjetividade e multiplicidade de processos e suportes que a constituem. Diversas instituições utilizam normas de descrição não específicas ou modelos especialmente desenvolvidos dentro de cada instituição para fazer face aos seus requisitos internos, bem como plataformas de gestão e descrição arquivística generalistas. Como refere Flores (2016), para cumprir a legislação e, ao mesmo tempo, romper paradigmas, é fundamental a utilização de uma plataforma de descrição arquivística *online*, normalizada, com navegação multinível, que satisfaça as exigências dos utilizadores bem como as necessidades de produção, de divulgação e de acesso à informação. Tendencialmente tem-se procurado recorrer a *softwares open source*, adaptados consoante as necessidades e objetivos das instituições. Uma das soluções é o AtoM, *software* livre e de código aberto, que permite que as instituições possam ter os seus arquivos *online*, cumprindo as normas do *Internacional Council on Archives* (ICA). Funcionando em ambiente *Web* e de fácil utilização, é flexível e adaptável às necessidades de cada instituição, independentemente da sua dimensão, suporta diversos tipos de coleções e permite a implementação de vários tipos de repositórios (Bushey, 2012). A possibilidade de agregar objetos digitais à descrição, viabilizando e facilitando a descrição de recursos ou suportes como a fotografia ou o vídeo, é também uma das vantagens desta plataforma.

O objetivo geral deste estudo consiste em analisar o AtoM enquanto plataforma de divulgação *online* de coleções de fotografia em instituições portuguesas, no sentido de avaliar a adequação da plataforma para a organização e disponibilização desta tipologia documental. Para tal, é realizada uma recolha bibliográfica sobre o tema e uma análise das instituições portuguesas que disponibilizam conteúdos fotográficos através desta plataforma. Espera-se concluir que o AtoM é uma mais-valia para a organização intelectual e disponibilização *online* de documentação fotográfica.

Metodologia

Para cumprir o objetivo estabelecido baseámo-nos numa abordagem metodológica teórica e empírica. A metodologia adotada consistiu, numa primeira fase, numa pesquisa exploratória do tema da descrição arquivística através de plataformas de código aberto, especificamente o AtoM. Procedeu-se a uma recolha da literatura científica nacional e internacional, com base numa pesquisa de carácter seletivo para identificar textos em repositórios e bases de dados *online* como a *Web of Science*

(*Information Science & Library Science*), a B-on (Biblioteca do Conhecimento Online), agregadora de repositórios como a EBSCO (*Library & Information Science Source*), RCAAP (Repositório Científico de Acesso Aberto em Portugal) e OASIS (Portal brasileiro de publicações científicas em acesso aberto), e a BRAPCI (Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação). Esta pesquisa decorreu durante Novembro e Dezembro de 2020, por títulos e assuntos, em artigos em acesso aberto e texto integral. Utilizaram-se os termos “ICA-AtoM” e “AtoM” conjugado com “archiv*” como delimitadores da pesquisa. Aplicou-se como restrição temporal 2006 (data de lançamento do ICA-AtoM) e 2020 e selecionaram-se os resultados que se enquadravam no domínio da Biblioteconomia e Ciência da Informação. Analisaram-se, também, as referências bibliográficas dos artigos recuperados, no sentido de identificar quais os autores mais referenciados.

Procedeu-se, seguidamente, a uma pesquisa para identificar instituições portuguesas utilizadoras do AtoM para descrição arquivística e divulgação *online* de conteúdos (levantamento feito em Novembro e Dezembro de 2020 através da página *Web* do AtoM (*AtoM: Open Source Archival Description Software*, sem data) e da bibliografia consultada).

Definiram-se as características a analisar para cada instituição, nomeadamente, o cumprimento dos campos estabelecidos pela ISAD(G), a identificação/ separação entre documentação fotográfica e outra, identificação de pontos de acesso (taxonomias/ vocabulários controlados) e disponibilização de objetos digitais. A análise da documentação disponibilizada dentro de cada instituição baseou-se na seleção aleatória de dois fundos ou coleções diferentes (quando existentes), partindo do geral (fundo ou coleção) para o particular (item ou documento simples). As categorias de análise definidas respeitam as regras propostas por Bardin (2008) para que uma análise seja válida: são homogéneas (no sentido em que não “misturam” assuntos), exclusivas (já que um mesmo elemento do conteúdo não pode ser classificado em duas categorias diferentes) e adequadas ou pertinentes (estão adaptadas ao material de análise e refletem as intenções da investigação). A categorização fez-se segundo o processo “por caixa”: define-se um sistema de categorias e repartem-se da melhor maneira possível os elementos consoante vão sendo encontrados.

Após esta tarefa recolheram-se e analisaram-se os dados e os resultados foram tratados quantitativamente através de análises estatísticas simples (ocorrências) e qualitativamente através da sua interpretação à luz dos objetivos propostos.

Resultados e discussão

Identificaram-se 17 instituições/projetos portugueses que utilizavam, à data da recolha dos dados, este *software* para descrição arquivística e divulgação *online* dos seus acervos (cf. Tabela). Destas, sete são arquivos municipais, quatro são arquivos ou projetos ligados a universidades e duas são arquivos históricos de instituições religiosas. As restantes correspondem ao arquivo de uma central sindical, ao arquivo de uma empresa privada, ao arquivo de uma fundação e a um arquivo histórico governamental. Este levantamento permite perceber que o *software* é indicado para a descrição arquivística de diferentes tipos de arquivos e documentação e a sua utili-

zação é transversal a vários tipos de entidades (instituições e projetos de carácter público, empresas privadas, fundações, partidos políticos e arquivos religiosos, entre outros), podendo ser utilizado para descrever documentação de uma única instituição ou como multi-repositório de várias instituições.

Tabela 1 – Instituições/projetos portugueses que disponibilizam conteúdos online através do AtoM

Instituição ou projeto	URL	Documentação o fotográfica
Arquivo Municipal de Albergaria a Velha	http://arquivo.cm-albergaria.pt	sim
Arquivo Municipal de Évora/ Arquivo Fotográfico	http://arqm.cm-evora.pt/	sim
Arquivo Municipal de Oliveira de Azeméis	https://arquivodigital.cm-oaz.pt	sim
Arquivo Municipal de Reguengos de Monsaraz	http://arquivocmrm.cm-reguengos-monsaraz.pt/	sim
Arquivo Municipal de Tavira	https://arquivo.cm-tavira.pt/	não
Arquivo Municipal de Vale de Cambra	http://arquivo.cm-valedecambra.pt	sim
Câmara Municipal de Arruda dos Vinhos	https://arquivo.arrudadosvinhos.com.pt/	sim
Arquivo Histórico da Igreja Lusitana	http://arquivo.igreja-lusitana.org	sim
Arquivo histórico da Santa Casa da Misericórdia do Porto	https://atom.scmp.pt/	não
Centro de Arquivo e Documentação CGTP-IN	http://cad.cgtp.pt/ica/index.php/	sim
Arquivo Histórico Fábrica Maceira – Liz	https://arquivo.cdi-maceiraliz.pt/	sim
Arquivo Fotográfico Teófilo Rego (Museu Casa da Imagem – Fundação Manuel Leão)	https://arquivofotografico.fmleao.pt/	sim
Arquivo Histórico do Ministério Público	https://arquivohistorico.ministeriopublico.pt/	não
Arquivo de História Social	http://ahsocial.ics.ulisboa.pt/atom	sim
INVENT.ARQ Inventários de arquivos de família, sécs. XV-XIX	https://inventarq.fcsh.unl.pt/index.php/	não
PAPIR – Plataforma de Arquivos Pessoais e de Instituições Religiosas	http://portal.cehr.ft.lisboa.ucp.pt/arquivos/index.php/	não
Universidade do Porto – Fundação Instituto Arquitecto José Marques da Silva	http://arquivoatom.up.pt	sim

Fonte: elaboração da autora, com base na consulta das fontes mencionadas

Selecionaram-se para análise as 12 instituições que descrevem e disponibilizam *online* coleções e/ou documentos fotográficos: seis arquivos municipais (Albergaria-a-Velha, Arruda dos Vinhos, Évora, Oliveira de Azeméis, Reguengos de Monsaraz e Vale de Cambra; três arquivos ligados a universidades e projetos de investigação, dois deles no âmbito na fotografia de arquitetura (Arquivo Fotográfico Teófilo Rego e Fundação Instituto Arquitecto José Marques da Silva) e outro no âmbito da história social, relacionado com o estudo dos movimentos operários, sindicais, estudantis, feministas, entre outros (Arquivo de História Social); um arquivo religioso (Arquivo Histórico da Igreja Lusitana); um arquivo de uma central sindical (Centro de Arquivo e Documentação CGTP-IN) e um arquivo que documenta a história de uma empresa privada (Arquivo Histórico Fábrica Maceira-Liz).

Verifica-se que a maioria das instituições analisadas incorpora os documentos fotográficos nos seus fundos ou coleções, descrevendo-os a par da outra documentação. Apenas quatro são dedicadas exclusivamente à gestão e descrição de documentação fotográfica: Arquivo Teófilo Rego (ATR), Câmara Municipal de Arruda dos Vinhos (CMAV), Arquivo Municipal de Vale de Cambra (AMVC) e Arquivo Fotográfico da Câmara Municipal de Évora (AFCME)¹.

A análise do *software* foi feita do ponto de vista do *front office*, isto é, do utilizador ou investigador. A aplicação é de utilização “amigável”, apresentando uma boa funcionalidade dos menus e qualidade na visualização dos objetos digitais. Existem várias opções de pesquisa (“pesquisa simples” e “pesquisa avançada”, com filtros) e todos os campos de descrição são pesquisáveis, característica que se revela de especial utilidade quando se pesquisam documentos fotográficos, já que a legenda ou título nem sempre é suficiente para descrever a imagem.

O AtoM está estruturado de acordo com a ISAD(G), definindo sete zonas de informação descritiva: “identificação”, “contexto”, “conteúdo e estrutura”, “condições de acesso e de utilização”, “documentação associada”, “notas” e “controlo da descrição”. Estas zonas abrangem 26 elementos, sendo que apenas seis são essenciais: “código de referência”, “título”, “nível de descrição”, “data”, “dimensão e suporte” e “produtor” (Tojal et al., 2002). De uma maneira geral, as instituições analisadas cumprem estes seis requisitos de descrição e respeitam a descrição multinível.

Observa-se que é possível organizar e descrever a documentação fotográfica utilizando os mesmos níveis de descrição utilizados para a restante documentação, do geral para o particular: “fundo” ou “coleção”, “série”, “documento simples” e “item”. O nível “série” é utilizado por dez instituições para descrever unidades de instalação que cobrem aspetos de uma mesma função, atividade ou assunto dentro de um fundo (Rosseau & Couture, 1994). Em descrição de fotografia pode consistir num conjunto de imagens que fazem sentido entre si, como por exemplo uma reportagem feita por um fotógrafo em determinada data, ou um conjunto de imagens agrupados mediante características definidas pelo autor, colecionador ou pela própria instituição de arquivo. As menores unidades documentais encontradas (“documento simples”, em oito instituições e “item”, em três) referem-se à espécie fotográfica propriamente dita, isto é, o negativo ou prova (analógica ou digital) descrita e disponibilizada. Uma descrição a este nível é, para muitas instituições, uma tarefa quase impossível, dado o número de espécies que detêm, optando as instituições, por isso, por descrever apenas ao nível da “coleção” ou “série”, dando assim um aspeto geral da mesma para que esta seja útil aos investigadores, e ir fazendo a descrição à medida das necessidades e disponibilidade dos técnicos. Esta tarefa pode ser potenciada pela possibilidade de importação e exportação de registos em vários formatos (EAD, EAC-CPF, CSV e SKOS), característica que se revela de extrema importância sobretudo para instituições que tenham os dados registados em tabelas ou outras

¹ O Arquivo Municipal de Évora (AME) optou por juntar na mesma plataforma o Arquivo Municipal (Arquivo Histórico e Intermédio) e o Arquivo Fotográfico. As descrições são independentes (embora possam ser relacionadas) e geridas por duas equipas diferentes, pelo que o AFCME foi considerado, para efeitos deste estudo, como instituição exclusivamente dedicada à documentação fotográfica.

bases de dados. Através da importação ou migração dos dados as instituições salvaguardam o trabalho realizado anteriormente, necessitando apenas, e em alguns casos, de o corrigir ou completar.

Todas as instituições analisadas atribuem título às descrições, em todos os níveis, do geral para o particular. Os títulos identificam objetivamente o representado na imagem e o(s) conjunto(s) em que ela se insere (“fundo”, “coleção” e “série”). Identificaram-se algumas descrições sem data, residuais. Como foi referido anteriormente, esta informação nem sempre está diretamente associada à imagem, pelo que as datas encontradas são, muitas vezes, inferidas a partir da análise da própria imagem e descritas de uma forma mais genérica (como por exemplo, “década de...”) ou disponibilizadas apenas nos níveis superiores.

No campo “dimensão e suporte”, na maioria das instituições descrevem-se os conjuntos documentais totais nos níveis superiores e individuais nos inferiores. Este campo permite o registo de informação relacionada com os processos fotográficos, suportes e formatos das espécies, fundamental para uma correta descrição e análise desta tipologia documental.

Na “zona do contexto”, todas as instituições preenchem o campo “produtor” e “entidade detentora”, levando à criação de registos de autoridade segundo as normas ISAAR(CPF). É no campo “produtor” que se regista o fotógrafo ou estúdio fotográfico, quando conhecido, e a entidade singular ou coletiva responsável pela acumulação da documentação.

As restantes zonas não são de preenchimento obrigatório e algumas instituições optam por não o fazer, mas tratando-se de descrição de fotografia o seu preenchimento é aconselhado, com vista a uma correta disponibilização e recuperação da informação. A “zona do conteúdo e estrutura”, onde se registam informações acerca do conjunto documental ou do documento isolado (e que permite uma descrição pormenorizada que vai para além do título, fundamental para a contextualização da imagem), sistema de organização ou estado de conservação foi preenchida, sobretudo, ao nível do “fundo” ou “coleção”, oferecendo uma panorâmica geral da documentação. Isto parece indicar uma tendência para evitar a redundância, já que a informação é disponibilizada nos níveis hierárquicos superiores. No entanto, nem sempre uma abordagem geral reflete a especificidade de cada documento fotográfico, que é único no tempo. Quando a análise, descrição e contextualização é feita individualmente obtém-se informação específica que contribui para um maior conhecimento do todo. A mesma tendência se verifica na “zona das condições de acesso e utilização”, cujo preenchimento é feito, sobretudo, ao nível superior.

É na “zona das notas” que se regista informação que não possa ser incluída em qualquer outra zona, como anotações, dedicatórias ou inscrições, fontes preciosas como auxiliares de informação para a identificação e contextualização de imagens, bem como identificadores alternativos, importantes para a manutenção da regra da organização original. Esta zona foi preenchida apenas por seis instituições e sobretudo ao nível do documento ou item. A “zona da documentação associada” parece ser a menos utilizada, com apenas três instituições a preencher os campos nos três níveis de descrição, embora esta zona seja fundamental para o relacionamento entre imagens, série ou instituição arquivística, bem como com outros documentos de arquivo ou publicações onde a imagem esteja representada.

O AtoM permite associar taxonomias na zona “pontos de acesso”. Identificaram-se três categorias para a definição de pontos de acesso, designadamente “Locais”, “Assuntos” e “Nomes”. Estas três categorias vão ao encontro da opinião de Sánchez Vigil (2006), que propõe que aquando da indexação de imagens se utilizem descritores onomásticos, temáticos, geográficos e temporais. Da mesma forma, estes campos podem ser utilizados para responder às quatro categorias de indexação propostas por Shatford Layne (1994), anteriormente indicadas, e verifica-se que a maioria das instituições lhes confere alguma importância. A categoria mais utilizada é a que se refere a “Locais”, identificando nove instituições que atribuem estes descritores nos três níveis de descrição. A categoria “Assuntos” é utilizada por sete instituições. Nos casos em que são utilizados, os pontos de acesso relativos a “Locais” e “Assuntos” podem ser organizados segundo uma hierarquia ou constituírem uma simples lista de termos, dependendo dos objetivos e práticas da instituição. A preocupação com a organização hierárquica verifica-se sobretudo ao nível dos descritores “Locais” (cinco instituições organizam hierarquicamente estes descritores definindo, por exemplo, uma cadeia que parte do país até à freguesia ou rua), enquanto que apenas duas demonstram esta preocupação ao nível do “Assunto”. De salientar que o AtoM permite a introdução de termos nestes campos durante o preenchimento, pelo que não é necessária a existência de uma lista pré-determinada para o fazer. À semelhança do que acontece com o campo “Produtor” os descritores utilizados no campo “Nomes” dão origem a um registo de autoridade segundo as normas ISAAR(CPF) e são usados para identificar a(s) pessoa(s) fotografada(s) e/ou o fotógrafo.

Os objetos digitais disponibilizados são, sobretudo, espécies fotográficas digitalizadas, revelando uma preocupação em descrever e divulgar imagens mais antigas e cuja preservação dos originais se torna premente.

Conclusão

O recurso a softwares *open source* para divulgação de conteúdos *online* por parte de instituições arquivísticas públicas e privadas vem ganhando terreno face aos softwares proprietários, destacando-se as vantagens a nível orçamental e facilidade de instalação e personalização, e o AtoM surge como uma das aplicações utilizadas para a descrição de acervos fotográficos. Não existindo um modelo único para descrição de fotografia, as instituições adaptam as normas em vigor, como a ISAD(G), para descrever estes documentos. As zonas de descrição disponíveis no AtoM são, assim, adequadas para a descrição de fotografia, já que proporcionam uma descrição multinível e permitem o registo dos dados essenciais para uma correta organização e disponibilização deste tipo de documentação. Para além dos elementos essenciais de identificação e descrição, os campos disponíveis permitem adicionar informações que enriquecem a descrição e análise das imagens, como sejam dedicatórias ou inscrições e outras informações que resultem de pesquisas feitas *a posteriori*. Igualmente, a possibilidade de incorporar “pontos de acesso”, isto é, um vocabulário controlado organizado sob três categorias principais, sistematiza e permite a recuperação da informação e satisfação das necessidades dos utilizadores, nomeadamente ao nível da otimização da obtenção dos resultados que procuram.

Estes pontos de acesso, organizados hierarquicamente ou simples listas de termos que podem ser inseridos livremente durante o processo de descrição, deverão ser alvo de uniformização, de forma a manter a coerência em relação a todos os descritores e disponibilizar uma lista de termos que seja útil ao utilizador, significando uma tarefa extra para o gestor ou administrador da base de dados. É de salientar a pertinência, para a descrição de documentos fotográficos, do ponto de acesso “Nomes” assim como do campo “Produtor”: como foi referido, o preenchimento destes campos gera um registo de autoridade segundo a norma ISAAR(CPF), permitindo elaborar e gerir uma lista de fotógrafos e/ou estúdios fotográficos, relacionando-os entre si através de relações familiares, sociais e profissionais. Isto facilita a criação de uma rede que poderá ser útil para o estudo da história da fotografia nacional e internacional, já que muitos dos fotógrafos eram, também, viajantes e registavam momentos dos países por onde passavam.

Do ponto de vista do utilizador, o AtoM é de utilização “amigável” e proporciona várias opções de pesquisa, facilitando a rápida recuperação da informação. A implementação do programa nas diferentes instituições permite uma consulta autónoma por parte do utilizador, *online*, libertando-o da obrigação da consulta presencial.

A opção pela divulgação de conteúdos informacionais para além dos limites da instituição, disponibilizando-as na *Web*, deve ter em conta a autenticação, identificação e a preservação da integridade dos registos. A representação digital da informação remete-nos para questões éticas, legais e ainda as difíceis questões da propriedade intelectual, da proteção da confidencialidade e da privacidade (direito à imagem). Ainda assim, a digitalização de acervos fotográficos e sua consequente disponibilização vão ao encontro das expectativas do público e é natural que a implementação destes sistemas potenciadores da visualização das fotografias depositadas em instituições traga efeitos significativos à própria coleção e ao seu uso (Silva, 2006), resultando numa mais valia tanto para o utilizador como para a instituição. Por outro lado, o acesso aos documentos custodiados pelas instituições deixa de ser possível apenas a investigadores e passa a ser aberto a qualquer cidadão, de forma rápida, em qualquer lugar e a qualquer momento, desde que exista uma ligação à internet.

Referências

- AtoM: Open Source Archival Description Software*. (sem data). <https://www.accessmemory.org/pt/>
- Bushey, J. (2012). ICA-AtoM: Open-source software for archival description. *Archivi & Computer*, 22(1), 10-25.
- Flores, D. (2016). A adoção do AtoM (ICA-AtoM) para descrição, difusão e acesso de documentos arquivísticos e as perspectivas de preservação digital e autenticidade com sua interconexão aos repositórios arquivísticos digitais confiáveis—RDC-Arq. *AtoM: Work in Progress*, 28-37.
- Pavão, L. (1997). *Conservação de Coleções de Fotografia*. Dinalivro.
- Rosseau, J., & Couture, C. (1994). *Os fundamentos da disciplina arquivística*. Publicações D. Quixote.

- Sánchez Vigil, J. M. (2006). *El documento fotográfico. História, usos, aplicaciones*. Ediciones Trea.
- Shatford Layne, S. (1994). Some issues in the indexing of images. *Journal of the American Society for Information Science*, 45(8), 583–588. [https://doi.org/10.1002/\(SICI\)1097-4571\(199409\)45:8<583::AID-ASI13>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1097-4571(199409)45:8<583::AID-ASI13>3.0.CO;2-N)
- Silva, R. (2006). Acervos fotográficos públicos: Uma introdução sobre digitalização no contexto político da disseminação de conteúdos. *Ciência da Informação*, 35(3), 194-200. <https://doi.org/10.1590/S0100-19652006000300018>
- Tojal, A. A., Frazão, A., Henriques, C., Sousa, J. B., Vieira, J., Borges, L. C., Lopes, L. D., Runa, L., Garcia, M., Farinha, M. do C. D., Fevereiro, M. I., & Azevedo, R. B. (2002). *ISAD(G): Norma geral internacional de descrição arquivística (tradução do Grupo de Trabalho para a Normalização da Descrição em Arquivo)*.

(Página deixada propositadamente em branco)

O IMPACTO DA INTELIGÊNCIA ARTIFICIAL NOS SERVIÇOS DE INFORMAÇÃO: INOVAÇÃO E PERSPETIVAS PARA AS BIBLIOTECAS

Liliana Isabel Esteves Gomes¹, Viviana Fernández Marcial²,
Miguel Nuno Marques dos Santos³

¹Faculdade de Letras da Universidade de Coimbra, liliana.gomes@fl.uc.pt,
ORCID iD 0000-0003-3786-2942

²Universidade da Coruña, viviana.fernandez@udc.es, ORCID iD 0000-0002-9277-266X

³Símbolo de Memória, miguelnuno@simbolodememoria.com,
ORCID iD 0000-0003-1136-2229

Resumo

Esta comunicação tem como objetivo analisar e explorar o impacto atual da Inteligência Artificial (IA) em serviços de informação, em particular nas bibliotecas. O propósito final é identificar as áreas em que se aplica e como esta tecnologia incide na inovação das bibliotecas e seus serviços. O quadro teórico-metodológico desta investigação, na área científica da Ciência da Informação, compreende a revisão da literatura, pelo que se elege uma metodologia qualitativa.

O impacto da IA é de primeira ordem nas tarefas relacionadas com o processamento da informação e com os serviços de pesquisa e referência. De modo crescente, o seu impacto abrange também outras áreas do trabalho técnico e reflete-se nos desafios profissionais de bibliotecários/profissionais da informação no presente e futuro. Conclui-se que a IA em bibliotecas e outros serviços de informação é uma realidade, que será incrementada nos próximos anos.

Palavras-chave: *Inteligência Artificial, Serviços de Informação, Bibliotecas, Ciência da Informação.*

Introdução

No mundo pós-industrial da internet é crescente o número de indivíduos que utilizam motores de pesquisa, apesar de frequentemente ficarem frustrados com o volume e a irrelevância dos resultados, pois não são a mesma coisa que o conhecimento que realmente desejam. Num âmbito geral, “compreender o uso das tecnologias na gestão e na organização da informação em suas dimensões científica, tecnológica, industrial, mercadológica, estratégica e social é fator fundamental em um cenário onde o Universo Digital se apresenta como relevante desafio na Sociedade em Rede” (Paletta & Silva, 2020, p. 164).

A Inteligência Artificial (IA) é uma inovação, com referências no âmbito dos sistemas e serviços de informação (Gomes & Fernández Marcial, 2019), desde a década dos oitenta do século passado.

Um texto de incontornável referência intitula-se “*Toward the library of the future*”, em que o seu autor projeta uma biblioteca com serviços avançados baseados em IA, em particular “*expert systems*”, e convida todos a visionar a biblioteca como “*an active inteligente knowledge server*” (Feigenbaum, 1989, p. 122).

A IA, tem diversas definições, podendo ser identificada como a inteligência demonstrada por uma entidade artificial (máquina/computador), que geralmente segue ou imita as características e analogias de pensamento da inteligência humana e realizar ações como os seres humanos, mediante a combinação de grandes quantidades de dados, algoritmos inteligentes e processamento rápido.

Desde a década de cinquenta, nomes como os de Alan Turing e John McCarthy são menções incontornáveis. Trata-se de um ramo da ciência interligado à Computação, às Neurociências, às Ciências Cognitivas, à Matemática, à Filosofia e à Psicologia, entre outras áreas do conhecimento, que tem adquirido um peso extraordinário em diversos setores da economia e na sociedade. “*Advances in artificial intelligence (AI) will transform modern life by reshaping transportation, health, science, finance, and the military*” (Grace, et al., 2018).

Nos últimos seis anos: o número de publicações relacionadas com IA no arXiv cresceu de 5.478 em 2015 para 34.736 em 2020; o número de publicações em periódicos de IA aumentou 34,5% de 2019 a 2020; em 2020, e pela primeira vez, a China ultrapassou os Estados Unidos na proporção de citações de periódicos de IA. A pesquisa do “*AI Index*” realizada em 2020 sugere que as melhores universidades do mundo aumentaram o seu investimento na educação em IA nos últimos quatro anos (Zhang, et al., 2021).

Apesar disto, é uma tecnologia de que muitos gestores de serviços de informação não têm plena consciência do seu impacto. Contudo, “*los trabajos tradicionales de bibliotecas y archivos se están también automatizando y telematizando a gran velocidad, de tal manera que las nuevas bibliotecas digitales funcionan cada vez con menos concurso de trabajo humano*” (García-Marco, 2013, p. 492).

Os denominados sistemas inteligentes lançaram uma luz de otimismo e inovação em processos técnicos como a catalogação e a indexação e projetaram-se sobre os serviços bibliotecários, tais como o serviço de referência. Pese embora este interesse, na área do trabalho biblioteconómico a IA ficou esquecida com a chegada de outras tecnologias, destacando-se a *World Wide Web* (WWW).

Neste contexto, a questão colocada por Bohyun Kim, diretora de tecnologia nas bibliotecas da Universidade de Rhode Island, em Kingston (Estados Unidos da América), mantém-se atual: “*Are libraries ready for intelligent machines that directly interact with human users? There certainly are opportunities for improving existing services and programs. But there are challenges as well*” (Eberhart, 2018) e riscos que advém de tecnologias como a IA, nomeadamente a tomada de decisões opaca, a discriminação baseada no género ou raça, a intrusão na vida privada ou o seu uso para fins criminosos (European Commission, 2020).

Na segunda década do século XXI, a IA volta novamente a ser aplicada no contexto das bibliotecas e de outros sistemas de informação. Agora, não com uma visão

prospetiva, mas sim como uma realidade. Com especial desenvolvimento por parte das bibliotecas americanas, o seu alcance não tardou a chegar a outras áreas geográficas.

Esta investigação tem como objetivo analisar e explorar o impacto atual da IA em serviços de informação, em particular nas bibliotecas, através da revisão da literatura. Optou-se por uma investigação de tipo exploratória e descritiva com abordagem de análise qualitativa. O propósito final é identificar as áreas em que se aplica e como esta tecnologia incide na inovação e competitividade das bibliotecas e seus serviços.

Metodologia

Esta investigação é de caráter teórico, mediante pesquisa e análise da literatura científica, de modo a responder às questões de investigação (QI): QI1: Qual é o impacto atual da IA nos serviços de informação? QI2: De que modo esta tecnologia incide na inovação das bibliotecas e seus serviços?

O quadro teórico-metodológico circunscreve esta investigação na área científica da Ciência da Informação (CI), trans e interdisciplinar (Gomes, 2020).

Considerou-se a recomendação de Griffey (2019, p. 6) na “modern library and information science, (...) using artificial intelligence as the very broad category and sticking with machine learning for referring to specific systems.” Bem como a seguinte relação entre os termos e a IA: “Machine learning produced AI programs whose performance is close to or even surpasses that of humans. (...) Deep learning utilizes an artificial neural network with (...) impressive results in many fields, such as computer vision, facial and speech recognition, natural language processing, machine translation, and customized recommendations” (Kim, 2019, p. 16).

A relevância das perguntas de investigação resulta da observação das mudanças tecnológicas dos últimos anos e, em particular, do impacto da IA na sociedade. Adicionalmente, os serviços de informação e, em particular as bibliotecas, parecem ainda não ter plena consciência da aplicabilidade, consequências e perspectivas desta tecnologia na prestação de serviços e gestão de recursos (Arlitsch, 2017).

A metodologia qualitativa eleita neste estudo compreende a revisão da literatura (Kitchenham, 2004) mediante pesquisa geral e, pesquisa específica, nas seguintes fontes: *EBSCO Discovery Service*, *Web of Science*, *Scopus* e *Library & Information Science Source*. A utilização de operadores booleanos permitiu concretizar a pesquisa em português, inglês e espanhol, para o período cronológico 2010-2020, com as palavras-chave: “Inteligência Artificial” AND “Ciência da Informação”; “Inteligência Artificial” OR “Machine Learning” OR “Deep Learning” AND “Bibliotecas”; “Inteligência Artificial” AND “Serviços de informação”.

Esta metodologia permitiu identificar os autores que publicaram sobre o tema em estudo e a relação das publicações ao longo dos anos. Para organização dos resultados da pesquisa utilizou-se a ferramenta *Mendeley*. Uma vez identificada a produção científica, procedeu-se à análise dos artigos com os aspetos mais relevantes para o tema em estudo (Stapleton, Carter, & Bredahl, 2020), ou seja, na área científica da CI social e aplicabilidade em bibliotecas, através da leitura inicial dos resumos e, posteriormente, do conteúdo integral, bem da exploração de bibliografia citada.

Resultados

Globalmente, as pesquisas na área de IA agrupam-se em torno de diversos tópicos.

China is very focused on the applied side of AI, with publications targeted towards the topics of computer vision, neural networks, planning and decision making, fuzzy systems, natural language processing, and knowledge representation. Europe is the most diverse, and the USA is very strong in the corporate sector, with a lot of research coming out of IBM and Microsoft (Lawlor, 2020).

No que respeita ao estudo da temática central desta investigação, esta tem-se intensificado ao longo dos últimos dez anos, por diversos autores, nomeadamente norte-americanos, mas também e de modo crescente, de outros países: Inglaterra, China, Canadá, Índia, Alemanha, Itália, França, Holanda e Brasil.

Bibliotecas e bibliotecários têm uma longa história de integração de novas tecnologias nos seus espaços e na prática profissional (ACRL, 2020). “Libraries express increased interest in the use of algorithmic methods. The reasons are many and include, but are not limited to, creating efficiencies in collection description, discovery, and access; freeing up staff time to meet evolving demands; and improving the overall library user experience” (Padilla, 2019, p. 9).

Na atualidade temos a IA a melhorar os serviços de muitas bibliotecas, nomeadamente:

- Abstracting and indexing services are much faster and more accurate.
- In information discovery and retrieval, a second discovery layer offers a wider range of answers to user queries. New products include *Quartolio*, a platform that plots connections across multiple research points, and *Yewno*, which finds connections the researcher may not have suspected.
- Feature detection and content extraction, in which a neural network extracts data and predicts outcomes effectively, such as a project at the University of Rochester that conducted space assessment to support decision-making for operations, design, and service delivery.
- Voice-user interface and chatbots, such as providing reference service through Amazon’s Alexa virtual assistant (...) (Kim, cit. por Eberhart, 2018).

A análise da produção científica que especifica a aplicabilidade de IA a serviços inovadores em bibliotecas permite identificar a necessidade de explicar e compreender o contributo desta tecnologia e de outras correlacionas nos referidos serviços, através de diversas iniciativas e projetos de investigação (Otterlo, 2016) neste campo. Concomitantemente destaca-se a sua aplicabilidade (Garcia-Febo, 2019) na otimização e/ou resolução de problemas ou procura de novas soluções ao nível do processamento, tratamento, armazenamento, organização e disponibilização de informação.

Também se assinala a discussão acerca de questões éticas e desafios para a atividade profissional e para os serviços tradicionais vs. inovadores em bibliotecas,

nomeadamente no ensino superior. De acordo com Wheatley e Hervieux (2019, p. 347), importa perceber que papel os bibliotecários vão desempenhar “in an AI-dominant future, as well as how libraries are responding to this change.”

Processamento de informação

De acordo com pesquisas recentes, os principais problemas dos utilizadores numa biblioteca física são a navegação e a localização de livros e tópicos/assuntos, apesar dos muitos avanços nos catálogos e no planeamento dos espaços. Além disso, as bibliotecas e os livros físicos exigem navegar fisicamente, os utilizadores deslocam-se à biblioteca e consultam os recursos impressos. Quem trabalha na biblioteca desempenha tarefas físicas como o transporte, a colocação de itens nas estantes e depósitos, a realocação após consulta, entre outras. Uma grande diferença para a pesquisa e consulta na internet, no geral, ou de coleções digitais, em particular.

Neste âmbito, a IA tem já diversas aplicações. No tratamento digital da informação, utiliza-se nomeadamente para solucionar problemas de processamento manual ou automatizado, mas também como apoio ou substituição daquele processamento mediante programação algorítmica convencional (Martins, 2010).

No que respeita à troca de informação entre diferentes sistemas, assinala-se a investigação que permitiu o desenvolvimento de um modelo genérico de ontologia, com potencial aplicação na indexação em bibliotecas, “this model will be a metadata for a knowledge base system to be used in different purposes of interest such as education applications (...) and enhancing the index tool in libraries to facilitate access to information collections” (Sawsaa, 2013).

Craig Boman, bibliotecário na Universidade de Miami, realiza pesquisa a partir de metadados de registos MARC e uso de *Machine Learning*, para criar um sistema automatizado de atribuição de cabeçalhos de assuntos formais/pontos de acesso. “The opportunities associated with new machine learning systems to reform large portions of library activities will be rich and varied” (Griffey, 2019).

Catherine Coleman, “digital research architect” nas Bibliotecas de Stanford, Califórnia, desenvolveu um relevante trabalho com “bibliographers, archivists, and catalogers to explore the possibilities of AI for metadata and collection development” (Garcia-Febo, 2019). O seu objetivo compreende a identificação e desenvolvimento de aplicações da IA “that will help us (the university) make our rich collections of maps, photographs, manuscripts, data sets and other assets more easily discoverable, accessible, and analyzable for scholars”. Conversas entre os profissionais da biblioteca e discussão acerca da IA são também desenvolvidas e a IA é aplicada em projetos de modo a tornar as coleções “more discoverable and analyzable. The [IA] studio is staffed by volunteer librarians and the library provides access to Yewno, discovery tool that provides a graphical display of the interrelationships between concepts” (Lawlor, 2020).

O crescimento e a crescente conectividade proporcionada pela Internet das Coisas (Aqarone & Luz, 2017), que beneficia do desenvolvimento de outras tecnologias como o *blockchain* e a IA, permite também identificar a aplicabilidade daquelas tecnologias em bibliotecas e o desenvolvimento de projetos, por exemplo em Itália,

nos Estados Unidos da América ou na Austrália, em “traditional library activities for any kind of library (cataloguing, collection development, information literacy) and in scholarly communication and academic libraries (scholarly publishing, peer review, assessment and evaluation) (Morrielo, 2019).

A nível editorial, no geral, e de publicação académica em particular, a IA e o *Machine Learning* são utilizados para melhorar os serviços de descoberta, a experiência geral de pesquisa dos utilizadores e a gestão de submissões de trabalhos e revisão por pares. “Elsevier is the first to characterize the field of AI in a comprehensive, structured manner using extensive datasets from their own and public sources”. Editores e editoras necessitam mudar para permanecerem relevantes no contexto digital atual, pois é tão importante a pesquisa e o acesso a conteúdos como a sua qualidade, e “academic publishers can deliver better user experiences by developing a product mindset underpinned by data “maturity” and new software paradigms, (...) defined by the publisher’s ability to store, manage, create, and use data to deliver value to their users” (Lawlor, 2020). A *ScholarOne Manuscripts*, da *Clarivate Analytics*, desenvolveu um projeto piloto que usa IA para ajudar os autores a enviar manuscritos de alta qualidade e também acelera o processo de submissão, revisão por pares e de avaliação editorial.

No que respeita à criação e gestão de coleções de imagens (gravuras, desenhos, fotografias, mapas), nomeadamente em bibliotecas digitais com valor histórico/patrimonial, identificam-se três componentes a considerar em resultado da aplicação da aprendizagem profunda ou *Deep Learning*: “1) identify and extract the iconography wherever it is found, either in images or in the printed documents, 2) transform, harmonise and enrich the descriptive metadata with the help of artificial intelligence, and 3) incorporate this data into a web application dedicated to iconographic research (Moreux, 2019, p. 5).

Não sendo recente a discussão sobre a aplicação da IA na área do património cultural e organizações sem fins lucrativos, esta tornou-se mais urgente na última década porque os custos do hardware necessário caíram drasticamente, enquanto as ferramentas de software tornaram-se mais prontamente disponíveis (Cordel, 2020). No campo específico das bibliotecas digitais, desde a década de noventa do século XX que são listadas as suas potencialidades:

Machine learning, together with intelligent object-centered techniques, can offer a valuable support when building intelligent digital libraries. Indeed, all the tasks related to information capture and semantic indexing can take advantage of the use of intelligent techniques and machine learning methods for layout analysis, document classification and understanding, while the integration of worldwide distributed digital libraries demands the definition of a standard query language for information retrieval. Moreover, machine learning techniques allow to infer user models from user interactions and this turns out to be useful to implement an adaptive interface” (Esposito et al., 1998, p. 18).

Identificam-se, também, propostas para a criação de bases de dados digitais, sem intervenção humana, que contenham informação sobre todos os livros de uma biblioteca, através do *Deep Learning*, nomeadamente mediante digitalização, Reco-

nhecimento Ótico de Carateres e Processamento de Linguagem Natural (Muley, et al., 2019).

Os sistemas que utilizam a IA podem ser treinados para reconhecer locais e pessoas em fotografias, algo extremamente útil para a organização e descrição da informação, bem como para a tornar visível na WWW e, ainda, solucionar o problema de acesso e fruição massiva da herança cultural. “Libraries can watch these systems as they develop, work with vendors, and create their own services and systems so that our values and ethics are baked into the technology at the outset” (Griffey, 2019).

Organização e gestão de serviços

De acordo com o relatório da “Society of College, National & University Libraries” (Pinfield, Cox & Rutter, 2017), quando questionados, os gestores de bibliotecas universitárias/acadêmicas estão bem cientes da necessidade de oferecer novos serviços, nomeadamente mediante a adoção de tecnologias avançadas, como impressão 3D, gamificação e reconhecimento de voz. Contudo, lidar com as dificuldades de adoção da IA é uma etapa necessária na transição das bibliotecas, no geral, de modo a terem relevância, assumindo novas funções e serviços, e evitar que se tornem obsoletas.

Embora diversas associações tenham começado a estimular o envolvimento das bibliotecas na IA, a aceitação tem sido lenta. Os principais exemplos de envolvimento em projetos ou laboratórios de IA são de bibliotecas universitárias/acadêmicas americanas, nomeadamente de Stanford, do MIT, da Universidade de Oklahoma (Oklahoma’s Digital Skills Hub), e da Universidade de Rhode Island (Wheatley & Hervieux, 2019, p. 354).

Contudo, a IA vem transformar muitas áreas do trabalho nos serviços e operações das bibliotecas. Tal como refere Kim (2019, p. 17):

(1) (...) the AI-powered digital assistant mediating a library user’s information search, discovery, and retrieval process, directly interacting with library systems and applications. (2) Many tasks in cataloging, abstracting, and indexing that are currently performed by skilled professionals may be automated by AI applications as they become more sophisticated. (3) A chatbot may take up part of the library’s reference or readers’ advisory service. (4) AI applications may be used to extract key information from a large number of documents or even information-rich visual materials such as maps and generate a summary to facilitate research. Libraries will need to keep a close eye on the developments in AI and carefully follow how it may affect the way people’s information-seeking, learning, and teaching activities, as well as the library’s traditional services and operation, are currently conducted.”

No âmbito de uma visão geral da IA conversacional, a mudança de um serviço de referência presencial para a esfera virtual, como se verificou em face da atual situação pandémica Covid-19, beneficia da implementação de “Bots” ou “Chatbots” para dar resposta às necessidades dos utilizadores 24 horas / 7 dias por semana, sem intervenção humana. A título exemplificativo, um “Chatbot allows librarian to create

multilingual virtual assistants that speak to users in their native tongue, eliminating language barriers and making the localization process easier, more effective, and more affordable” (Panda & Chakravarty, 2021, p. 130).

Partindo da discussão acerca das tecnologias emergentes que, integradas, formam a nova geração de bibliotecas, as “Smart Libraries”, pretende-se colmatar as falhas decorrentes da prestação de serviços e acompanhar as necessidades dos utilizadores em constante mudança (Gul & Bano, 2019). Como resultado, a futura biblioteca será um espaço de biblioteca inteligente, sistema de recursos inteligente e atendimento ao cliente inteligente.

A IA pode também melhorar o nível do serviço de uma biblioteca digital, de acordo com um dos estudos desenvolvido em bibliotecas digitais chinesas (Li, et al., 2019).

No que respeita à interação entre aspetos físicos e digitais das bibliotecas públicas, a utilização de sensores e a tecnologia de smartphones auxilia os bibliotecários e utilizadores em várias atividades. A inovação baseada em dados é o resultado do projecto holandês BLIIPS – “Books and Libraries: Intelligence and Interaction through Puzzle and Skinnerboxes”, para que as bibliotecas públicas possam ser “more intelligent by data-driven optimization”. Esta abordagem compreendeu quatro elementos essenciais: “data gathering in physical spaces using sensors, retail approaches in current public libraries, intelligent algorithms for data analysis and experimentation with data-driven strategies” (Otterlo, 2016, p. 287). O seu objetivo principal é registar como os utilizadores interagem com a biblioteca física e a sua coleção, para os ajudar a navegar até livros “possivelmente interessantes” e para auxiliar os gestores da biblioteca na ordenação física das coleções de modo mais flexível.

Ferramentas cada vez mais disponibilizadas por bibliotecas, beneficiam da IA, nomeadamente impressão 3D, realidade virtual e robótica (Ylipulli & Luusua, 2019).

No Instituto de Tecnologia de Massachusetts (MIT), Chris Bourg, diretor das bibliotecas, refere o trabalho com investigadores de IA/ML na universidade, para analisar tarefas e fluxos de trabalho da biblioteca que podem ser aprimorados pela IA, pois “it is important for academic libraries to make their collections accessible to AI tools like Alexa so that when someone asks a voice assistant for information, reputable scholarly literature is available” (Garcia-Febo, 2019).

A Cambridge Public Library, em Massachusetts (EUA), instalou em parceria com as *MIT Libraries* e o *Harvard metaLAB*, um “Laughing Room”, uma sala artificialmente inteligente cujo propósito é alertar para o impacto da vigilância e da IA na vida privada (Garcia-Febo, 2019).

A Universidade de Rhode Island (Kingston, EUA) criou um laboratório de IA localizado na sua biblioteca, em que os estudantes aprendem num ambiente colaborativo multidisciplinar, com os seguintes objetivos: “(1) educating students and faculty about AI’s rapidly developing capabilities, (2) facilitating interdisciplinary collaboration in AI research, and (3) promoting active discussion about AI’s social implications as its core mission” (Kim, 2019, p. 18).

Em síntese, a IA afeta os sistemas de informação, num ambiente generalizado de informação disponível na WWW, e a sua utilização é muito útil na competitividade e inovação dos serviços.

De modo não exaustivo, são já uma realidade: serviços que disponibilizam resumos de publicações e indexação, mais rápidos e precisos; busca e recuperação de informação através de novas plataformas que conectam vários pontos de pesquisa; interfaces de voz e *chatbots* que fornecem serviços de referência; bibliotecas inteligentes; classificação de conteúdos com recurso a ontologias; criação automatizada de metadados; bases de dados digitais de livros em que o *Deep Learning* é usado na digitalização, na *Optical Character Recognition* e no Processamento de Linguagem Natural; dados socioculturais extraídos de arquivos e bibliotecas através de *Machine Learning* (Eberhart, 2018; Cox & Rutter, 2019; Rimland, 2019; Gul & Bano, 2019; Muley, et al., 2019; Jo & Gebru, 2020).

Mudanças para a profissão

Trabalhos técnicos favorecidos pelo emprego da IA são a classificação, a criação de metadados ou o processamento em linguagem natural. Também se destaca, a outro nível, a implementação de agentes conversacionais que, através de interfaces desenhados para estimular a comunicação, enriquecem a interação humano-máquina (Talley, 2016).

Em bibliotecas académicas da área jurídica, assinala-se a discussão acerca da utilização de agentes inteligentes e IA, bem como os respetivos usos, atuais e futuros, benefícios e desvantagens (Talley, 2016).

Na atualidade, “the intelligent consulting robots and the book inventory robot have been applied in some libraries of China” (Wang, Qin & Huang, 2020).

Oportunidades e ameaças para bibliotecas universitárias e públicas resultam do desenvolvimento da IA e da crescente interação humano-computador, pelo que a análise e reflexão acerca desta temática verte-se nas publicações científicas (Arlitsch & Newell, 2017; Tavosanis, 2017; Massis, 2018).

Em particular, destaca-se a avaliação dos efeitos da automatização de tarefas e outros usos da IA e de sistemas robóticos para a atividade de bibliotecários/gestores de informação e os desafios da profissão no futuro próximo (Yildiz & Yildirim, 2018; Cox, Pinfield & Rutter, 2019).

Recorde-se que a investigação realizada na Universidade de Oxford, em 2017, identificava a profissão de bibliotecário como uma das que poderia brevemente desaparecer. Frey e Osborne (2017) “estimate the probability of the replacement by computers of “library technicians” as 99%, “Library assistants, clerical” 95%, archivists 76% and librarians 65%” (Cox, Pinfield & Rutter, 2019, p. 419). No futuro próximo, “the only way to continue in this profession would be to keep the professional knowledge up to date as well as to follow the technological developments in areas such as computers, communication, and the internet” (Yildiz & Yildirim, 2018, p. 6).

Contudo, de acordo com Cox, Pinfield e Rutter (2019), nas bibliotecas universitárias/académicas identificam-se como prementes a curadoria de dados, a aquisição e implementação de ferramentas de IA que auxiliem na pesquisa, na seleção de recursos/informação e a literacia em dados. Novas funções para bibliotecários/profissionais da informação emergem a partir da IA, nomeadamente fornecimento e aquisição de

conteúdo licenciado ou em Acesso Aberto para o trabalho em IA, pesquisa de ferramentas para explorar conteúdo e dados, monitorização da qualidade do material da coleção, auxiliar os utilizadores na navegação no mundo digital, a avaliar conteúdos e a gerir a sua privacidade digital, colaborar no desenho e implementação de infraestruturas tecnológicas de acordo com o perfil dos utilizadores, informar e formar em alfabetização de dados mediante o desenvolvimento de competências para a avaliação de ferramentas de IA, seus resultados e proteção de dados pessoais.

As bibliotecas da Universidade Estadual da Pensilvânia, uma universidade pública localizada nos Estados Unidos, adotaram microcredenciais de alfabetização informacional entre os alunos. “Microcredentials – transferable forms of metadata-encoded, performance-based educational credits – are not new (...). What is different at Penn State is that to help manage the sudden volume of badge submissions, librarians have turned to a new ally: artificial intelligence (AI)” (Rimland, 2019). Neste caso, destaca-se o papel da/o bibliotecária/o ao nível da literacia informacional e coordenação de tecnologias de aprendizagem.

No que respeita à ligação entre a tecnologia e o conhecimento da CI, Jo & Gebru (2020) consideram que “a new specialization should be formed within ML that is focused on methodologies for data collection and annotation (...). Specifically, for sociocultural data, parallels can be drawn from archives and libraries. (...) By showing data collection practices from another field, we encourage ML research to be more cognizant and systematic in data collection and draw from interdisciplinary expertise”.

Em relação à informação disponível na WWW, os profissionais da CI têm um importante e novo papel a considerar ao nível da literacia informacional e de dados, pois a mediação algorítmica é responsável por decidir quais as informações a que temos e teremos acesso e quais permanecerão invisíveis, “according to the economic interests of the companies that control the platforms we visit on the internet, acting as obstacle to the prospects of informational diversity and autonomy that are fundamental in free and democratic societies” (Bezerra & Almeida, 2020).

Conclusões

A IA é reconhecida como uma das tecnologias de ponta nos últimos anos, e muitos países têm considerado o seu desenvolvimento a um nível estratégico nacional, como é o caso dos Estados Unidos e da China. As mudanças que daí advêm para a sociedade serão universais e as bibliotecas não são exceção.

O fenómeno contemporâneo da “migração” da informação para o digital, de que somos testemunhas, é de importância capital para a CI, bem como os desafios de aplicabilidade da IA aos serviços de informação. Reconhece-se que estamos a assistir a uma revolução tecnológica que avança rapidamente num mundo globalizado. Em CI considera-se premente para os profissionais da informação o conhecimento e uso de ferramentas de IA.

Os resultados deste trabalho permitem especificar a aplicabilidade da IA a bibliotecas, com referências a outros serviços de informação, uma realidade que será certamente incrementada nos próximos anos.

A pesquisa atual sobre a aplicação de IA a bibliotecas cobre todos os aspectos da gestão e serviços da biblioteca. Apresentam-se, mediante referências e exemplos ilustrativos, as diversas áreas em que se verifica o impacto da IA e tecnologias correlatas. Para trabalho posterior considera-se uma prospeção sobre o futuro e as implicações a médio e longo prazo da utilização e crescente implementação desta tecnologia.

Do trabalho realizado apresentamos as seguintes conclusões: o impacto da IA é de primeira ordem nas tarefas relacionadas com o processamento da informação, não só no trabalho que pode ser considerado mais mecânico, como a catalogação, mas também para os que exigiam maior intervenção humana, como a indexação, a atribuição de metadados ou a descrição de conteúdo; os efeitos da AI são visíveis na própria prestação do serviço, com resultados na transformação dos espaços físicos e na gestão de recursos.

As mutações que a IA traz para o mundo das bibliotecas supõem uma alteração significativa no seu papel mediador, dada a confluência das mudanças já referidas e as alterações no comportamento informacional. Bibliotecas e serviços de informação precisam de algo mais que aceitar a IA, é necessária uma posição ativa e crítica. Neste sentido a IFLA (2020) incita à colaboração das bibliotecas e suas associações com expertos/especialistas em IA, para que possam desenvolver-se aplicações específicas para uso interno dos serviços ou em resposta às necessidades dos utilizadores, nomeadamente a criação de serviços acessíveis, e para trabalhar na linha da literacia digital e alfabetização dos utilizadores naquela tecnologia.

Referências bibliográficas

- ACRL (2020). 2020 top trends in academic libraries: A review of the trends and issues affecting academic libraries in higher education. (2020). *College & Research Libraries News*, 81(6), 270-278. <https://doi.org/10.5860/crln.81.6.270>
- Arlitsch, K., & Newell, B. (2017). Thriving in the Age of Accelerations: A Brief Look at the Societal Effects of Artificial Intelligence and the Opportunities for Libraries. *Journal of Library Administration*, 57(7), 789-798. <https://doi.org/10.1080/01930826.2017.1362912>
- Bezerra, A. C., & Almeida, M. A. (2020). Rage against the machine learning: a critical approach to the algorithmic mediation of information. *Brazilian Journal of Information Studies: Research trends*, 4(2), 6-23. <https://revistas.marilia.unesp.br/index.php/bjis/article/view/9934>
- Cordell, R. (2020). *Machine Learning + Libraries: A Report on the State of the Field*. <https://labs.loc.gov/static/labs/work/reports/Cordell-LOC-ML-report.pdf?loclr=blogsig>
- Cox, A. M., Pinfield, S., & Rutter, S. (2019). The intelligent library: Thought leaders' views on the likely impact of artificial intelligence on academic libraries. *Library Hi Tech*, 37(3), 418-435. <https://doi.org/10.1108/LHT-08-2018-0105>
- Eberhart, G. M. (2018). An AI Lab in a Library – Why artificial intelligence matters. *American Libraries Magazine*. <https://americanlibrariesmagazine.org/blogs/the-scoop/ai-lab-library>
- Esposito, F., Malerba, D., Semeraro, G., Fanizzi N., & Ferilli, S. (1998). Adding Machine Learning and Knowledge Intensive Techniques to a Digital Library Service. *International Journal on Digital Libraries*, 2, 1 3-19. <https://doi.org/10.1007/s007990050033>

- European Commission (2020). *White Paper on Artificial Intelligence – A European approach to excellence and trust*. <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12270-White-Paper-on-Artificial-Intelligence-a-European-Approach/public-consultation>
- Feigenbaum, E.A. (1989). Toward the library of the future. *Long Range Planning*, 22(1), 118-123. [https://doi.org/10.1016/0024-6301\(89\)90059-9](https://doi.org/10.1016/0024-6301(89)90059-9)
- Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254-280. <https://doi.org/10.1016/j.techfore.2016.08.019>
- García-Febo, L. (2019). Exploring AI – How libraries are starting to apply artificial intelligence in their work. *American Libraries Magazine*. <https://americanlibrariesmagazine.org/2019/03/01/exploring-ai>
- García-Marco, F.-J. (2013). Educación y aprendizaje de la información y la documentación: raíces, desafíos y líneas de acción. *El profesional de la información*, 22(6), 489-504. <http://dx.doi.org/10.3145/epi.2013.nov.01>
- Gomes, L. E. (2020). Ciência da Informação: fundamentos e perspectivas da área científica. In M. B. Marques & L. E. Gomes (Coord.), *Ciência da Informação: visões e tendências* (pp. 89-113). <https://doi.org/10.14195/978-989-26-1896-8>
- Gomes, L. I. E., & Fernández Marcial, V. (2019). Sistema de Informação: abordagem conceitual e metodológica. *Bibliotecas. Anales de Investigación*, 15(3), 395-404. <http://revistas.bnjm.cu/index.php/BAI/article/view/111>
- Grace, K., et al. (2018). *When Will AI Exceed Human Performance? Evidence from AI Experts*. <https://arxiv.org/abs/1705.08807>
- Griffey, J. (ed.) (2019). Artificial Intelligence and Machine Learning in Libraries. *Library Technology Reports*, 55, 1. <https://doi.org/10.5860/ltr.55n1>
- Gul, S., & Bano, S. (2019). Smart libraries: an emerging and innovative technological habitat of 21st century. *The Electronic Library*, 37(5), 764-783. <https://doi.org/10.1108/EL-02-2019-0052>
- IFLA (2020). *IFLA Statement on Libraries and Artificial Intelligence*. <https://www.ifla.org/publications/ifla-statement-on-libraries-and-artificial-intelligence>
- Jo, E. S., & Gebru, T. (2020). *Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning*. <https://arxiv.org/abs/1912.10389>
- Kim, B. (2019). AI and Creating the First Multidisciplinary AI Lab. In J. Griffey, (ed.), *Artificial Intelligence and Machine Learning in Libraries* (pp. 16-20). <https://doi.org/10.5860/ltr.55n1>
- Kitchenham, B. (2004). *Procedures for performing systematic reviews*. Keele University. <http://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf>
- Lawlor, B. (2019). An overview of the NFAIS Conference: Artificial intelligence: Finding its place in research, discovery, and scholarly publishing. *Information Services & Use*, 39(4), 249-280. <https://doi.org/10.3233/ISU-190068>
- Martins, A. L. (2010). Potenciais aplicações da Inteligência Artificial na Ciência da Informação. *Informação & Informação*, 15(1), 1-16. <http://dx.doi.org/10.5433/1981-8920.2010v15n1p1>
- Muley, G., et al. (2019). Use of Deep Learning in Digitization of Libraries. *Helix*, 9(3), 5000-5003. <https://doi.org/10.29042/2019-5000-5003>
- Moreux, J.-P. (2019). Recherche d'images dans les bibliothèques numériques patrimoniales et expérimentation de techniques d'apprentissage profond. *Documentation et bibliothèques*, 65(2), 5-27. <https://doi.org/10.7202/1063786ar>

- Otterlo, M. van (2016). Project BLIIPS: Making the Physical Public Library more Intelligent through Artificial Intelligence. *Qualitative & Quantitative Methods in Libraries*, 5(2), 287-300. <http://www.qqml-journal.net/index.php/qqml/article/view/341/339>
- Padilla, T. (2019). *Responsible Operations: Data Science, Machine Learning, and AI in Libraries*. OCLC Research. <https://doi.org/10.25333/xk7z-9g97>
- Paletta, F., & Silva, A. (Org.) (2020). *Série Tecnologia e Organização da Informação: contribuições para a Ciência da Informação*. <https://doi.org/10.31219/osf.io/tfqnk>
- Panda, S., & Chakravarty, R. (2021). Implementing Conversational AI in Libraries: A Practical Approach. In S. Bhattacharjee (ed.), *Impact of COVID-19 in Academic Institutions* (pp. 124-145). <https://doi.org/10.5281/zenodo.4751040>
- Pinfield, S., Cox, A., & Rutter, S. (2017). *Mapping the Future of Academic Libraries: a Report for SCONUL*. <https://sconul.ac.uk/sites/default/files/documents/SCONUL%20Report%20Mapping%20the%20Future%20of%20Academic%20Libraries.pdf>
- Rimland, E. (2019). The Making of a Microcredential – Penn State University Libraries evaluates badge steps with help from artificial intelligence. *American Libraries Magazine*. <https://americanlibrariesmagazine.org/2019/01/02/making-microcredential-psu-libraries/>
- Stapleton, J., Carter, C., & Bredahl, L. (2020). Developing systematic search methods for the library literature: Methods and analysis. *The Journal of Academic Librarianship*, 46 (5), Article 102190. <https://doi.org/10.1016/j.acalib.2020.102190>
- Wang, B., Qin, Y., & Huang, Q. (2020). The Prospective of Library Spurred by Artificial Intelligence-China Survey. In *2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)* (pp.175-179). doi: 10.1109/AEMCSE50948.2020.00044
- Wheatley, A., & Hervieux, S. (2019). Artificial intelligence in academic libraries: An environmental scan. *Information Services & Use*, 39, 4, 347-356. DOI: 10.3233/ISU-190065
- Yildiz, M., & Yildirim, B. F. (2018). The Effects of Artificial Intelligence and Robotic Systems on Librarianship. *Turkish Librarianship*, 32(1) 26-32. <https://doi.org/10.24146/tkd.2018.29>
- Zhang, D., et al. (2021). *The AI Index 2021 Annual Report*. Stanford University: AI Index Steering Committee, Human-Centered AI Institute.

(Página deixada propositadamente em branco)

TRAS LA ICONOGRAFÍA E ICONOLOGÍA DE LAS REPRESENTACIONES DEL LIBRO EN LAS ARTES: DISEÑO DE UNA HERRAMIENTA PARA SU DESCRIPCIÓN Y CATALOGACIÓN

Carlos Díaz-Redondo¹, José Antonio Frías²,
Elvira-Julieta Miguélez-González³

¹Departamento de Biblioteconomía y Documentación, Facultad de Traducción y Documentación, Universidad de Salamanca, charlierdiaz@usal.es,
ORCID iD 0000-0002-1391-3429

²Departamento de Biblioteconomía y Documentación, Facultad de Traducción y Documentación, Universidad de Salamanca, frias@usal.es,
ORCID iD 0000-0002-5425-8950

³Departamento de Biblioteconomía y Documentación,
Facultad de Traducción y Documentación, Universidad de Salamanca, ejulieta@usal.es,
ORCID iD 0000-0003-3923-494X

Resumo

Se presenta el diseño preliminar de “LiberIC”, una herramienta de descripción y catalogación que permite recopilar, controlar, contextualizar, clasificar y acceder a las múltiples representaciones del libro en las artes pictóricas. Siguiendo una metodología de corte exploratorio y descriptivo, se trabajó con una muestra formada por dos mil piezas pictóricas. A través de la observación directa de estas pinturas, se detectaron aquellas características y particularidades que podían aportar significación al libro. La información extraída se ordenó posteriormente en grandes categorías de información. Partiendo de los principios de descripción y organización del conocimiento de otros estándares de catalogación internacionales ampliamente consolidados, estas categorías se fueron depurando y ajustando. Se desarrolló así un modelo descriptivo que dispone la información en torno a nueve áreas y cerca de cien campos de descripción, que permite en último término recuperar las catalogaciones a través de diferentes puntos de acceso. Con el fin de profundizar en las formas y los significados del libro, tres de estas áreas se diseñaron atendiendo a los principios de la iconografía e iconología moderna, que establecen tres fases diferenciadas en el análisis iconológico de los objetos en el arte. La herramienta que se presenta suple un vacío en la literatura científica y abre nuevas vías de estudio, especialmente en el campo de la Historia del Libro, la Encuadernación, también en la Historia del Arte. El modo en que se han estructurado los datos permite además generar un modelo de descripción extrapolable a otras temáticas pues, al sustituir los campos de descripción vinculados al libro por campos de descripción relativos a otra clase de objeto, se genera un sistema capaz de describir y catalogar cualquier tipo de iconografía.

Palavras-chave: Descripción y catalogación, Historia del libro, Iconografía, Iconología, Libros en la pintura.

Introdução

Las Ciencias de la Información y Documentación y las Artes han ido de la mano, parejas, desde la antigüedad. Si bien una y otra disciplina se encargan de espacios de conocimiento diferentes, mantienen una relación inseparable y apenas explorada hasta el momento. Desde que el ser humano dio el salto creador, las piezas producidas por su mano se han convertido en objetos documentales que pueden ser descritos y catalogados. De la misma manera, desde que existen los libros, éstos han quedado reflejados en todo tipo de manifestaciones creativas, principalmente en la pintura, envueltos en una forma y un significado concreto, según el periodo y el influjo artístico del momento.

Para acometer una primera aproximación a la comprensión de esta forma y este significado, resulta necesario acudir a los primeros tratados de iconografía e iconología (Ripa, 1593). En ellos, se aportan acepciones muy limitadas que tratan la representación del libro bien como una obvia simbología del conocimiento, bien como un mero elemento figurativo. Durante el siglo XVIII, se observa una tendencia clara de recoger y repetir estos mismos conceptos, esta vez en forma de diccionarios y enciclopedias (Gravelot, 1791; Prezel, 1779; Richardson, 1779; Noel, 1811; Guénebault, 1845; Castellanos, 1850; Wessely, 1874; Verneuil, 1897). Tampoco la literatura científica del siglo XX y principios del XXI (Chevalier y Gheerbrant, 1986; Tervarent, 2002; Cirlot, 2006; Revilla, 2007) proporciona demasiadas nuevas vías de comprender la forma y el significado de estas representaciones librarias en el arte, pues el libro se continúa tratando como algo anecdótico, sin prestar atención a la posibilidad de determinar nuevos significados siguiendo los principios de la iconografía e iconología moderna (Panofski, 1939 y 1955).

Con todo, en los últimos años, estas manifestaciones librarias en la pintura han comenzado a cobrar importancia, empleándose como fuente de información para reconstruir técnicas y estilos de ornamentación de la encuadernación (McGrath, 2012 y 2015), para profundizar en la significación de símbolos y atributos de la iconografía religiosa (Pérez Suescun, 2013; Feo, 2019), para analizar los hábitos de lectura en determinados segmentos de población (Bernárdez, 2007), incluso para desentrañar el metadiscursio construido por el mundo del libro empleado como ilustración en las ediciones de fondo antiguo (Rozzo, 2016). Ante la ingente cantidad de producción pictórica en la que el libro es retratado, muchos de estos estudios emplean como muestra de trabajo una serie de obras descontextualizadas, a veces inconexas, sustentando los análisis en descripciones arbitrarias que obvian una gran cantidad de aspectos que aportan matices de significado.

Hoy, en pleno siglo XXI, diversas instituciones culturales internacionales de gran prestigio, como la Fundación Paul Getty (Harpring, 2019), están trabajando en la creación de vocabularios controlados relacionados con la iconografía con el fin de favorecer el estudio de los significados que los objetos cobran en el arte. Es el caso de la herramienta Getty Iconography Authority (IA), incardinada en el proyecto

Cultural Objects Name Authority (CONA). No obstante, la búsqueda del término “libro” en este tipo de bases de datos arroja escasos resultados, encontrándonos siempre registros con información muy superficial que, si bien se interrelacionan perfectamente con los registros de autoridad de otros objetos culturales, en ningún caso presentan campos de descripción concretos que permitan profundizar en sus significados iconográficos e iconológicos.

Objetivos

Con el fin de suplir este vacío, se propuso como objetivo el diseño de una herramienta de descripción y catalogación que permitiese recopilar, controlar, contextualizar, clasificar y acceder a las múltiples representaciones pictóricas del libro. Se perseguía dotar a la comunidad investigadora de un instrumento que normalizara las descripciones y facilitara de ese modo la investigación en dos vertientes. Por un lado, su técnica de construcción y ornamentación. Por otro, los significados iconográficos e iconológicos que este objeto ha adquirido en las sociedades que nos han precedido.

Metodología

Siguiendo una metodología de corte exploratorio y descriptivo, se partió de una muestra de trabajo intencional, formada por dos mil representaciones pictóricas, en cuya escena aparecía la representación de, al menos, un libro. Estas obras, producidas en Europa entre los siglos I y XX, fueron recopiladas de los principales museos europeos a través de diferentes fuentes, ya fuese la exploración de inventarios, catálogos o de las muestras virtuales de sus sitios web.

Mediante la observación directa de estas piezas, unas veces in situ y otras a través de copias fotográficas, se fueron detectando todas aquellas características y particularidades que podían aportar significación al libro, ya fuera por sí mismo, por su relación con la escena o por su interacción con el discurso de la pieza pictórica. La información extraída se ordenó posteriormente en grandes categorías de información. Tomando como punto de partida los principios de descripción y organización del conocimiento de otros estándares de catalogación internacionales ampliamente consolidados, como ISBD(C) para patrimonio bibliográfico, ISAD(G) para patrimonio documental, y OA 3.0 u ObjectID para objetos culturales, estas categorías se fueron depurando y ajustando en torno a una serie de áreas de información, compuestas a su vez por campos de descripción.

Resultados

Así pues, se crearon 9 grandes áreas en las que la información se estructura en torno a cerca de cien campos de descripción que, siempre que resulta posible, son alimentados mediante parámetros normalizados y que, en muchos casos, suponen además un punto de acceso para recuperar la catalogación. Estos campos recogen

desde aspectos básicos, como las características técnicas de la obra pictórica, el movimiento artístico y la categoría a la que pertenece, descriptores libres y códigos de tesaurus; hasta otros aspectos más complejos, como la presencia de un espacio librario en la escena o útiles relacionados con la lectura y/o escritura, el nivel de definición del objeto representado, su posición en el encuadre, su forma de presentación, la tipología y el estilo de encuadernación que presenta, posibles paratextos, la función del libro en la escena, su asociación e interacciones con el entorno, su cometido iconográfico en el discurso artístico de la pieza o la propuesta de interpretación iconológica que se desprende de la catalogación.

A este respecto, es necesario recalcar que las áreas 3, 4 y 5, sobre el objeto, la escena y su interpretación, resultan de especial interés para la investigación, pues se diseñaron atendiendo al método iconográfico de Panofsky (1939 y 1955). Este método nos habla de tres fases diferenciadas en el análisis iconológico de los objetos en el arte: la fase pre-iconográfica (es decir, el estudio de la forma pura del objeto), la fase iconográfica (esto es, el estudio de esa forma en relación con su entorno), y finalmente la fase iconológica (o sea, la determinación de los valores simbólicos que se desprenden de esa forma en ese contexto).

A continuación, se presenta el desglose de campos en cada área de acuerdo con las tablas que ilustran los siguientes subapartados. En ellas, la columna “Campo” recoge el nombre del campo de descripción y un código alfabético entre corchetes que identifica su contenido y funciona como un alias, de cara a facilitar la implementación de la herramienta en lenguaje SQL. En la columna “Tipo” se especifican las características del campo conforme a las siguientes abreviaturas: OB: Obligatorio; CN: Condicional; OP: Optativo; R: Repetible; NR: No Repetible; PA: Punto de acceso. En la columna “Alcance” se detalla la finalidad del campo y algunas directrices para su cumplimentación. Finalmente, en la última columna “Parámetros de caso”, se dispone, a modo de ejemplo, el desglose de la catalogación de una obra con una rica iconografía libraria, concretamente “La Magdalena leyendo”, de Rogier von der Weyden.

El control del registro catalográfico

**Tabela 1. LiberIC: Disposición de campos en Área 0 [A0]:
Control del registro.**

Campos	Tipo	Alcance	Parámetros de caso
Número de control del registro [NCR]	OB NR	Código alfanumérico correlativo, asignado por orden de entrada en el sistema, que identifica de manera unívoca cada registro catalográfico.	LBC00001
Fecha de creación del registro [FCR]	OB NR	Fecha de creación del registro catalográfico. Se redactará siguiendo el formato AAAA-MM-DD.	2021-07-11
Fecha de modificación del registro [FMR]	CN R	Fecha de modificación del registro catalográfico. Se redactará siguiendo el formato AAAA-MM-DD. A continuación, entre paréntesis, se realizará una nota que indique qué	2021-07-13 (Se añade material fotográfico)

		cambios se han realizado en la modificación. Debido al nivel de especialización requerido para realizar la descripción, solamente la persona que ha creado el registro puede realizar cambios en él.	
Responsabilidad [RES]	OB NR	Identidad de la persona que realiza la catalogación, reseñada mediante sus iniciales.	CDR
Entrada/s relacionada/s [ERE]	CN R	Número/s de control de posible/s registro/s catalográficos relacionado/s, por ejemplo, en caso de tratarse de la catalogación de una obra compleja compuesta por varias obras simples o de la catalogación de una obra derivada de una original, cuando esta última también se haya catalogado. Cuando en una pintura aparece la representación de varias iconografías librerías que deben recogerse, cada una de ellas se catalogará independientemente y se vincularán sus registros utilizando este campo.	—
Citación del registro [CTR]	OB NR	Referencia de citación del registro catalográfico que identifica el título, el autor y la fecha de creación de la obra, el número de registro catalográfico, y la URL donde se encuentra alojado.	<i>La Magdalena leyendo, de Rogier von der Weyden, hacia 1435-1438.</i> Registro LBC00001. En: LiberIC: Iconología y Técnica del Libro en las Artes. Universidad de Salamanca. En línea: liberic.usal.es/LBC00001

Fuente: Elaboración propia (2021)

La localización y el acceso a la obra pictórica

Tabela 2. LiberIC: Disposición de campos en Área 1 [A1]: Localización y acceso.

Campos	Tipo	Alcance	Parámetros de caso
Propietario [PRO]	OB NR PA	Nombre del propietario, normalizado según VIAF o cualquier otro fichero de autoridades que se considere pertinente. En caso de obras pertenecientes a particulares, se empleará la expresión “Colección particular” y se colocará entre paréntesis y a continuación un nombre que designe al propietario. Si este último se desconoce, el nombre se sustituirá por el término “Desconocido”.	National Gallery (Londres)
Localización física [LOF]	OB NR	Denominación normalizada según VIAF o cualquier otro fichero de autoridades que se considere pertinente. En caso de obras pertenecientes a particulares, se	National Gallery (Londres)

		empleará “Colección particular” y se colocará entre paréntesis y a continuación el nombre de la ciudad, si se conoce, y el país. Si estos últimos datos se desconocen, se sustituirán por la expresión “Localización desconocida”.	
Acceso digital [ADI]	OP R	Si el propietario es una institución que mantiene un espacio web en el que se puede acceder a una ficha descriptiva de la obra, se reseñará el identificador uniforme de recursos (URI o URL). Si es posible, se empleará un acortador de enlaces.	Bit.ly/LBC00001

Fuente: Elaboración propia (2021)

La identificación y el contexto cultural de la obra pictórica

**Tabela 3. LiberIC: Disposición de campos en Área 2 [A2]:
Identificación y contexto de la obra.**

Campos	Tipo	Alcance	Parámetros de caso
Título [TIT]	OB NR PA	Título de la obra, en castellano.	La Magdalena leyendo
Variante/s de título [VTI]	OP R PA	Otros posibles títulos por los que la obra es conocida, incluyendo su título en el idioma original.	The Magdalene Reading
Autor [AUT]	OB NR PA	Autoridad del autor de la obra, normalizada según VIAF o cualquier otro fichero de autoridades que se considere pertinente. En caso de autores cuya identidad se desconoce, se empleará la expresión “Autor desconocido”. En caso de autores que no han querido firmar su obra, se empleará la expresión “Autor anónimo”. Para otras casuísticas, podrán emplearse expresiones tales como “Taller de...”, “Círculo de...”.	Weyden, Rogier van der
Fechas asociadas al autor [FAU]	OB NR	Fechas de nacimiento y muerte asociadas al autor. Si las fechas son aproximadas, se utilizará el término “Hacia” seguido del año. En caso de desconocerse, se indicarán las fechas de actividad, empleando el término “Activo”. Si estas también se desconocen, se empleará la expresión “Fechas desconocidas”.	Hacia 1399-1464
Lugar asociado al autor [LAU]	OB NR PA	Lugar de nacimiento del autor. En caso de desconocerse, se empleará la expresión “Lugar desconocido”.	Tournai, Bélgica
Pertenencia [PER]	CN NR	Pertenencia del autor a una corriente o escuela artística.	Escuela primitiva flamenca

Tipo de obra [TOB]	OB NR	Tipología de obra en función de su unicidad. Será “Obra simple” si se concibió como una sola. Por el contrario, será “Obra compleja” si pertenece a un conjunto definido, como es el caso de dípticos, trípticos, retablos, series, etc.	Obra compleja
Tipo de creación [TCR]	OB NR	Tipología de creación en función de su originalidad. Si se trata de una creación original, se utilizará la expresión “Obra original”. En caso de ser una copia de una obra original, se utilizará la expresión “Obra derivada” seguida entre paréntesis de los datos identificativos de la obra original.	Obra original
Tipo de autoría [TAU]	OB NR	Tipología de la autoría en función del conocimiento de su creación. Si la autoría de la obra está documentada se utilizará la expresión “Obra documentada”. En caso de ser una atribución, se utilizará la expresión “Obra atribuida”.	Obra documentada
Fecha de creación de la obra [FCO]	OB NR	Fecha de creación de la obra, expresada en años. Para consignar fechas aproximadas, se empleará el término “Hacia” seguido del año de creación. Para consignar una horquilla temporal, se utilizará el término “Entre” seguido de los años. Podrán utilizarse asimismo las expresiones “Antes de...” o “Después de...” cuando solamente se conoce este dato.	Entre 1435-1438
Época [EPO]	OB NR PA	Época o periodo histórico de creación de la obra, en siglos, en números arábigos.	Siglo 15
Movimiento artístico [MAR]	OB NR	Movimiento artístico en el que, por su data de creación, se puede encuadrar la obra.	Renacimiento
Ámbito geográfico de creación la obra [AGC]	OB NR	Lugar de creación de la obra. Si la información se desconoce, se utilizará la expresión “Lugar desconocido”.	Bruselas, Bélgica
Técnica y soporte [TYS]	OB NR	Técnica y soporte utilizados en la creación de la obra.	Óleo sobre tabla
Dimensiones [DIM]	OB NR	Medidas de la obra expresadas en altura por anchura, en centímetros.	62,2 x 54,4 cm
Representación libraria [RLI]	OB NR	Tipología de la(s) representación(es) libraria(s) que figura(n) en la obra en función de su unicidad. Si solamente aparece una, se empleará la expresión “Representación única”. Si por el contrario aparecen dos o más representaciones, se utilizará la expresión “Representación múltiple”.	Representación única

La descripción del objeto: exploración de la fase preiconográfica

**Tabela 4. LiberIC: Disposición de campos en Área 3 [A3]:
Descripción del objeto (Fase preiconográfica).**

Campos	Tipo	Alcance	Parámetros de caso
Definición [DEF]	OB NR	Nivel de definición pictórica que el artista ha dado al objeto. Cuando el libro resulta una mancha sin detalles, se utilizará la expresión “Baja definición”. Cuando el libro presenta una forma definida pero no permite observar con claridad los detalles que le son propios, se utilizará la expresión “Media definición”. Si el libro aparece representado con gran cantidad de detalles, se utilizará la expresión “Alta definición”.	Alta definición
Posición [POS]	OB NR	Posición que el objeto ocupa dentro del plano general de la obra.	Posición central
Tipo de representación [TRE]	OB NR	Si la representación del libro responde a las características propias del momento en que se crea la representación, se utilizará la expresión “Representación contemporánea”. En caso de tratarse de una representación que no responde al momento de creación (por ejemplo, el libro representado como un libro cuando el formato códex aún no existía), se utilizará la expresión “Representación figurada”.	Representación figurada
Forma de presentación [FPR]	OB NR	Modo en el que el objeto aparece presentado.	Libro abierto
Material del cuerpo [MCU]	CN NR	Material de composición del cuerpo. Si el material es visible pero no resulta posible diferenciarlo, se utilizará el término “Indeterminado”.	Pergamino
Tipo de texto [TXT]	CN NR	Técnica de presentación del texto. Solamente se consignará en caso de que el libro se presente abierto. Podrá ser manuscrito o impreso. Y si el texto es visible pero no resulta posible diferenciarlo, se utilizará el término “Indeterminado”.	Texto manuscrito
Espacio textual [ETX]	CN NR	Tipología de la caja o espacio textual en función del número de columnas. Solamente se consignará en caso de que el libro se presente abierto. Si el espacio textual es visible pero no resulta posible diferenciarlo, se utilizará el término “Indeterminado”.	Dos columnas
Elementos sustentados [ELS]	CN NR	Tipología de los elementos sustentados. Solamente se consignará en caso de que el libro se presente abierto. Si los	Tintas de escritura Tintas pictóricas

		elementos sustentados son visibles pero no resulta posible diferenciarlos, se utilizará el término “Indeterminado”.	
Legibilidad [LEG]	CN NR	Legibilidad del texto del libro. Solamente se consignará en caso de que el libro se presente abierto. Si el texto que presenta el libro es interpretable se utilizará el término “Texto legible”. Si por el contrario no lo es, se utilizará el término “Texto ilegible”. Si el texto es visible pero no resulta posible determinar su legibilidad, se utilizará el término “Indeterminado”.	Texto ilegible
Transcripción [TNC]	CN NR	Transcripción del texto que contiene el libro. Solamente se consignará en caso de que el libro se presente abierto y el texto resulte legible. Se redactará literalmente y entre comillas.	–
Materia [MAT]	CN R	Temática o materia sobre la que versa el libro en función del texto. Solamente se consignará en caso de que el libro se presente abierto y el texto resulte legible. Si fuera necesario, se emplearán hasta cuatro materias. Para normalizar, se elegirán de una lista de encabezamientos de materia.	–
Elementos auxiliares del texto [EAT]	CN R	Elementos auxiliares que acompañan al texto. Solamente se consignará en caso de que el libro se presente abierto. Se reseñarán escogiendo entre un vocabulario controlado redactado previamente. Si aparecen elementos auxiliares pero no puede determinarse qué tipos son, se utilizará el término “Indeterminado”.	Letras capitulares simples
Tipología de encuadernación [TEN]	CN NR	Tipo de encuadernación del libro, en función de su aparente montaje. Se reseñarán escogiendo entre un vocabulario controlado redactado previamente. Si no resulta posible apreciarlo, se utilizará el término “Indeterminado”.	Encuadernación cosida sobre nervios
Material de recubrimiento [MCB]	CN NR	Material con el que aparentemente están realizadas las cubiertas del libro. Si no resulta posible apreciarlo, se utilizará el término “Indeterminado”.	Recubrimiento de tela
Color de cubierta [CCU]	CN NR	Color predominante de las cubiertas. Si no resulta posible apreciarlo, se utilizará el término “Indeterminado”.	Blanco
Montaje del lomo [MLO]	CN NR	Se reseñará el tipo de montaje que se observa en el lomo. Si no resulta posible apreciarlo, se utilizará el término “Indeterminado”.	Lomo hueco
Forma del lomo [FLO]	CN NR	Forma de lomo en función de su montaje. Si este elemento es visible pero	Indeterminado

		no resulta posible apreciarlo correctamente, se utilizará el término “Indeterminado”.	
Cabezadas [CAB]	CN NR	Tipo de cabezadas que presenta el libro. Si este elemento es visible pero no resulta posible apreciarlo correctamente, se utilizará el término “Indeterminado”.	Cabezada simple
Cortes [COR]	CN NR	Presentación de los cortes del libro. Si este elemento es visible pero no resulta posible apreciarlo correctamente, se utilizará el término “Indeterminado”.	Cortes dorados
Elementos auxiliares de la encuadernación [EAU]	CN R	Si el libro presenta elementos que pueden considerarse auxiliares de la encuadernación, se reseñarán todos ellos partiendo de un vocabulario controlado redactado previamente.	Camisa de tela Cierres de manecilla Portaseñuelos
Ornamentación exterior [OEX]	CN NR	Se reseñará si el libro contiene ornamentación exterior. Si no es posible apreciarlo, se utilizará el término “Indeterminado”.	Indeterminado
Técnicas de ornamentación [TOR]	CN R	Si hay ornamentación exterior, se reseñará la(s) técnica(s) recreada(s) que tendrían que haberse utilizado. Si la ornamentación es visible pero no resulta posible determinar las técnicas, se utilizará el término “Indeterminado”.	-
Herramientas de ornamentación [HOR]	CN R	Si hay ornamentación exterior en el libro, se anotarán las herramientas que deberían haberse utilizado para crearla. Si la ornamentación es visible pero no resulta posible determinar las herramientas que deberían haberse utilizado, se utilizará el término “Indeterminado”.	-
Estilo de ornamentación [EOR]	CN NR	Si hay ornamentación en el libro y esta encaja dentro de un estilo definido, se reseñará cuál escogiendo entre un vocabulario controlado. Si no resulta posible determinar un estilo, se utilizará el término “Indeterminado”.	Rica tela
Soporte de lectura y/o escritura [SLE]	CN NR	Si el libro aparece representado apoyado sobre otro elemento que hace las veces de soporte de lectura y/o escritura, se reseñará cuál.	-
Manifestación bibliotecaria [MBI]	CN R	En caso de que el libro presente cualquier tipo de manifestación que permita ponerlo en relación con el oficio bibliotecario, se reseñará en qué consiste.	-

Fuente: Elaboración propia (2021)

La descripción de la obra: exploración de la fase iconográfica

**Tabela 5. LiberIC: Disposición de campos en Área 4 [A4]:
Descripción de la obra (Fase iconográfica).**

Campos	Tipo	Alcance	Parámetros de caso
Categoría pictórica [CTP]	OB NR PA	Género o categoría pictórica en la que se puede encuadrar la obra en función de su temática.	Escena religiosa
Acotación de género pictórico [ADG]	CN R	Acotación del género o categoría pictórica que puede profundizar en la temática de la obra.	Sacra conversazione
Descripción de conjunto [DCO]	CN NR	En caso de tratarse de una obra compleja formada por varias obras simples, se realizará una descripción breve del conjunto en texto libre.	La obra es uno de los tres fragmentos, concretamente el que correspondía a la parte inferior derecha, de lo que debió de ser un gran retablo de la Virgen y el Niño acompañados por diversos santos, pintado por Rogier van der Weyden para una iglesia en Bruselas. El Museo Nacional de Estocolmo conserva un dibujo del retablo original datado a finales del siglo XVI. Gracias a él, otros fragmentos del conjunto se han identificado en el Museo Calouste Gulbenkian de Lisboa. Según Campbell (2004, p. 50), la obra en su conjunto se titulaba “La madre de Dios y el Niño, con un Santo Obispo, San Juan Bautista y San Juan Evangelista” y fue uno de los primeros encargos del pintor.
Descripción de la escena [DES]	OB NR	Breve descripción de la escena en texto libre.	La obra representa a una mujer joven, de rasgos dulces y piel pálida, ataviada con un velo o toca blanca y un vestido verde, atuendo propio de la moda del siglo XV. Se encuentra sentada sobre un cojín en una sala medieval, sujetando entre sus manos un libro abierto, que parece ser un Libro de Horas. La

			expresión del rostro y la posición de su cuerpo, con la cabeza y las piernas curvadas alrededor de las rodillas dobladas, enfatizan su absorción en la lectura. El tarro de perfume para ungir, que aparece colocado en primer plano, permite identificarla con Santa María Magdalena. Tras ella, puede apreciarse parte del cuerpo de un hombre, que se ha identificado con San José sosteniendo un rosario, mientras otro personaje se arrodilla ante ella, dejando ver los dedos de sus pies por debajo de su túnica roja.
Elementos iconográficos de relevancia [EIR]	CN R	Se indicarán otros elementos iconográficos presentes en la obra.	Tarro de perfumes Rosario
Iconclass [ICC]	OB R PA	Se utilizarán hasta cuatro código(s) normalizado(s) según el tesoro Iconclass que recojan la iconografía presente.	11HH(MARY MAGDALENE)4 11Q51
Descriptorios [DCR]	OB R PA	Se escogerán hasta cuatro descriptorios que definan con claridad la obra, de modo que el registro pueda ser recuperado en una búsqueda simple por palabra clave.	María Magdalena, Santa Lectura Encuadernación en rica tela Libro manuscrito
Espacio librario [ELI]	CN NR	Si el libro aparece representado formando parte de un espacio librario, se reseñará cuál.	–
Materiales de escritura [MAE]	CN R	Si la escena presenta materiales y útiles de escritura, como plumas, tinteros, etc., se reseñarán cuáles.	–

Fuente: Elaboración propia (2021)

La interpretación del objeto: Construcción de la Fase iconológica

Tabela 6. LiberIC: Disposición de campos en Área 5 [A5]: Interpretación del objeto (Fase iconológica).

Campos	Tipo	Alcance	Parámetros de caso
Misión del objeto en la escena [MIE]	OB NR	Misión que el libro cumple dentro de la escena general. Si el libro actúa como elemento de refuerzo para una idea, se utilizará la expresión “Elemento de refuerzo”. Si el libro	Elemento de refuerzo

		actúa como sustitución de otra idea, se utilizará la expresión “Elemento de sustitución”. Si por el contrario, el libro no es más que un mero componente figurativo, se utilizará la expresión “Elemento nulo”.	
Soporte de otro elemento [SOE]	CN R	Si el libro actúa a su vez como un soporte de otro(s) elemento(s) iconográfico(s) presente(s) en la obra, se reseñará cuál(es) partiendo de un vocabulario controlado.	–
Interacción con un personaje [ICP]	CN NR	En caso de que el libro forme parte del discurso propio de un personaje principal, se indicará su identidad utilizando VIAF o cualquier otro catálogo de autoridades que se considere pertinente. Se considerará el personaje principal aquel sobre el que recaiga el peso temático de la obra. Si la identidad se desconoce, se utilizará la expresión “Personaje desconocido”.	María Magdalena, Santa
Tipo de interacción [TDI]	CN NR	En caso de que el libro acompañe a un personaje, se indicará el tipo de interacción que el objeto guarda con el personaje al que acompaña. Si el personaje principal está contacto directo con el libro, se utilizará la expresión “Interacción directa”. Si el personaje mantiene un contacto indirecto con el libro, se utilizará la expresión “Interacción indirecta”.	Interacción directa
Género del personaje [GEN]	CN NR PA	En caso de que el libro acompañe a un personaje principal, se indicará su género aparente. En caso de figuras celestiales, como es el caso de ángeles, se utilizará la expresión “Ser asexual”.	Mujer
Ocupación del personaje [OPE]	CN NR PA	En caso de que el libro acompañe a un personaje principal, se indicará su profesión u ocupación empleando la forma plural. Si no se conoce, se utilizará la expresión “Profesión desconocida”.	Santos
Acto de vinculación del objeto con el personaje [VIN]	CN NR PA	Solamente en caso de que la relación con el personaje principal sea directa, se indicará qué acto define aquella vinculación.	Acto de lectura
Interacción con un personaje secundario [IPS]	CN R	En caso de que el libro forme parte del discurso propio de un segundo personaje, se indicará su identidad utilizando VIAF o cualquier otro catálogo de autoridades. Se considerará(n) personaje(s) secundario(s) aquel(los) que	–

		acompaña(n) al personaje principal. Si la identidad se desconoce, se utilizará la expresión “Personaje desconocido”.	
Función iconográfica [FIC]	OB NR	Función iconográfica que el libro cumple dentro de la escena.	Atributo
Línea/s iconográfica/s [LIC]	OB R PA	Se escogerá una o varias línea(s) del diccionario de líneas iconográficas que se ha desarrollado a tal efecto.	Santos – María Magdalena
Propuesta de interpretación iconológica [PIC]	OB NR	Se redactará en texto libre una breve propuesta subjetiva sobre la interpretación iconológica que el objeto cumple con respecto a la escena en la que se representa, partiendo de las líneas iconográficas seleccionadas anteriormente.	El color blanco, que tradicionalmente se interpreta como significado de humildad, divinidad y pureza, se repite en el tarro de perfume para ungir sobre el suelo, en la camisa que viste la encuadernación sobre las rodillas de la Santa y en la toca o velo que cubre su cabeza. Esto sugiere una relación triangular ascendente y directa entre estos elementos que puede explicarse como la humildad para renunciar a la vida pecaminosa a través de Jesucristo, la ascunción de Dios a través de la lectura de su palabra, y las ideas de pureza que estos actos han elevado o generado en el pensamiento o conciencia de la Santa.

Fuente: Elaboración propia (2021)

Las notas

Tabela 7. LiberIC: Disposición de campos en Área 6 [A6]: Notas.



Campos	Tipo	Alcance	Parámetros de caso
Nota de control del registro [NA0]	OP R	Información relativa al control del registro que no tuvo cabida en su área.	–
Nota de localización y acceso [NA1]	OP R	Información relativa a la localización y acceso que no tuvo cabida en su área.	–
Nota de identificación y contexto de la obra [NA2]	OP R	Información relativa a la identificación de la obra que no tuvo cabida en su área.	–
Nota de descripción del objeto [NA3]	OP R	Información relativa a la descripción del objeto que no tuvo cabida en su área.	–
Nota de descripción de la obra	OP R	Información relativa a la descripción de la obra que no tuvo cabida en su	–

[NA4]		área.	
Nota de la interpretación del objeto [NA5]	OP R	Información relativa a la interpretación del objeto que no tuvo cabida en su área.	–
Nota de documentación fotográfica [NA7]	OP R	Información relativa a la documentación fotográfica que no tuvo cabida en su área.	–
Nota de bibliografía y fuentes [NA8]	OP R	Información relativa a la bibliografía y fuentes que no tuvo cabida en su área.	–

Fuente: Elaboración propia (2021)

La documentación fotográfica

Tabela 8. LiberIC: Disposición de campos en Área 7 [A7]: Documentación fotográfica.

Campos	Tipo	Alcance	Parámetros de caso
Imagen [FOT]	OB NR	Fotografía de la obra, preferiblemente en formato JPEG o PNG.	 <p>LBC00001.png</p>
Detalle [FDE]	OB R	Fotografía de detalle de la representación del libro, preferiblemente en formato JPEG o PNG.	 <p>LBC00001Detalle.png</p>
Suministrador [SUM]	OB NR	Se indicará la fuente de la que se ha extraído la fotografía y su(s) detalle(s). En caso de haberse	National Gallery (Londres)

		extraído de cualquier tipo de material bibliográfico, se utilizará el encabezamiento principal de la publicación seguido del año de edición entre paréntesis, desarrollándose obligatoriamente esta información en el Área 8, dentro del campo de referencias bibliográficas.	
Copyright [COP]	OB NR	Se reseñará el nombre de la persona o institución que detenta los derechos de la propiedad intelectual de la fotografía irá precedido del símbolo ©. Cuando la fotografía esté extraída de la página web de un museo, preferiblemente se utilizará el texto normalizado que la institución demande. En caso de fotografías que se encuentren en dominio público, se utilizará la expresión “Obra en dominio público”.	© Copyright The National Gallery, London 2021
Licencia [LCN]	CN NR	Si la fotografía está sujeta a derechos de reproducción, se detallará el tipo de licencia de reproducción, si la hubiere.	CC BY-NC-ND 4.0

Fuente: Elaboración propia (2021)

Las fuentes de información

Tabela 9. LiberIC: Disposición de campos en Área 8 [A8], Bibliografía y fuentes.

Campos	Tipo	Alcance	Parámetros de caso
Referencias bibliográficas [REF]	OB R	Listado de referencias bibliográficas, ordenadas alfabéticamente, que hayan sido utilizadas para la realización de la catalogación. Para su redacción, se emplearán las normas APA.	Campbell, Lorne (2004). Van der Weyden. Londres: Chaucer Press.

Fuente: Elaboración propia (2021)

Conclusões

Por un lado, si bien LiberIC permite describir ampliamente una iconografía libraria presente en la obra pictórica, nos encontramos con algunas particularidades que requieren un tratamiento diferente. Por ejemplo, ¿qué hacemos cuando una misma obra pictórica presenta varias representaciones librarias y ambas cumplen funciones iconográficas o iconológicas diferentes? En este punto, se optó por tratar las representaciones del libro de forma analítica. Es decir, en estos casos, se realizaría una

catalogación individual para cada una de las representaciones y se vincularían ambos registros utilizando el campo de Entradas relacionadas [ERE] en el Área 0. Además, con la mirada puesta en el futuro, nos encontramos también afrontando nuevos retos. ¿Qué hacemos con las representaciones del libro en otras manifestaciones artísticas, tal como la escultura, en las que las representaciones del libro pasan de ser de 2D a 3D y comenzamos a observar la importancia de otros factores como la deformación plástica de los libros para mejorar aspectos como la perspectiva de visión? A este respecto, si el modelo de descripción prospera, será preciso añadir campos que valoren aspectos como el volumen de la materia.

Por otro lado, con el fin de convertir esta herramienta en un instrumento útil para el investigador, desde el inicio se diseñó con vistas a implementarlo en una base de datos que toma el nombre del modelo de descripción: “LiberIC: Iconografía e Iconología del Libro en las Artes” y que pronto estará en línea. En estos momentos, nos encontramos traduciendo el modelo al lenguaje SQL e implementando la base de datos. Una vez lanzada la herramienta, quedará abierta a la colaboración con otras instituciones culturales, especialmente museos, pues son estos los que custodian las fuentes primarias de las que se sirve este modelo.

En definitiva, la herramienta que se presenta suple un vacío en la literatura científica y abre nuevas vías de estudio especialmente en el campo de la Historia del Libro, la Encuadernación, también en la Historia del Arte. Sintetiza en un solo recurso la enorme complejidad técnica, iconográfica e iconológica de las representaciones del libro en las artes. Facilita, asimismo, una lectura mucho más profunda y completa al adaptar en sus campos de descripción los principios de la iconografía moderna. Se trata, además, de un modelo extrapolable a otras temáticas pues, si se sustituyen los campos de descripción de las áreas 3, 4 y 5 por campos de descripción relativos a otra clase de objeto, se genera un sistema capaz de describir y catalogar cualquier tipo de iconografía. Finalmente, al normalizar y fragmentar en campos todas estas descripciones, éstas se convierten en un conjunto de metadatos y por tanto, en potenciales datos abiertos para otras instituciones.

Referências

- BERNÁRDEZ, Asunción (2007). Pintando la lectura: mujeres, libros y representación en el siglo de oro. *Edad de Oro*, XXVI, 67-89.
- CASTELLANOS de Losada, Basilio Sebastián (1850). *Compendio del sistema alegórico y diccionario manual de la iconología universal*. Madrid: Imprenta de D. B. González.
- CHEVALIER, Jean; Gheerbrant, Alain (1986). *Diccionario de los símbolos*. Barcelona: Herder.
- CIRLOT, Juan Eduardo (2006). *Diccionario de símbolos*. 10.^a edición. Madrid: Siruela
- GRAVELOT, Hubert-François; Cochin, Charles-Nicolas (1791). *Iconologie par figures, ou Traité complet des allegories, emblemes, etc.: ouvrage utile aux artistes, aux amateurs, et pouvant servir à l'éducation des jeunes personnes*. 3 volúmenes. París: Chez Le Pan.
- FEO, Michele (2019). *Cosa leggeva la Madonna? Quasi un romanzo per immagini*. Firenze: Polistampa.
- GUÉNEBAULT, Louis-Jean (1845). *Dictionnaire Iconographique des Monuments de l'Antiquité Chrétienne et du Moyen Age...* 2 volúmenes. París: Ch. Porquet.

- HARPRING, Patricia (2019). Getty Iconography Authority: Introduction and Overview [Documento de trabajo del Getty Vocabulary Program]. J. Paul Getty Trust. Accesible en: <https://www.getty.edu>
- MCGRATH, Anthony (2012). *Books in Art: the meaning and significance of images of books in Italian religious painting, 1250-1400* [Doctoral Thesis]. Sussex: University.
- (2015). Using religious art as pictorial evidence for Medieval Book History. *Book History*, 18, 33-47.
- NOËL, Fr. (1810). *Dictionnaire de la Fable*. Paris: Le Normant.
- PANOFSKY, Erwin (1939). *Estudios sobre Iconología*. Madrid: Alianza, 2008.
- (1955). *El significado en las artes visuales*. Madrid: Alianza, 2008.
- PÉREZ Suescun, Fernando (2013). Los evangelistas en el scriptorium... *Funciones y prácticas de la escritura*. Madrid: Universidad Complutense; Escalona: Ayuntamiento, 193-200.
- PREZEL, M. de (1779). *Dictionnaire iconologique ou introduction a la connoissance des peintures...* Paris: Chez Hardoium.
- REVILLA, Federico (2007). *Diccionario de Iconografía y Simbología*. 5.ª edición ampliada. Madrid: Cátedra.
- RICHARDSON, George (1779). *Iconology, or a collection of emblematical figures...* (2 vols.) London: G. Scott.
- RIPA, Caesar (1593). *Iconologia, overo, Descrittione dell'Imagini universali*. Roma: Gli heredi di Gio. Gigliotti.
- ROZZO, Hugo (2016). *Iconologia del libro nelle edizioni dei secoli XV e XVI*. Udine: Forum.
- TERVARENT, Guy de (2002). *Atributos y Símbolos en el Arte Profano. Diccionario de un Lenguaje Perdido*. Barcelona: Ediciones del Serbal.
- VERNEUIL, Maurice Pillard (1897). *Dictionnaire des symboles, emblèmes & attributs*. París: Librairie Renouard.
- WESSELY, J. E. (1874). *Iconographie Gottes und der Heiligen*. Leipzig: T. O. Weigel.

MUSEUS BRASILEIROS COM COLEÇÕES DE ARTE: ANÁLISE DOS INSTRUMENTOS DE REGISTRO PARA BENS CULTURAIS

Camila Aparecida da Silva¹

¹Programa de Pós-Graduação em Ciência da Informação (PPGCI/ECA/USP),
Universidade de São Paulo, Brasil, apsilva.camila@gmail.com,
ORCID iD 0000-0003-3568-592X

Resumo

Este trabalho buscou identificar os instrumentos de registro e as diretrizes de documentação adotadas pelos museus brasileiros com coleções de arte. O diagnóstico foi feito com dados brutos do Cadastro Nacional de Museus de 2010, pesquisa estatística de maior abrangência no Brasil e realizada pelo Instituto Brasileiro de Museus. Para os objetivos deste trabalho, esses dados foram tabulados e analisados em detalhes, a partir dos quais foi possível definir o perfil dos museus com coleções de arte no Brasil que inclui: a) o total de museus de arte e o total de museus com coleções híbridas que englobam obras artísticas; b) o total de bens culturais de Artes Visuais presentes nesses museus; c) os instrumentos de registro por eles utilizados; e d) o total de museus com coleções de arte que fazem registro de suas coleções, bem como aqueles que não o fazem. A análise indicou que 544 museus híbridos com coleções de arte registram seus acervos, contra 133 que não o fazem, e que 88 museus de arte do Brasil utilizam pelo menos um tipo de instrumento para o registro de suas coleções, contra 11 que não fazem uso de nenhum instrumento. Contudo, a partir dos dados estatísticos do Cadastro Nacional de Museus, não foi possível verificar as diretrizes e recomendações adotadas por esses museus, o que revela que esse aspecto da documentação de museus não foi considerado relevante pelos órgãos nacionais. O cenário identificado aponta para a necessidade de capacitação de pessoal, de estatísticas periódicas sobre os museus com coleções de arte no Brasil, bem como para a urgência de elaboração de políticas institucionais de documentação, nas quais o procedimento de catalogação deve estar inserido. Essas políticas e procedimentos são elementos indispensáveis para a implantação, manutenção e aprimoramento das atividades de registro de bens culturais de Artes Visuais voltadas à padronização de processos. O desenvolvimento de um programa de acreditação que inclua a adoção de procedimentos mínimos, entre eles o de catalogação, como critério para o recebimento de certificação pelo cumprimento de boas práticas pode ser uma alternativa para o estabelecimento do registro de objetos museológicos apoiados em padrões de dados atualizados e reconhecidos internacionalmente.

Palavras-chave: catalogação, documentação em museus, museu de arte, padrões de dados, representação da informação.

1 Introdução

Este trabalho teve como objetivo discutir e contribuir para o uso de instrumentos de registro baseados em diretrizes atualizadas e reconhecidas internacionalmente, de modo a possibilitar a otimização do acesso informacional de obras de arte presentes em museus brasileiros. O problema do trabalho concentrou-se na identificação dos instrumentos de registro utilizados pelos museus brasileiros com coleções de arte, em vista das dificuldades enfrentadas pelos profissionais de museus no tratamento e processamento informacional de acervos museológicos que, como consequência, limitam a recuperação e o acesso às informações e também a extroversão de obras de arte para diversos públicos nacionais e internacionais. Essas dificuldades foram observadas em estudos de casos apresentados em conferências nacionais e no diálogo com a comunidade de profissionais que atuam diariamente nessas instituições.

A busca pela identificação dos instrumentos e diretrizes foi feita por meio da análise de dados estatísticos de 2010 oriundos do Cadastro Nacional de Museus (CNM) e fornecidos pelo Instituto Brasileiro de Museus (IBRAM) em agosto de 2018 (Instituto Brasileiro de Museus, 2010). O IBRAM, órgão nacional criado em 2009, é responsável pela gestão de 30 instituições museológicas e tem como uma de suas atribuições promover boas práticas de documentação para o setor museológico do país. O resultado da análise mostrou-se importante para traçar o perfil dos museus brasileiros com coleções de arte e acredita-se que será útil para o planejamento de programas institucionais com foco na documentação de coleções artísticas.

1.1 Estatísticas para museus brasileiros com coleções de arte

Estatísticas com dados atualizados sobre os museus brasileiros com coleções de arte são importantes para a elaboração e manutenção de programas institucionais. Os dados do CNM possuem uma amostra significativa de museus brasileiros participantes. Contudo, eles não cobrem as especificidades de cada tipologia de museu, como Artes Visuais. Nesse sentido, esta pesquisa buscou, ainda que a partir dos dados brutos do CNM fornecidos pelo IBRAM que remontam à 2010, conhecer o perfil dos museus de arte, principalmente no que diz respeito ao registro de suas coleções. O registro de obras de arte deve, pelo menos, ter um nível básico de informação que permita a identificação e a localização de cada obra em um acervo. Um registro é composto por metadados definidos e selecionados especificamente para a tipologia em foco. Metadados agrupados em um esquema são úteis para a descrição padronizada de informações sobre os objetos e auxiliam no atendimento de solicitações feitas por consulentes internos e externos aos museus em questões como o total de obras criadas na última década e seus respectivos títulos ou como o total de obras que representam um determinado assunto em sua composição, apenas para dar alguns exemplos. Para esse fim, catalogadores utilizam recursos de representação descritiva e representação temática de modo que informações sobre os objetos sejam organizadas, facilmente acessíveis e recuperáveis.

Apesar das transformações tecnológicas que possibilitam o uso de sistemas e catálogos *on-line*, parte dos museus brasileiros faz registro de seus acervos em papel

com o uso, por exemplo, de fichas catalográficas. No entanto, independente do meio utilizado para a descrição de coleções – seja manual ou informatizado –, recomenda-se o uso de diretrizes para essa atividade. Na tentativa de conhecer as diretrizes em uso pelos museus de arte no Brasil, encontramos dificuldades de acesso a dados específicos sobre a tipologia Artes Visuais. Portanto, conhecer o total de museus com coleções de arte no país e saber o total de bens culturais sob sua guarda para, então, compor um quadro geral dessas características foi primordial. Neste trabalho, compreendemos bens culturais musealizados como “o conjunto de testemunhos culturais e naturais que se encontram sob a proteção de instituições museológicas” (Lei n. 11.906, 2009).

O levantamento de material para este trabalho incluiu guias de museus e pesquisas estatísticas que hoje encontram-se desatualizadas. A publicação *Museus em Números* (Instituto Brasileiro de Museus, 2011a, 2011b) com os resultados do CNM não apresentou os dados por tipologias de museus, de modo a impossibilitar o diagnóstico sobre os museus de arte. Por essa razão, foi solicitado ao IBRAM os dados brutos do CNM para que fosse possível tabulá-los para os objetivos desta análise. Não foram encontradas outras estatísticas nacionais para museus mais abrangentes do que a do CNM, que teve 1.500 instituições participantes. A Pesquisa Anual de Museus (PAM), por exemplo, que se iniciou em 2014 e foi descontinuada no ano seguinte, contou com 993 instituições (Instituto Brasileiro de Museus, 2014, ca. 2015). E a plataforma MuseusBr (<http://museus.cultura.gov.br/>), apesar de mais recente, não abrange uma categoria específica para museus com coleções de arte. Foram também consultados o guia de museus publicado pela Editora da Universidade de São Paulo (Universidade de São Paulo, 2000), elaborado antes do CNM, e o guia de museus do IBRAM (Instituto Brasileiro de Museus, 2011d), composto apenas por uma lista de museus. Embora desatualizados, para este trabalho optou-se por usar exclusivamente os dados brutos do CNM de 2010 fornecidos em agosto de 2018 pelo IBRAM por ser o recurso com informações sobre registro de acervo com o maior número de participantes dentre os recursos levantados.

1.2 Questionário do Cadastro Nacional de Museus: registro de bens culturais

O CNM foi realizado a partir de um Questionário aplicado nas instituições museológicas do Brasil e cobriu tópicos como acesso ao público, atividades, caracterização física, acervo, entre outros. Para desenhar o perfil dos museus com coleções de arte no Brasil, analisamos as respostas dos museus na seção denominada “Acervo”. Os itens do Questionário do CNM (Instituto Brasileiro de Museus, 2011c) relativos ao registro do acervo compreendem o total de bens culturais que compõem o acervo; se esse total é aproximado ou exato (e quais os números); a tipologia do acervo; se o acervo é “registrado/documentado”, com as opções de resposta “sim” e “não”, e quais os instrumentos de registro utilizados, incluindo o “Livro de registro” (com a indicação do número de bens culturais registrados), a “Ficha de catalogação/registo” (com o número de bens culturais catalogados), a “Documentação fotográfica” (com o número de bens culturais fotografados) e o “*Software/Programa*”

informatizado” (com o número de bens culturais inseridos). Instrumentos de registro são ferramentas utilizadas para representar informações sobre cada objeto presente em uma coleção. Essas informações, quando descritas de forma consistente, permitem e facilitam sua recuperação e acesso pelos usuários.

Entre as tipologias do acervo, as opções no Questionário do CNM eram Antropologia & Etnografia, Arqueologia, Artes Visuais, Ciências Naturais & História Natural, Ciência & Tecnologia, História, Imagem & Som, Virtual e Outros. Na categoria “Outros” havia as opções Biblioteconômico, Documental e Arquivístico. Para cada uma dessas tipologias era preciso incluir o total de bens culturais pertencente à instituição. Neste trabalho, dividiu-se os museus com coleções de arte em dois grupos, ou seja, foram considerados como museus híbridos aqueles que possuem mais de uma tipologia de coleções em seu acervo, sendo uma delas de Artes Visuais, e como museus de arte aqueles com coleções formadas exclusivamente por obras de arte. De acordo com a tabulação realizada, identificamos que no Brasil as tipologias predominantes nos museus híbridos, tal como denominado neste trabalho, são História e Imagem & Som. Esse aspecto é importante para a definição de projetos que precisam definir, por exemplo, metadados para obras de arte e para objetos museológicos de História e de Imagem & Som que levem em consideração suas especificidades, assim como as características em comum entre essas tipologias.

Foram observados problemas no Questionário do CNM como a falta de definição para “acervo registrado/documentado” e para cada um dos instrumentos de registro, que se definidos, poderiam facilitar o entendimento das perguntas pelas instituições e evitar dificuldades no preenchimento do Questionário, de modo a garantir informações mais precisas. Para o instrumento *Software/Programa informatizado*, por exemplo, há respostas que variam entre nomes de *softwares*, incluindo editores de texto, e nomes de linguagens de programação, aspecto também apontado por Miranda (2016).

2 Os museus brasileiros com coleções de arte, seus bens culturais e instrumentos de registro

Norteados pela pergunta inicial deste trabalho relacionada à identificação de diretrizes de documentação e instrumentos de registro utilizados pelos museus brasileiros para catalogar suas coleções de arte, e após a análise dos dados do CNM, foi possível traçar o perfil dos museus brasileiros com coleções de arte. Esse perfil inclui o total de museus com coleções de arte, o total de bens culturais da tipologia Artes Visuais e os instrumentos de registro utilizados. Essa questão de investigação mostra-se relevante para a teoria e para a prática de documentação em museus porque permite identificar e tornar de conhecimento público os instrumentos em uso para bens culturais de arte e, conseqüentemente, contribuir para o seu aprimoramento, bem como para a promoção de boas práticas de registro e de processos de documentação apoiados em diretrizes atualizadas e reconhecidas internacionalmente pela comunidade museológica.

Em 2010, o IBRAM mapeou 3.025 museus no Brasil, dos quais 1.500 participaram do CNM. A tabulação que realizamos com os dados do CNM (Instituto Brasileiro de Museus, 2010) indicou a existência de 776 museus brasileiros com coleções de arte, sendo 677 museus híbridos e 99 museus de arte (Figura 1). Os 677 museus híbridos que

possuem coleções de arte estão distribuídos nas cinco regiões do país, sendo elas Centro-Oeste, Nordeste, Norte, Sudeste e Sul. A maior concentração está no Sudeste, com 251 museus híbridos, e a menor concentração está no Norte, com 30 museus híbridos, conforme visto na Figura 1(a). Do mesmo modo, a Figura 1(b) apresenta a distribuição dos 99 museus de arte nas Unidades Federativas (UF) do território brasileiro. Conforme se observa na Figura 1(b), a maior concentração está na região Sudeste do país, economicamente mais desenvolvida e formada pelas UFs de São Paulo, Minas Gerais, Rio de Janeiro e Espírito Santo, que juntas totalizam 37 museus de arte.

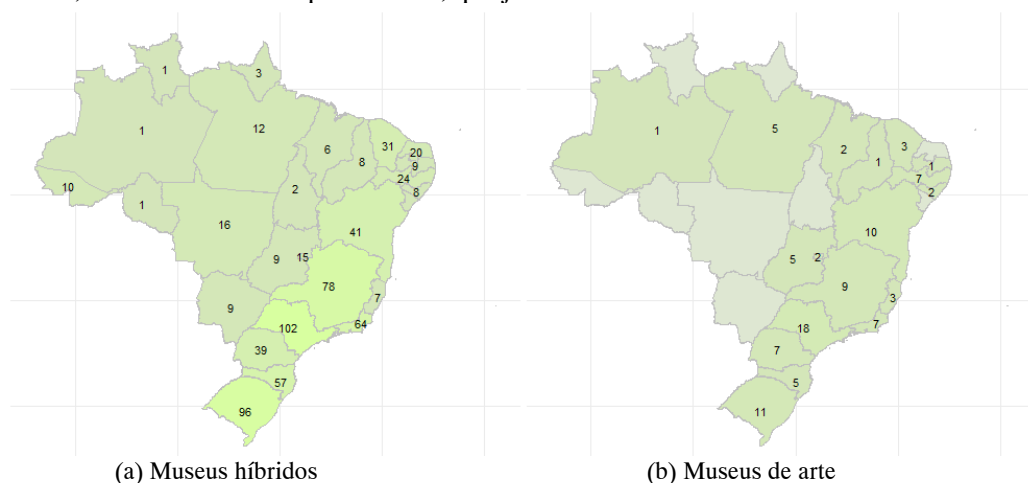


Figura 1. Museus com coleções de arte no Brasil, 2010.

Referência: Adaptado de Silva (2020)

Ao considerar todas as tipologias presentes nos 677 museus híbridos do Brasil, há um total de 26.205.685 bens culturais, dos quais 801.043 são bens culturais de Artes Visuais. Dentre esses 677 museus, 133 declararam não registrar suas coleções, o que totaliza 57.735 bens culturais de arte sem informações disponíveis para acesso aos usuários (Tabela 1).

Tabela 1. Bens culturais de Artes Visuais não registrados pelos museus híbridos no Brasil, 2010.

Região	Museus híbridos com coleções de arte	Total de bens culturais	Museus híbridos que registram suas coleções	Museus híbridos que não registram suas coleções	Total de bens culturais de artes	Bens culturais de arte não registrados em nenhum instrumento	
Sudeste	251	6.803.378	194	57	602.095	29.580	4,91%
Sul	192	2.228.443	165	27	91.717	13.865	15,11%
Nordeste	155	9.192.937	128	27	83.613	12.564	15,02%
Centro-Oeste	49	2.621.676	35	14	15.554	1.563	10,04%
Norte	30	5.359.251	22	8	8.064	163	2,02%
Brasil	677	26.205.685	544	133	801.043	57.735	7,20%

Referência: Elaborada pela autora com com dados do CNM (Instituto Brasileiro de Museus, 2010).

Já o total de bens culturais nos museus de arte é de 182.398, sendo que 70.059 estão localizados em instituições do Sudeste. Dentre esses 99 museus, 11 declararam não registrar suas coleções, o que totaliza 8.906 bens culturais não registrados em nenhum dos quatro instrumentos considerados no Questionário do CNM (Tabela 2). Portanto, havia no Brasil até o ano de 2010, pelo menos 66.641 bens culturais de Artes Visuais não registrados nos 776 museus com essa tipologia. Além disso, entre os museus híbridos que declararam catalogar suas coleções, não é possível afirmar a partir dos dados do CNM o número exato de bens culturais de arte ainda não processados, uma vez que alguns museus possivelmente fazem uso simultâneo dos instrumentos de registro, o que indica que o total de bens culturais de Artes Visuais não registrados é provavelmente superior a 66.641 para o período analisado.

Tabela 2. Bens culturais de Artes Visuais não registrados pelos museus de arte no Brasil, 2010.

Região	Museus de arte	Total de bens culturais	Museus de arte que registram suas coleções	Museus de arte que não registram suas coleções	Bens culturais de arte não registrados em nenhum instrumento	
Sudeste	37	70.059	34	3	1.049	1,49%
Sul	23	38.804	21	2	3.132	8,07%
Nordeste	26	63.966	20	6	4.725	7,38%
Centro-Oeste	7	4.224	7	0	0	0%
Norte	6	5.345	6	0	0	0%
Brasil	99	182.398	88	11	8.906	4,88%

Referência: Elaborada pela autora com dados do CNM (Instituto Brasileiro de Museus, 2010).

Enquanto nos museus de arte o instrumento de registro mais comum é a Ficha de catalogação, nos museus híbridos é o Livro de registro. Todavia, o CNM não informa se o Livro de registro é um instrumento ainda em uso ou se serve apenas como fonte de referência, como observa-se em algumas instituições. Entre os quatro tipos de instrumentos considerados no CNM, pode-se considerar a Ficha de catalogação/registo como o instrumento mais adequado, se apoiada em diretrizes atualizadas e reconhecidas internacionalmente que recomendam metadados para a identificação e a localização de cada objeto em uma coleção. O *Software/Programa* informatizado, se destinado à catalogação de objetos museológicos e se baseado em diretrizes, também pode ser um instrumento adequado para o registro de coleções. Nos museus híbridos, o *Software* é o instrumentos menos utilizado. Entre os 544 museus híbridos que afirmaram registrar o acervo, 273 usam Livro de registro, 256 usam Ficha de catalogação, 215 fazem Documentação fotográfica e 130 utilizam *Software*. Já entre os 88 museus de arte que afirmaram registrar o acervo, 42 usam Ficha de catalogação, 39 fazem Documentação fotográfica, 35 utilizam *Software* e 30 usam Livro de registro. Esses dados refletem o uso simultâneo dos instrumentos de registro.

3 Diretrizes para registro de objetos museológicos e programas de acreditação

Além de identificar os instrumentos de registro para bens culturais utilizados pelos museus brasileiros com coleções de arte, o objetivo principal deste trabalho era conhecer os padrões de estrutura de dados e diretrizes por eles adotados; todavia, esses tópicos não foram abordados no CNM. Padrões de dados, também conhecidos como *standards* na literatura americana, são úteis para assegurar consistência na representação de informações sobre os objetos e para auxiliar na padronização dos registros. Há os padrões de estrutura de dados, isto é, esquemas de metadados para objetos museológicos; padrões de valor de dados, que incluem vocabulários controlados, e padrões de conteúdo de dados que fornecem regras para o formato de informações registradas (Harpring, 2013; Gilliland, 2016).

Diretrizes apresentam recomendações para a realização de procedimentos, tal como o procedimento de catalogação presente na norma *Standard Procedures for Collections Recording Used in Museums* (SPECTRUM) (Collections Trust, 2014). Para a adoção de diretrizes e padrões de dados voltados para o registro de objetos museológicos, ponto central deste trabalho, o investimento de pessoal é importante (Harpring, 2020) para garantir consistência nas atividades de documentação; além disso, os padrões podem ser implantados em sistemas manuais ou informatizados. No caso de sistemas informatizados, existem *softwares* gratuitos para o gerenciamento de informações de coleções museológicas, como o Tainacan (<https://tainacan.org/>) e o Arches (<https://www.archesproject.org/>). Seja em sistemas manuais ou informatizados, é essencial conhecer o acervo para estabelecer prioridades e critérios e, então, elaborar programas e políticas institucionais. Na busca por modelos internacionais, a acreditação de museus pode ser uma maneira de estimular a adoção desses recursos; um exemplo é o programa do *Arts Council England* (2019), que consiste em estabelecer procedimentos mínimos para o gerenciamento de coleções e credenciar museus que os implementaram.

Por fim, outro aspecto a ser considerado para o incentivo de boas práticas para o setor são as colaborações entre instituições de diferentes países que possam ter desafios em comum no que se refere à catalogação e temas correlatos, por meio, por exemplo, de associações como o Comitê Internacional para Documentação do Conselho Internacional de Museus (CIDOC ICOM) ibero-americano, denominado *Documentando Museu Iberoamericano* (DOMINO) (<http://cidoc.mini.icom.museum/working-groups/domino/>), cuja primeira reunião ocorreu em dezembro de 2020 com aproximadamente trinta participantes de países como o Chile, México, Brasil, Espanha e Portugal, pois, assim cria-se um espaço amplo de discussão para profissionais de museus. Nessa primeira reunião, foi discutida a tradução para o português e o espanhol de materiais relevantes para o registro de objetos de museus, entre outros tópicos (DOMINO Working Group, 2020). Nesse sentido, outro projeto em andamento é a tradução para o português do *Art and Architecture Thesaurus* (AAT) do *Getty Research Institute*, um padrão de valor de dados, para o qual a tradução em espanhol está disponível *on-line*. Do mesmo modo, a realização de conferências e o oferecimento de treinamentos sobre documentação em museus exercem papel

importante para o diálogo entre pares e para a disseminação de conhecimento, a exemplo do workshop ministrado por Hagedorn-Saupe, M. & A. Ermert (2020).

A elaboração de um panorama sobre as condições locais no que diz respeito às coleções de arte presentes em museus brasileiros, bem como as iniciativas de traduções de recursos e a adoção de padrões de dados, visam o avanço do gerenciamento informacional sobre os acervos desses museus, sem mencionar discussões atuais de recursos relacionados à *Digital Art History* (DAH), que permitem o acesso às informações de arte via tecnologia digital e que podem transformar o modo de se fazer pesquisa em História da Arte (Pugh, 2020), campo a ser explorado pelos museus locais e em trabalhos futuros.

4 Considerações finais

Ponto crucial para a documentação em museus, infelizmente as diretrizes para o registro de bens culturais não foram abordadas no Questionário do CNM. Em vista disso, este trabalho alerta para a necessidade de estatísticas atualizadas e periódicas sobre os padrões de dados e diretrizes em uso pelos museus brasileiros com coleções de arte para ampliar o conhecimento sobre o perfil desses museus e, por conseguinte, com vistas à possibilidade de criação de um programa de acreditação, tal como realizado pelo *Arts Council England*, que ao menos estimule a adoção de procedimentos mínimos de documentação em museus. Este trabalho permitiu, por outro lado, a identificação do total de museus do Brasil com coleções de arte, de seus bens culturais e dos instrumentos de registro por eles utilizados, informações essenciais para nortear a elaboração e o andamento de projetos institucionais para a documentação de coleções museológicas. Foram apontadas a importância da capacitação de pessoal, do estabelecimento de parcerias entre museus e de iniciativas de traduções de materiais e de recursos voltados à comunidade ibero-americana para a ampliação do diálogo sobre documentação em museus em países com falantes de português e espanhol. Buscou-se contribuir, particularmente, para a implementação e aperfeiçoamento de boas práticas de registro e de gestão de informação de coleções artísticas no Brasil. No futuro, cabe analisar detalhadamente plataformas mais recentes como a *MuseusBr*, à medida que mais dados sobre os museus são nela adicionados e tendo em conta que tal plataforma apresenta uma forma diferente de classificar os museus quando comparada CNM. Por último, acreditamos que a elaboração e publicação de guias atualizados e específicos de museus de arte e de museus híbridos com coleções de arte é um projeto que se faz necessário para a extroversão desses acervos brasileiros.

Referências

- Arts Council England. (2019). *Accreditation Guidance*. London. https://www.artscouncil.org.uk/sites/default/files/download-file/Accreditation_Guidance_Mar_2019_0.pdf
- Collections Trust. (2014). *SPECTRUM 4.0: Padrão para gestão de coleções de museus do Reino Unido/Collections Trust*. Secretaria de Estado da Cultura; Associação de Amigos do Museu do Café; Pinacoteca do Estado de São Paulo. Gestão e documentação de acervos). Tradução da obra “SPECTRUM 4.0: the UK Collections Management Standard. https://issuu.com/sisem-sp/docs/spectrum_pt_net
- DOMINO Working Group. (2021, dezembro). In CONFERÊNCIA DO CIDOC, Genebra, virtual. <https://cidoc.mahgeneve.news/en/conference/>
- Gilliland, A. J. (2016). Setting the Stage. In M. Baca (ed.). *Introduction to Metadata* (3rd ed.). Getty Publications. <http://www.getty.edu/publications/intrometadata/setting-the-stage/>
- Hagedorn-Saupe, M. & A. Ermert (2020). *O desenvolvimento de normas para a gestão da informação em museus*. Workshop online, outubro de 2020. <https://www.bad.pt/eventos/event/workshop-o-desenvolvimento-de-normas-para-a-gestao-da-informacao-em-museus-sessao-online/>
- Harpring, P. (2013). *Introduction to Controlled Vocabularies: Terminology for art, architecture, and other cultural works* (2nd ed.). Getty Research Institute.
- Harpring, P. (2020). *Vocabularies for Access to History and New Knowledge: Focus on the Getty Vocabularies*. In: EuroMediterranean Conference, Cyprus, 2020, virtual. https://www.euromed2020.eu/application/files/3616/0430/8327/WS1_Patricia_Harpring.pdf
- Instituto Brasileiro de Museus. (2010). *Cadastro Nacional de Museus*. Dados fornecidos pelo IBRAM em 24 ago. 2018 via Sistema Eletrônico do Serviço de Informação ao Cidadão, conforme Lei de Acesso à Informação (LAI), Brasil. Lei n.º 12.527/2011. Brasília, DF.
- Instituto Brasileiro de Museus. (2011a). *Museus em números*. Brasília: DF, v. 1. https://www.museus.gov.br/wp-content/uploads/2011/11/museus_em_numeros_volume1.pdf
- Instituto Brasileiro de Museus. (2011b). *Museus em números*. Brasília: DF, v. 2A. https://www.museus.gov.br/wp-content/uploads/2011/11/Museus_em_Numeros_Volume_2A.pdf
- Instituto Brasileiro de Museus. (2011c). Questionário de cadastramento. In Instituto Brasileiro de Museus. *Museus em números*. Brasília: DF: Instituto Brasileiro de Museus, v. 1, pp. 21-32. https://www.museus.gov.br/wp-content/uploads/2011/11/museus_em_numeros_volume1.pdf
- Instituto Brasileiro de Museus. (2011d). *Guia dos museus brasileiros*. Brasília: DF. <http://www.museus.gov.br/guia-dos-museus-brasileiros>
- Instituto Brasileiro de Museus. (2014). *Pesquisa Anual de Museus (PAM)*. Brasília, DF: Instituto Brasileiro de Museus. <https://www.museus.gov.br/wp-content/uploads/2016/10/PAM-2014-Resultados-ASCOM.pdf>
- Instituto Brasileiro de Museus. (ca. 2015). *Pesquisa Anual de Museus (PAM)*. <https://www.museus.gov.br/pesquisa-anual-de-museus/>
- Lei n. 11.906, de 20 de janeiro de 2009. (2009, 21 janeiro). Cria o Instituto Brasileiro de Museus – IBRAM, cria 425 (quatrocentos e vinte e cinco) cargos efetivos do Plano Especial de Cargos da Cultura, cria Cargos em Comissão do Grupo-Direção e Assessoramento Superiores – DAS e Funções Gratificadas, no âmbito do Poder Executivo Federal, e dá outras providências. *Diário Oficial da União*, Casa Civil, Brasília, DF. http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2009/Lei/L11906.htm
- Miranda, R. M. (2016). Acervo em rede: interligando o patrimônio museológico brasileiro. *Anais do 3.º Seminário Serviços de Informação em Museus: colecionar e significar –*

- documentação de acervos e seus desafios*, 2014, SESC Bom Retiro. São Paulo: Pinacoteca do Estado de São Paulo, pp. 117-128.
- Pugh, E. (2020). Art History Now: Institutional Change and Scholarly Practice. *International Journal for Digital Art History*, 4, pp. 3.47-3.59. <https://doi.org/10.11588/dah.2019.4.63448>
- Silva, C. A. (2020). *Esquema de metadados para descrição de obras de arte em museus brasileiros: uma proposta*. 646 p. [Tese de Doutorado, Escola de Comunicações e Artes da Universidade de São Paulo]. <https://doi.org/10.11606/T.27.2020.tde-01032021-162722>
- Universidade de São Paulo. (2000). *Guia de Museus Brasileiros*. EDUSP, Imprensa Oficial do Estado.

GLOSSÁRIO TERMINOLÓGICO DEFINITÓRIO PARA SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO

Luander Falcão¹, Maristela Sanches Lima Mesquita²,
Benildes Coura M. S. Maculan³

¹Universidade Federal de Minas Gerais, *luanderfalcao@yahoo.com.br*,
ORCID iD 0000-0003-2417-6345

²Universidade Federal de Minas Gerais, *mesquita@ufop.edu.br*,
ORCID iD 0000-0003-4880-5599

³Universidade Federal de Minas Gerais, *benildes@gmail.com*,
ORCID iD 0000-0003-4303-9071

Resumo

(Introdução) Trata da criação de um glossário terminológico definitório para o domínio dos Sistemas de Organização do Conhecimento (SOC). A formação em Ciência da Informação, em específico no que tange à Organização da Informação, exige, além da compreensão de seus fundamentos e metodologias, entender os instrumentos de representação que são denominados de Sistemas de Organização do Conhecimento (SOC). **(Problema)** No recorte dos SOC como instrumentos de representação, verificou-se a carência de definições compiladas sobre os seus diferentes tipos, que explicitem suas características de maneira mais ampla. Na formação em CI, essa lacuna tem provocado ambiguidades que, às vezes, dificultam o entendimento desses instrumentos pelos discentes. **(Objetivo)** O objetivo deste estudo é o desenvolvimento de um glossário terminológico definitório, que possa esclarecer a terminologia sobre SOC, a partir da padronização das definições que serão elaboradas. **(Metodologia)** A metodologia empregada é caracterizada como de um estudo exploratório, descritivo e aplicado, utilizando como fonte de coleta de dados e terminologia a ISKO Encyclopedia of Knowledge Organization (IEKO). Os fundamentos teórico-metodológicos têm como fundamento a Linguística de Corpus para a construção do glossário. **(Resultados)** A partir de um recorte nos principais SOC na área de Ciência da Informação, tendo por base as normas NISO Z39.19 e a ISO 25964 (2011; 2013) é criado o glossário proposto que não constitui um produto definitivo e fechado; está aberto a sugestões e inclusões e pode ser acessado em sua hospedagem (URL: <https://appcode.com.br/socs/vocab/>). **(Conclusão)** Espera-se que o glossário terminológico criado permita melhorar o entendimento dos discentes sobre os distintos SOC, ao compilar as definições, imprimindo maior semântica a elas.

Palavras-chave: Glossário terminológico, Sistemas de Organização do Conhecimento, Representação do Conhecimento.

Introdução

O Sistema de Organização do Conhecimento (SOC) é um termo genérico, que nomeia uma gama de ferramentas, tais como: glossários, classificações, taxonomias, tesouros, ontologias, entre outros. No âmbito da Ciência da Informação, são ferramentas utilizadas para a representação da informação de um dado domínio, e têm como propósito organizar o conhecimento e codificá-lo em uma linguagem legível não só por humanos, mas também por sistemas computacionais. São caracterizados por diferentes estruturas e funções específicas, tendo em comum servir de apoio à organização da informação, facilitando a sua gestão e recuperação (Mazzocchi, 2018). Também dão apoio num mundo virtual de informação cada vez mais diversificado e populoso de dados, com tendência a um crescimento vertiginoso, por meio de múltiplos recursos, que vem expandindo de forma descentralizada, tornando-se um imenso repositório de documentos onde a recuperação da informação relevante torna-se quase impraticável (Souza; Alvarenga, 2004).

A perspectiva de uma comunicação sem ruídos, para que haja um entendimento compartilhado de terminologia e sentido, torna-se chave importante na recuperação da informação relevante para o usuário que busca conhecimento especializado. Nesse contexto, destaca-se os glossários, que têm como função elucidar definições sob a perspectiva de seus conceitos (Morais, Ramalho & Souza, 2019) e facilitar a interoperabilidade semântica, que permite a padronização de interpretação de significados das informações.

Zeng (2008) e Hodge (2000) consideram os glossários como um tipo de SOC que possuem grau semântico simples, relacionados na categoria de “Lista de termos”, e Guarino (2006) os considera um tipo de “ontologia leve”. McGuinness (2003) os define como listas alfabéticas de termos, juntamente com suas definições, que têm o propósito de eliminar ambiguidades (Mazzocchi, 2018). Segundo Vaz, Oliveira e Pierozzi Junior (2017, p. 83), os glossários representavam no passado uma “reunião de anotações, antes interlineares (glosas), sobre o sentido de palavras antigas ou obscuras encontradas nos textos”, e, atualmente, apresentam-se como “um conjunto organizado de termos de uma área de conhecimento e seus significados e definições”, e, numa conotação ainda mais relevante e atual, são conhecidos como “ferramentas de representação do conhecimento, juntamente com lista de termos e vocabulários controlados”.

A compilação de glossários vale-se da Linguística de Corpus (LC), que é o estudo de uma língua com base em evidências empíricas, que são obtidas a partir da exploração e análise de um dado *corpus* (Berber Sardinha, 2004), que é um conjunto homogêneo dessa tal língua. Tagnin (2009) destaca que se pode criar glossários empregando a LC como metodologia (terminologia baseada em *corpus*) ou como abordagem (terminologia direcionada pelo *corpus*).

Para a LC, um *corpus* é uma coletânea de textos, necessariamente em formato eletrônico, compilados e organizados segundo critérios ditados pelo objetivo de pesquisa a que se destina. O formato eletrônico permite que esses textos sejam investigados e analisados automaticamente, com o uso de ferramentas computacionais específicas (Tagnin, 2005, p. 21).

O autor ressalta que os textos utilizados para compor o *corpus* devem ser em linguagem natural, ou seja, não terem sido elaborados com o intuito de uma análise linguística.

Neste estudo optou-se pela criação de um glossário terminológico definitório, com o apoio da Linguística de Corpus, tendo em vista atender à demanda de informações sobre SOC de discentes em formação na Ciência da Informação. Destaca-se que se verificou a carência de definições compiladas sobre os diferentes tipos de SOC considerados como instrumentos de representação da informação, que explicitem suas características e funções, de maneira mais completa. Na formação em CI, essa lacuna tem provocado ambiguidades que, às vezes, dificultam o entendimento e o aprendizado desses instrumentos pelos discentes. Assim, este estudo pretende responder à seguinte questão: Como facilitar a compreensão dos termos utilizados na formação em Ciência da Informação, em específico no que tange aos conceitos dos diferentes tipos de SOC aos seus discentes?

O glossário será desenvolvido a partir de um *corpus* composto por textos da ISKO Encyclopedia of Knowledge Organization (IEKO), de onde serão extraídos os termos mais representativos e suas definições (Finatto, 2002). A identificação dos termos e excertos definitórios é realizada por meio de *softwares* específicos, que auxiliam o pesquisador, reduzindo o tempo de execução, custos e mão-de-obra especializada (Wendt, 2011). Os contextos definitórios são identificados a partir da citação de um termo técnico por um autor, quando na publicação científica encontram-se termos novos ou pouco conhecidos e quando o autor indica um termo equivalente em outra língua, sendo que dessa forma não é colocada a sua definição (De Lucca, 2006).

Metodologia

Para elaboração do glossário terminológico definitório foi aplicada a metodologia da Linguística de Corpus, adotando a abordagem da terminologia baseada em *corpus*. Nessa perspectiva, o *corpus* é formado pelos textos que compõem a ISKO Encyclopedia of Knowledge Organization (IEKO), em especial, daqueles que tratam sobre os SOC. Optou-se pela seleção de um *corpus* de pequeno porte, com menos de 80 mil palavras (Berber Sardinha, 2000), pois este estudo se trata de um glossário para um domínio específico, caracterizando-se como um trabalho terminológico, e, não, lexicográfico. A seleção de textos dá o norte para a elaboração das definições dos tipos de SOC que fazem parte do recorte trabalhado.

A identificação da lista dos termos que representam os SOC do recorte trabalhado, e dos excertos definitórios, é realizada por meio do *software* Freeware Concordance Program AntConc, que extrai, automaticamente, todos os sintagmas nominais. Os termos são, então, analisados e definidos, formando a base estrutural do glossário terminológico. Como insumo tecnológico para a edição do glossário, foi utilizado o *software* gestor de tesouros TemaTres é uma “aplicação web para a gestão de representações formais de conhecimento, tesouros, taxonomias e vocabulários multilíngues.” (disponível no site <https://www.vocabularyserver.com/>).

Os procedimentos metodológicos seguiram as etapas sugeridas por Krieger e Finatto (2004): (1) Planejamento do Trabalho; (2) Reconhecimento Terminológico e

Preparação Inicial; (3) Listagem de Termos; (4) Registro de Dados; (5) Fase Final, detalhadas a seguir.

(1) Planejamento do Trabalho: a) resolver questões relativas ao *corpus*; b) determinar a macroestrutura e a microestrutura do glossário; c) elaborar definições; d) elaborar texto introdutório ao glossário.

(2) Reconhecimento Terminológico e Preparação Inicial: estabelecer os princípios básicos para a construção do glossário, a partir da compilação do *corpus* a ser trabalhado.

(3) Listagem de Termos: estabelecer o conjunto de termos para compor o glossário.

(4) Registro de Dados: elaborar fichas terminológicas para compor a documentação do glossário.

(5) Fase Final: finalizar a proposta de elaboração do glossário.

Resultados

Os resultados são apresentados por etapas.

(1) Planejamento do Trabalho

a) resolver questões relativas ao *corpus*: compor o *corpus* para análise terminológica, para observar aspectos morfológicos (classe gramatical), sintáticos (função) e discursivos (contexto). O *corpus* foi formado pelos textos que compõem a ISKO Encyclopedia of Knowledge Organization (IEKO), em especial, daqueles que tratam sobre os SOC, na língua portuguesa. Para analisar esse *corpus* considerou-se utilizar o *software Freeware Concordance Program AntConc*.

O AntConc é um software freeware, uma ferramenta multiplataforma executável em versões para o Windows, Linux e Macintosh. Seu arquivo tem aproximadamente 4 Mb, por isso é considerado leve e dispensa a necessidade de instalação, o que torna a utilização mais fácil, sem a necessidade de um especialista na área. O modelo utilizado foi o AntConc 3.5.8, versão mais atual para Windows disponível em sua homepage onde constam suporte com manual, tutoriais em vídeo e grupos de discussão. Sua interface é simples e intuitiva com diferentes opções de análises que podem ser abertas na mesma janela (por uma aba). Nesse processo é possível descobrir a ocorrência de uma palavra, como e quantas vezes de sua incidência, permitindo realizar as principais técnicas de linguística de *corpus*, tais como frequência de palavras (*word frequencies*), colocações (*collocation*), concordância (*concordance*), *n-grams* e comparação de corpora a qualquer tipo de texto. A análise de *corpus* permite ver padrões de uso gramatical e outros, sendo útil para testar intuições sobre textos e /ou triangular resultados de outros métodos digitais (Barreiros, 2017). Foram obtidas 2.904 *Types* (número de palavras contadas apenas uma vez) e 19.913 *tokens* (número total de palavras no *corpus*). Na aba *concordance* (que mostra os resultados em formato *Key Word in Context – KWIC*), foi possível analisar as palavras linkadas ao texto do próprio *corpus*, assim, verificando o seu contexto.

b) determinar a macroestrutura e a microestrutura do glossário: a macroestrutura foi organizada em ordem alfabética, sem subentradas (cada termo é um novo verbete), para facilitar a busca de informações, tendo em vista que o público-alvo são os

discentes em formação, e, assim, não se pressupõe ser especialista na área. Já a microestrutura (estrutura interna dos verbetes), é formada com informações julgadas pertinentes aos discentes (público-alvo), tais como a definição do termo, equivalente em língua estrangeira, outros equivalentes (se houver), e a sigla do termo (se houver). Também são inseridas remissivas, que direcionam o leitor a outros verbetes (que foram utilizados na definição do termo; relação associativa), aos superordenados e subordinados (relação hierárquica) e aos seus equivalentes (relação de equivalência), que “representam uma opção de ampliar o uso pragmático do instrumento uma vez que auxilia o leitor a recuperar, de forma rápida e objetiva outras informações sobre o tema tratado” (Marini, 2013, p. 81), e torna o glossário com maior semântica, caracterizando-o como um SOC, segundo exemplo na letra “c”:

c) elaborar definições: foram planejadas as características para a definição dos termos, considerando que o termo deve ser observado no contexto em que é empregado (Almeida, 2006). Sendo assim, as definições são elaboradas a partir dos excertos extraídos do corpus, e, caso necessário, outras obras são consultadas, com o intuito de tornar a definição mais clara, precisa e completa. Neste glossário, a definição foi elaborada englobando os três tipos básicos: i) definição lexicográfica (explicita os significados e os seus empregos, sem redundância); ii) definição enciclopédica (fornece um conjunto de informações suficientes para entender o termo no contexto de uso); iii) definição terminológica (delimita o conceito de especialidade, representado pelo termo, dentro do sistema nocional, 1996). A seguir um exemplo:

Macroestrutura (entrada do verbete)

SISTEMAS DE CLASSIFICAÇÃO

Microestrutura (definição do termo)

Tipo de SOC de natureza conceitual, que pode ser criado em arranjos hierárquicos ou facetados, com notações numéricas, alfabéticas ou mistas, concebidos geralmente como sistemas gerais (abrangendo todos os campos do conhecimento), para qualquer domínio de objetos, assuntos ou conceitos, composto por conjunto de classes nas quais um determinado universo de coisas possa ser dividido de tal forma que as coisas que compartilham certas características, consideradas semelhantes, sejam reunidas em uma classe e distinguidas das coisas que não compartilham essas características. As classes são formadas sempre de acordo com alguma ordem, princípio, concepção, propósito ou interesse específico, ou com uma combinação desses elementos, obedecendo a uma certa ordenação em uma série, cuja sequência é determinada de acordo com algum princípio de organização.

Relação de equivalência (remissiva)

usado para (inglês): classification schemes

usado para (outros): esquema de classificação

Relação hierárquica (remissiva)

termo genérico (superordenado): sistemas de organização do conhecimento

termo específico (subordinado): sistemas de classificação bibliográfica; sistemas de classificação para gestão de registros

Relação associativa (remissiva)

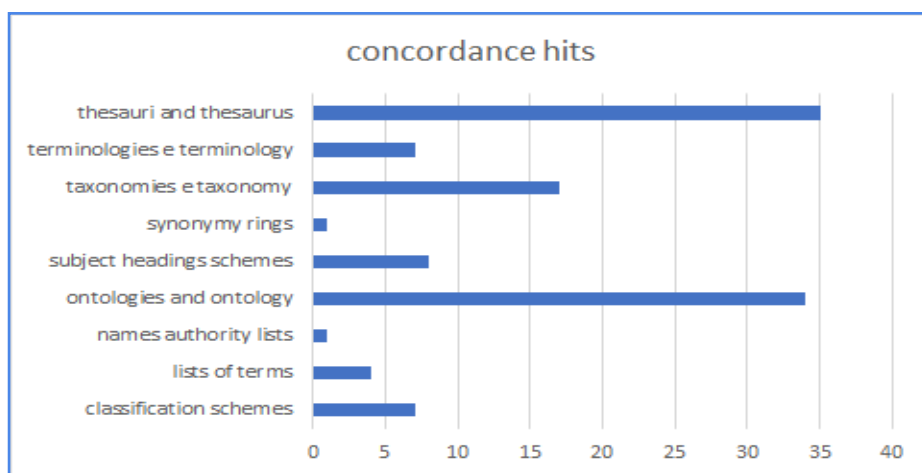
termos associados: classificação (processo); arranjos hierárquicos; arranjos facetados; notações numéricas; notações alfabéticas; notações mistas

d) elaborar texto introdutório ao glossário: o texto introdutório está em fase de elaboração, devido às diversas mudanças que foram necessárias, mas ela irá apresentar as informações que permite ao público-alvo entender a abordagem adotada na construção do glossário, assim como os limites de sua representação.

(2) Reconhecimento Terminológico e Preparação Inicial: estabeleceu-se como princípios básicos para a construção do glossário, a saber: atender às necessidades de um público-alvo (no caso, os discentes em formação em Ciência da Informação), a utilização de dados confiáveis da enciclopédia mantida pela ISKO internacional, e garantir que os textos do *corpus* sejam realmente representativos sobre a temática dos tipos de SOC, sob a perspectiva da Ciência da Informação. Para analisar o *corpus* utilizando o AntConc foi necessário converter os textos em .txt, e, após, foi feita a compilação dos excertos (textos) selecionados, que ficou com um total menor do que 80 mil palavras, o que o caracteriza como um *corpus* pequeno.

(3) Listagem de Termos: os critérios para a seleção dos termos se norteou na ocorrência deles no *corpus*, e que eles fossem considerados como um tipo de SOC, sob a abordagem da Ciência da Informação, ou seja, um instrumento de representação da informação, dentro do conjunto de instrumentos que são citados na norma NISO Z39.10 (2005, R2010), que são: listas de termos, anel de sinônimos, taxonomia e tesouro; e na norma ISO 25964 (2011; 2013), a saber: tesouro, sistemas de classificação, sistema de classificação para gestão de registros, taxonomia, lista de cabeçalho de assunto, ontologias, terminologias, lista de autoridades e anel de sinônimos. Os termos candidatos (figura 1) foram buscados no *corpus*, e aqueles que possuíam excertos definitórios, no próprio *corpus*, foram selecionados para compor o glossário, representando o recorte estabelecido para este estudo. Os termos candidatos sem excertos definitórios no próprio *corpus* foram, a princípio, descartados. A obtenção das equivalências, em português, dos termos do recorte foi efetuada por meio de buscas na literatura brasileira da área da Ciência da Informação, na língua portuguesa, fora do *corpus*.

Figura 1 – Frequência de concordância dos termos candidatos no *corpus*



(4) Registro de Dados: para cada termo do recorte de estudo, foi feita uma entrada no glossário, elaborando-se uma ficha terminológica na qual se registraram todos os dados e informações que foram necessários para a composição dos verbetes, que foram armazenados como documentação do glossário. A ficha terminológica foi elaborada com os seguintes campos: identificação do termo (ID único), termo de entrada, data de coleta, fonte do termo, categoria gramatical (masculino, feminino ou neutro, representada pelas letras M, F ou N, respectivamente), área temática, definição, fonte da definição, contexto do termo no *corpus* (excertos definitórios; exemplo de aplicação real do termo junto ao seu contexto, em um dos textos do *corpus*), fonte do contexto (texto no *corpus*), remissivas a termos equivalentes, inclusive em outros idiomas (traduções conceituais, com a fonte), remissivas a termos superordenados, remissivas a termos subordinados, remissivas a termos associados, notas para informação não prevista, responsável pelo preenchimento. Modelo de ficha terminológica com as informações dos campos a serem preenchidos, usada no trabalho:

Quadro 1 – Ficha terminológica do glossário

ID do termo:	ID único
Termo e categoria gramatical () Descritor preferido () Descritor não-preferido	Cat. Gramatical: usar M, F ou N Preferido: entrada em português Não-preferido: equivalentes em inglês e outros
Relação de Equivalência Use (USE)	Indicar o termo preferido
Usado Para (UP)	Indicar remissivas aos termos não preferidos em português
Usado Para (UP em inglês)	Indicar remissivas aos termos não preferidos em inglês
Relação hierárquica Termo Genérico (TG)	Indicar remissivas a ao termo em superordenação
Termo Específico (TE) –	Indicar remissivas ao(s) termo(s) em subordinação
Relação Associativa (TR)	Indicar remissivas ao(s) termo(s) em associação
Definição (Def.)	Definição do termo preferido – formulado, preferencialmente, a partir dos excertos do <i>corpus</i>
Nota de Escopo (NE)	Notas explicativas quanto ao escopo do termo preferido
Nota Histórica (NH)	Nota histórica referente ao termo preferido
Fonte de dados principal	<i>Corpus</i> da Encyclopedia of Knowledge Organization (IEKO)
Fonte de dados secundária	NISO Z39.10 (2005, R2010); ISO 25964 (2011; 2013)
Nome do coletor	Nome do responsável pela ficha
Data da entrada	Data da coleta (dd/mm/aaaa)

Fonte: dados do estudo.

(5) Fase Final: das fichas terminológicas foram extraídas as informações mais relevantes para o público-alvo (discentes em formação em Ciência da Informação), extraindo-se dados tais como: categoria gramatical, fonte da definição e exemplo retirado do *corpus*, para evitar sobrecarga de informações. Após a seleção das informações para constar nos verbetes, as entradas do glossário são organizadas em ordem alfabética. Neste momento, também é redigida a introdução do glossário, que vai apresentar todas as informações pertinentes a ele. Por fim, a forma de apresentação do glossário da seguinte maneira: a) cada entrada aparece somente com primeira letra maiúscula, à esquerda e em negrito; b) em seguida, vem a definição, na qual são grifados em negritos os termos em remissivas externas, para remeter a outras entradas, e termos associados; c) por fim, aparece a equivalência em outro idioma, também marcada em negrito.

Figura 2 – Estrutura do glossário no software TemaTres



Fonte: dados do estudo, TemaTres, 2021

Introdução do glossário: O Glossário Terminológico Definitório para Sistemas de Organização do Conhecimento foi criado para atender à demanda de informações sobre SOC, de discentes em formação, na Ciência da Informação. O tema versa sobre as definições dos diferentes tipos de SOC considerados como instrumentos de representação da informação, explicitando suas características e funções, de maneira mais completa, facilitando o entendimento e o aprendizado desses instrumentos pelos discentes.

Remissivas e abreviaturas utilizadas

Para permitir maior clareza e eficiência no uso do glossário, foram utilizadas as remissivas:

Simbologia – significado

Termo relacionado (TR) – Indica um tipo de relacionamento (não hierárquico)

Termo geral (TG) – Indica um termo superordenado

Termo específico (TE) – Indica um termo subordinado (ao Termo Geral ou parte do Termo Geral)

Use (USE) – Indica o termo preferido.

Usado por (UP) – Indica um termo não-preferido, referente a uma equivalência terminológica (usada somente para o controle de sinônimos). Não deve ser utilizada na representação temática.

Nota de escopo ou nota explicativa (NE) – Explica o contexto de uso de um termo ou apresenta alguma nota significativa para a adoção do termo.

Nota Histórica (NH) – Nota histórica referente ao termo preferido

Inglês (ing.) – para o termo equivalente na língua inglesa

Definição (DF) – Definição do termo preferido

Fonte de dados (FD) – fonte do *corpus*

Este glossário não constitui um produto definitivo e fechado; está aberto a sugestões e inclusões e pode ser acessado em sua hospedagem (URL: <https://app.code.com.br/socs/vocab/>). Trata-se de um glossário que propõe registrar a terminologia sobre os sistemas e organização do conhecimento (SOC), enquanto um instrumento de representação da informação, sob a abordagem da Ciência da Informação, utilizando como corpus os excertos definitórios contidos na Encyclopedia of Knowledge Organization (IEKO).

Considerações finais

Este estudo alcançou o seu objetivo ao criar um glossário terminológico definitivo sobre Sistemas de Organização do Conhecimento (SOC), para apoiar o entendimento sobre a temática, a partir de um *corpus* composto por textos da ISKO Encyclopedia of Knowledge Organization (IEKO). Todos os termos que compõem o glossário são acompanhados da sua definição, de maneira a limitar o sentido do conceito, evitando, assim, ambiguidades. O percurso de desenvolvimento do glossário evidenciou a relevância de se delinear a definição e todas as equivalências para um mesmo termo de entrada, já que alguns dos termos trabalhados fazem parte da terminologia de especialidade de outras áreas do conhecimento, além da Ciência da Informação.

Agradecimentos

A última co-autora agradece ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela bolsa de Produtividade em Pesquisa (PQ-2), Processo número 303650/2019-2.

Referências bibliográficas

- ANSI/NISO Z39.19-2005 (R2010), de 25 de julho de 2005. Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. Baltimore: National Information Standards Organization.
- Almeida, G. M. B. A. (2006). Teoria Comunicativa da Terminologia e a sua prática. São Paulo: Alfa.
- Barreiros, L. L. S. (2017). O uso de ferramentas computacionais na elaboração do Vocabulário de Eulálio Motta: AntConc e FLEx. A Cor das Letras. Feira de Santana, v. 18, n. 2, p. 216-241.
- Berber Sardinha, T. (2000). Lingüística de Corpus: Histórico e problemática. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 16(2), 323–367. <https://doi.org/10.1590/S0102-44502000000200005>
- Finatto, M. J. B. (2002). O papel da definição de termos técnico-científicos. *Revista da ABRALIN. Rio de Janeiro. Vol. 1, n. 1 (jul. 2002)*, p. 73-97. <https://lume.ufrgs.br/handle/10183/184253>
- Guarino, N., & Guizzardi, G. (2006). *In the Defense of Ontological Foundations for Conceptual Modeling*. 18, 13.
- Hodge, G. (2000). *Systems of knowledge organization for digital libraries: Beyond traditional authority files*. Digital Library Federation, Council on Library and Information Resources.
- ISO 25964-1:2013, de 04 de março de 2013. Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies. Vernier: International Organization for Standardization.
- Krieger, M. D. G. (2004). Finatto, Maria José Bocorny. Introdução à terminologia.
- Larivière, L. (1996). Comment formuler une définition terminologique. *Meta: journal des traducteurs / Meta: Translators' Journal*, 41(3), 405-418. <https://doi.org/10.7202/003401ar>
- Lucca, J. L. D. (2006). Identificação de padrões recorrentes no discurso técnico e científico para a extração automática de candidatos a contextos definitórios em língua portuguesa. *Intercâmbio. Revista do Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem. ISSN 2237-759X, 15(0)*, Article 0. <https://revistas.pucsp.br/index.php/intercambio/article/view/3683>
- Marini, S. (2013). *Da tradução terminológica em glossário temático na área de saúde suplementar*. Brasília: Departamento de Línguas Estrangeiras e Tradução.
- Mazzocchi, F. (2018). *Knowledge organization system (KOS): an introductory critical account*. *Ko Knowledge Organization*, 45(1), 54-78. Recuperado 2 de abril de 2021, de <https://www.nomos-elibrary.de/10.5771/0943-7444-2018-1-54/knowledge-organization-system-kos-an-introductory-critical-account-volume-45-2018-issue-1>
- Mcguinness, D., Fikes, R., Stein, L., & Hendler, J. (2003). *DAML-ONT: An ontology language for the semantic web* (p. 93).
- Morais, M. W., Souza, J. L., & Ramalho, R. A. S. (2019, junho 16). *Representação de um microtesauro em SKOS: Um estudo metodológico* [Preprint]. <http://eprints.rclis.org/38636/>
- Sardinha, T. B. (2004). *Lingüística de corpus*. Editora Manole Ltda.
- Souza, R. R., & Alvarenga, L. (2004). A Web Semântica e suas contribuições para a ciência da informação. *Ciência da Informação*, 33(1), 132–141. <https://doi.org/10.1590/S0100-19652004000100016>
- Tagnin, S. E. O. (2005). O jeito que a gente diz: expressões idiomáticas e convencionais- inglês e português. São Paulo: Disal.

- Vaz, G., Oliveira, L. M., & Pierozzi, I., Jr. (2017). Visualização de glossário em sistemas de recuperação de informação (Glossary visualization in information retrieval systems) [In Portuguese]. *Proceedings of the 11th Brazilian Symposium in Information and Human Language Technology*, 83–92. <https://www.aclweb.org/anthology/W17-6611>
- Wendt, I. S., & Vieira, R. (2011). *Extração de contextos definitórios a partir de textos em língua portuguesa* (Dissertação de mestrado). Faculdade de Informática – Pontifícia Universidade Católica, Rio Grande do Sul, RS, Brasil. <http://tede2.pucrs.br/tede2/handle/tede/5206>
- Zeng, M. L. (2008). Knowledge Organization Systems (KOS). *KNOWLEDGE ORGANIZATION*, 35(2-3), 160-182. <https://doi.org/10.5771/0943-7444-2008-2-3-160>

(Página deixada propositadamente em branco)

IMPORTÂNCIA DO CONTROLE DE VOCABULÁRIO EM UM REPOSITÓRIO INSTITUCIONAL: ESTUDO DE CASO

Isidoro Gil-Leiva¹, Isaque Katahira², Mariângela Spotti Lopes Fujita³

¹Universidad de Murcia, isgil@um.es, <https://orcid.org/0000-0002-7175-3099>

²Universidade Estadual Paulista “Júlio de Mesquita Filho”, Unesp,
isaque.katahira@fatec.sp.gov.br, <https://orcid.org/0000-0001-5800-9890>

³Universidade Estadual Paulista “Júlio de Mesquita Filho”, Unesp,
mariangela.fujita@unesp.br, <https://orcid.org/0000-0002-8239-7114>

Resumo

Os Repositórios Institucionais armazenam expressivas quantidades de recursos e documentos digitais relacionados a produções acadêmicas e/ou científicas. Assim, a busca e o acesso temático por meio de assuntos ou palavras-chave desempenha um papel muito importante para facilitar a busca e o acesso as informações armazenadas nos Repositórios Institucionais. Para que a recuperação de documentos ocorra de modo efetivo, faz-se necessário padronizar esse sistema de indexação, para minimizar os ruídos intrínsecos da linguagem natural, como ambiguidades, sinônimas e inadequações de registros e ampliar a convergência entre a linguagem natural dos usuários e os vocabulários controlados disponibilizados no arquivamento e organização dos repositórios. Neste contexto, o estudo aqui apresentado têm como **objetivos**: analisar o Repositório Institucional da Unesp do ponto de vista da utilização de vocabulário em linguagem natural e em linguagem controlada; analisar a lista de assuntos disponível no Repositório; e explorar possibilidades oferecidas pela análise de *logs* de buscas dos usuários para atualização de vocabulários controlados. **Metodologia**: Para tanto, realizou-se uma pesquisa bibliográfica no Banco de Dados *Scopus* sobre repositórios e o uso de *logs* para criar e atualizar vocabulários controlados. Posteriormente, realizamos uma análise da lista alfabética em linguagem natural disponibilizada pelo Repositório da Unesp aos seus usuários para auxiliar na busca por informações. Por fim, foi realizada a coleta de *logs* de busca dos usuários do Repositório Institucional da Unesp. O conjunto de dados inicial utilizado na análise continha 480.742 registros de *logs*, armazenados no período de 01/07/2019 a 31/12/2019. **Resultados/ Conclusões**: a análise realizada, revela que o Repositório Unesp adota a linguagem natural sem nenhum controle de vocabulário ou uso de instrumentos de padronização; nas listas de palavras-chave indexadas dos documentos, observamos que não há controle de vocabulário o que resulta em uma lista com variações terminológicas que podem prejudicar a recuperação da informação. Dentre os problemas estão variações de singular ou plural, digitação com maiúscula ou minúscula, uso de aspas e outros sinais como hífen. Através da análise realizada neste estudo, análise de *logs*, foi possível obter, uma lista de expressões de busca contendo palavras simples, palavras compostas ou frases para possibilitar a criação ou atualização de um

vocabulário controlado. A metodologia aqui apresentada mostrou-se eficaz pois, permitiu uma grande depuração de expressões de busca de usuários que produziu um corpus terminológico proveniente diretamente da linguagem utilizada nesse sistema. A lista de palavras-chave e a análise de logs podem oferecer possibilidades de criação e manutenção de ferramentas de controle de vocabulário, desde a lista alfabética de palavras-chave e até tesouros, mas precisam passar por rigoroso controle terminológico e linguístico para o aproveitamento.

Palavras-chave: Vocabulário controlado, Análise de logs, Repositórios institucionais, Avaliação de repositórios.

Introdução

Os Repositórios Institucionais (RI) tiveram surgimento após a Declaração de Berlim, em 2003, para facilitar o acesso ao conhecimento e divulgação do conhecimento científico e fortalecer o trabalho desenvolvido nas universidades (Serrano Vicente, Melero Melero e Abadal, 2014). Assim, os RIs foram criados com a tripla missão de serem instrumentos digitais de coleta, conservação e divulgação da gestão do conhecimento, de ensino e de pesquisa.

O repositório institucional Unesp (repositorio.unesp.br) da Universidade Estadual Paulista “Júlio de Mesquita Filho”, implantado em 2013, utiliza o software DSpace e o padrão Dublin Core para a definição de um conjunto de metadados com a finalidade de atender necessidades de representação e divulgação referentes à produção científica dos pesquisadores (Vidotti, Bastos, Ferreira, Grisoto, Assumpção e Silva, 2015). Sua implantação foi possível com a aplicação dos processos de coleta, conversão e importação automática dos registros referentes a produção científica de pesquisadores vinculados à Universidade indexada, inicialmente coletadas nas bases de dados *Web of Science* e *Scopus* e publicada em periódicos da *Scientific Electronic Library Online* (SciELO) (Assumpção, Silva, Ferreira e Bastos, 2014) e, na sequência, na *PubMed*, Currículos da Plataforma Lattes e registros MARC 21 importados da Biblioteca Digital de Teses e Dissertações (C@THEDRA) e também da Biblioteca Digital de Trabalhos de Conclusão de Curso (C@PELO) da Unesp (Vidotti, Bastos, Grisoto, Arakaki e Ferreira, 2016).

Atualmente, o Repositório Unesp armazena expressiva quantidade de recursos e documentos digitais de diferentes tipos de produção além da científica e acadêmica, tais como, a produção administrativa, artística, comemorativa e técnica. Apesar de dispor de interface de busca e de “filtros” diversos para refinamento das buscas, a linguagem natural é utilizada sem uso de ferramentas de controle de vocabulário tanto para a representação de conteúdos quanto para a busca e recuperação.

Diante do exposto, justifica-se a proposição desta investigação no sentido de buscar recursos que colaborem para o conhecimento do vocabulário usado pelos usuários no Repositório Institucional da Unesp, 5.º maior repositório brasileiro e 28.º maior repositório mundial, posto que, a partir dos dados extraídos, será possível analisar o comportamento informacional desses usuários e propor melhorias para garantir a recuperação das informações desejadas (WEBMOTRICS, 2021).

Todo documento é produzido para comunicar, porém, para isso, não basta que seja publicado, é indispensável a sua descoberta e a sua leitura pelos usuários, fatores associados à qualidade dos metadados registrados. Na busca e descoberta de informações pelo usuário, o acesso temático por meio de assuntos ou termo de busca desempenha um papel essencial. Para descobrir tais informações, o sistema de indexação deve estar minimamente padronizado para reportar ao usuário documentos que estão relacionados aos termos de busca utilizados pelo usuário. Portanto, a padronização do sistema de indexação busca minimizar os ruídos intrínsecos da linguagem natural, tais como ambigüidades, sinonímias e inadequações de registros. Ademais, cabe promover pesquisas que ampliem a convergência entre a linguagem natural dos usuários e os vocabulários controlados disponibilizados para arquivamento e organização, de forma que a recuperação possa ser efetiva. Neste sentido, a *Transaction Log Analysis* (TLA), análise de *logs* de transações dos usuários ao acessarem os RIs, método adotado neste artigo, torna-se uma ferramenta promissora, visto que permite a padronização e atualização dos RIs, baseados no percurso investigativo do usuário.

Durante a operação e uso dos bancos de dados, são gerados registros de transações (ações executadas). Penniman (2008, p. 20) apontou Parker e Paisley (1966) como aqueles que identificaram o potencial da Transaction Log Analysis, análise de registros de transações (doravante, *log analysis*) para estudar o comportamento dos usuários durante a busca por informações de forma precisa e discreta ao mesmo tempo. Penniman também escreveu que o primeiro trabalho centrado no usuário no final dos anos 1960 e início dos anos 1970 foi realizado em sistemas de busca de banco de dados como RECON, Medline ou ERIC da NASA. E no final dos anos 1970 e no início dos anos 1980, a OCLC começou a analisar os registros do catálogo de acesso público (OPAC) para entender seu uso e melhorar sua implementação e *design*. O relatório de Meister e Sullivan (1967) intitulado “Avaliação das reações do usuário a um protótipo de sistema de recuperação de informação on-line” é considerado o primeiro trabalho utilizando a metodologia de análise de log para conhecer o comportamento do usuário de forma discreta a ser utilizada no desenvolvimento de um banco de dados da NASA. Além da análise de registros, eles coletaram dados adicionais por meio de observação, entrevistas e questionários com os usuários. Analisaram mais de seis mil logs de buscas coletados durante seis semanas na Base de Dados RECON para saber a frequência de uso, a velocidade de recuperação online em relação ao sistema offline, a opinião do usuário e o número de buscas nos campos autor, assunto, data, fonte corporativa ou número de referência.

Peters (1993, p. 42) afirmou que a análise de log se refere tanto a uma forma de monitorar o sistema quanto a uma forma de observar, geralmente de forma discreta, o comportamento do usuário no próprio sistema. Por sua vez, Jansen, Isak e Spink (2008) observaram que a análise de logs pode ser associada a métodos de investigação que cobrem várias subcategorias, como análise de logs da web, de bancos de dados específicos para diversas áreas e ambientes variados de pesquisa que têm a missão de disponibilizar materiais aos usuários a partir de sua solicitação. Para esses autores, a análise de logs é um método de medição não intrusivo que permite a análise indireta de conteúdo e secundária em: I) interação dos usuários com o sistema para obtenção da informação; II) termos ou frases das consultas de pesquisa ou

nos resultados obtidos e III) dados quantitativos já coletados por terceiros para realizar uma análise diferente dos mesmos. A propósito da análise de conteúdo, mesmo que indireta, tem função importante para a criação e manutenção de vocabulários controlados, em especial de tesouros, porque os registros de consultas poderão contribuir com novos termos ou termos sinônimos e, assim, melhorar a formulação de buscas por assunto.

Os logs têm sido usados também para melhorar os mecanismos de pesquisa na web (Pu & Chien, 2004; Lee, Huang & Chao, 2007; Castellani et al., 2009) ou para identificar os termos usados pelos usuários para atualizar um vocabulário controlado (Xiaozhong, et al., 2008; Dixon & McGowan, 2010; Vallez, Pedraza-Jimenez, Codina, Blanco & Rovira, 2015), entre outros.

Neste contexto, o trabalho aqui apresentado tem um triplo objetivo:

- Analisar o Repositório Unesp do ponto de vista da utilização de vocabulário em linguagem natural e em linguagem controlada;
- Analisar a lista de assuntos disponível no Repositório Unesp, e finalmente;
- Explorar possibilidades oferecidas pela análise de logs para atualização de vocabulários controlados.

Metodologia

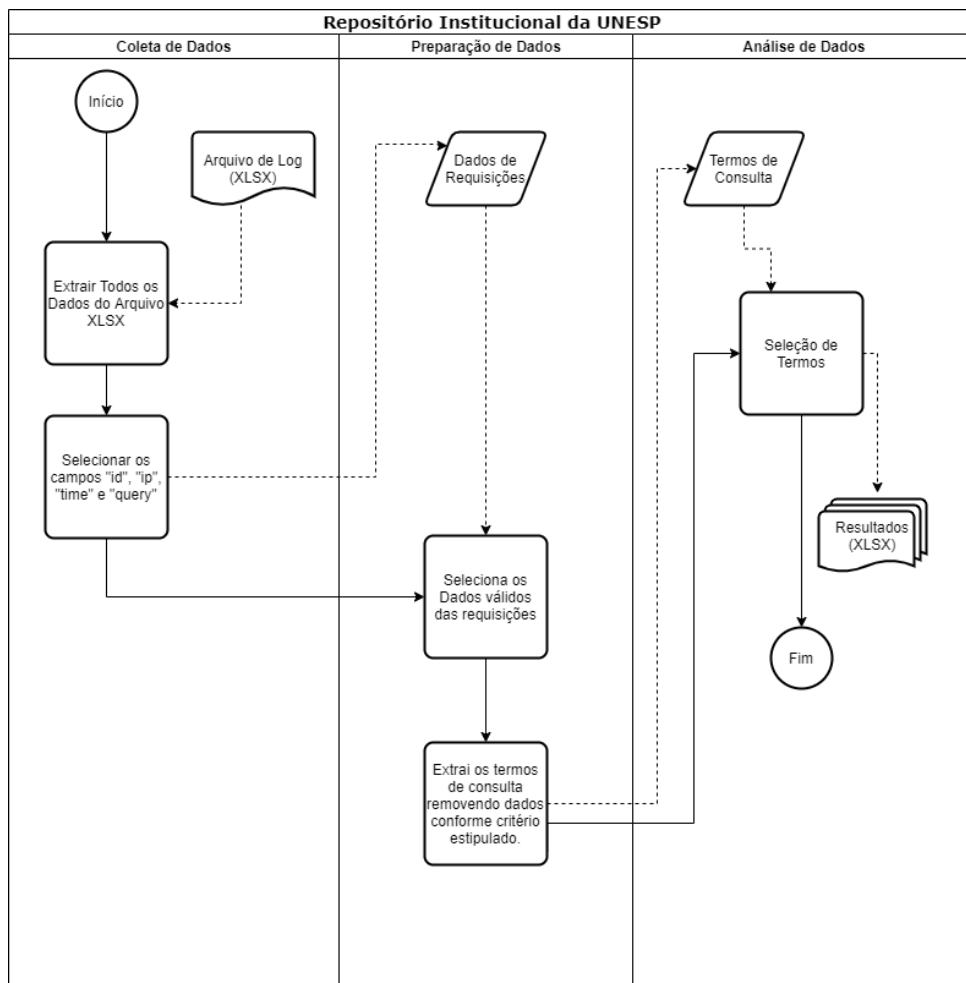
Realizamos, levantamento bibliográfico no banco de dados *Scopus* sobre o uso de repositório e análise de *logs* para criar e atualizar vocabulário controlado combinando os termos ‘*institutional repository*’, ‘*headings*’, ‘*terminology*’ e ‘*subject*’. Posteriormente, realizamos uma análise da lista alfabética de assuntos em linguagem natural que o repositório da UNESP disponibiliza aos seus usuários para buscarem informações no repositório.

Após, ao considerar TLA como registro de informações que permitem verificar e analisar a ocorrência de determinados eventos, a coleta de *logs* de busca dos usuários do RI da UNESP permitiu a extração da linguagem natural empregada pelos usuários para realizar buscas conforme apresentado no fluxo de trabalho de análise de logs (Figura 1) que se constitui no método idealizado para esta investigação composto de três fases sequenciais: coleta, preparação e análise de dados.

A coleta de dados foi realizada entre os dias 01 de julho ao dia 31 de dezembro de 2019 pelos responsáveis do Repositório Institucional UNESP que disponibilizaram um total 480.742 requisições de usuários registradas em logs da base Solr, de todas as áreas do conhecimento, exportável pelo formato XLSX..

Em posse dos dados exportados os arquivos referenciais foram submetidos a um processo de codificação dos endereços IP para realizar o processamento inicial e preservar o anonimato com relação à identidade dos usuários. O conceito de sessão fixado na análise deste conjunto de dados incorpora um limite de tempo de 24 horas, o intervalo permite unificar os registros dentro deste limite de tempo, mas, por outro lado, implica que o mesmo IP pode aparecer em dois dias diferentes sem necessariamente implicar um mesmo usuário, equipamento ou local de conexão. Portanto, a análise é feita em um recorte de caráter estático, sem entrar na dinâmica do processo.

Figura 1. Fluxo de trabalho de análises de logs



Fonte: elaborado pelos autores

A preparação dos dados com a finalidade de obter termos ou conjunto de termos válidos na atualização de vocabulários controlados utilizou a linguagem *Python* (<https://www.python.org>), para realizar duas etapas de limpeza dos dados coletados dos logs de busca dos usuários. Tal linguagem torna exequível a análise de grandes conjuntos de dados. Contudo, o conjunto inicial de dados apresentava 480.742 registros de logs. Assim, a linguagem *Python* foi utilizada para tornar viável o carregamento e análise dos dados empregados nesta pesquisa. A primeira etapa de limpeza dos dados foi realizada através da exclusão de: I) todos os registros sem nenhum caractere; II) os registros feitos apenas com a utilização de filtros; III) registros indicativos apenas de códigos; IV) registros repetidos; V) registros somente com caracteres numéricos; VI) registros com menos de três caracteres incluindo os espaços; VII) registros com mais de seis palavras.

Tendo em vista os critérios de limpeza descritos, utilizou-se algumas expressões regulares para automatizar a exclusão de registros irrelevantes. Para especificar um conjunto de sequências as expressões regulares “são utilizadas para detectar sua ocorrência em textos em linguagem natural, buscando ser o mais generalista possível” (Ferreira, 2019, p. 45), de modo a permitir a identificação e a automatização do processamento de um grande volume de textos. Conforme Jurafsky e Martin (2000, p. 3, tradução nossa), “formalmente, uma expressão regular é uma notação algébrica que caracteriza um conjunto de símbolos, essas notações podem ser usadas em todas as linguagens de programação para facilitar o processamento de corpus compostos por textos”, podendo auxiliar tanto na limpeza de dados como na identificação de padrões de caracteres irrelevantes.

Quadro 1. Expressões regulares em *Python* para limpeza inicial dos logs

```
r'^d' : PatternMode.REGEX,
r'^[^\w]+\w*' : PatternMode.REGEX,
r'[A-Z]\. [A-Z]\.' : PatternMode.REGEX,
r'[A-Z]\. [A-Z]\' : PatternMode.REGEX,
r'&\w*' : PatternMode.REGEX,
r'\w*' : PatternMode.REGEX,
```

Fonte: Elaborado pelos autores

Sequencialmente, foram realizados outros procedimentos da segunda etapa de limpeza dos dados, a partir da observação da lista resultante da primeira etapa de limpeza: I) unificação de termos nas formas maiúscula e minúscula; II) retirada do padrão formado com indicação de apenas duas iniciais separadas por ponto, tipo “J. D.”; III) remoção dos registros de nomes próprios pela identificação com nomes presentes na Classificação de gênero em nomes brasileiros, com base em dados do CENSO de 2010¹; IV) remoção de substantivos próprios referentes a cidades brasileiras, de acordo com dados do IBGE² e por fim, VI) seleção dos termos candidatos a criação e atualização de vocabulários controlados.

A fase de análise dos dados utilizou os dados resultantes das duas etapas de limpeza realizadas na fase de preparação dos dados para discussão com a literatura na terceira categoria de análise do próximo item, “Possibilidades oferecidas pela análise de logs”, tendo em vista o terceiro objetivo enunciado por esta investigação.

¹ Disponível: <https://brasil.io/dataset/genero-nomes/nomes/>. Acesso em: 23 mar. 2021.

² Disponível em: http://blog.mds.gov.br/redesuas/wp-content/uploads/2018/06/Lista_Munic%C3%A0prios_com_IBGE_Brasil_Versao_CSV.csv. Acesso em: 31 mar. 2021.

Resultados e discussão

Os resultados obtidos da análise de uso de controle de vocabulário, vocabulário controlado, linguagem natural e lista alfabética de assuntos do RI-UNESP bem como da aplicação de método de análise de logs, foram analisados e discutidos com a literatura revisonada na perspectiva de três categorias de análise atribuídas conforme objetivos desta investigação: Controle de vocabulário, Lista alfabética de assuntos disponível no RI-UNESP e Possibilidades oferecidas pela análise de logs.

1. Controle de vocabulário

Após as análises realizadas no RI-UNESP, verificamos que a representação temática por palavras-chave adotada é a linguagem natural dos autores, disponibilizada nos registros coletados, assim como no arquivamento sob demanda e no auto-arquivamento pelos autores de teses e dissertações. O repositório não fornece ao usuário, quando este vai inserir documentos, ferramentas de controle de vocabulário para favorecer a consistência e o controle de qualidade dos registros, embora esse tipo de ferramenta seja considerada comum em sistemas de informação.

A partir do problema observado, possíveis soluções devem ser buscadas como, por exemplo, o uso dos repositórios institucionais do *Digital Commons*, que, atualmente, são utilizados por mais de seiscentas instituições, entre os recursos do *Digital Commons* estão: um menu expansível, aproximadamente mil categorias organizadas em três níveis que facilitam a adição de conteúdo nos repositórios pelos usuários e, a partir da mesma, a referida taxonomia em linguagem controlada proporciona fácil acesso aos recursos (Bundza, 2014); o uso de vocabulários controlados como os utilizados nos repositórios da *Texas A&M University*, nestes os vocabulários controlados são usados para garantir a consistência e a precisão dos metadados (Ho & Stokes, 2019); e o uso de vocabulários para tópicos, que são utilizados no repositório DigitalHub, este vocabulário permite localizações geográficas e preenchimento automático de nomes durante a entrada de dados (Ilik, Hebal, Olson, Wishnetsky, Pastva, Kubilius & Holmes, 2018).

2. Lista alfabética de assuntos disponível no RI-UNESP

Nas listas de palavras-chaves indexadas dos documentos, observamos que não há controle de vocabulário. As palavras-chaves coletadas dos metadados estão disponíveis em uma lista alfabética para consulta dos usuários durante a busca, mas não há mecanismos de controle, o que resulta uma lista com variações que podem prejudicar a recuperação da informação. Dentre os problemas estão variações de singular/plural, digitação com maiúscula ou minúscula, uso de aspas e outros sinais como hífen. Destacamos que a alteração em qualquer caractere registrado pode influenciar nos resultados obtidos durante a recuperação.

Através da análise de palavras-chave dos autores no repositório da UNESP, percebemos inconsistências na recuperação como, por exemplo, ‘Profissão Docente’ que retornou 1 documento de busca, ‘Profissão docente’ com o retorno de 10 documentos e ‘profissão docente’ que retornou 1 documento; além de problemas de

compatibilidade de codificação, por exemplo, a palavra-chave dos autores foi indexada por ‘“campos rupestres”’, no qual os códigos “ e ” representam os caracteres de aspas duplas. Em consulta ao artigo original, a palavra-chave empregada pelos autores é *campos rupestres* (sem aspas), porém ao coletar os dados desse artigo, o sistema do Repositório Institucional da Unesp indexou como ‘“campos rupestres”’. Desse modo, quando o usuário buscar exatamente pela palavra-chave identificada pelos autores, *campos rupestres*, não haverá correspondência no repositório, e portanto, não haverá recuperação do documento.

Verificamos exemplos no Quadro 2, que ressaltam a não realização de tratamento de linguagem natural no Repositório UNESP para diminuir essas variações.

Quadro 2. Exemplo de erros detectados

Letras ou palavras sem sentido	B[1]; P[1]; S[1]; V%[1]
Maiúsculas-Minúsculas	Árvores tropicais [1] / árvores tropicais [1]; C-BAND [4] / C-band [23] C-Fos [2] / C-fos [1] / c-Fos [2] E-Commerce [2] / E-commerce [2] / e-commerce [1] Rabbit [45] / rabbit [82] Y Chromosome [3] / Y chromosome [7]
Singular / Plural	Kaleidoscope [2] / Kaleidoscopes [3]
Acentos	Karate-do [1] / Karate-dô [1]
Outros	Karl Marx [2] / Karl Marx (1818-1883) [1] N potencial mineralizável [1] / N potencialmente mineralizável [2] Y chromosome [7] / Y-chromosome [2] “Candidatus” Liberibacter spp. “Desenvolvimento Pacífico” “Ethos” della mobilità del servizio sociale brasiliano “Geração de 45” “Kissing spines” “Mãe Cristina” Municipal Nursery “Ovsynch” “Pivô” para a Ásia “São Paulo do School” Program “Upconversion” “de novo” assembly “Órfãos do Eldorado” #ErectileDysfunction #ProstateCancer #elenão #elesim –Active Learning –Aprendizagem Ativa –Educação Médica

Fonte: Elaborado pelos autores a partir do Repositório Unesp

Para evidenciar alguns dos ruídos descritos no Quadro 2, a Figura 2 mostra a localização dos erros detectados na página do Repositório UNESP quando da listagem por temas.

Figura 2: Exemplo localização de erros na lista por tema

Fonte: <https://repositorio.unesp.br/browse?type=subject>

Alguns dos problemas visualizados na Figura 2 poderiam ser resolvidos com certa facilidade. No momento da análise a lista alfabética de termos do Repositório Unesp tem 310.134 palavras-chave dos autores, das quais provavelmente mais de 30% devem ser eliminadas mediante rotinas implementadas ou através de algum software de limpeza de dados como o OpenRefine, por exemplo, que permite a identificação e resolução de usos inconsistentes do singular/plural, erros ortográficos simples, diferenças de espaçamento ou capitalização. Neste estudo, empregamos três rotinas de limpeza simples para eliminar palavras-chave com menos de três letras, palavras idênticas presentes em maiúsculas e minúsculas e palavras-chave com números. Através deste procedimento, foram eliminadas 44.705 palavras-chave, o que representou 14,4% do total, ficando a lista posteriormente com 265.429 palavras-chave de autores.

3. Possibilidades oferecidas pela análise de logs

A Tabela 1 apresenta um resumo dos trabalhos realizados e os resultados obtidos nos logs de busca dos usuários do RI UNESP.

Tabela 1 – Especificação das buscas dos usuários

Registro de consultas dos usuários	
Período analisado	31-07-2019/31-12-2019
Número de logs (expressões de busca)	480.742
Número de expressões de busca após a primeira etapa	50.968
Número de expresiones de busca limpas após a segunda etapa	25.451

Fonte: Elaborado pelos autores

Por expressões de busca, entendemos tudo o que foi escrito pelos usuários para realizar uma busca no RI UNESP, desde uma simples palavra (beleza), várias palavras (centro de processamento de dados) ou uma frase (vazamento de hipoclorito no tecido periapical).

O período selecionado de seis meses foi considerado adequado, pois permitiu que um número suficiente de buscas fosse coletado e analisado para os fins deste trabalho. Foram processadas 480.742 buscas digitalizadas pelos usuários. Muitos dos logs continham informações que não eram úteis. Após o trabalho de depuração realizado (Tabela 2), obtivemos 25.451 termos únicos de busca de usuários, o que representa apenas 5,29% dos logs iniciais que foram extraídos do sistema.

Tabela 2 – Número de buscas excluídas de acordo com cada critério estabelecido

Exemplos de linhas de busca eliminadas	
Exclusão de buscas com menos do que 3 caracteres	203
Exclusão de busca por apresentar nomes próprios	11.342
Exclusão de busca por apresentar nomes de cidades	124
Exclusão de busca por apresentar somente números	12
Exclusão de busca de termos repetidos por maiúscula/minúscula	25.517

Fonte: Elaborado pelos autores

Os dados contidos na Tabela 2 indicam que o problema mais recorrente, dentre os identificados, foi a duplicidade de registros de consultas dos usuários diferenciadas apenas por iniciais maiúsculas ou minúsculas, critério que gerou a exclusão de 25.517 registro de consultas dos usuários.

Essa metodologia de análise de log permitiu uma grande depuração que produziu um corpus terminológico proveniente diretamente da linguagem utilizada pelos usuários de forma precisa e discreta ao mesmo tempo (Quadro 3).

Quadro 3 – Depuração de logs de busca de registros de consultas dos usuários

Lista sem depuração	Processo de limpeza	Lista depurada
Abacate – Efeito da radiação	Abacate – Efeito da radiação	ABACATE – EFEITO DA
Abacate – efeito da radiação	Abacate – efeito da radiação	RADIAÇÃO
Abate comparativo	Abate comparativo	ABATE COMPARATIVO
abate comparativo	abate comparativo	ACEITABILIDADE
Aceitabilidade	Aceitabilidade	ACEITAÇÃO
aceitabilidade	aceitabilidade	ATTITUDES
Aceitação	Aceitação	
aceitação	aceitação	
Atitudes	Atitudes	
56	56	
W'	W'	

Fonte: elaborado pelos autores

Do nosso ponto de vista, tal lista de palavras-chave poderia ser utilizada por uma equipe de trabalho que deseja construir ou atualizar um vocabulário controlado para o RI UNESP. Obviamente, esta equipe de trabalho também poderia utilizar a lista de assuntos (palavras-chave dos autores) do repositório para essas tarefas de construção ou atualização, se previamente refinada.

Conclusões

A análise realizada revela que o Repositório Unesp adota a linguagem natural sem nenhum controle de vocabulário ou uso de instrumentos de padronização. Nas listas de palavras-chaves indexadas dos documentos, observamos que não há controle de vocabulário. As palavras-chaves coletadas dos metadados estão disponíveis em uma lista alfabética para consulta dos usuários durante a busca, mas não há mecanismos de controle, o que resulta uma lista com variações que podem prejudicar a recuperação da informação. Dentre os problemas estão variações de singular/plural, digitação com maiúscula ou minúscula, uso de aspas e outros sinais como hífen, frases inteiras, inconsistências determinadas por diferentes grafias e codificações atribuídas pelo sistema acompanham as palavras.

Através da análise de logs realizada neste estudo foi possível obter uma lista de expressões de busca contendo palavras simples, palavras compostas ou frases para possibilitar a criação ou atualização de um vocabulário controlado. A metodologia aqui apresentada mostrou-se eficaz pois, permitiu uma grande depuração de expressões de busca de usuários que produziu um corpus terminológico proveniente diretamente da linguagem utilizada nesse sistema. A lista de palavras-chave e a análise de logs podem oferecer possibilidades de criação e manutenção de ferramentas de controle de vocabulário, desde a lista alfabética de palavras-chave e até tesouros, mas precisam passar por rigoroso controle terminológico e linguístico para o aproveitamento.

Referências

- Assumpção, F. S., Silva, R. E.; Ferreira, J. A. & Bastos, F. M. (2014). A Conversão de registros na implantação de repositórios institucionais: o caso do repositório institucional Unesp. In XVIII *Seminário Nacional de Bibliotecas Universitárias*. Belo Horizonte, MG: UFMG.
- Bundza, M. (2014). The choice is yours! Researchers assign subject metadata to their own materials in institutional repositories. *Cataloging and Classification Quarterly*, 52(1)110-8. doi:10.1080/01639374.2013.852439.
- Castellani, S. et al. (2009). Creation and maintenance of query expansion rules. In: *International Conference on Enterprise Information Systems* (pp. 819-30). Berlin: Springer.
- Dixon, B. & McGowan, J. (2010). Enhancing a taxonomy for health information technology: An exploratory study of user input towards folksonomy. In: *Studies in health technology and informatics*. 160 (pp. 1055-9). doi: 10.3233/978-1-60750-588-4-1055.
- Ferreira, H. H. (2019). *Processamento de linguagem natural e classificação de textos em sistemas modulares*. (Trabalho de Conclusão de Curso). Instituto de Ciências Exatas, Universidade de Brasília, Brasília, DF, Brasil.
- Ho, J. & Stokes, C. (2019). Core metadata element recommendations for institutional repositories at texas A&M university libraries. *Journal of Library Metadata*, (19), 1-28. Doi:10.1080/19386389.2019.1651499.
- Ilik, V., Hebal, P., Olson, A., Wishnetsky, S., Pastva, J., Kubiilius, R. & Holmes, K. (2018). DigitalHub: a repository focused on the future. *Medical Reference Services Quarterly*, 37(1), 31-42. doi:10.1080/02763869.2018.1404386
- Jansen, B. J, Spink, A. & Taksa, I. (2008). Web log analysis. In Jansen, B. J, Spink, A. & Taksa, I. (Eds.). *Research and methodological foundations of transaction log analysis* (pp. 1-17). Hershey: Information Science Reference.
- Jurafsky, D., & Martin, J.H. (2000). Speech and language processing – an introduction to natural language processing, computational linguistics, and speech recognition. *Prentice Hall series in artificial intelligence*. Disponível em: <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>. Acesso em 27 de maio de 2021.
- Lee, H.-M., Huang, C.-C. & Chao, C.-Y. (2007). Association thesaurus construction for interactive query expansion based on association rule mining. *Journal of Information Science and Engineering*, (23)2, 617-627.
- Meister, D. & Sullivan, D.J. (1967). *Evaluation of user reactions to a prototype on-line information retrieval system*, report to NASA by the Bunker-Ramo Corporation under contract No. NASA-1369, Report No. NASA CR-918, 1967. ERIC ED 019 094.
- Parker, E. B., & Paisley, W. J. (1966). Research for psychologists at the interface of the scientist and his information system. *American Psychologist*, 21(11), 1061-1071.
- Penniman, W. D. (2008). Historic perspective of log analysis. In Jansen, B. J, Spink, A. & Taksa, I. (Eds.). *Web log analysis* (pp. 18-38). Hershey: Information Science Reference.
- Peters, T.A. (1993). The history and development of transaction log analysis. *Library Hi Tech*, 11(2), 41-66.
- Pu, H.-T. & Chien, L.-F. (2004). Integrating log-based and text-based methods towards automatic web thesaurus construction. *Proceedings of the American Society for Information Science and Technology*, (41)1, 463-471.
- Serrano Vicente, R., Melero Melero, R. Y. & Abadal, E. (2014). Indicadores para la evaluación de repositorios institucionales de acceso abierto. *Anales de Documentación*, (17)2. doi: 10.6018/analesdoc.17.2.190821

- Vallez, M., Pedraza-Jimenez, R., Codina, L., Blanco, S. & Rovira, C. (2015). Updating controlled vocabularies by analysing query logs. *Online Information Review*, (39), 870-84. doi:10.1108/OIR-06-2015-0180.
- Vidotti, S.A.B.G., Bastos, F. M., Ferreira, J. B., Grisoto, A. P., Assumpcao, F. S. & Silva, R. E. (2015). Metadata reuse to populate an institutional repository: procedures applied in Unesp institutional repository. In *International Conference on Dublin Core & Metadata Applications (DC-2015)* (v. 1. p. 234-235). Washington: DCMI, ASIS&T – Association for Information Science and Technology.
- Vidotti, S.A.B.G., Bastos, F. M., Grisoto, A. P., Arakaki, F. A. & Ferreira, J. B. (2016). Coleta automática para povoamento de repositórios digitais: conversão de registros utilizando XSLT. *Tendências da Pesquisa Brasileira em Ciência da Informação*, (9), 1-21
- Xiaozhong, L. et al. (2008). Automatic semantic mapping between query terms and controlled vocabulary through using WordNet and Wikipedia. *Proceedings of the American Society for Information Science and Technology*, (45)1, 1-10.
- WEBMOTRICS, 2021. TRANSPARENT RANKING: All Repositories (May 2021) | Ranking Web of Repositories (webometrics.info)

(Página deixada propositadamente em branco)

LA ORDENACIÓN POR RELEVANCIA DE LOS RESULTADOS DE BÚSQUEDA DE INFORMACIÓN CIENTÍFICA: POSICIÓN Y FRECUENCIA. ESTUDIOS DE CASO

Isidoro Gil-Leiva

Universidad de Murcia, isgil@um.es, <https://orcid.org/0000-0002-7175-3099>

Resumen

Los recursos de información bibliográfica pueden albergar cientos de miles o millones de documentos por lo que se ha trabajado intensamente en el campo de la recuperación de información para propiciar que los usuarios encuentren documentos relevantes para sus necesidades de información. Un tema de gran interés ha sido la ordenación por relevancia de los documentos recuperados tras una consulta usado extensamente en catálogos, bases de datos bibliográficas, repositorios y buscadores web. **Objetivos:** Se pretende verificar la validez de un sistema de ponderación basado en la posición que ocupan los términos en los documentos y su frecuencia para ordenar por relevancia los resultados de las búsquedas. Para ello, se utiliza el sistema de ordenación por relevancia de SISA en relación a la ordenación por relevancia en las Bases de datos LISA, SCOPUS y ERIC. **Métodos:** Tras la indización automática en SISA de 200 artículos científicos publicados en la *Revista Española de Documentación Científica* y 135 artículos sobre Educación publicados en *PROFILE Issues in Teachers' Professional Development*, se ejecutaron varias búsquedas en el módulo de recuperación de SISA para lograr listas de documentos ordenados por relevancia. Las mismas consultas fueron planteadas en las Bases de datos LISA, SCOPUS y ERIC y las listas de documentos ofrecidas por las bases de datos y SISA fueron comparadas. **Resultados:** Aunque en algunos casos existe coincidencia en la posición que ocupan los documentos en los diferentes sistemas, se producen diferencias significativas en la ordenación de los resultados, e incluso LISA, SCOPUS y ERIC no presentan en sus listas ordenadas documentos que en SISA ocupan posiciones destacadas. **Conclusiones:** La metodología presentada para ordenar por relevancia los resultados de las búsquedas ofrece resultados satisfactorios, si bien es necesario seguir investigando para aplicarla en entornos más reales para terminar de confirmar su efectividad.

Palabras-chave: Ordenación por relevancia, Base de datos LISA, Base de datos SCOPUS, Base de datos ERIC, Indización automática, SISA.

Introducción

Las Bases de datos SCOPUS y Web Of Science (en adelante WoS) albergan más de 80.000.000 de documentos cada una de ellas. Además, de acuerdo a las cifras del *World Bank Open Data* se publicarán en 2021 más de 3.000.000 de artículos científico-técnicos, cifra que parece quedar corta ya que, solamente SCOPUS incorpora cada año casi tres millones de documentos. Por otro lado, Google estimó en agosto 2010 que en nuestra historia moderna ya se habían publicado alrededor 130.000.000 de libros y según la UNESCO cada año se publican más de 2 millones de libros al año. Una cantidad ingente de publicaciones nuevas cada año. En cualquier caso, por millones o cientos de miles de documentos que contenga un recurso de información existen mecanismos para seleccionar aquellos documentos de interés para un usuario. Las bases de datos y repositorios con registros descriptivos paulatinamente están siendo complementadas con textos completos. Por tanto, las estrategias de búsqueda de información pueden abarcar el texto completo, ampliando así los límites de los metadatos descriptivos al texto íntegro.

Catálogos y bases de datos ofrecen diversas formas de ordenar los resultados de las búsquedas ejecutadas. Las opciones mínimas suelen ser tres: ‘Relevancia’, de ‘Más reciente a más antiguo’ y de ‘Más antiguo a más reciente’, si bien, por lo general ofrecen más opciones. Por defecto, al ejecutar una búsqueda los resultados se ordenan por ‘Relevancia’ o bien de ‘Más reciente a más antiguo’.

La ordenación y presentación de los resultados de una búsqueda ha sido un asunto central en investigaciones sobre recuperación de información, sobre todo con la aparición de Internet y los motores de búsqueda. A partir de una consulta determinada, a los documentos recuperados se les aplica alguna fórmula para conceder a cada documento recuperado una puntuación que servirá para establecer una lista de clasificación en orden descendente presentada al usuario. Por tanto, podemos entender por ‘relevancia’, la relación existente entre el término o frase usada en una búsqueda con la información contenida en los registros bibliográficos o en el texto completo de los documentos.

Numerosas son las investigaciones efectuadas sobre la ordenación por relevancia. Así, Cutler, Shih y Meng (1997), Robertson, Zaragoza y Taylor (2004), Hu et al. (2005), Beel y Gipp (2009), Gelernter, Cao y Carbonell (2010), Wang et al. (2012) y Pang et al. (2017) han tenido en cuenta la estructura de los documentos como títulos, frases de anclaje, cuerpo de páginas webs, tamaño de los títulos, primeras 100 palabras, etc. para calcular la puntuación de relevancia. También hay otras formas que dirigen la ordenación de los resultados como las citas recibidas (Rovira, Guerrero y Codina, 2018) o la diversidad en los resultados mostrados (Chang, Haoyun y Rijke, 2020), entre otras muchas.

SISA es un sistema de indización automática que produce términos de indización a partir de reglas fundamentadas principalmente en la posición que ocupan los términos en los textos (Gil-Leiva, 2017). Tras el procesamiento automático de un documento se recopilan datos sobre la posición y frecuencia de los términos, de tal manera, que esta valiosa información además de usarse para generar términos de indización es reutilizada durante el proceso de recuperación para asignar puntuaciones a los términos y ordenar así los documentos por relevancia.

En la Figura 1 se muestra el módulo de recuperación tras la realización de una consulta en SISA. La puntuación por relevancia para cada documento se muestra en la parte derecha en recuadros en rojo.

Recuperación

The screenshot displays the SISA search interface. At the top, the search term 'bibliotecas publicas' is entered in a search box. Below the search box, there are options for 'Texto completo +' and 'Texto completo'. A 'Buscar' button is visible. Below the search results, the text 'Documentos encontrados: 30' is shown. The search results are sorted by 'Ponderación por Posición'. Two results are visible, each with a red box highlighting its relevance score: 'Relevancia: 53,3' and 'Relevancia: 49'. The first result is 'Sobre la organización de espacios en las Bibliotecas Publicas del Estado de España' and the second is 'Modelizando el uso de las bibliotecas publicas a través de sus características: estudio comparativo entre España y Finlandia para los préstamos y las vistas'.

Figura 1. Búsqueda en el Módulo de recuperación de SISA por el término ‘bibliotecas públicas’

Desde que fue implementado en SISA el procedimiento para la ordenación por relevancia de los resultados de una búsqueda no había sido evaluado. El trabajo que se presenta aquí pretende averiguar la efectividad del sistema de ordenación de SISA en relación a la ordenación por relevancia ofrecido por las Bases de datos LISA, SCOPUS y ERIC, tres bases de datos bibliográficas de referencia internacional.

Materiales y metodología

La base de datos LISA contiene tanto metadatos descriptivos como textos completos de casi cuatrocientas cincuenta publicaciones de acceso abierto sobre Library & Information Science. Las tres opciones que ofrece LISA para la ordenación de los resultados son: ‘Más antiguo en primer lugar’, ‘Más reciente en primer lugar’ y ‘Relevancia’. Usa por defecto, la ordenación por relevancia para la presentación de los resultados. La base de datos SCOPUS no alberga textos completos, pero sí registros (metadatos descriptivos y referencias bibliográficas usadas por los autores) de múltiples disciplinas. Tras una consulta el usuario puede ordenar la lista de documentos recuperados de nueve maneras diferentes, entre ellas la relevancia. Por defecto los ordena del más reciente a más antiguo. Por último, la Base de datos ERIC dedicada a la Educación la patrocina el Instituto de Ciencias de la Educación del Departamento de Educación de los Estados Unidos. Proporciona acceso a casi

dos millones registros bibliográficos de publicaciones periódicas y no periódicas desde 1966 hasta el presente. Ofrece las siguientes opciones para ordenar los resultados de las búsquedas: 'Fecha más reciente', 'Fechas más antigua', 'Autor', 'Fuente' y 'Relevancia'. Los resultados los ordena por defecto por 'Relevancia'.

Como se ha mencionado, los datos disponibles en SISA tras el procesamiento de los documentos permiten asignar una puntuación numérica a cada palabra o frase empleada en una búsqueda que es reutilizada para establecer la clasificación de los documentos por 'Relevancia' en el módulo de recuperación del sistema (Figura 1). SISA brinda la posibilidad de ordenar los resultados de una búsqueda ejecutada sobre metadatos descriptivos o sobre el texto completo de tres formas: por Posición y Frecuencia, por Frecuencia y por TFIDF. En la Tabla 1 se resume el proceso para lograr la lista ordenada por relevancia.

Tabla 1. Puntuación y ordenación por relevancia en SISA

1	T= Título; R=Resumen; PC=Palabras clave; DE=Descriptores; EP=Epígrafe; PP=Primer párrafo; OT=Otros párrafos; TT=Título tabla; TF=Título figura; CO=Conclusiones; RE=Referencias.												
2	T	R	PC	DE	EP	PP	OP	TT	TF	CO	RE		
	6	0,7	10	15	0,8	0,6	0,2	0,8	0,8	0,6	0,5		
3		T	R	PC	DE	EP	PP	OP	TT	TF	CO	RE	
	D1	1x6	2x0,7	1x10	1x15	0x0,8	9x0,6	26x0,2	0x0,8	0x0,8	6x0,6	6x0,5	
	D2	1x6	1x0,7	1x10	1x15	4x0,8	12x0,6	25x0,2	2x0,8	1x0,8	5x0,6	11x0,5	
	D3	0x6	0x0,7	0x10	0x15	0x0,8	3x0,6	4x0,2	0x0,8	0x0,8	x0,6	2x0,5	
	...												
4		T	R	PC	DE	EP	PP	OP	TT	TF	CO	RE	Total
	D1	6	1,4	10	15	0	5,4	5,2	0	0	3,6	3	49,6
	D2	6	0,7	10	15	3,2	7,2	5	1,6	0,8	3	5,5	59
	D3	0	0	0	0	0	1,8	0,8	0	0	0	1	3,6
	...												
5		T '...'	T 'bibliotecas públicas'				T '...'	T '...'	T '...'				
	D1	7,7	49,6				16,2	51,1	29,5				
	D2	2,2	59				57,9	0,8	13,1				
	D3	22,6	3,6				2,7	3,6	77,9				
	D4	0,5	54,3				28,2	6,7	2,3				
...					
6	Búsqueda				D2	D4	D1	D3	D	D	D		
	bibliotecas públicas				59	54,3	49,6	3,6		

La leyenda de la tabla anterior es la siguiente: 1. Partes de los documentos que son tenidos en cuenta y por tanto, reciben una puntuación si un término o frase aparece en esas posiciones. 2. Puntuaciones establecidas actualmente para cada posición. Esos valores podrían ser modificados si se considera conveniente. 3. Matriz con la posición-frecuencia para el término ‘bibliotecas públicas’; 4. Matriz con las puntuaciones obtenidas por el término ‘bibliotecas públicas’; 5 Matriz con la puntuación de los términos para la colección de documentos; 6. Puntuación logrado por el término de búsqueda ‘bibliotecas públicas’ usado en la ordenación por relevancia.

De acuerdo a la Tabla 1, la búsqueda con el término ‘bibliotecas públicas’ produce la siguiente lista de documentos ordenados por relevancia:

- 1.º Documento 2
- 2.º Documento 4
- 3.º Documento 1
- 4.º Documento 3

A continuación se detalla la metodología seguida en la ejecución de los experimentos.

1. Selección de las fuentes. La selección de los dominios y los corpus ha estado sujeto a tres condicionantes: dominios en los que hubiera disponible un vocabulario controlado que pudiera ser usado por SISA, la existencia de artículos científicos en formato procesable por SISA y por último, que dichos artículos científicos estuvieran indizados en al menos tres bases de datos accesibles. Teniendo en cuenta estas premisas fueron seleccionados los dominios de Documentación y Educación. La colección de documentos en Documentación se conformó con 200 artículos publicados en la Revista Española de Documentación Científica entre 2013 y 2021 y el corpus de Educación con 135 artículos publicados en la revista colombiana *PROFILE Issues in Teachers' Professional Development* entre 2016 y 2021 disponibles en la plataforma Redalyc. La razón por la que se eligieron estas dos revistas es porque SISA procesa los formatos usados por estas revistas. En el caso de la revista española porque SISA reconoce el formato XML empleado y en el caso de la colombiana porque también reconoce el formato HTML JATS manejado por las revistas albergadas en la plataforma Redalyc.

También fueron usados dos vocabularios controlados para la indización automática de los documentos que precisa realizar SISA para posteriormente obtener las puntuaciones que permiten la ordenación de los documentos por relevancia tras la realización de las búsquedas. Un vocabulario controlado sobre Documentación facilitado por indizadores de la Base de datos INDICES y adaptado para ser manejado por SISA y otro vocabulario sobre Educación generado a partir del Tesauro ERIC 2021 disponible en la web del *Educational Resource Information Center*. Y por último, fueron seleccionadas las Bases de datos LISA, SCOPUS y ERIC porque están indizadas las dos revistas elegidas y por tanto, permiten ejecutar búsquedas y ordenar los resultados por relevancia.

2. Las ecuaciones de búsqueda fueron ajustadas al contenido de cada base de datos. Así por ejemplo, como la Base de datos LISA alberga el texto completo de los

artículos, el campo ‘Texto completo’ fue uno de los elegidos para ejecutar las búsquedas. En la Base de datos SCOPUS como incluye el campo ‘Referencias’ citadas por los autores se recurrió al campo ‘All’ para realizar las búsquedas. De este modo, las dos búsquedas ejecutadas en cada base de datos fueron las siguientes:

Búsquedas en LISA:

1.^a (“bibliotecas públicas”) AND pub (“Revista Española de Documentación Científica”) AND py (2017-2021) = 16 documentos

2.^a (“bibliometric indicators”) AND pub(Revista Espanola de Documentacion Cientifica) AND PY (2013-2021) = 31 documentos

Búsquedas en SCOPUS:

1.^a (SRCTITLE (revista AND española AND de AND documentacion AND cientifica)) AND ((ALL (“public libraries”) OR ALL (“bibliotecas publicas”)) AND PUBYEAR > 2016) = 13 documentos

2.^a (ALL (“bibliometric indicators”) AND SRCTITLE (revista AND española AND de AND documentacion AND cientifica)) AND PUBYEAR > 2012 = 31 documentos

Búsqueda en ERIC:

1.^a SO profile AND YR 2016-2021 AND TX efl = 40 documentos.

2.^a SO profile AND YR 2016-2021 AND TX cooperative learning = 11 documentos.

Tras la ejecución de cada búsqueda los resultados siempre se ordenaron por Relevancia y la lista resultante fue guardada en un documento para su análisis y comparación posterior.

3. Indización de los dos corpus por parte de SISA. Se indizaron con SISA los 200 artículos de la Revista Española de Documentación Científica. Posteriormente, se usó el módulo de recuperación de SISA para realizar las búsquedas por *bibliotecas publicas*, *indicadores bibliometricos*, *efl* (education foreign language) y *cooperative learning*. Los documentos recuperados fueron ordenados por una las opciones disponibles en SISA: Relevancia (PosiciónxFrecuencia), es decir, teniendo en cuenta en qué partes de los documentos aparecen los términos, así como el número de apariciones en cada una de las partes (títulos, resúmenes, palabras clave de autor, epígrafes, primeros párrafos, otros párrafos, títulos de tablas, títulos de gráficos, conclusiones y referencias).

4. Tabulación y análisis de los datos.

Esta metodología nos ha permitido realizar tres experimentos. En cada experimento entra en juego una determinada base de datos y SISA. De este modo, tenemos LISA-SISA (Experimento 1), SCOPUS-SISA (Experimento 2) y ERIC-SISA (Experimento 3).

Resultados

Experimento 1: Base de datos LISA y SISA

Búsqueda 1. Se usó el término ‘bibliotecas públicas’ en la Base de datos LISA y en SISA sobre los metadatos y el texto completo. En LISA los resultados se ordenaron por Relevancia y en SISA fueron ordenados por Relevancia (PosiciónxFrecuencia).

Tabla 2. Ordenación de los resultados LISA-SISA de la búsqueda 1.^a

	LISA		SISA	P
1.	Sobre la organización		1. Sobre la organización	54,2
2.	Análisis presencia		2. Modelizando	49,3
3.	Contacto		3. Las colecciones	46,8
4.	Modelizando		4. Contacto	40,7
5.	Bibliotecas y empo		5. Análisis presencia	40,5
6.	Actualidad		6. El comportamiento	40,3
7.	La promoción		7. Análisis prestación	37,2
	¿ ?			

En la ordenación realizada por LISA y por SISA han coincidido en presentar en el puesto número uno el mismo documento. Si bien lo más significativo es que SISA muestra en lugares destacados (puestos 3.^o, 7.^o y 8.^o) documentos que LISA no presenta en los dieciséis documentos recuperados y ordenados.

Búsqueda 2. El término ‘indicadores bibliométricos’ fue empleado en la Base de datos LISA y en SISA usando de nuevo los metadatos (título, resumen, palabras clave, descriptores) y el texto completo. En LISA los resultados fueron ordenados por Relevancia y en SISA por Relevancia (PosiciónxFrecuencia).

Tabla 3. Ordenación de los resultados LISA-SISA de la búsqueda 2.^a

	LISA		SISA	P
1.	Estudio de la relación		1. Estudio de la relación	31,9
2.	Ranking de revistas		2. Los sectores	28,1
3.	Deskop		3. Análisis de la relación	25,8
4.	Investigación		4. Deskop	25,7
...			...	
10.	La relación		13. Ranking de revistas	2,6
...			...	
15.	Análisis de la relación		19. Investigación	0,8
...			...	
22.	Los sectores		21. La relación	0,6
...			...	
25.				

De nuevo LISA y SISA han mostrado el mismo documento en la primera posición, si bien, ahí acaban las semejanzas ya que varios documentos que ocupan posiciones muy destacadas en SISA aparecen en puestos retrasados en LISA y vice-versa, algún documento que SISA retrasa LISA lo ubica en las primeras posiciones. Por tanto, SISA ha producido una ordenación un tanto invertida en relación a la listada por la Base de datos LISA.

Experimento 2: Base de datos SCOPUS y SISA

Búsqueda 1. El término ‘bibliotecas públicas’ fue empleado en los campos ‘título’, ‘resumen’, ‘palabras clave’, ‘descriptores’ y ‘referencias bibliográficas’ tanto en la Base de datos SCOPUS como en SISA.

Tabla 4. Ordenación de los resultados SCOPUS-SISA de la búsqueda 2.^a

SCOPUS		SISA		P
1.	Bibliotecas	1.	Sobre la organización	36,4
2.	Sobre la organización	2.	El comportamiento	35,7
3.	Modelizando	3.	Análisis presencia	34,7
4.	Análisis presencia	4.	Modelizando	33,9
...		5.	Bibliotecas	32,9
...		6.	Contacto	32,7
...	¿ ?	7.	Análisis prestación	32,4

De la Tabla 4 lo más llamativo de la ordenación de SISA es la presencia de documentos en posiciones destacadas (2.^o, 7.^o y 8.^o) mientras que en la Base de datos SCOPUS no fueron listados entre los dieciséis documentos recuperados y ordenados.

Búsqueda 2. El término ‘indicadores bibliométricos’ se usó en los campos ‘título’, ‘resumen’, ‘palabras clave’, ‘descriptor’ y ‘referencias bibliográficas’ tanto en la Base de datos SCOPUS como en SISA.

Tabla 5. Ordenación de los resultados SCOPUS-SISA de la búsqueda 2.^a

SCOPUS		SISA		P
1.	Deskop	1.	Estudio de la relación	31,5
2.	Porcentaje	2.	Los sectores	26,5
3.	El efecto Cajal	3.	Deskop	25,7
4.		4.	Análisis de la relación	25
...		4.	El efecto Cajal	25
8.	Estudio de la relación	5.	Porcentaje	10,7
9.	Análisis de la relación	6.		
...		...		
15.	Los sectores	...		

SISA ha vuelto a ordenar los resultados de la búsqueda de manera un tanto diferente. Una vez más SISA coloca en las primeras posiciones documentos que SCOPUS relega a posiciones no tan destacadas.

Experimento 3: Base de datos ERIC y SISA

Para la búsqueda 1.^a en la Base de datos ERIC se usó el término ‘writing skills’ en el campo ‘Texto completo’. Los documentos recuperados en ERIC se ordenan por defecto por Relevancia. En SISA también se usó la opción de búsqueda en texto completo para ejecutar la búsqueda y de nuevo se ordenaron los documentos recuperados por Relevancia (Posición x Frecuencia). En la Tabla 6 se muestran las listas de algunos de los documentos ordenados por ERIC y SISA.

Tabla 6. Ordenación de los resultados ERIC-SISA de la búsqueda 1.^a

ERIC		SISA	
1	Fostering EFL	1	Self and Peer 46,7
2	The Impact of Assessment	2	The Impact of Oral 40,7
3	The Use of Plotagon	3	The Use of Plotagon 32,4
4	Self and Peer	4	Fostering EFL 23,8
5	Using Cooperative	5	Using Cooperative 21,7
6	Daily 6: An approach	6	Chilean Pre-Service 15,7
7	Theater for Language	7	Didactic Model 0,7
8	Language Learning Shifts	8	
9	The Impact of Oral	9	
10	Chilean Pre-Service	10	
11	Didactic Model	11	
			¿ ?

Lo más destacable de la ordenación dispuesta en la Tabla 6 es que ERIC ubicó en posiciones destacadas los documentos 2, 6, 7 y 8 que SISA no recuperó y por tanto, no llegó a presentar en su lista.

Para la búsqueda 2.^a en la Base de datos ERIC se usó el término ‘*teacher education*’ en el campo ‘Texto completo’. En SISA se usó la opción de búsqueda en texto completo para ejecutar la búsqueda y de nuevo se ordenaron los documentos recuperados por Relevancia (Posición x Frecuencia). La Tabla 7 presenta los resultados.

Tabla 7. Ordenación de los resultados ERIC-SISA de la búsqueda 2.^a

ERIC		SISA	
1	A Practice-Based	1	A Practice-Based 38,8
2	Analysis of MA	2	Scaffolding the 31,7
3	Teachers' Assessment	3	Teachers' Assessment 25,7
4	The Impact of Assessment	4	Empowering Pre-Service 25
5	Pedagogy of Experience	5	A Reflection on 25
6	From Utopia	6	Foreign Language 25
7	Beliefs of Two	7	Novice Non-Native 23,1
8	English Language	8	Pedagogy of Experience 23,1
9	Becoming Language	9	Contributions of 21,7
10	English Language	10	From Utopia 21,7
11	Differences between	11	Policies for English 21,7
12	Emotions Attributions	12	The Tensions Between 15
13	Profiles of Students	13	Becoming Language 10
14	Re-Signifying Teacher	14	The Relevance of 8,1
15	The Relevance of	15	Problem-Based Learning 6
16	Scaffolding the	16	A Learning Experience 0,7
17	Elementary School	17	Analysis of MA 0,7
18	Chilean Pre-Service	18	Chilean Pre-Service 0,7
19	Problem-Based Learning	19	English Language 0,7
20	Employment of	20	Pedagogical and Research 0,7
			¿ ?

En esta ocasión lo más llamativo es que el documento 2 fue ordenado por ERIC en segunda posición mientras que SISA lo hace en el puesto 17; y que SISA ha ordenado en los puestos 4, 5, 6 y 7 documentos que ERIC no ha recuperado.

Consideramos que las dos razones principales por las cuales se producen las diferencias en la ordenación de los resultados por parte de LISA, SCOPUS y ERIC y de

SISA son las siguientes. En primer lugar, hay que señalar que existe disparidad en los metadatos usados en la bases de datos. Así pues, mientras que LISA maneja títulos, resúmenes, palabras clave, descriptores, referencias bibliográficas y el texto completo, en cambio SCOPUS no contiene los textos completos o ERIC no maneja palabras clave, referencias bibliográficas ni textos completos. Aunque las búsquedas ejecutadas se han intentado ajustar al contenidos de cada base de datos, esta cuestión hay que tenerla en cuenta. En segundo lugar, tanto las bases de datos bibliográficas como SISA albergan su propia indización, así pues, para un determinado documento en el campo 'Subject' de estas bases de datos y en el campo 'Descriptores' de SISA podría haber términos de indización diferentes. Por tanto, como de alguna manera en la ordenación por relevancia se tiene en cuenta el contenido del campo 'subject' puede provocar resultados diferentes en las ordenaciones. Y en tercer lugar, en la ordenación de los resultados intervienen también los criterios establecidos por cada sistema para producir la lista de documentos presentada al usuario. En estos experimentos conocemos los criterios de ordenación manejados en SISA pero desconocemos los de las tres bases de datos.

Conclusiones

Las bases de datos bibliográficas contienen decenas de miles o millones de registros, por lo que es habitual efectuar una búsqueda y que los documentos recuperados sean varios cientos o incluso miles, número que no permite una completa revisión, de ahí la importancia de un buen sistema de ordenación que posicione en los primeros lugares a los documentos más relevantes para cada búsqueda. Para poder llevar a cabo los experimentos descritos se ha tenido que construir un entorno controlado con apenas doscientos artículos sobre Documentación y ciento treinta y cinco en Educación. En este entorno de 'laboratorio' la ordenación por relevancia de SISA ha obtenido resultados satisfactorios en relación a las Bases de datos LISA, SCOPUS y ERIC, si bien, sería conveniente emprender estudios más amplios en ambientes lo más reales posible para verificar la validez de la propuesta presentada para ordenar los resultados de las búsquedas por la ponderación de los términos de los documentos por posición y frecuencia.

Referências

- Beel, J. & Gipp, B. (2009). Google Scholar's ranking algorithm: An introductory overview. *Proceedings of the 12th International Conference on Scientometrics and Informetrics*, 1 (p. 230-242). Rio de Janeiro (Brazil), July 2009.
- Chang L., Haoyun, F. & Rijke, M. (2020). Cascading Hybrid Bandits: Online Learning to Rank for Relevance and Diversity. *In Fourteenth ACM Conference on Recommender Systems (RecSys '20)*, September 22-26, 2020, Brazil. ACM, New York, NY, USA. doi.org/10.1145/3383313.3412245.
- Cutler, M., Shih, Y., & Meng, W. (1997). Using the Structure of HTML Documents to Improve Retrieval. *USENIX Symposium on Internet Technologies and Systems*, p. 241-251.
- Gelernter, J., Cao, D. & Carbonell, J. (2010). Studies on Relevance, Ranking and Results Display. arXiv preprint arXiv:1006.4535.

- Gil-Leiva, Isidoro. (2017). SISA: Automatic indexing system for scientific articles. Experiments with location heuristics rules versus TF-IDF rules. *Knowledge Organization*, 44(3), 139-162. DOI:10.5771/0943-7444-2017-3-139.
- Hu, Y., Xin, G., Song, R., Hu, G., Shi, S., Cao, Y. & Li, H. (2005). Title extraction from bodies of html documents and its application to web page retrieval. In Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (250-257). Salvador, Brazil, August 15-19.
- Pang, L., Lan, Y., Guo, J., Xu, J., Xu, J., Cheng, X. (2017). DeepRank: A New Deep Architecture for Relevance Ranking in Information Retrieval. arXiv:1710.05649.
- Robertson, S. E., Zaragoza, H., & Taylor, M. J. (2004). Simple bm25 extension to multiple weighted fields. In *Proceedings of the thirteenth ACM international conference on Information and knowledge management* (42-49). New York, NY, USA: ACM. DOI: 10.1145/1031171.1031181.
- Rovira, C, Guerrero-Solé, F., Codina, L. (2018). Received citations as a main SEO factor of *Google Scholar* results ranking. *El profesional de la información*, (27)3, 559-569. DOI: 10.3145/epi.2018.may.09.
- Wang, C., Bi, K., Hu, Y., Li, H., & Cao, G. (2012). Extracting search-focused key n-grams for relevance ranking in web search. In Proceedings of the Fifth International Conference on Web Search and Web Data Mining (WSDM 2012) (p. 343-352). Seattle, WA. DOI: 10.1145/2124295.2124338.

(Página deixada propositadamente em branco)

MODELO SEMÂNTICO DE INFORMAÇÕES CORPORATIVAS PARA TRATAMENTO DE OBJETOS MULTIMÍDIA

Sergio de Castro Martins¹

¹Universidade Federal do Rio de Janeiro (UFRJ), sergio.martins@ufrj.br, ORCID iD 0000-0003-2406-6648.

Resumo

O objetivo deste trabalho consiste no desenvolvimento de um modelo de sistema de informações que utiliza tecnologias semânticas para aplicação a objetos multimídia em ambientes corporativos. Para tanto, o modelo considera ontologias expressas em vocabulários como Schema, FOAF, DC e OA Annotation, além de utilizar a infraestrutura existente em grande parte dos ambientes corporativos. A aplicação do modelo visa dar potencialidade de representação e recuperação aos metadados de objetos de informações não estruturados, como as multimídias, que compõem grande parte do ecossistema de informações das organizações. A abordagem metodológica utilizada é o *Design Science*, visando oferecer soluções a problemas práticos e, ao mesmo tempo, problemas de conhecimento. O modelo propõe-se como referência para a construção de artefatos de informação considerando as seguintes etapas: Ambientação, Declaração de parâmetros, Destilação dos dados, Enriquecimento de metadados e Armazenamento. Nas modelagens, o resultado apresentado sugere que o modelo é capaz de representar e recuperar semanticamente objetos multimídia heterogêneos e permite a aplicabilidade a qualquer tipo de organização comercial.

Palavras-chave: Enriquecimento Semântico de Metadados, Organização da Informação e do Conhecimento, Gestão de Conteúdo Corporativo, Web Semântica, Multimídia.

Introdução

Um dos grandes desafios da Gestão de Conteúdo Corporativo (*ECM – Enterprise Content Management*) é o gerenciamento de objetos de informação não estruturados, especialmente multimídias. Na atualidade, para além do entorno web, tais objetos também se constituem numa significativa parcela da informação organizacional. Objetos multimídia oferecem dificuldades adicionais para sua representação e recuperação, visto que são objetos que se modificam em tempo de sua execução e a interpretação de seu conteúdo requer uma imersão em sua narrativa que poucos métodos automatizados mostram-se tão eficientes como a interpretação humana. Além disso, o enlace com outros objetos com entidades abstratas também se apresenta como um aspecto problemático adicional.

Nos últimos anos, várias iniciativas têm sido instituídas neste sentido, sobretudo aquelas advindas das áreas de inteligência artificial, como o PNL – *Processing Natural Language* (Eisenstein, J., 2018) e o GAN – *Generative Adversarial Networks* (Creswell, A. et. al., 2018). Conquanto tais técnicas venham se aprimorando com o tempo, uma outra iniciativa tem se mostrado bastante útil para a representação e recuperação de objetos de informação, sobretudo aqueles de natureza não estruturada: as tecnologias provenientes do entorno da Web Semântica (Chu, 2010; Hitzler, P., 2021).

A Web Semântica oferece padrões e infraestrutura em nível de metadados que permitem uma complexa teia de representações para o enriquecimento do poder de recuperação de objetos de informação. Esta teia, denominada ontologia (Berners-Lee, 2001), visa dotar o objeto com propriedades que permitem uma recuperação “inteligente” de informações. Tal “inteligência” consiste no enriquecimento semântico dos metadados dos objetos de informação, utilizando ontologias como vocabulários semânticos formalizados que permitem representar inúmeros aspectos descritivos de tais objetos e suas relações com outros objetos e entidades. Este poder representativo baseia-se no padrão RDF – *Resource Framework Description* – um padrão que pretende representar as relações entre sujeitos e objetos mediante triplas descritivas. As triplas podem ser replicadas quantas vezes forem necessárias para representar descrições e relações entre objetos de e entidades concretas ou abstratas. Os metadados podem ser caracterizados de acordo com seus tipos, conforme a figura 1:



Figura 1: Tipos de metadados (Fisher e Sheth, 2004).

A infraestrutura proposta pela Web Semântica prevê a adoção de um identificador universal único para cada conceito ou entidade concreta ou abstrata, bem como a adoção de vocabulários semânticos baseados no padrão expressivo RDF para representar uma modelagem de domínio. Tais aspectos permitem inferências de alto nível por motores de busca especialmente preparados para estas interpretações – agentes inteligentes – que retornam informações mais contextualizadas mediante os enlaces e relações inferidos. Para que a inferência seja efetiva, é necessária uma modelagem dos metadados mediante utilização de vocabulários semânticos – ou ontologias – que permitam descrever todos os aspectos necessários ao objeto e suas diversas

relações. Esse enriquecimento semântico inserido nos metadados dos objetos permite um alto poder de integração não somente *a priori*, mas também *a posteriori*, potencializando a capacidade de inferência automática a ser realizada pelos agentes.

No entorno corporativo, as tecnologias semânticas permitem a criação de um ecossistema de informações mais inteligente de modo a aumentar o poder informativo dos objetos de informação, sobretudo os multimídias, dotando-os de alta capacidade de enlace a outros objetos e entidades. Esta prática apresenta-se como especialmente útil no entorno organizacional, visto que permite fatores como descoberta de dados, reutilização e retorno mais efetivo de informações sobre dados, documento, atores e processos.

Metodologia

Para os fins deste trabalho, pretende-se descrever um modelo de construção de um ecossistema semântico que faça uso de vocabulários e que se utilize também da infraestrutura de processamento e armazenamento de informações típicos de um ambiente corporativo. A abordagem metodológica baseia-se no *Design Science*. De acordo com Valverde, Toleman e Cater-Steel, “A abordagem *Design Science* tem um histórico de fornecer resultados úteis na avaliação de construções e modelos em sistemas de informação” (VALVERDE, TOLEMAN E CATER-STEEL, 2009, P. 211, tradução nossa). Também Wieringa (2009) sustenta que esta metodologia pode ser utilizada não somente para resolver problemas práticos, mas também problemas de conhecimento. A abordagem *Design Science* apresenta-se como uma das possibilidades de tratamento do problema de representação e recuperação inteligente de informações por permitir o pleno desenvolvimento de um modelo que contemple não somente aspectos teóricos como também práticos. Segundo Hevner et. al. (2004), o *Design Science* possui as seguintes orientações: 1) Design como artefato; 2) Relevância do problema; 3) Avaliação do Design; 4) Contribuição da pesquisa; 5) Rigor da pesquisa; 6) Design como processo de busca; 7) Comunicação da pesquisa junto à sua comunidade de interesse.

Considerando a abordagem acima, o modelo proposto pretende situar-se no escopo da Gestão de Conteúdo Corporativo (ECM), visando permitir a aplicação de inteligência semântica aos objetos multimídia do entorno corporativo. O modelo tem como escopo geral:

- Âmbito do Gestão de Conteúdo Corporativo, sob a perspectiva do conteúdo;
- Foco no caráter informativo e dinâmico dos objetos para tomada de decisões;
- Utilização da infraestrutura informacional e tecnológica pré-existente;
- Utilização de padrões de vocabulários, metodologias e ferramentas de licença livre;
- Estabelecimento de um modelo de ecossistema semântico com base na web.

Da mesma forma, o modelo não possui como foco os seguintes aspectos:

- Curadoria e custódia de objetos;
- Segurança e preservação dos objetos;

- Aspectos de gestão arquivística de objetos com caráter comprobatório o de mero registro;
- Construção de ontologias de domínio ou de aplicação;
- Utilização de ontologias prontas pré-existentes.

No que se refere aos requisitos de construção e aplicação do modelo, estes podem ser assim descritos:

Tabela 1: Requisitos de infraestrutura do modelo

Requisitos Técnicos de Infraestrutura para Implementação	Aplicação
Sistemas de Armazenamento: <ul style="list-style-type: none"> • Data Lake • BD NoSQL • Data Pond • Data Mart • Data Warehouse • Bases Multimídia • Base de Taxonomia Corporativa 	Armazenamento em várias instâncias dos objetos de informação operacionalizados pelo modelo, como os objetos multimídia.
Ferramentas <ul style="list-style-type: none"> • Sistemas de Informação que suportem tecnologias semânticas • Ferramentas de Anotação e Enriquecimento Semântico de Objetos 	Inserção de anotações e vocabulários semânticos aos objetos; As ferramentas devem possuir funcionalidades referentes ao processamento dos padrões exigidos, como o MPEG-7 e padrões de metadados do entorno da Web Semântica.

Fonte: O autor

O modelo proposto prevê as etapas de entrada de dados, destilação, enriquecimento semântico e saída/armazenamento, que podem ser entendidos segundo a tabela abaixo:

Tabela 2: Etapas do processo de destilação e enriquecimento semântico dos objetos de informação

Entrada	Entrada Entrada de objetos no Data Lake ou Lago de Dados: os objetos são armazenados no seu estado bruto, sem qualquer restrição de esquema ou estrutura. Inscrição de Metadados de baixo-nível Metadados automáticos inseridos no objeto: no momento da criação ou da captura dos objetos de informação, metadados de baixo-nível são associados e inseridos automaticamente nos objetos.
----------------	---

<p>Destilação</p>	<p>Reconhecimento de entidades Processo de destilação ou extração de conteúdo dos objetos multimídia: Ferramentas de decomposição de padrão MPEG-7 reconhecem de maneira automática vários recursos inerentes à mídia, representando aspectos diversos nos metadados criados no momento do reconhecimento.</p> <p>Comparação Ferramentas de decomposição de padrão MPEG-7 são executadas; comparações são feitas com os metadados de negócios proveniente de fontes como Repositórios Multimídia ou a Taxonomia Corporativa de modo a reconhecer entidades e outros aspectos para extração de conteúdo.</p> <p>Destilação Reconhecimento de conteúdo: inserção de metadados de baixo-nível dos elementos reconhecidos no objeto; Saída em formato WSDL, XML Schema, dentre outros.</p> <p>Alocação Uma vez realizada a atribuição dos metadados de baixo-nível, o conteúdo é alocado nas lagoas de dados (Data Ponds).</p>
<p>Enriquecimento Semântico</p>	<p>Enriquecimento Semântico Nos Data Ponds serão realizados os procedimentos de enriquecimento semântico, com utilização dos vocabulários destacadas pelo modelo para suportar as descrições e referências semânticas relacionadas ao objeto. O enriquecimento semântico dos objetos é manual, isto é, necessita da participação humana para o processo. Para o enriquecimento semântico serão utilizados nos objetos de informação vocabulários diversos, como o FOAF, Schema, DC, dentre outros que se façam necessários para a descrição dos aspectos requeridos e necessários. Serão inscritos vocabulários que representem atores, processos e outros objetos, juntamente com as anotações dos metadados de negócios.</p> <p>Decomposição Partes dos objetos poderão ser decompostas.</p>
<p>Saída</p>	<p>Realocação Uma vez enriquecidos, serão decompostos nos Data Marts, bancos de dados ou na web.</p>

Fonte: O autor

Os elementos do ecossistema de informações corporativas que o modelo visa representar são os seguintes:

1. **Objetos de Informação:** Objetos de informação são os elementos básicos não somente do ecossistema informacional corporativo, mas também um dos ativos mais essenciais ao funcionamento de uma organização. Dos insumos de informação saem produtos e serviços para o consumidor, além de vários processos internos e externos à organização.

2. **Atores:** Os atores são a razão de ser dos objetos de informação; eles são os produtores, veiculadores, processadores e consumidores de tais objetos. Também são os atores que executam os processos de negócios, mantendo uma organização funcional. Os atores podem ser pessoas, departamentos ou mesmo instituições.

3. **Processos:** Os processos são os movimentos padronizados, regulados e dinâmicos que mantêm as organizações em funcionamento. Geralmente são formalizados em procedimentos, etapas ou esquemas de modo a dar objetividade máxima à realização de tarefas. Os processos necessitam de atores e objetos de informação para serem executados.

Estes elementos possuem partes e relações interdependentes no ecossistema de informações corporativo, que podem assim ser definidas, conforme a tabela 3:

Tabela 3: Elementos constituintes do modelo

Objeto de Informação	<ul style="list-style-type: none"> ▪ Objeto Principal (Vídeo e Áudio) ▪ Partes Constituintes
Atores	<ul style="list-style-type: none"> ▪ Pessoas ▪ Departamentos ▪ Organizações
Processos	<ul style="list-style-type: none"> ▪ Processos ou Atividades de Negócio

Fonte: O autor

MODELAGEM CONCEITUAL

Os elementos acima, bem como seus componentes, formam o escopo do ecossistema semântico que o modelo pretende representar. Desta maneira, pretende-se criar uma cadeia semântica entre estes elementos, interligando-os mediante relações. Cada um destes elementos possui classes e várias propriedades ou atributos selecionados como requisitos para representação por vocabulários semânticos e anotações. Os vocabulários que compõem os requisitos descritivos de objetos e relações do modelo apresentam-se nas tabelas abaixo e, em negrito, encontram-se os atributos e elementos obrigatórios de interlace do registro, funcionando como chaves de associação.

Tabela 4: Atributos selecionados para classe de Vídeo

OBJETOS DE INFORMAÇÃO			
VÍDEO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Tem Tipo	Tipo do Objeto	<i>DCMI MovingImage</i>	DC; OA Annotation
Tem Identificador	Identificador único do objeto	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Título	Título do Objeto	<i>Title</i>	DC; OA Annotation
Tem Formato	Formato	<i>HasFormat</i>	DC; OA Annotation
Tem Duração	Tempo de exibição	<i>Duration</i>	DC; OA Annotation
Tem Data	Data de criação	<i>Date</i>	DC; OA Annotation

Tem Lugar	Local do conteúdo	<i>ContentLocation</i>	Schema.org/RDF; OA Annotation
Tem Transcrição	Transcrição do conteúdo	<i>Transcription</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	DC; OA Annotation
Tem Período de Tempo	Cobre período de tempo	<i>TemporalCoverage</i>	Schema.org/RDF; OA Annotation
Tem Espaço Geográfico	Cobre espaço geográfico	<i>SpatialCoverage</i>	Schema.org/RDF; OA Annotation
Tem Parte	Possui objeto-filho	<i>HasPart</i>	Schema.org/RDF; OA Annotation
Faz Referência A	Referência a outra coisa	<i>Mentions</i>	Schema.org/RDF; OA Annotation
Tem Assunto Principal	Assunto Principal do Objeto	<i>PrimaryTopic</i>	FOAF; OA Annotation
Tem Assunto	Assunto do conteúdo	<i>About</i>	DC; OA Annotation
Tem Criador	Criador do objeto	<i>Creator</i>	DC; OA Annotation
Tem Contribuidor	Co-autores do objeto	<i>Contributor</i>	DC; OA Annotation
Tem Produtor	Produtor do objeto	<i>Producer</i>	DC; OA Annotation
Tem Fonte	Organização do Criador	<i>SourceOrganization</i>	Schema.org/RDF; OA Annotation
Tem Texto Associado	Conteúdo textual	<i>Text</i>	Schema.org/RDF; OA Annotation
Tem Posição	Objeto inteiro ou sequencial	<i>Position</i>	Schema.org/RDF; OA Annotation
Tem Licença	Licença de uso	<i>Rights</i>	DC; OA Annotation
Tem Comentário	Comentários ou anotações	<i>Comment</i>	DC; OA Annotation
Tem Restrições de Acesso	Controle de acesso ao objeto	<i>ConditionsOfAccess</i>	Schema.org/RDF; OA Annotation
Tem Patrocinador	Patrocinador do objeto	<i>Sponsor</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Pessoas no Objeto	<i>AgentClass – Person</i>	DC; Schema.org/RDF; OA Annotation
Tem Grupo	Grupo ou Departamento	<i>Department</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto ou processo	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave do Objeto	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Tem Público-alvo	Atores interessados	<i>TargetAudience</i>	Schema.org/RDF; OA Annotation
Tem Tipo de Uso	Material do uso pretendido	<i>LearningResourceType</i>	Schema.org/RDF; OA Annotation
Tem Local de Produção	Local de produção do objeto	<i>RecordedAt</i>	Schema.org/RDF; OA Annotation
Tem Propósito	Ação pretendida pelo objeto	<i>PotentialAction</i>	Schema.org/RDF; OA Annotation
Tem Evento	Se relaciona a evento	<i>Event</i>	Schema.org/RDF; OA Annotation

Tabela 5: Atributos selecionados para segmento de Vídeo

OBJETOS DE INFORMAÇÃO			
SEGMENTO DE VÍDEO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Tem Tipo	Tipo do Objeto	<i>DCMI MovingImage</i>	DC; OA Annotation
Tem Identificador	Identificador único do segmento	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Título	Título do Objeto	<i>Title</i>	DC; OA Annotation
Tem Formato	Formato	<i>HasFormat</i>	DC; OA Annotation
Tem Duração	Tempo de exibição	<i>Duration</i>	DC; OA Annotation
Tem Data	Data de criação	<i>Date</i>	DC; OA Annotation
Tem Data de Início	Data inicial	<i>StartTime</i>	Schema.org/RDF; OA Annotation
Tem Data de Fim	Data final	<i>EndTime</i>	Schema.org/RDF; OA Annotation
Tem Lugar	Local do conteúdo	<i>ContentLocation</i>	Schema.org/RDF; OA Annotation
Tem Transcrição	Transcrição do conteúdo	<i>Transcription</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	DC; OA Annotation
Tem Período de Tempo	Cobre período de tempo	<i>TemporalCoverage</i>	Schema.org/RDF; OA Annotation
Faz Parte	Possui objeto-pai	<i>IsPartOf</i>	Schema.org/RDF; OA Annotation
Faz Referência A	Referência a outra coisa	<i>Mentions</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto do conteúdo	<i>About</i>	DC; OA Annotation
Tem Criador	Criador do objeto	<i>Creator</i>	DC; OA Annotation
Tem Contribuidor	Co-autores do objeto	<i>Contributor</i>	DC; OA Annotation
Tem Produtor	Produtor do objeto	<i>Producer</i>	DC; OA Annotation
Tem Fonte	Organização do Criador	<i>SourceOrganization</i>	Schema.org/RDF; OA Annotation
Tem Texto Associado	Conteúdo textual	<i>Text</i>	Schema.org/RDF; OA Annotation
Tem Posição	Objeto inteiro ou sequencial	<i>Position</i>	Schema.org/RDF; OA Annotation
Tem Licença	Licença de uso	<i>Rights</i>	DC; OA Annotation
Tem Comentário	Comentários ou anotações	<i>Comment</i>	DC; OA Annotation
Tem Restrições de Acesso	Controle de acesso ao objeto	<i>ConditionsOfAccess</i>	Schema.org/RDF; OA Annotation
Tem Patrocinador	Patrocinador do objeto	<i>Sponsor</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Pessoas no Objeto	<i>AgentClass – Person</i>	DC; Schema.org/RDF; OA Annotation
Tem Grupo	Grupo ou Departamento	<i>Department</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto ou processo	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	Schema.org/RDF; OA Annotation

Tem Palavras-chave	Palavras-chave do Objeto	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Tem Público-alvo	Atores interessados	<i>TargetAudience</i>	Schema.org/RDF; OA Annotation
Tem Tipo de Uso	Material do uso pretendido	<i>LearningResourceType</i>	Schema.org/RDF; OA Annotation
Tem Propósito	Ação pretendida pelo objeto	<i>PotentialAction</i>	Schema.org/RDF; OA Annotation
Tem Evento	Se relaciona a evento	<i>Event</i>	Schema.org/RDF; OA Annotation

Fonte: O autor

Tabela 6: Atributos selecionados para classe de Áudio

OBJETOS DE INFORMAÇÃO			
ÁUDIO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Tem Tipo	Tipo do Objeto	<i>DCMI Sound</i>	DC; OA Annotation
Tem Identificador	Identificador único do objeto	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation DC
Tem Título	Título do Objeto	<i>Title</i>	DC; OA Annotation
Tem Formato	Formato	<i>HasFormat</i>	DC; OA Annotation
Tem Duração	Tempo de exibição	<i>Duration</i>	DC; OA Annotation
Tem Data	Data de criação	<i>Date</i>	DC; OA Annotation
Tem Lugar	Local do conteúdo	<i>ContentLocation</i>	Schema.org/RDF; OA Annotation
Tem Transcrição	Transcrição do conteúdo	<i>Transcription</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	DC; OA Annotation
Tem Período de Tempo	Cobre período de tempo	<i>TemporalCoverage</i>	Schema.org/RDF; OA Annotation
Tem Espaço Geográfico	Cobre espaço geográfico	<i>SpatialCoverage</i>	Schema.org/RDF; OA Annotation
Tem Parte	Possui objeto-filho	<i>HasPart</i>	Schema.org/RDF; OA Annotation
É assunto Principal	Assunto Principal do Objeto	<i>PrimaryTopic</i>	FOAF; OA Annotation
Faz Referência A	Referência a outra coisa	<i>Mentions</i>	Schema.org/RDF; OA Annotation
Tem Assunto Principal	Assunto Principal do Objeto	<i>PrimaryTopic</i>	FOAF; OA Annotation
Tem Assunto	Assunto do conteúdo	<i>About</i>	DC; OA Annotation
Tem Criador	Criador do objeto	<i>Creator</i>	DC; OA Annotation
Tem Contribuidor	Co-autores do objeto	<i>Contributor</i>	DC; OA Annotation
Tem Produtor	Produtor do objeto	<i>Producer</i>	DC; OA Annotation
Tem Fonte	Organização do Criador	<i>SourceOrganization</i>	Schema.org/RDF; OA Annotation
Tem Texto Associado	Conteúdo textual	<i>Text</i>	Schema.org/RDF; OA Annotation
Tem Posição	Objeto inteiro ou sequencial	<i>Position</i>	Schema.org/RDF; OA Annotation
Tem Licença	Licença de uso	<i>Rights</i>	DC; OA Annotation

Tem Comentário	Comentários ou anotações	<i>Comment</i>	DC; OA Annotation
Tem Restrições de Acesso	Controle de acesso ao objeto	<i>ConditionsOfAccess</i>	Schema.org/RDF; OA Annotation
Tem Patrocinador	Patrocinador do objeto	<i>Sponsor</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Pessoas no Objeto	<i>AgentClass – Person</i>	DC; Schema.org/RDF; OA Annotation
Tem Grupo	Grupo ou Departamento	<i>Department</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto ou processo	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave do Objeto	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Tem Público-alvo	Atores interessados	<i>TargetAudience</i>	Schema.org/RDF; OA Annotation
Tem Tipo de Uso	Material do uso pretendido	<i>LearningResourceType</i>	Schema.org/RDF; OA Annotation
Tem Local de Produção	Local de produção do objeto	<i>RecordedAt</i>	Schema.org/RDF; OA Annotation
Tem Distinções	Desambiguação de sons	<i>Disambiguating Description</i>	Schema.org/RDF; OA Annotation
Tem Propósito	Ação pretendida pelo objeto	<i>PotentialAction</i>	Schema.org/RDF; OA Annotation
Tem Evento	Se relaciona a evento	<i>Event</i>	Schema.org/RDF; OA Annotation

Fonte: O autor

Tabela 7: Atributos selecionados para segmento de Áudio

OBJETOS DE INFORMAÇÃO			
SEGMENTO DE ÁUDIO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Tem Tipo	Tipo do Objeto	<i>DCMI Sound</i>	DC; OA Annotation
Tem Identificador	Identificador único do segmento	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Título	Título do Objeto	<i>Title</i>	DC; OA Annotation
Tem Formato	Formato	<i>HasFormat</i>	DC; OA Annotation
Tem Duração	Tempo de exibição	<i>Duration</i>	DC; OA Annotation
Tem Data	Data de criação	<i>Date</i>	DC; OA Annotation
Tem Data de Início	Data inicial	<i>StartTime</i>	Schema.org/RDF; OA Annotation
Tem Data de Fim	Data final	<i>EndTime</i>	Schema.org/RDF; OA Annotation
Tem Lugar	Local do conteúdo	<i>ContentLocation</i>	Schema.org/RDF; OA Annotation
Tem Transcrição	Transcrição do conteúdo	<i>Transcription</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	DC; OA Annotation
Tem Período de Tempo	Cobre período de tempo	<i>TemporalCoverage</i>	Schema.org/RDF; OA Annotation

Tem Espaço Geográfico	Cobre espaço geográfico	<i>SpatialCoverage</i>	Schema.org/RDF; OA Annotation
Faz Parte	Possui objeto-pai	<i>IsPartOf</i>	Schema.org/RDF; OA Annotation
Faz Referência A	Referência a outra coisa	<i>Mentions</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto do conteúdo	<i>About</i>	DC, OA Annotation
Tem Criador	Criador do objeto	<i>Creator</i>	DC, OA Annotation
Tem Contribuidor	Co-autores do objeto	<i>Contributor</i>	DC, OA Annotation
Tem Produtor	Produtor do objeto	<i>Producer</i>	DC, OA Annotation
Tem Fonte	Organização do Criador	<i>SourceOrganization</i>	Schema.org/RDF; OA Annotation
Tem Texto Associado	Conteúdo textual	<i>Text</i>	Schema.org/RDF; OA Annotation
Tem Posição	Objeto inteiro ou sequencial	<i>Position</i>	Schema.org/RDF; OA Annotation
Tem Licença	Licença de uso	<i>Rights</i>	DC, OA Annotation
Tem Comentário	Comentários ou anotações	<i>Comment</i>	DC, OA Annotation
Tem Restrições de Acesso	Controle de acesso ao objeto	<i>ConditionsOfAccess</i>	Schema.org/RDF; OA Annotation
Tem Patrocinador	Patrocinador do objeto	<i>Sponsor</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Pessoas no Objeto	<i>AgentClass – Person</i>	DC; Schema.org/RDF; OA Annotation
Tem Grupo	Grupo ou Departamento	<i>Department</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto ou processo	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem palavras-chave	Palavras-chave do Objeto	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Tem Público-alvo	Atores interessados	<i>TargetAudience</i>	Schema.org/RDF; OA Annotation
Tem Tipo de Uso	Material do uso pretendido	<i>LearningResourceType</i>	Schema.org/RDF; OA Annotation
Tem Local de Produção	Local de produção do objeto	<i>RecordedAt</i>	Schema.org/RDF; OA Annotation
Tem Distinções	Desambiguação de sons	<i>DisambiguatingDescription</i>	Schema.org/RDF; OA Annotation
Tem Propósito	Ação pretendida pelo objeto	<i>PotentialAction</i>	Schema.org/RDF; OA Annotation
Tem Evento	Se relaciona a evento	<i>Event</i>	Schema.org/RDF; OA Annotation

Fonte: O autor

Tabela 8: Classes e subclasses selecionados para representação de Atores

ATORES			
Classe	Subclasse	Descrição	Vocabulário
Agente			Schema.org/RDF; OA Annotation
	Organização	Subclasse de Agente; Organização ou Instituição	Schema.org/RDF; OA Annotation
	Departamento	Subclasse de Agente; Subclasse de Organização; Departamento ou Área	Schema.org/RDF; OA Annotation
Pessoa		Subclasse de Agente; Subclasse de Organização ou Instituição; Subclasse de Departamento; Indivíduo	Schema.org/RDF; OA Annotation

Fonte: O autor

Tabela 9: Atributos selecionados para a subclasse Organização

ATORES			
ORGANIZAÇÃO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Faz Parte De	Compõe “Agente”	<i>IsPartOf:</i> <i>Agent</i>	Schema.org/RDF; OA Annotation
Tem Identificador	Identificação única da organização	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Nome	Nome da organização	<i>Name</i>	Schema.org/RDF; OA Annotation
Tem departamentos ou Áreas	Divisões e departamentos internos	<i>HasPart:</i> <i>Department</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Funcionários e colaboradores	<i>HasPerson:</i> <i>Pessoas</i>	Schema.org/RDF; OA Annotation
Tem Endereço	Endereço da organização	<i>Location</i>	Schema.org/RDF; OA Annotation
Tem Email	E-mail da organização	<i>Email</i>	Schema.org/RDF; OA Annotation
Tem Página Web	Página web da organização	<i>Url</i>	Schema.org/RDF; OA Annotation
Tem Telefone	Telefone da organização	<i>Telephone</i>	Schema.org/RDF; OA Annotation
Tem Logotipo	Logomarca da organização	<i>Logo</i>	Schema.org/RDF; OA Annotation
Atua em Ramo de Negócio	Sector de atuação	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave da organização	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Tem Publicações	Publicações em mídia impressa ou digital	<i>Publisher</i>	DC; Schema.org/RDF; OA Annotation
Produz Informação	Produz objetos de Informação	<i>Creator</i>	DC, OA Annotation
Possui Processos	Possui processos de negócios	<i>SubjectOf</i>	DC; Schema.org/RDF; OA Annotation

Fonte: O autor

Tabela 10: Atributos selecionados para a subclasse Departamento

ATORES			
DEPARTAMENTO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Faz Parte De	Compõe “Agente”	<i>IsPartOf:</i> <i>Agent</i> <i>Organization</i>	Schema.org/RDF; OA Annotation
Tem Identificador	Identificação única do Departamento	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Nome	Nome do Departamento	<i>Name</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Funcionários e colaboradores	<i>HasPerson:</i> <i>Pessoas</i>	Schema.org/RDF; OA Annotation
Tem Email	E-mail do Departamento	<i>Email</i>	Schema.org/RDF; OA Annotation
Tem Telefone	Telefone do Departamento	<i>Telephone</i>	Schema.org/RDF; OA Annotation
Atua em Área de Negócio	Área de negócio	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Publicações	Publicações em mídia impressa ou digital	<i>Publisher</i>	DC; Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do departamento	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave do departamento	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Produz Informação	Produz objetos de Informação	<i>Creator</i>	DC, OA Annotation

Fonte: O autor

Tabela 11: Atributos selecionados para a subclasse Pessoa

ATORES			
PESSOA			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Faz Parte De	Compõe “Agente”	<i>IsPartOf:</i> <i>Agent;</i> <i>Organization;</i> <i>Department</i>	Schema.org/RDF; OA Annotation
Tem Identificador	Identificação única da pessoa	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Nome Principal	Nome da pessoa	<i>FirstName</i>	FOAF; OA Annotation
Tem Sobrenome	SobrenomeLastName		FOAF; OA Annotation
É membro	Pertence a organização e departamento	<i>MemberOf</i>	Schema.org/RDF; OA Annotation
Se relaciona com Pessoas	Funcionários, colaboradores e outros	<i>Knows</i>	FOAF; OA Annotation
Tem Email	E-mail da pessoa	<i>Email</i>	Schema.org/RDF; OA Annotation

Tem Telefone	Telefone da pessoa Departamento	<i>Telephone</i>	Schema.org/RDF; OA Annotation
Tem Criador	Criador do objeto	<i>Creator</i>	DC, OA Annotation
É Colaborador	Colaborador na criação de objetos	<i>Contributor</i>	DC, OA Annotation
É Produtor	Produtor de objetos	<i>Producer</i>	DC, OA Annotation
Atua em Área de Negócio	Atua em projetos e processos	<i>Project</i>	FOAF; OA Annotation
Se relaciona com Área de Negócio	É relacionado a processos ou áreas de negócios	<i>Project</i>	FOAF; OA Annotation
Atua em Projeto ou Processo	Participou de Projeto ou Processo	<i>CurrentProject</i>	FOAF; OA Annotation
Atua em Projeto ou Processo	Participa de Projeto ou Processo	<i>PastProject</i>	FOAF; OA Annotation
Tem capacitação	Conhece atividades, tarefas e processos	<i>KnowsAbout</i>	Schema.org/RDF; OA Annotation
Tem Função	Possui cargo ou função	<i>HasOccupation</i>	Schema.org/RDF; OA Annotation
É assunto de Objetos	Referenciado por objetos	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Registro em Vídeo	Tem registro de vídeo	<i>MovingImage</i>	DC; Schema.org/RDF; OA Annotation
Tem Foto	Tem registro de imagem	<i>StillImage</i>	DC; Schema.org/RDF; OA Annotation
Tem Registro em Áudio	Tem registro de som	<i>Sound</i>	DC; Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do ator	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave do ator	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation

Fonte: O autor

Tabela 12: Atributos selecionados para a classe Processos

PROCESSOS			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Tem Identificador	Identificação única do processo	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Compõe uma Organização	Compõe atividades de negócios da organização	<i>Organization</i>	Schema.org/RDF; OA Annotation
Compõe um Departamento	Compõe atividades de negócios de departamento	<i>Department</i>	Schema.org/RDF; OA Annotation
Tem Responsável	Responsáveis pelo processo	<i>Sponsor</i>	Schema.org/RDF; OA Annotation
Tem Colaborador	Possui equipes e colaboradores	<i>Contributor; Employee</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Pessoas envolvidas	<i>Person; Duns</i>	FOAF; Schema.org/RDF; OA Annotation
Tem Duração	Tempo de exibição	<i>Duration</i>	DC; OA Annotation

Tem Data	Data de criação	<i>Date</i>	DC; OA Annotation
Tem Data de Início	Data inicial	<i>StartTime</i>	Schema.org/RDF; OA Annotation
Tem Data de Fim	Data final	<i>EndTime</i>	Schema.org/RDF; OA Annotation
Se Relaciona a Outro Processo	É composto por outros processos menores	<i>SuperEvent</i>	Schema.org/RDF; OA Annotation
Se Relaciona a Outro Processo	Compõe outro processo maior	<i>SubEvent</i>	Schema.org/RDF; OA Annotation
Há apresentação de trabalhos	Atividades relacionadas ao processo	<i>WorkFeatured</i>	Schema.org/RDF; OA Annotation
Há execução de tarefas	Tarefas relacionadas ao processo	<i>WorkPerormed</i>	Schema.org/RDF; OA Annotation
Se relaciona com Área de Negócio	É relacionado a processos ou áreas de negócios	<i>Project</i>	FOAF; OA Annotation
Tem propósito	Possui função nos negócios	<i>PotentialAction</i>	Schema.org/RDF; OA Annotation
Tem assunto	Se relaciona a assunto ou trabalho	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do processo	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave do ator	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Tem Propriedades Adicionais	Possui valores e referências adicionais do processo	<i>AdditionalProperty</i>	Schema.org/RDF; OA Annotation
É composto por objetos de informação	Necessita de objetos para ser efetuado	<i>Text; StillImage; MovingImage; Sound</i>	DC; Schema.org/RDF; OA Annotation
Produz Objetos de Informação	Objetos de informação como produto do processo	<i>Text; StillImage; MovingImage; Sound</i>	DC; Schema.org/RDF; OA Annotation

Fonte: O autor

Resultados

De acordo com a modelagem, os seguintes resultados foram percebidos:

- Os objetos multimídia podem ser decompostos em estrutura e conteúdo;
- O processo de destilação e enriquecimento semântico necessitam de intervenção humana;
- A utilização de vocabulários semânticos requer um ecossistema ambientado na web;
- Os vocabulários semânticos utilizados para declarações de requisitos e atributos de objetos multimídia selecionam aspectos descritivos não exaustivos;
- O modelo é aberto, isto é, passível de atualização, revisões, reconfigurações e adaptações sempre que necessário;

O modelo, ademais, apresentou as seguintes características:

- Flexibilidade e adaptabilidade aos mais diversos contextos de negócios, mediante a possibilidade de atualização, substituição ou adição de novos vocabulários;
- Baixo custo de implementação, por prescindir de ontologias prontas e aproveitar a infraestrutura local;
- Escalabilidade, podendo ser utilizado em ecossistemas informacionais de diferentes tamanhos;
- Fácil implementação;
- Possibilidade de revisão, atualização e melhorias contínuas.

Além disso, apresentou as seguintes limitações:

- O modelo opera de maneira mais eficiente em organizações cujo ecossistema informacional tenha base na web;
- O modelo, embora criteriosamente descritivo em seus requisitos, etapas e processos, ainda não foi implementado ou testado em cenário real. Cenários reais podem apresentar variáveis e fatores não previstos ou considerados pelo modelo;
- No tratamento de objetos multimídia, o modelo tal como desenhado não contempla os requisitos e necessidades de empresas específicas de mídia ou comunicação de massa, como redes de televisão, rádios e outros ramos de broadcast e cinematografia, que requerem especificações mais detalhadas, mas sim os objetos multimídia mais comuns do ambiente corporativo.

Conclusões

Os resultados mostraram que o modelo, considerando sua estrutura, é consistentemente adaptável e flexível, podendo ser aplicado a diversos tipos de organizações. Também permite alterações e escalabilidade. Para ser eficiente, como apontado, o ecossistema de informações corporativas necessita estar baseado em ambiente web, de modo a haver enlaces e retorno de informações úteis fora do próprio entorno organizacional. Também ao permitir inferências autônomas por motores de busca especialmente preparados, o uso de tecnologias semânticas podem ser considerados aplicações de inteligência artificial aos objetos de informação. Visando sua adaptação a contextos específicos, o modelo requer aplicações e pesquisas mais extensivas.

Referências

- Berners-Lee, T. (2001). The Semantic Web. *Scientific American*. 284 (5), 34-43.
- Brisebois, R.; Abran, A.; Nadembega, A. (2017). A Semantic Metadata Enrichment Software Ecosystem based on Metadata and Affinity Models. *International Journal of Information Technology and Computer Science*, 8, 1-13, aug 2017.
- Chu, H. (2010). *Information representation and retrieval in the digital age*. 3. ed. Medford, NJ: ASIST Monographic Series.

- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative Adversarial Networks: An Overview. *IEEE Signal Processing Magazine*, 35(1), 53-65.
- Eisenstein, J. (2018). *Natural Language Processing*. Cambridge (MA): The MIT Press.
- Fisher, M., Sheth, A. (2004). Semantic Enterprise Content Management. In: SINGH, M. P. *The practical handbook of internet computing*. (Computer and Information Science Series). Boca Raton, FL: Chapman & Hall/CRC.
- Hevner et. al. (2004). Design Science in Information Systems Research. *MIS Quarterly*, vol. 28 n. 1, pp. 75-105.
- Hitzler, P. (2021). A review of the semantic web field. *Communications of the ACM*, 64(2), 76-83.
- Inmon, W. (2016). *Data Lake Architecture: Designing the Data Lake and Avoiding the Garbage Dump*. Basking Ridge NJ, Technics Publications.
- Martins, S. de C. (2019). *Modelo conceitual de ecossistema semântico de informações corporativas a ser aplicado em objetos multimídia*. Tese de Doutorado. Universidade Federal Fluminense. Programa de Pós-Graduação em Ciência da Informação da Universidade Federal Fluminense. Niterói: UFF.
- Schmitz, P. (2002). Multimedia goes corporate. *IEEE Multimedia*, 9(3), 18-19.
- Valverde, R.; Toleman, M.; Cater-Steel, A. (2009). Design Science: A Case Study in Information Systems Re-Engineering. in: Cater-Steel, A., Al-Hakim, L. *Information Systems: Research Methods, Epistemology, and Applications*. Hershey, NY: Information Science Reference.
- W3C. (2008). Best Practice Recipes for Publishing RDF Vocabularies. *W3C Working Group Note*, 28 August 2008.
- Wieringa, R. (2009). Design science as nested problem solving. *Proceedings of the 4th int. conf. on design science research in information systems and technology*. ACM.
- Wieringa, R. (2014). *Design Science Methodology for Information Systems and Software Engineering*. Berlin: Springer-Verlag.

(Página deixada propositadamente em branco)

A INCLUSÃO DAS MULHERES NA BRIGADA MILITAR NO RS: A ANÁLISE DO DOMÍNIO COMO SUBSÍDIO PARA CONSTRUÇÃO DE UMA TAXONOMIA

Carine Melo Cogo Bastos¹; Thiago Henrique Bragato Barros²

¹ Mestranda UFRGS/PPGCIN; carinecogo@yahoo.com.br;
ORCID iD 0000-0001-5305-2486

² Professor Doutor UFRGS/PPGCIN; bragato.barros@ufrgs.br;
ORCID iD 0000-0001-7439-5779

Resumo

Este artigo tem como tema a análise de domínio sobre a inclusão de mulheres como policiais militares na Instituição Brigada Militar do RS no ano de 1985. Ao conhecer o histórico da trajetória das mulheres, este trabalho objetiva mapear a presença das mulheres na Brigada Militar do Estado do Rio Grande do Sul a partir de documentos arquivísticos na atividade de polícia numa organização militar. Para a realização deste trabalho, adotou-se o referencial teórico da organização do conhecimento de Barros e Sousa (2016, 2019), Bräscher (2010) e da análise do domínio de Hjørland (2002, 2008) e Tennis (2012). Além disso, discorreu-se sobre a representação do conhecimento e o sistema de organização do conhecimento. Foram adotados, como corpus desta análise, boletins internos da Companhia Feminina Militar do período de 1985 a 1993. Por meio dessa documentação, puderam ser verificadas formas de representar o conhecimento através de um sistema de organização do conhecimento para fins de experimentação por meio de uma taxonomia. Em relação aos procedimentos metodológicos, a pesquisa foi aplicada com abordagem qualitativa, e seus objetivos são de caráter exploratório, apresentando um enfoque mais metodológico, ou seja, mostrando como aplicar taxonomias no contexto dos documentos de arquivos. Quanto aos procedimentos de pesquisa, é um estudo de caso com pesquisa documental, e os instrumentos de coleta de dados definidos para a pesquisa foram a análise dos documentos selecionados como fontes primárias e a análise do domínio como forma de contextualizar a construção da taxonomia. Para o desenvolvimento da taxonomia, foi elaborada a análise do domínio da inclusão desse grupo de mulheres dentro da instituição militar com o intuito de que esse conhecimento possa ser representado, bem como contribuir com a organização do conhecimento ao trabalhar gênero e registros de arquivo de uma instituição militar.

Palavras-chaves: Análise do domínio, organização do conhecimento, taxonomia, mulheres, arquivos militares.

Introdução

Na Arquivologia, as necessidades de acesso à informação estão vinculadas diretamente às dificuldades que os profissionais tiveram e têm no decorrer dos anos para estruturar esquemas de classificação e fazer uma recuperação eficaz da informação. Ao diferenciar informação de conhecimento, dizemos que informação é algo pontual, tem um prazo de vida curto, serve para o hoje, mas não nos garante nenhuma utilidade daqui a um tempo. Já o conhecimento tem um propósito e significado diferentes de informação, sua utilização é atemporal, agrega valor e se transforma em benefício real, possuindo o poder de até mesmo mudar a forma de pensar das pessoas. A Organização do Conhecimento se aproxima da Arquivologia quando esta busca estudar esquemas de classificação para representação do conhecimento, pois os instrumentos elaborados nos arquivos catálogos, guias e inventários possuem o mesmo objetivo que a própria Organização do Conhecimento, que é facilitar o acesso, a recuperação da informação e a gestão do conhecimento para os usuários. Cada domínio tem suas especificidades e sua trajetória e devem ser tratados de maneira específica, visto que apresentam diferentes formas de interpretação, as quais dependem muito do objeto a ser analisado. A análise do domínio tem como finalidade elaborar informações coletivas e estruturas de conhecimento de comunidades discursivas. Baseado na teoria de Hjørland (2002), em que a riqueza da análise fica por conta da combinação de uma ou mais abordagens, pode-se identificar que este estudo está relacionado com as abordagens 6 e 7, que são, respectivamente, estudos históricos e estudos de gênero/tipologias documentais. Tais domínios contam a história das primeiras mulheres em uma instituição de polícia militar, evidenciando o gênero feminino numa atividade secular masculina através de registros documentais produzidos no meio militar.

Metodologia

A metodologia adotada consiste na análise documental calcada em análise do domínio, sendo esta elaborada como um subsídio para a pesquisa ao utilizar o caráter descritivo da teoria de Tennis (2012), que oferece aos seus leitores uma declaração clara sobre a sua definição, escopo, alcance e finalidade, pois lista formas de descrever o conhecimento como uma maneira de auxiliar no processo de modelagem para o sistema de organização do conhecimento a ser definido para a instituição Brigada Militar. Este estudo possui um enfoque metodológico, ou seja, busca apresentar a forma de como aplicar taxonomias no contexto dos documentos de arquivo, com a intenção de que possa ser aplicada em outros contextos arquivísticos.

Resultados

Sendo assim, tem-se abaixo a análise do domínio como um resultado inicial da pesquisa.

Inclusão da Mulher na Brigada Militar do RS Definição: Iniciadas as mudanças na legislação federal e no Exército Brasileiro com a portaria de 1977, as mulhe-

res foram aceitas na corporação militar. A crescente inserção feminina nas mais diversas áreas de trabalho, as exigências nas legislações, bem como a portaria do Estado Maior do Exército em 10 de junho de 1977, que garantiu a participação das mulheres nas fileiras do Exército, fez com que, em fevereiro de 1984, as mulheres fossem admitidas no Exército. Nesse contexto histórico, político e social, a Brigada Militar do Estado do RS vê-se com a necessidade de evoluir e adequar-se às mudanças emergentes em função da realidade que surge face às transformações da sociedade. Dessa forma, em 8 de janeiro de 1985, pela Lei Estadual n.º 7.977, cria-se a 1.ª Companhia de Polícia Militar Feminina, sancionada pelo governador Jair Soares, que preconiza que as mulheres da Companhia de Polícia Militar Feminina (Cia PM Fem), a qual era subordinada operacionalmente ao Comando de Policiamento da Capital, para emprego exclusivamente na atividade-fim da Corporação, ainda de acordo com a Lei 7.977, só poderiam chegar ao posto máximo do Quadro Especial de Oficiais de Polícia Militar Feminina como Capitão Policial Militar Feminino.

Escopo e alcance: Para esta análise de domínio, serão utilizados os boletins internos gerados em função das atividades das Companhias Femininas. Tais informações são registradas e encadernadas, publicadas internamente, onde são divulgadas as portarias, resoluções do Conselho Superior, atos de pessoal, resultados de editais, entre outros atos administrativos essenciais para conhecimento da comunidade interna. A documentação a ser analisada refere-se ao período de 1985 a 1993, gerada no decorrer das atividades administrativas das mulheres dentro da Instituição Militar, a partir do momento em que elas passam efetivamente a trabalhar como policiais militares.

Extensão: Todos os conceitos e suas relações que serão identificados no conteúdo dos registros dos arquivos, ou seja, os boletins internos. **Exclusão:** Não será considerada nenhuma outra documentação com data posterior a 1993, em função de que, nesse ano, as Companhias Femininas deixaram de trabalhar separadamente, sendo unificadas à atividade de policial militar. **Rótulo:** Inclusão da Mulher na Brigada Militar do RS. **Foco e Especialização:** Serão retirados os termos mais comuns e presentes no acervo com o objetivo de mapear a presença das mulheres na Brigada Militar do Estado do Rio Grande do Sul a partir de documentos arquivísticos (Boletins Internos) da instituição militar, bem como de analisar a contribuição da Organização do Conhecimento com vistas a construir uma taxonomia que será usada para indexar os boletins internos. **Propósito:** Contribuir na representação da construção da história das mulheres policiais militares por meio dos documentos arquivísticos acerca da inclusão destas na Brigada Militar, além de trazer questões de gênero e a memória documental, que serão representadas a partir de um sistema de organização do conhecimento. Como objetivo principal desta pesquisa, tem-se: mapear a presença das mulheres na Brigada Militar do Estado do Rio Grande do Sul a partir de documentos arquivísticos. Os objetivos específicos, por sua vez, são elencados a seguir: (1) conhecer o histórico de inclusão das mulheres na Brigada Militar do RS, (2) identificar as temáticas oriundas dos boletins internos (1985-1993), (3) analisar o *corpus da pesquisa* de forma a trazer elementos para a construção da memória institucional e social e (4) construir uma taxonomia que será usada para indexar os boletins internos. A análise do domínio também está subsidiada nos estudos sobre a Brigada Militar dos seguintes autores: Calazans (2003), Flôres (2012) e Ribeiro (2011).

Os arquivistas atuam como intermediários e organizadores desse conhecimento quando elaboram instrumentos de busca e estruturam sistemas para representar as informações dos documentos e suas temáticas, pois não há razão para armazenar e organizar informação se esta não estiver relacionada com o objetivo de difundir o conhecimento para gerar novos conhecimentos. As classificações em arquivo não são uniformes, visto que estão automática e diretamente relacionadas à organização produtora desses documentos. Dessa maneira, os arquivistas fazem uso de legislações e organogramas do órgão que os auxiliam nessa atividade. De acordo com Barros e Sousa (2019), a base para a construção de sistemas de organização do conhecimento são as estruturas organizacionais, uma vez que elas oferecem para os arquivistas e profissionais da informação subsídios para modelagem desses sistemas de organização do conhecimento, contribuindo para melhores esquemas de classificação no desenvolvimento do fazer arquivístico e produzindo melhores resultados para representar, disseminar e acessar as informações e o uso do conhecimento por parte dos usuários finais. O referencial teórico sobre a Organização do Conhecimento e a Arquivologia é dos autores: Barité (2001), Barros e Sousa (2016, 2019), Bräscher (2010), Dias (2011), Guimarães et al. (2012) Negreiros (2008) e Sousa (2017).

Conclusão

A grande maioria das instituições busca soluções que resolvam seus problemas com desorganização dos seus arquivos. Com o enorme volume documental acumulado e a falta de acessibilidade aos documentos arquivísticos, recebem propostas de sistemas informatizados que prometem soluções milagrosas em curto espaço de tempo, os quais, em grande parte, negligenciam a estrutura que o acervo produziu, utilizou e acumulou. Esse é um dos problemas que os Sistemas de Organização do Conhecimento podem auxiliar a sanar. Além de buscar identificar as temáticas oriundas dos documentos gerados, em função das atividades que surgiram no decorrer da inclusão feminina, esta pesquisa também mapeou a presença das mulheres na Brigada Militar do Rio Grande do Sul a partir dos documentos arquivísticos. A classificação de documentos baseada no princípio de proveniência e auxiliada pelo contexto de produção documental seria a base para a elaboração de esquemas de representação do conhecimento. Esses sistemas são ferramentas que auxiliarão os usuários a encontrar a informação de maneira mais rápida, eficaz e segura, pois foram elaborados levando em consideração o contexto de geração da informação registrada em documento de arquivo. A classificação hierárquica é muito utilizada em esquemas de classificação na Arquivologia, auxiliando os usuários, nas taxonomias, a compreender como o conhecimento pode ser categorizado. Quando representa conceitos, oferece um mapa e serve como guia para os processos de conhecimento. A taxonomia foi a solução escolhida por atender à demanda de maneira sucinta, visto que consiste em um sistema de organização do conhecimento dos mais simples, sendo menos estruturado e podendo ser aplicado em qualquer universo e contexto organizacional, além de ser uma ferramenta de gestão que auxilia a realizar a organização em si com um olhar voltado ao usuário e para quem realmente utiliza

e precisa da informação. Nesse contexto, os usuários têm papel fundamental, pois auxiliam na classificação das informações e na modelagem dos sistemas de organização do conhecimento em conjunto aos profissionais da informação. Na Arquivologia, é a partir dos documentos de arquivo que podemos elaborar representações sobre o que é abordado nestes como uma forma de garantir a recuperação da informação, pois é necessário tornar mais eficientes o armazenamento e o acesso às informações.

Essa relação da Arquivologia com a Organização do Conhecimento nos trará melhores respostas às demandas institucionais para que os arquivistas possam auxiliar na representação do conhecimento das mais diversas comunidades, grupos ou Instituições.

Referências

- BARITÉ, M. (2001). Organización del conocimiento: Un nuevo marco teórico-conceptual en Bibliotecología y Documentación. In K. Carrara (Org.), *Educação, Universidade e Pesquisa* (pp. 35-60). FAPESP.
- Barros, T. H. B., & Sousa, B. T. R. R. (2019). Organização do Conhecimento e Arquivologia: Abordagens metodológicas. *Informação & Informação*, 24(2), 76-92. https://www.uel.br/revistas/uel/index.php/informacao/article/view/38290/pdf_1
- Barros, T. H. B. (2016). A Indexação e a Arquivística: Aproximações iniciais no universo teórico da organização e representação do conhecimento. *Encontros Bibli*, 21, 33-44. https://brapci.inf.br/_repositorio/2016/07/pdf_a262f021a4_0000020034.pdf
- Bräscher, M., & Café, L. (2010). Organização da informação ou organização do conhecimento? In Marilda L. G. Lara, & J. Smit. (Orgs.), *Temas de pesquisa em Ciência da Informação no Brasil* (pp. 85-102). Escola de Comunicação e Artes/USP.
- Calazans, M. E. (2003). *A constituição de mulheres em policiais: Um estudo sobre policiais femininas na Brigada Militar do Rio Grande do Sul* [Dissertação, Universidade Federal do Rio Grande do Sul]. Lume. <https://lume.ufrgs.br/handle/10183/2940>
- Dias, C. C., & Alvarenga, L. (2011). Análise de domínio organizacional na perspectiva arquivística: Um estudo baseado na metodologia proposta por Designing na Implementing Recordkeeping Systems, DIRKS. *Ciência Da Informação*, 40(2), 180-191. <http://revista.ibict.br/ciinf/article/view/1309>
- Flôres, R. J. (2012) *A inclusão feminina na Brigada Militar* [Trabalho de Conclusão de Curso, Pontifícia Universidade Católica do Rio Grande do Sul].
- Guimarães, J. A. C., González, J. A. M., & Alencar, M. F. (2012). *A análise documental no universo científico dos ENANCIBS: Elementos para uma análise de domínio*. [Anais]. Encontro Nacional de Pesquisa em Ciência da Informação, Rio de Janeiro, RJ, Brasil, 13. <http://hdl.handle.net/123456789/1952>
- Guimarães, J. A., Pinho, F. A., & Ferreira, G. M. (2012). Relações teóricas da Organização do Conhecimento com as abordagens de 69 catalogação de assunto, indexação e análise documental: Uma análise de domínio da revista. *Scire. Scire*, 18(2), 31-41. <http://hdl.handle.net/11449/114698>
- Hjørland, B. (2002). Domain analysis in information science: Eleven approaches – traditional as well as innovative. *Journal of Documentation*, 58(4), 422-462. https://pdfs.semanticscholar.org/7c9e/40b40f900232b6039bd3819d43db78d538a2.pdf?_ga
- Hjørland, B. (2008). What is Knowledge Organization (KO)? *Knowl. Org.*, 35(2), 86-101. 10.5771/0943-7444-2008-2-3-86

- Negreiros, R. N. (2008). Uma compreensão dos fundos arquivísticos sob o enfoque da análise de domínio. *Informação & Informação*, 13(2), 79-95. 10.5433/1981-8920.2008v13n2p79
- Ribeiro, L. C. (2011). *História das polícias militares no Brasil e da Brigada Militar no Rio Grande do Sul*. [Anais]. Simpósio Nacional de História, São Paulo, SP, Brasil, 36.
- Sousa, R. T. B., & Araújo, R. H, Jr. (2017). A indexação e criação de taxonomias para documentos de arquivo: Proposta para a expansão do acesso e integração das fontes de informação. *Brazilian Journal of Information Science: Research Trends*, 11(4), 47-56. <https://doi.org/10.36311/1981-1640.2017.v11n4.08.p47>
- Tennis, J. T. (2012). Com o que uma análise do domínio se parece no tocante a sua forma, função e gênero. *Brazilian Journal of Information Science: Research Trends*, 6(1), 3-15. <https://doi.org/10.36311/1981-1640.2012.v6n1.02.p3>

OS NATIVOS DIGITAIS E MODERNIZAÇÃO DAS BIBLIOTECAS UNIVERSITÁRIAS: TENDÊNCIA OU MODISMO?

Thais Batista Zaninelli¹, Sandra Gomes de Oliveira Reis², Ana Luise Moura³

¹Universidade Estadual de Londrina, *tbz@uel.br*,
ORCID iD <https://orcid.org/0000-0001-5865-3109>

²Universidade Estadual de Londrina, *sandrareismga@gmail.com*,
ORCID iD <https://orcid.org/0000-0001-6177-493X>

³Universidade Estadual de Londrina, *ana.mperes@outlook.com*,
ORCID iD <https://orcid.org/0000-0002-1500-6320>

Resumo

Este trabalho tem como finalidade apresentar o resultado de um estudo piloto realizado com acadêmicos, usuários das bibliotecas universitárias, conhecidos na literatura como nativos digitais, por nascerem entre 1980 e 1995. Uma característica importante dessa geração é que já nasceram no mundo da web ou ciberconectados e não conhecem outra forma de realizar as suas atividades e de se relacionarem que não seja por meio da conexão. Dessa forma, as bibliotecas universitárias precisam desenvolver novos produtos e serviços, inovar, objetivando atender às novas demandas desse usuário digital. Para tanto, foram entrevistados 57 estudantes de graduação de diversos cursos de uma universidade estadual do norte do Paraná, para verificar se o anseio por inovação desses usuários é uma necessidade real – tendência – ou se é modismo. Como resultado da análise das informações, por um lado, se consegue identificar algumas novas necessidades e desejos dos nativos digitais no que tange aos produtos e serviços informacionais oferecidos pelas bibliotecas universitárias e, por outro lado, já se percebe uma tendência dessas unidades em investir em inovação.

Palavras-chave: Nativos digitais, Geração Z, Bibliotecas universitárias, Inovação, Serviços de informação.

Introdução

A Biblioteca Universitária (BU) está associada às atividades acadêmicas, auxiliando na formação profissional e no progresso da sociedade e sofrem influência das transformações ocorridas na sociedade, muitas relacionadas ao uso de novas tecnologias, que interferem nos serviços informacionais oferecidos. (Santa Anna, 2015; Silveira, 2014).

Embora, tradicionalmente ligadas a dar suporte ao ensino e extensão por intermédio de seus acervos, as BUs estão modificando este estereótipo buscando agregar valor ao conjunto de produtos e serviços por elas ofertado a comunidade acadêmica. As BUs, “como fornecedoras de insumos de valor estratégico no processo de crescimento e modernidade, devem aprender a atuar neste ambiente mutável, adequando seus serviços às novas exigências de seus usuários” (Rocha; Gomes, 1993, p. 142).

Isso se justifica uma vez que as bibliotecas, como outros setores, também enfrentam novas pressões econômicas e sociais para que os serviços de informação sejam capazes satisfazer constantemente as necessidades de informações reais (explícitas ou implícitas), bem como atendam as expectativas dos usuários (Kostagiolas et al., 2011), em especial os usuários incluídos na geração dos Nativos Digitais.

Para atender às necessidades de informação da geração digital, os nativos digitais, também conhecidos como Geração Z, nascidos de 1995 a 2010, geração que não sabe o que é viver sem a internet e as tecnologias, nasceram imersas nesse mundo e são imediatistas, não sabendo diferenciar o mundo off-line do on-line. Geração veloz, autodidata e empreendedora.

O objetivo deste estudo foi entender a visão que esses usuários, que utilizam os serviços do Sistema de Bibliotecas da Universidade Estadual de Londrina (SB/UEL) que é composto por quatro unidades, se realmente desejam mudanças no formato dos produtos e serviços da biblioteca, para assim utilizar com maior frequência esse ambientes.

Nativos digitais, serviços de informação e as bibliotecas universitárias

São muitas as discussões em torno das gerações, os termos Veteranos, Nativos Digitais, Geração Polegar, Millenions, Geração Z, Y e X, Baby Boomers e Geração Alfa têm ganhado notoriedade nas discussões entre pesquisadores (Tapscoot, 2010; Briscoe, Hall & Mayrhofer, 2012; Azevedo, 2016; Sinek, 2020 como citado em Penso, 2020) devido às grandes mudanças de comportamento de cada geração causado pelos aparatos tecnológicos, entretanto, não cabe aqui uma profunda discussão no âmbito de cada geração.

O foco deste estudo é estudar os serviços de informação no contexto dos nativos digitais, que caracterizam a maior parte dos universitários. A geração dos nativos digitais (Cerbasi & Barbosa, 2009) compreende sujeitos nascidos depois do ano 2000 até os dias atuais. Também conhecida como a Geração Z ou Next (Tapscoot, 2010), agrupam pessoas que nasceram de 1998 até a data presente.

Essa geração, que tem atualmente entre 21 a 23 anos de idade, são indivíduos preocupados em conectividade com os demais de forma permanente. Maffesoli (1998) diz que as tribos, ou seja, toda a geração, passam a ser permeadas pela rede e o ritmo a ser seguido é ditado pela tecnologia quando se torna decisiva para marcar a personalidade de uma época. O jovem da geração Z tem um lado independente muito forte, pois através da mobilidade que seu telefone concede é possível conhecer o mundo, assistir a séries e filmes de qualquer lugar, ouvir música de variados artistas e conhecer pessoas.

Dessa forma, procurando adaptar-se às novas demandas informacionais dessa geração, as BUs instituem novas formas de conexão entre seus usuários e sistemas,

oferecendo serviços de informação inovadores, de modo a objetivar a disseminação da informação que se encontra em seus ambientes, conforme discorrem Araújo, Oliveira e Bezerra (2017). Para isso, de acordo com os autores, é necessário que as BUs se mantenham atualizadas no que diz respeito às inovações tecnológicas, apropriando-se dessas ferramentas visando um aprimoramento constante de seus serviços, satisfazendo assim seus usuários reais e potenciais.

O estudo de Santos, Gomes e Duarte (2016) apresenta uma pesquisa no qual identificou-se por meio de estudos nacionais e internacionais a adoção e em certa medida um avanço, na utilização dos dispositivos de comunicação da web social pelas BUs. Entre os dispositivos de comunicação da web social, como facebook, Orkut, Myspace e Ning pela BUs, entre as BUs no Brasil, Estados Unidos e Espanha, realizado por Aguiar e Silva em 2010 que também investigaram as informações disponibilizadas e o tipo de interação que é estabelecida com os usuários.

Metodologia

A foi pesquisa desenvolvida no âmbito das atividades do grupo de pesquisa Inovação em Serviços de Informação – INSERI e teve como objetivo identificar se os estudantes universitários, considerados na literatura como Nativos Digitais têm necessidades e desejos de serviços informacionais mais inovadores.

Após uma extensa revisão de literatura, seguido de um estudo piloto realizado com acadêmicos de diferentes cursos da universidade, considerados na literatura como nativos digitais. (Coelho, 2012; Lanzi, Ferneda & Vidotti, 2013).

Em outubro de 2019 foi administrado questionários, presencialmente no campus com 57 estudantes envolvendo questões relacionadas aos produtos e serviços informacionais já existentes e sobre o desejo por inovação nessas unidades.

Resultados

Tradicionalmente ligadas a dar suporte ao ensino e extensão por intermédio de seus acervos, podemos inferir que as BUs estão buscando modificar este estereótipo por meio de pesquisas que ajudam seus gerentes a entenderem o conjunto de produtos e serviços que os acadêmicos desejam. Os entrevistados desta pesquisa tinham entre 17 e 29 anos, considerada na literatura como geração dos nativos digitais e alguns da Geração Z ou Next. A finalidade do uso da biblioteca para estudar foi indicado por 81% e/ou realizar empréstimos de materiais 67%. Frequência de utilização são no mínimo duas vezes por semana (36%).

Há uma insatisfação em relação ao ambiente físico da biblioteca, em especial com relação à falta de climatização (34%) e se dá pois os prédios são em espaço não adequados, já que são salas de aula adaptadas.

Relativamente ao tipo de modernização e inovação poderia transformar o ambiente mais agradável, as respostas foram variadas e surpreendentes (Figura 1).

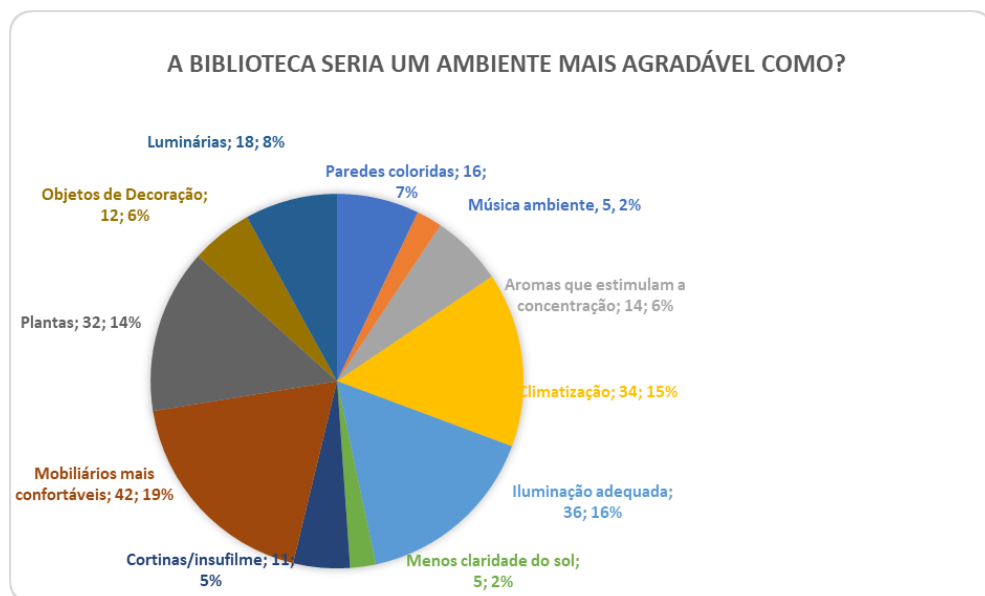


Figura 1 – Inovação na Biblioteca

Fonte: Desenvolvido pelos autores (2020).

Se os serviços oferecidos (Tabela 1) são bem divulgados, foi outro questionamento, 88% afirmam que não são, assim os serviços são subutilizados, já que os usuários não os conhecem. 82% afirmaram não acessar as redes sociais, mesmo tendo um comportamento informacional digital.

Produto/Serviço	Descrição
Catálogo da biblioteca	Consulta / Renovação e reserva / Tutoriais
Acervos digitais	Acervos UEL / Portal de periódicos da CAPES / Outros acervos
Produtos	Catálogo na publicação / Geração automática de ficha catalográfica / Pesquisa em bases de dados / Comutação bibliográfica / Normalização de documentos / Aquisição
Capacitação	Treinamentos / Agenda de treinamentos
Galeria de fotos	Exposições
Normas de apresentação de trabalhos	Modelos padrão UEL / Apostilas para normalização / Conheça outras normas

Tabela 1 – Serviços oferecidos pelo Sistema de Bibliotecas da UEL

Fonte: Desenvolvido pelos autores (2020).

Questionados sobre quais são as inovações que a biblioteca não oferece, mas que seriam um atrativo e um desejo deles, solicitamos que indicassem tudo que entendiam como inovação nos serviços e no ambiente físico da biblioteca (Figura 2).

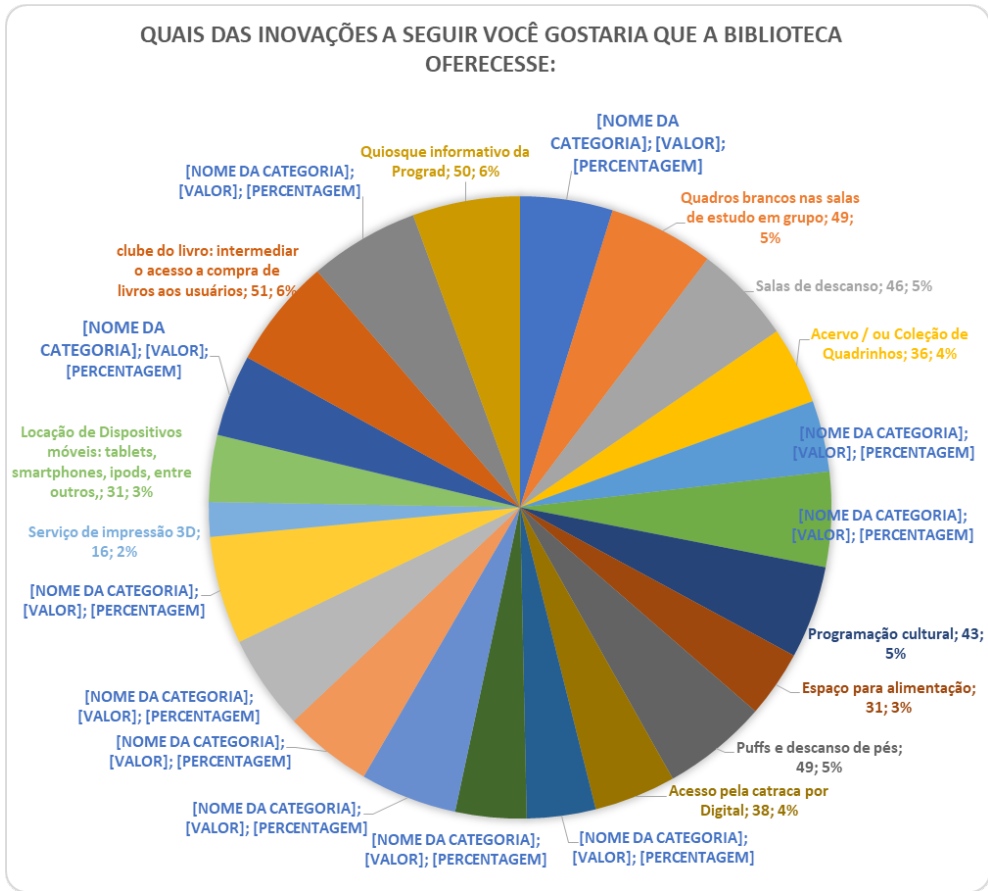


Figura 2 – Inovação no Sistema de Bibliotecas

Fonte: Desenvolvido pelos autores (2020).

Todas as opções de inovação foram bem aceitas pelos respondentes, já que são serviços que envolvem transformações digitais e tecnológicas e até mesmo uma mudança no conceito de serviços de biblioteca.

Conclusões

Conhecer o perfil do público e oferecer produtos e serviços que atendam às suas necessidades contribui para a melhoria no relacionamento com o usuário, trazendo benefícios à unidade através da gestão eficiente de recursos e criando um ambiente propício para a inovação e criatividade, com desenvolvimento de novas potencialidades e na formação de novos públicos (Bragança et.al., 2016). Por um lado, esse trabalho identificou, algumas novas necessidades e desejos em produtos e serviços informacionais oferecidos pelas bibliotecas universitárias e, por outro lado, já se percebe uma tendência dessas unidades em investir em inovação, de ofertar serviços, produtos de informação. Isso é fundamental e indispensável para que as novas

demandas dos usuários atuais e futuros sejam compreendidas pelos bibliotecários e dirigentes tendo em vista que o futuro tende a ser cada vez mais digital e o papel dos bibliotecários é de extrema relevância para essa adaptação.

Referências

- Azevedo, D. (2016). *Online full time: a sociabilidade das gerações y & z no cotidiano tecnológico*. [Dissertação de Mestrado, Universidade Federal Fluminense].
- Araújo, A. R. S., Oliveira, R. M. F. S., & Bezerra, M. G. (2017). *Serviços de informação em bibliotecas universitárias: estudo comparativo entre bibliotecas de instituições de ensino superior da cidade de Juazeiro do Norte, Ceará*. Anais do Vigésimo Sétimo Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação. <https://rbbd.febab.org.br/rbbd/article/view/822/860>.
- Bortolin, S., & Vignoli, R. G. (2014). A biblioteca escolar e as mediações com a geração polegar. *Biblioteca Escolar Em Revista*, 2 (2), 45-59. <http://www.revistas.usp.br/berev/article/view/106598>.
- Briscoe, J. P., Hall, D. T., & Mayrhofer, W. (2012). *Careers around the world: individual and contextual perspectives*. Routledge.
- Bragança, F. C., Zaccaria, R. B., Giuliani, A. C., & Pitomba, T. C. D. T. (2016). Marketing, criatividade e inovação em unidades de informação. *Revista Brasileira de Marketing*, 15 (2), 237-245. <https://www.redalyc.org/pdf/4717/471755312007.pdf>.
- Cerbasi, G., & Barbosa, C. (2009). *Mais tempo, mais dinheiro*. Thomas Nelson Brasil.
- Coelho, P. M. F. (2012). Os nativos digitais e as novas competências tecnológicas. *Texto Livre: Linguagem e Tecnologia*, Belo Horizonte, 5(2), 88-95. <http://periodicos.letras.ufmg.br/index.php/textolivre/article/view/2049/7254>.
- Kostagiolas, P., Margiola, A., & Avramidou, A. (2011). A library management response model against the economic crisis: The case of public libraries in Greece. *Library Review*, 60 (6), 486-500.
- Lanzi, L. A. C., Ferneda, E., & Vidotti, S. A. B. G. (2013). *A biblioteca escolar e a geração nativos digitais: construindo novas relações*. São Paulo: Cultura Acadêmica.
- Maffesoli, M. (1998). *O Tempo das Tribos: O Declínio do individualismo nas sociedades de massa*. Forense-universitária.
- Penso, A. (2020, 02 Março). *Como a geração Z pode nos fazer mudar?* <https://ofuturodascoisas.com/como-a-geracao-z-pode-nos-fazer-mudar/>.
- Rocha, E. C., & Gomes, S. H. A. (1993). Gestão da qualidade em unidades de informação. *Ciência da Informação*, 22 (2), 142-52.
- Santa Anna, J. (2015). A biblioteca universitária no presente: de labirinto à encruzilhada em busca da biblioteca híbrida. *Revista ACB: Biblioteconomia em Santa Catarina*, 20(1), 6-18. https://revista.acbsc.org.br/racb/article/view/982/pdf_117.
- Santos, R. R., Gomes, H. F., & Duarte, E. N. (2016). Processo dialógico entre bibliotecários e usuários: reflexão em torno da utilização dos dispositivos de comunicação da web social. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 21 (45), 115-129. <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2016v21n45p115>.
- Silveira, N. F. (2014). Evolução das bibliotecas universitárias: information commons. *Revista ACB: Biblioteconomia em Santa Catarina*, 19(1), 69-76. https://revista.acbsc.org.br/racb/article/view/923/pdf_88.
- Tapscott, D. (2010). *A hora da geração digital: como os jovens que cresceram usando a internet estão mudando tudo, das empresas aos governos*. Agir Negócios.

OS BENEFÍCIOS DA FOLKSONOMIA COM CURADORIA PARA A RECUPERAÇÃO DA INFORMAÇÃO EM BIBLIOTECAS UNIVERSITÁRIAS

Aline Araújo¹, Camilla Oliveira²

¹Universidade do Porto, alinekaroline4@gmail.com

²Universidade do Porto, olivercami2@hotmail.com

Resumo

Desde os seus primórdios, a biblioteca procurou se adaptar às mudanças sociais, culturais e tecnológicas da sua época. Como instituição, ela tem um grande papel social como difusora de informação e conhecimento. Em busca de tornar a informação mais acessível para os utilizadores, as bibliotecas vêm empregando novas formas de tratar a informação tornando mais flexíveis os seus antes rígidos sistemas de catalogação e indexação. Uma nova tendência nesse caminho é o uso em conjunto dos vocabulários controlados com a folksonomia, também conhecida como marcação colaborativa, com o intuito de expandir o acesso e a recuperação da informação aos utilizadores sem entendimento técnico e transformar a biblioteca em um ambiente mais democrático e inclusivo. Apesar do uso da folksonomia possibilitar melhor acessibilidade aos recursos informacionais da biblioteca e torná-la mais democrática e orientada para as necessidades dos utilizadores, a folksonomia também apresenta algumas falhas como ambiguidade, variação na granularidade e sinonímia. Uma das soluções encontradas para amenizar essas lacunas é o uso da folksonomia com curadoria. Alguns dos métodos utilizados nessa abordagem incluem o uso de folksonomias estruturadas, filtros de marcação que utilizam métricas de consenso e mapeamento automatizado em ontologias. Este trabalho tem como objetivo principal apontar os benefícios da folksonomia com curadoria em bibliotecas universitárias e como seu emprego pode ajudar a amenizar as falhas encontradas no uso tradicional da folksonomia. A metodologia empregada consiste em uma revisão de literatura com base em artigos científicos, recuperados nas bases Library, Information Science & Technology Abstracts (LISTA), Qualitative and Quantitative Methods in Libraries (QQML) e Portal de Periódicos da CAPES, utilizando-se o recorte temporal de cinco anos (2016-2020). Este trabalho justifica-se pela relevância da folksonomia para o acesso e recuperação da informação, pois apesar da folksonomia ser tema recorrente na literatura publicada, a folksonomia com curadoria mostra-se como alternativa válida para amenizar as suas falhas e discordâncias. Como resultados pretende-se demonstrar com base na análise da literatura recente como a curadoria em folksonomia pode melhorar a precisão, funcionalidade e escalabilidade dos sistemas híbridos em bibliotecas, focando-se nas bibliotecas universitárias e nos seus repositórios digitais. Para que isso ocorra, são necessárias algumas mudanças nas orientações e práticas dominantes dentro

das bibliotecas, como a adoção de uma estrutura distribuída ao invés de uma centralizada, o controle deve ser focado no utilizador e o processamento de certas etapas das informações deve ser feito por pessoas e não por computadores.

Palavras-chave: folksonomia, marcação colaborativa, curadoria, biblioteca.

Introdução

Desde os seus primórdios, a biblioteca procurou se adaptar às mudanças sociais, culturais e tecnológicas de sua época. Como instituição, ela tem um grande papel social como difusora de informação e conhecimento. Em busca de tornar a informação mais acessível para os utilizadores, as bibliotecas vêm empregando novas formas de tratar a informação tornando mais flexíveis os seus antes rígidos sistemas de catalogação e indexação. Uma nova tendência nesse caminho é o uso em conjunto dos vocabulários controlados com a folksonomia, também conhecida como marcação colaborativa, com o intuito de expandir o acesso e a recuperação da informação aos utilizadores sem entendimento técnico e transformar a biblioteca em um ambiente mais democrático e inclusivo.

O termo folksonomia, criado em 2004 por Thomas Vander Wal, refere-se ao sistema no qual os utilizadores aplicam tags públicas a itens online, as quais facilitam a organização e recuperação desses itens (Kipp et al., 2017).

O uso da folksonomia torna-se uma alternativa que visa expandir a acessibilidade informacional e melhorar a recuperação da informação por parte de seus utilizadores. No entanto, apesar das vantagens apresentadas pela folksonomia, ela também mostra algumas falhas, como: ambiguidade, variação na granularidade e sinonímia. Surge então a problemática a ser respondida: como amenizar as falhas trazidas pelo uso da folksonomia tradicional de maneira a melhorar a precisão, funcionalidade e escalabilidade dos sistemas híbridos em bibliotecas universitárias e nos seus repositórios digitais?

Uma das soluções encontradas para diminuir essas lacunas e tornar os sistemas informacionais mais eficazes é o uso da folksonomia com curadoria. Sendo assim, o objetivo principal deste trabalho é apontar os benefícios da folksonomia com curadoria em bibliotecas universitárias e como seu emprego pode ajudar a amenizar as falhas encontradas no uso tradicional da folksonomia.

Através de uma revisão bibliográfica baseada na literatura científica dos últimos cinco anos, esse trabalho traz alguns exemplos de sistemas de folksonomia com curadoria que demonstram como a curadoria pode melhorar os próprios sistemas, aprimorando a precisão, funcionalidade e escalabilidade desses sistemas híbridos em bibliotecas.

Metodologia

A metodologia empregada consiste em uma pesquisa de caráter exploratório e qualitativa, tendo como foco a revisão de literatura com base em artigos científicos recuperados nas bases Library, Information Science & Technology Abstracts (LISTA), Qualitative and Quantitative Methods in Libraries (QQML) e Portal de

Periódicos da CAPES. O recorte temporal da pesquisa foi delimitado em cinco anos (2016-2020). Depois de aplicados os critérios adequados, foram selecionados nove artigos científicos com relevância para a realização desta pesquisa.

Para a recuração dos artigos foram utilizados os termos “folksonomy”, “curated folksonomy” e “collaborative tagging”. Levando-se sempre em consideração a questão norteadora deste trabalho: como é possível amenizar as falhas apresentadas pela folksonomia tradicional e melhorar a precisão, funcionalidade e escalabilidade dos sistemas híbridos em bibliotecas universitárias?

Este trabalho justifica-se pela relevância da folksonomia para o acesso e recuperação da informação, pois apesar da folksonomia ser tema recorrente na literatura publicada, a folksonomia com curadoria mostra-se como alternativa válida para amenizar as suas falhas e discordâncias. Como resultados pretende-se demonstrar com base na análise da literatura recente como a curadoria em folksonomia pode melhorar a precisão, funcionalidade e escalabilidade dos sistemas híbridos em bibliotecas universitárias e seus repositórios digitais.

Resultados

Apesar do uso da folksonomia possibilitar melhor acessibilidade aos recursos informacionais das bibliotecas universitárias e torná-las mais democráticas aos seus utilizadores, a folksonomia tradicional apresenta algumas falhas como ambiguidade, variação na granularidade e sinonímia. Uma das soluções encontradas para amenizar essas lacunas é o uso da folksonomia com curadoria nos sistemas informacionais dentro das bibliotecas universitárias.

Em contraste com folksonomias clássicas ou não regulamentadas, folksonomias com curadoria carregam as particularidades dos utilizadores como ponto de partida e faz uso de especialistas ou da tomada de decisão coletiva para identificar e aliviar problemas com sinonímia e homógrafos. Os princípios básicos de uma folksonomia com curadoria são os seguintes: os utilizadores criam tags; algum agente intencionalmente combina marcas sinônimas e / ou diferencia tags homográficas; o recall e a precisão são aprimorados. Ao contrário do vocabulário tradicional controlado de cima para baixo, a maior parte dos termos em uma folksonomia com curadoria são orientados pela atividade do utilizador e o desenho da organização do conhecimento segue a ação do utilizador. Folksonomias com curadoria são muito usadas para coleções geradas por utilizadores que estão crescendo rapidamente, para as quais o desenho de classificação profissional é muito caro e em que os utilizadores são motivados e aptos a envolver-se na organização do trabalho (Bullard, J., 2019, pag 9-10).

No sistema de curadoria por combinação de tags se obedece um senso de sinonímia onde a combinação é norteadora pela regra de que as tags devem ser combinadas apenas quando elas equivalem tanto no significado quanto no uso. Outro exemplo é o Stack Overflow, um site de pergunta e resposta para programação de computadores, onde tags são utilizados na recuperação de questões relacionadas e oferece páginas de informações básicas. O Stack Overflow “pede ajuda” aos utilizadores para controlar as tags de folksonomia. Nele, os utilizadores com os melhores rankings

algorítmicos no site conseguem modificar as informações das páginas de tags, identificar tags que são sinônimos de outras e opinar a favor ou contra as conexões. Nessas duas formas de curadoria os princípios básicos são os mesmos, ou seja, os utilizadores orientam a criação de novas tags e as tags sinônimas são transformadas em equivalentes para a recuperação. Um terceiro exemplo de curadoria é o AO3. Esse sistema é o mais restritivo e seletivo no que diz respeito à participação dos utilizadores. Nele, uma equipe de voluntários (~200) passa por um processo de recrutamento e de treinamento antes de poder ter acesso e permissões para usar a interface de folksonomia. Os voluntários não nomeiam ou votam nas decisões dentro da folksonomia, em vez disso, cada pessoa é responsável por uma seção do site e tem autonomia para fazer as mudanças necessárias (Bullard, J., 2019, pag 10-12).

Esses três exemplos de sistemas de folksonomia com curadoria demonstram como a curadoria pode melhorar os próprios sistemas. A escolha por determinado desenho de sistema pode afetar na precisão, funcionalidade e escalabilidade dos sistemas. A folksonomia com curadoria, como em qualquer abordagem de conhecimento organizacional, esta sujeita à análise e avaliação de suas propriedades. A escolha pela curadoria possibilita encontrar soluções para a folksonomia que vão contra as tendências atuais dentro da organização do conhecimento: desenho de sistemas distribuídos ao invés de centralizados, controle feito pelo utilizador em vez de profissional e processamento humano em vez de processado por computadores (Bullard, J., 2019, pag 13).

Como apontam Yu e Chen (2020), a associação e mapeamento de folksonomias podem enriquecer significativamente os vocabulários e melhorar a qualidade dos metadados de assunto existentes, bem como aliviar a ambiguidade e a complexidade causada pela folksonomia não controlada na área de sistemas de informação colaborativos.

Nos repositórios digitais das bibliotecas universitárias as falhas geralmente presentes na folksonomia tradicional podem ser amenizadas através do uso de uma folksonomia assistida com o apoio aos utilizadores através de um conjunto de termos existentes, de modo a auxiliá-los no momento da escolha dos termos mais adequados para os descritores que representarão seus trabalhos na plataforma digital (Santos, 2013). Isso pode amenizar as falhas de ambiguidades e sinonímia e tornar mais eficiente a recuperação da informação.

Conclusões

A folksonomia com curadoria oferece a capacidade de implementar a análise humana no design de vocabulário controlado por especialistas possibilitando aos utilizadores a oportunidade de revisar a folksonomia, interpretando as tags como sinônimos ou algo relacionado. Os mesmos sistemas que criam coleções geradas pelo utilizador e suas folksonomias desordenadas podem gerar vocabulários controlados, dadas as escolhas de desenhos de sistemas específicos. O uso da curadoria na folksonomia viabiliza compreender o campo da organização do conhecimento de uma forma mais ampla, evidenciando as suas noções de poder, responsabilidade e a possibilidade de representar uma pluralidade de vozes (Bullard, J., 2019, pag 4).

A folksonomia com curadoria envolve uma validação do julgamento humano e experiência de domínio, o que vai de acordo com as abordagens da organização do conhecimento que reconhecem o ser humano como um instrumento de análise, interpretação e responsabilidade. Assim, a curadoria encoraja a uma mudança de foco desde a implementação do sistema de organização do conhecimento até a revisão como práticas essenciais (Bullard, J., 2019, pag 17).

Visto que o processo de curadoria é fundamentalmente colaborativo e com necessidades tecnológicas que demandam participação e responsabilidade, o seu estudo e implementação demanda conexões apoiadas por computador, crowdsourcing e design participativo. Esse modelo também sugere uma oportunidade para o estudo da organização do conhecimento dentro do cenário atual da heteromação, da objeção à automação e domínio algorítmico (Bullard, J., 2019, pag 17).

Por fim, atentamos para o reduzido número de pesquisas em torno do tema da folksonomia com curadoria aplicada na prática às bibliotecas universitárias e seus sistemas informacionais, como os repositórios digitais, o que se tornou limitante para um maior aprofundamento de idéias e coleta de dados. Os autores apontam para a necessidade e importância de serem feitas mais pesquisas na área da folksonomia com curadoria, especialmente em bibliotecas universitárias, não só pela sua relevância como instrumento de acesso e recuperação de informação, mas como uma forma de estimular uma maior participação dos utilizadores.

Referências

- Bullard, J. (2019). Curated Folksonomies: *Three Implementations of Structure through Human Judgment*. Knowledge Organization, 45(8), 643-652. DOI: 10.5771/0943-7444-2018-8-643.
- Cevolini, Alberto. (2019). *The Memory of Tags*. 37. 211-222. DOI: 10.3828/indexer.2019.28.
- Choi, Namjoo & Harper, Lindsey M. (2019). *Public libraries and the social web: a review and analysis of the existing literature*. Journal of Documentation. ahead-of-print. DOI: 10.1108/JD-04-2019-0073.
- Einasto, O. (2019). *Library Communication in the Era of Digimodernism*. Qualitative & Quantitative Methods in Libraries, 8(2), 159–168. DOI: 142303572.
- Gupta, Shweta & Professor, Pandey. (2019). *Perception and Use of Semantic and Social Features of Digital Libraries among Library and Information Professionals and Users: A Comparative Study*.
- Johansson, S., Golub, K. (2019). LibraryThing for Libraries: How Tag Moderation and Size Limitations Affect Tag Clouds Knowledge organization, 46(4): 245-259
- Katagi, Sachin & Gala, Bhakti. (2020). *Social Tags of Select Books Written by Mahatma Gandhi: A Comparative Study of Library Thing Tags and OCLC Fast Subject Headings*. DESIDOC Journal of Library & Information Technology. 40. 382-387. DOI: 10.14429/djlit.40.1.15138.
- Kipp, M.E., Beak, J. and Choi, I. (2017). *Motivations and intentions of Flickr users in enriching flick records for library of congress photos*. Journal of the Association for Information Science and Technology, Vol. 68 No. 10, 2364-2379.
- Samanta, K. (2020). *User Generated Social Tags Versus Librarian Generated Subject Headings, A Comparative Study in the Domain of History*. DESIDOC Journal of Library & Information Technology, 40(03), 176-184.

- Santos, R. F. (2013). *Representação da Informação em repositórios digitais: implicações da Folksonomiano processo de recuperação da memória documental*. 70f. Monografia. Universidade Federal do Rio Grande do Norte, Natal, 70p.
- Yu, Wei & Chen, Junpeng. (2020). *Enriching the library subject headings with folksonomy*. The Electronic Library. ahead-of-print. DOI: 10.1108/EL-07-2019-0156.

CARACTERIZAÇÃO CONCEITUAL DO TRATAMENTO TEMÁTICO DA INFORMAÇÃO NA CIÊNCIA DA INFORMAÇÃO NO BRASIL: ANÁLISE DA PRODUÇÃO CIENTÍFICA NO PERÍODO DE 2001 A 2020

Lais Pereira de Oliveira¹, Maria Cláudia Cabrini Grácio²
e Daniel Martínez-Ávila³

¹Universidade Federal de Goiás, laispereira2@ufg.br,
ORCID iD 0000-0001-9092-4204

²Universidade Estadual Paulista, cabrini.gracio@unesp.br,
ORCID iD 0000-0002-8003-0386

³Universidad Complutense de Madrid, danima27@ucm.es,
ORCID iD 0000-0003-2236-553X

Resumo

Aborda o Tratamento Temático da Informação em sua perspectiva teórica, atrelada ao universo da Organização do Conhecimento. Objetiva explicitar a essência caracterizadora do tratamento temático da informação a partir das definições, presentes na literatura científica brasileira, que versam sobre o assunto. Enquanto investigação descritiva de abordagem qualitativa, aplica a revisão sistemática da literatura e a análise de conteúdo, como técnicas de coleta e análise dos dados. Tem como *corpus* de investigação, os artigos de periódicos levantados na Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação, tendo por base os critérios de aderência temática ao assunto e da presença de definições a seu respeito no texto. Os resultados demonstram que o tratamento temático da informação encontra distintas caracterizações na literatura brasileira, embora sobressaia sua perspectiva processual, assim como o entendimento de seu caráter detido ao assunto. Conclui-se que o tratamento temático da informação é, em essência, designado pelo aspecto temático que lhe é característico.

Palavras-chave: Organização do conhecimento, tratamento temático da informação, artigos científicos.

Introdução

Enquanto tema de investigação na Ciência da Informação (CI), o Tratamento Temático da Informação (TTI) corporifica-se em uma dimensão aplicada de trato informacional por assunto, que visa apontar o conteúdo temático (Monteiro & Giraldes, 2008), por meio de sua identificação, de seu processamento e de sua disponibilização (Rodrigues & Cervantes, 2015). Destarte, a instância de assunto é peculiar à

área de Biblioteconomia e Ciência da Informação, na qual este elemento é examinado e se torna objeto de classificação e indexação (Hjørland, 2017; Oliveira, Grácio & Martínez-Ávila, 2020).

A delimitação da abordagem temática, na literatura da área, teve em Antony Charles Foskett um dos principais agentes, quando da publicação de *The subject approach to information*, em 1969 (Guimarães, 2008; Café & Sales, 2010; Hider, 2020; Oliveira, Grácio & Martínez-Ávila, 2020), posto que ele não conseguiu combinar a abordagem do assunto com a catalogação descritiva, manifestando-se em prol de um tratamento mais amplo em organização do conhecimento (OC) (Hider, 2020). Entretanto, houve contributo de outros teóricos ao desenvolvimento do TTI, tais como Cutter, Mills, Barbosa, Piedade e Coates (Hickey, 1976; Sales, 2012).

O enfoque sobre a dimensão temática também contou com certo condicionamento histórico, na medida em que esforços passaram a ser feitos, entre o final do século XIX e o início do século XX, para facilitar a abordagem do assunto (Hider, 2020), cuja complexidade de representação começou a ser mais amplamente percebida (Pando, 2018). Mormente, a explosão documental pós Segunda Guerra Mundial incitou a transição mais incisiva da cultura de guarda e conservação para outra do acesso, embora com forte ênfase sobre o controle bibliográfico de coleções de ciência e tecnologia e sua recuperação de forma precisa (Sousa, 2013; Hider, 2020).

Na perspectiva brasileira, o TTI é definido como macroprocesso responsável por apresentar os atributos de assunto dos documentos (Oliveira, Grácio & Martínez-Ávila, 2020). Também é visto como dimensão da organização da informação mediadora entre o documento e o usuário (Braz & Carvalho, 2017) e, enquanto área que abrange a determinação e a especificação do assunto (Bräscher & Guimarães, 2018). No contexto da organização do conhecimento, o tratamento temático da informação encontra subsídio teórico (Barité, 2001) à sua consecução, sendo responsável pela disseminação de uma perspectiva mais técnica em organização do conhecimento, a partir dos estudos gerados desde os anos de 1960 (Sales, 2017).

Logo, em seu cerne caracterizador de conteúdo, o TTI dialoga historicamente com a OC. Essa consonância é explicada por Guimarães (2001), segundo o qual, em termos biblioteconômicos, a organização do conhecimento se insere justamente em seu aspecto temático, mormente, lidando diretamente com a vertente de conteúdo (Oliveira & Martínez-Ávila, 2019). Dito de outro modo, o tratamento temático da informação encontra na OC o universo teórico necessário a sua concretização (Guimarães, 2008).

Outrossim, áreas relacionadas à Organização do Conhecimento e da Informação são responsáveis por refletirem sobre a fruição da informação, prescindindo de investigações sobre seu percurso (Gracioso, Martínez-Ávila & Simões, 2019). Tal aporte de estudos se faz necessário com o próprio TTI, que carece de mais investigações e de sedimentação conceitual (Guimarães & Sales 2010; Pando, 2018), já que não recebeu a mesma atenção que o tratamento descritivo (Pando, 2018), tendo sido, na verdade, relegado e negligenciado (Hutchins, 1977; Oliveira, Grácio & Martínez-Ávila, 2020), ao longo dos anos. Mesmo nos currículos de Biblioteconomia, outras atividades de OC, como indexação e resumos, foram abordadas de forma espaçada e tardia (Hider, 2020).

É, todavia, na literatura brasileira de Ciência da Informação, que o tratamento temático da informação encontra sustentação, uma vez que aqui se institucionalizou como objeto de investigação (Gracioso, Martínez-Ávila & Simões, 2019). Nesse entendimento, a pesquisa objetiva explicitar a essência caracterizadora do tratamento temático da informação a partir das definições, presentes na literatura científica brasileira, que versam sobre o assunto. Parte do pressuposto que emanam, da literatura, diferentes caracterizações ao processo, em razão de seu desenvolvimento teorizante multifacetado e das variantes em seu entorno estrutural. Considera, também, a recente difusão do termo, de forma sistemática (Pando, 2018), assim como a diversidade de concepções teóricas (Guimarães, Danuello & Menezes, 2003) a seu respeito.

Desse modo, a investigação se estabelece sobre a problemática teorizante do TTI, em sua busca por evidência conceitual e ascepção individual no contexto bibliotecônico-informacional. Nesse intento, emprega revisão sistemática da literatura e análise de conteúdo.

Metodologia

A investigação é descritiva e de abordagem qualitativa. Emprega, respectivamente, como técnicas de coleta e análise dos dados, a revisão sistemática da literatura e a análise de conteúdo. O *corpus* de investigação é composto por 58 artigos de periódicos da área de Ciência da Informação publicados no período de 2001 a 2020, versados no tema tratamento temático da informação e que trazem definições do assunto ao longo do texto, apresentados no Quadro 1.

Esse corpus é resultante do levantamento realizado na Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI), utilizando o termo de busca “tratamento temático da informação” e seus correlatos análise de assunto”, “análise documentária”, “análise temática”, “descrição de conteúdo”, “descrição temática”, “organização temática” e “representação temática”, em dezembro de 2020. Adotou-se filtro de busca por termo composto, com uso de aspas, pesquisa sobre os campos de título, resumo e palavras-chave e, refinamento por ano de publicação, sobre o período de 2001 a 2020.

Ano	Autoria	Título
2001	LARA, Marilda Lopes Gínez de.	O Unicórnio (o Rinoceronte, o Ornitorrinco...), a Análise Documentária e a Linguagem Documentária.
	NAVES, Madalena Martins Lopes.	Estudo de fatores interferentes no processo de análise de assunto.
2003	FUJITA, Mariângela Spotti Lopes.	A identificação de conceitos no processo de análise de assunto para indexação.
2004	DIAS, Eduardo Wense.	Análise de assunto: percepção do usuário quanto ao conteúdo de documentos.
2007	BASTOS, Dilza Ramos; CAMPOS, Maria Luiza de Almeida.	Análise documentária de crônicas jornalísticas: uma busca de diretrizes teóricas e metodológicas.
	MAIMONE, Giovana Deliberali; GRACIOSO, Luciana de Souza.	Representação temática de imagens: perspectivas metodológicas.

2008	GUIMARÃES, José Augusto Chaves.	A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO).
2010	FUJITA, Mariângela Spotti Lopes; BOCCATO, Vera Regina Casari; RUBI, Milena Polsinelli.	O contexto da indexação para a catalogação de livros em abordagem sociocognitiva.
	GUIMARÃES, José Augusto Chaves; SALES, Rodrigo de.	Análise documental: concepções do universo acadêmico brasileiro em Ciência da Informação.
2011	CAVALCANTI, Hugo Carlos; CARVALHO, Maria Auxiliadora.	A informação na música impressa: elementos para análise documental e representação de conteúdos.
	LARA, Marilda Lopes Ginez de.	Conceitos de organização e representação do conhecimento na ótica das reflexões do Grupo Temma.
2012	DAL'EVEDOVE, Paula Regina; FUJITA, Mariângela Spotti Lopes.	Teoria e prática em catalogação de assunto: a sistematicidade do processo em contexto de bibliotecas universitárias pela perspectiva profissional.
	GUIMARÃES, José Augusto; SALES, Rodrigo de; GRÁCIO, Maria Cláudia Cabrini.	A dimensão interdisciplinar da análise documental nos contextos brasileiro e espanhol no âmbito da organização do conhecimento.
	NEVES, Dulce Amélia de Brito.	Representação temática da informação e mapas cognitivos: interações possíveis.
	REDIGOLO, Franciele Marques; ALMEIDA, Carlos Cândido de.	Algumas contribuições da perspectiva filosófico-semiótica de Peirce para a análise de assunto.
	SOUSA, Brisa Pozzi de; ALMEIDA, Carlos Cândido de.	Um olhar semiótico sobre o processo de indexação: a questão da representação e do referente.
	SOUSA, Dulce Elizabeth Lima de; BANDEIRA, Pablo Matias; SILVA, Márcio Bezerra da.	A representação temática em documentos arquivísticos: o caso da indexação documental realizada pelos alunos de Arquivologia da UFPB.
	TONELLO, Isângela M. S.; LUNARDELLI, Rosane Suely Álvares; ALMEIDA JÚNIOR, Oswaldo Francisco de.	Palavras-chave: possibilidades de mediação da informação.
2013	DAL'EVEDOVE, Paula Regina; FUJITA, Mariângela Spotti Lopes.	Estudo sociocultural da comunidade discursiva do tratamento temático da informação em bibliotecas universitárias.
	SOARES, Neila Celia; LUNARDELLI, Rosane Suely Álvares.	A representação da informação no âmbito das instituições de ensino superior: em foco os relatórios finais da Assessoria de Estatística do Centro de Ciências Exatas da Universidade Estadual de Londrina.
	SOUSA, Brisa Pozzi de.	Representação temática da informação documentária e sua contextualização em biblioteca.
	SOUSA, Brisa Pozzi de; FUJITA, Mariângela Spotti Lopes.	A classificação bibliográfica no contexto do tratamento temático da informação: um estudo com o protocolo verbal individual em bibliotecas do Instituto Federal de Educação, Ciência e Tecnologia (IF'S).
2014	LIMA, Gercina Ângela Borém de Oliveira; MACULAN, Benildes Coura Moreira dos Santos.	Análise de assunto a partir de uma perspectiva histórica do ARIST.
	MELO, Maria Antônia Fonseca; BRÄSCHER, Marisa.	Requisitos Funcionais para Dados de Autoridade Assunto (FRSAD): entidades, atributos e relacionamentos.

	REDIGOLO, Franciele Marques.	O processo de análise de assunto na catalogação em bibliotecas universitárias para proposta de normalização: estudo de observação com protocolo verbal.
2015	CATARINO, Maria Elisabete; CERVANTES, Brígida Maria Nogueira; ANDRADE, Ilza Almeida de.	A representação temática no contexto da web semântica.
	RODRIGUES, Maria Rosemary; CERVANTES, Brígida Maria Nogueira.	Análise de assunto e mapas conceituais: semelhanças nos processos.
2016	ALVES, Roberta Caroline Vesú; NEVES, Dulce Amélia de Brito; FUJITA, Mariângela Spotti Lopes; MORAES, João Batista Ernesto de.	Estratégias metacognitivas para análise de assunto: aspectos teóricos de superestrutura e esquemas sobre textos literários infanto-juvenis.
	FERREIRA, Ana Carolina; MACULAN, Benildes Coura Moreira dos Santos.	Organização da informação jurídica: proposta de modelo de leitura técnica de acórdãos no contexto do controle externo.
	MACHADO, Viviane Faria; ALBUQUERQUE, Ana Cristina de.	A representação temática da xilogravura: o processo de indexação da Coleção Paulo Menten.
	MEDEIROS, Graziela Martins de; VITAL, Luciane Paula; BRÄSCHER, Marisa.	Tratamento temático da informação em documentos arquivísticos: estudo dos anais da ISKO e do GT2 do ENANCIB.
2017	BRAZ, Márcia Ivo; CARVALHO, Evanise Souza de.	Práticas em tratamento temático da informação: interfaces de ensino e aprendizagem.
	CERVANTES, Brígida Maria Nogueira; SUENAGA, Cynthia Maria Kiyonaga; RODRIGUES, Maria Rosemary.	Os conceitos no tratamento da informação arquivística: unidade basilar para a compreensão do conteúdo documental.
	FERREIRA, Ana Carolina; MACULAN, Benildes Coura Moreira dos Santos.	Indexação de acórdãos no contexto dos tribunais de contas: estudos preliminares para a elaboração de um modelo de leitura técnica.
	FERREIRA, Ana Carolina; MACULAN, Benildes Coura Moreira dos Santos; ALMEIDA, Josiana Florêncio Vieira Régis de.	Fundamentos teóricos da representação temática da informação jurídica.
	LOPES, Felipe Vlxio; FERREIRA, Adne Laís Mendonça.	A análise documentária como ferramenta no trabalho de genealogia da base FamilySearch.org.
	REDIGOLO, Franciele Marques; SILVA, Marli Vitor da.	A representação temática como mediadora implícita da informação em bibliotecas universitárias.
	SANTOS, Charlene Maria dos; PINHO, Fabio Assis.	Aplicação do percurso temático e figurativo em literatura de cordel.
	SANTOS, Raimunda Fernanda dos.	Indexação em repositórios digitais: uma abordagem sobre o metadado assunto da Biblioteca Digital de Monografias da UFRN.
	VIEIRA, Ana Paula da Fonseca; OLIVEIRA, Laís Pereira de; CUNHA, Tatielle Marques.	Incurções sobre o tratamento temático da informação: estudo da política de indexação em bibliotecas universitárias goianas.
2018	BRÄSCHER, Marisa; GUIMARÃES, José Augusto Chaves.	Tratamento temático da informação (TTI): influência dos paradigmas físico, cognitivo e social em artigos de revisão de literatura no período de 1966-1995.
	BRENTAN JUNIOR, Elisio Custodio; MARTINS, Bianca Rodrigues; SANTOS NETO, João Arlindo dos.	A mediação cultural e a análise de assunto: mais que discursos, unindo comunidades.
	FERREIRA, Ana Carolina; MACULAN, Benildes Coura Moreira dos Santos.	Análise de assunto de acórdãos jurisprudenciais.

	GOMES, Thulio Pereira Dias.	Desafios e perspectivas para a organização de charges.
	SANTOS, Raimunda Fernanda dos; FELIPE, Carla Beatriz Marques.	A formação dos(as) bibliotecários(as) frente às novas possibilidades de representação da informação: análise das propostas pedagógicas dos cursos de Biblioteconomia do Norte e do Nordeste do Brasil.
2019	CARVALHO, Cláudia Pereira de Jesus; ALMEIDA, Carlos Cândido de.	Coleções de multimeios: reflexões sobre o tratamento temático da informação a partir do conceito de tradução intersemiótica.
	FEITOSA, Kézia de Lira; SALCEDO, Diego Andres.	Selos comemorativos brasileiros: uma proposta de representação do conhecimento.
	LUNARDELLI, Rosane Suely Álvares; MOLINA, Letícia Gorri; TONELLO, Izângela M. S.; BARBOSA, Aretusa Marques.	A análise da informação e seu ensino nos curso de Arquivologia e Biblioteconomia da Universidade Estadual de Londrina: relato de experiência.
	OLIVEIRA, Lais Pereira de; MARTÍNEZ-ÁVILA, Daniel.	Tratamento temático da informação: análise sobre abordagens e enfoques em artigos científicos da área de Ciência da Informação.
	RABELO, Camila Regina de Oliveira; BENTES PINTO, Virgínia.	Tendências nos estudos de representação temática da informação: uma revisão integrativa dos artigos científicos indexados na BRAPCI.
	SUNDSTRÖM, Admeire da Silva Santos; MORAES, João Batista Ernesto de; ALBUQUERQUE, Ana Cristina.	Filme de ficção para a Ciência da Informação: um estudo sobre as abordagens de organização e representação temática.
2020	DANTAS, Esdras Renan Farias; SAMPAIO, Débora Adriano; ALBUQUERQUE, Maria Elizabeth Baltar Carneiro de.	Avaliação da consistência de descritores: a representação da informação relacionada à temática responsabilidade social nas dissertações do PPGCI-UFPB.
	DIAS, Célia da Consolação.	Representação temática de imagens: reflexões acerca dos subsídios da indexação manual e do reconhecimento de imagens.
	DIAS, Danilo Camargo; MOREIRA, Walter; ALVES, Rachel Cristina Vesú.	A representação temática de imagens digitais da NASA no Flickr: as contribuições dos sistemas de organização do conhecimento.
	LAGES, Sandra Regina Moitinho; LUNARDELLI, Rosane Suely Álvares.	A representação temática da COVID-19 na literatura de cordel.
	OLIVEIRA, Lais Pereira de.	Por uma abordagem temática de organização da informação.
	OLIVEIRA, Lais Pereira de; GRÁCIO, Maria Cláudia Cabrini; MARTÍNEZ-ÁVILA, Daniel.	A expressão “Tratamento Temático da Informação” em artigos de periódicos nacionais: análise da ocorrência e de suas variantes designativas.
	VIEIRA, Keitty Rodrigues; BRÄSCHER, Marisa; SILVA, Eva Cristina Leite da; KARPINSKI, Cezar.	A Escola de Chicago e a dimensão temática da informação.

Quadro 1. Artigos sobre tratamento temático da informação integrantes do estudo.

O total de trabalhos refletiu a seleção sistemática sobre os registros obtidos durante a busca na BRAPCI. De um conjunto de 264 retornados, foram selecionados 186, diante das duplicatas e das categorias não enquadradas como artigo de texto completo, publicado em revista brasileira por autor brasileiro, restando, por fim, 76, que efetivamente tratavam do tema e, 58, que realmente o definiam no decorrer do texto.

Adotou-se tal metodologia visando atender ao propósito de estabelecimento da essência caracterizadora do TTI com base nas definições do assunto nos artigos da área da CI. E, em essência, dentro da perspectiva assumida pelo tema no Brasil. Destarte, sobrepõe-se aspecto condicionante à caracterização do tratamento temático da informação e, provisionalmente, a seu entendimento teórico na forma como é expresso nas publicações.

Resultados

Substancialmente, a partir da prospecção sobre os 58 artigos de periódicos brasileiros da área de Ciência da Informação, versados no tema TTI, foi possível constatar distintas caracterizações. As 7 mais recorrentes são expressas em um total de 35 trabalhos, como observado na Figura 1:

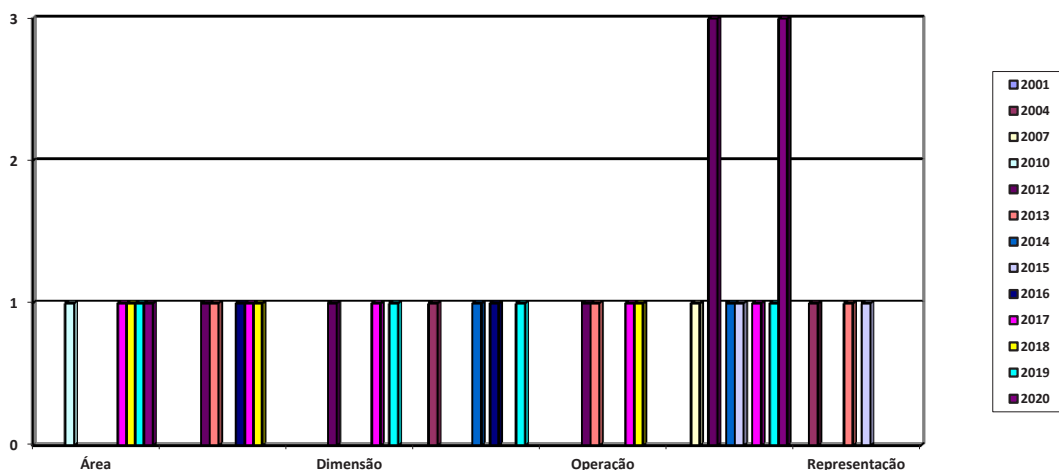


Figura 1. Caracterização do TTI.

De acordo com a Figura 1, predominam concepções de tratamento temático da informação enquanto processo (11 trabalhos). Além disso, o TTI também é visto como área (5), atividade (5), etapa (4), operação (4), dimensão (3) e representação (3). Nota-se, pois, uma tentativa assertiva de caracterização pelos autores dos artigos, sobretudo por sua distinção processual.

Embora não representadas na Figura 1, observam-se junto aos 23 artigos que completam o corpus analítico de 58, 13 concepções de TTI, que surgem em menor número. É o caso de abordagem, ação, análise, caráter, conceito, contributo, expressão, forma, mediador, prática, procedimento, temática e tratamento, sinalizadas, cada qual, em uma única publicação. Nesse sentido, ampliam-se as perspectivas acerca do tratamento temático da informação, a ponto de ser evidenciado em sua via contributiva e mediadora, por exemplo, o que ajuda a ampliar a compreensão sobre seus diversos atributos.

Também entre os 23 artigos restantes, constata-se outros 10 que apresentam mais de uma caracterização para o tratamento temático da informação ao longo do texto, a saber: área, atividade e operação (1); área e atividade (2); linha, conjunto, operação e atividade (1); linha e forma (1); processo, atividade e etapa (1); processo, atividade, operação e etapa (1); processo, etapa, atividade e tema (1); processo e forma (1); e processo, técnica, estágio, etapa e termo (1). Aqui, observa-se certa pulverização das concepções em TTI pelos teóricos brasileiros, na medida em que um mesmo texto sinaliza para mais de um caráter a seu respeito, o que permite enxergá-lo nas várias nuances assumidas enquanto atividade prática ou área de estudo, embora sem muita assertividade nessa delimitação por um mesmo autor.

Pontualmente, dentro do conjunto de 35 artigos com caracterizações mais recorrentes acerca do TTI, foi possível rotular instâncias definidoras, como agrupado no Quadro 2:

Caracterizações	Instâncias definidoras	Autoria
Área	concernente ao assunto	FUJITA, BOCCATO e RUBI
	de estudo	CERVANTES, SUENAGA, RODRIGUES
	desafiadora	BRÄSCHER e GUIMARÃES
	interdisciplinar e técnica	RABELO e BENTES PINTO
	de organização e representação de assuntos	DANTAS, SAMPAIO, e ALBUQUERQUE
Atividade	em busca da melhor representação do conteúdo	TONELLO, LUNARDELLI e ALMEIDA JÚNIOR
	intelectual	SOUSA e FUJITA
	de mediação	FERREIRA e MACULAN
	de representação	BRAZ e CARVALHO
	intelectual	FERREIRA e MACULAN
Dimensão	centrada no tratamento temático da informação	GUIMARÃES, SALES e GRÁCIO
	do tratamento da informação	VIEIRA, OLIVEIRA e CUNHA
	organizativa aplicada	OLIVEIRA, GRÁCIO e MARTÍNEZ-ÁVILA
Etapa	do tratamento temático da informação	DIAS
	de determinação do conteúdo	REDIGOLO
	da análise de conteúdo	ALVES, NEVES, FUJITA e MORAES
	que resulta na representação	SUNDSTRÖM, MORAES e ALBUQUERQUE
Operação	base	NEVES
	de tratamento temático da informação	DAL'EVEDOVE e FUJITA
	de tratamento temático da informação	SANTOS
	de tratamento temático da informação	SANTOS e FELIPE
Processo	amplo	BASTOS e CAMPOS
	concernente ao conteúdo	SOUSA, BANDEIRA e SILVA
	relevante e essencial	REDIGOLO e ALMEIDA
	anterior à recuperação	SOUSA e ALMEIDA
	descritivo complexo	MELO e BRÄSCHER
	conectado com a organização e representação do conhecimento	RODRIGUES e CERVANTES
	de descrição de dados	LOPES e FERREIRA
	com finalidade de geração de produtos documentais	LUNARDELLI, MOLINA, TONELLO e BARBOSA
	de transformação do conteúdo	VIEIRA, BRÄSCHER, SILVA e KARPINSKI

	de grande envergadura	OLIVEIRA
	de organização	DIAS
Representação	de conteúdos	LARA
	temática da informação	SOARES e LUNARDELLI
	do conteúdo	CATARINO, CERVANTES e ANDRADE

Quadro 2. Instâncias definidoras dos artigos sobre tratamento temático da informação.

Assim, em conformidade com o Quadro 2, observa-se uma ampla compreensão do TTI enquanto processo de análise, descrição, determinação, organização, representação e tratamento temático, do conteúdo ou assunto. Aspecto esse reforçado em: Fujita, Boccato e Rubi; Dantas, Sampaio e Albuquerque; Tonello, Lunardelli e Almeida Júnior; Guimarães, Sales e Grácio; Dias; Redigolo; Alves, Neves, Fujita e Moraes; Dal'Evedove e Fujita; Santos; Santos e Felipe; Sousa, Bandeira e Silva; Vieira, Bräscher, Silva e Karpinski; Lara; Soares e Lunardelli; e, Catarino, Cervantes e Andrade.

Percebem-se, ainda, instâncias indicativas do caráter desafiador, interdisciplinar, técnico, intelectual, mediador, aplicado, básico, amplo, relevante, essencial, descritivo, complexo, transformador e de grande envergadura do TTI, bem como enquanto área de estudo. Desse modo, constata-se o cuidado de alguns teóricos em adjetivar essa dimensão temática de trato informacional.

Vale ressaltar que, pela amplitude das definições selecionadas em cada artigo – mais de uma em boa parte deles –, o rol aqui estabelecido reflete uma das dimensões observadas durante a análise de conteúdo, concernente ao que é evidenciado pelos autores como caráter principal constitutivo e que exprime o tratamento temático da informação. Nessa medida, outros aspectos precisam ser explorados posteriormente, com aprofundamento sobre os demais atributos derivados das definições, tais como: processos associados ao TTI, contexto de desenvolvimento, objetivos, função principal e operações base.

Conclusões

A investigação buscou explicitar a essência caracterizadora do tratamento temático da informação a partir das definições presentes na literatura científica contemporânea da Ciência da Informação nos periódicos brasileiros. Infere-se, a partir disso, que há distintas caracterizações para o tema no âmbito dos artigos científicos publicados em periódicos brasileiros, apesar do predomínio de sua concepção processual. Também sobressai seu caráter detido ao assunto, via ações de análise, descrição, determinação, organização, representação e tratamento.

O pressuposto segundo o qual emanam, da literatura, diferentes caracterizações do TTI, em razão de seu desenvolvimento teorizante multifacetado e das variantes em seu entorno estrutural, foi parcialmente confirmado. Demanda, a partir de então, o aprofundamento sobre outros aspectos derivados das definições dos teóricos brasileiros para o que se entende por tratamento temático da informação, inclusive no que se refere à sua associação ao universo da organização da informação ou do conhecimento.

Portanto, uma implicação da investigação é o indicativo inicial de compreensão do TTI enquanto processo voltado à dimensão temática. Com efeito, outras variantes auxiliarão na ampliação desse entendimento, a começar pelos demais atributos presentes nas definições, por si só resultantes de diferentes concepções, a depender da vertente de formação e da ênfase de estudo do autor.

Destacam-se, enquanto limitações da investigação, a ênfase sobre artigos de periódicos, publicados em âmbito brasileiro, considerando que outras tipologias de publicações podem denotar caracterizações de outra ordem e, mesmo, sob influência de diferentes correntes e linhas teóricas, inclusive dentro das perspectivas estrangeiras em Biblioteconomia e Ciência da Informação. Reitera-se, portanto, a excepcionalidade teórica da dimensão temática de trato informacional, aqui investigada, como concebida no entorno científico brasileiro.

Estudos futuros podem se dedicar à explicitação do caráter do tratamento temático da informação a partir de seus processos relacionados e operações derivadas. Recomenda-se, ademais, o estabelecimento de suas atividades centrais, de acordo com o que apresentam as definições, que fazem da dimensão temática um eixo mais compreendido e facilmente reconhecido na prática.

Referências

- Barité, M. (2001). Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. In *Educação, Universidade e Pesquisa*. Marília: UNESP-Marília-Publicações; São Paulo: FAPESP (p. 35-60).
- Bräscher, M. & Guimarães, J. A. C. (2018). Tratamento temático da informação (TTI): influência dos paradigmas físico, cognitivo e social em artigos de revisão de literatura no período de 1966-1995. *Liinc em Revista*, 14(2), 241-258. Recuperado em 17 junho, 2021, de <http://revista.ibict.br/liinc/article/view/4347/3949>
- Braz, M. I., & Carvalho, E. S. (2017). Práticas em tratamento temático da informação: interfaces de ensino e aprendizagem. *Revista Brasileira de Biblioteconomia e Documentação*, 13(esp.), 2496-2509. Recuperado em 17 junho, 2021, de <https://rbbd.febab.org.br/rbbd/article/view/975/951>
- Café, L. M. A., & Sales, R. (2010). Organização da informação: conceitos básicos e breve fundamentação teórica. In *Passeios pelo bosque da informação: estudos sobre representação e organização da informação e do conhecimento*. Brasília: IBICT (p. 115-129).
- Gracioso, L. S., Martínez-Ávila, D., & Simões, M. G. M. (2019). “Tratamento Temático da Informação” na pesquisa brasileira em Ciência da Informação: percursos e relações. *Scire*, 25(2), 23-24. Recuperado em 28 janeiro, 2021, de <https://www.iberid.eu/ojs/index.php/scire/article/view/4647/4239>
- Guimarães, J. A. C. (2001). Perspectivas de ensino e pesquisa em organização do conhecimento em cursos de Biblioteconomia: uma reflexão. In *Educação, Universidade e Pesquisa*. Marília: Unesp-Marília-Publicações; São Paulo: FAPESP (p. 61-72).
- Guimarães, J. A. C. (2008). A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). *Revista Ibero-americana de Ciência da Informação*, 1(1), 77-99. Recuperado em 26 janeiro, 2021, de <https://periodicos.unb.br/index.php/RICI/article/view/940/815>
- Guimarães, J. A. C., Danuello, J. C., & Menezes, P. J. (2003). Ensino de tratamento temático da informação (T.T.I.) nos cursos de Biblioteconomia do Mercosul: uma análise de capacitação e produção científica docente com vistas ao delineamento de políticas integradas

- para área. In Encontro Nacional de Pesquisa em Ciência da Informação: *Anais do V Encontro Nacional de Pesquisa em Ciência da Informação*. Belo Horizonte: Escola de Ciência da Informação da UFMG. Recuperado em 29 janeiro, 2021, de http://repositorios.questoeseemrede.uff.br/repositorios/bitstream/handle/123456789/147/ENAN023_Guimar%C3%A3es_Danuelo_Menezes.pdf?sequence=1
- Guimarães, J. A. C., & Sales, R. (2010). Análise documental: concepções do universo acadêmico brasileiro em Ciência da Informação. *DataGramaZero – Revista de Ciência da Informação*, 11(1). Recuperado em 26 janeiro, 2021, de <http://www.brapci.inf.br/index.php/article/view/0000008345/a42485a4423fe3279bedcd5a919ee61e/>
- Hickey, D. J. (1976). Subject analysis: an interpretive survey. *Library Trends*, 25, 273-291. Recuperado em 31 janeiro, 2021, de <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.172.9861&rep=rep1&type=pdf>
- Hider, P. (2020). Origins of the knowledge organization field. In *ISKO Encyclopedia of Knowledge Organization*. Recuperado em 29 janeiro, 2021, de <https://www.isko.org/cyclo/origins>
- Hjørland, B. (2017). Subject (of documents). *Knowledge Organization*, 44(1), 55-64. Recuperado em 29 janeiro, 2021, de <https://www.isko.org/cyclo/subject>
- Hutchins, W. J. (1977). On the problem of ‘aboutness’ in document analysis. *Journal of Informatics*, 1(1), 17-35. Recuperado em 27 janeiro, 2021, de <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.126.8936&rep=rep1&type=pdf>
- Monteiro, S. D., & Giraldes, M. J. C. (2008). Aspectos lógico-filosóficos da organização do conhecimento na esfera da Ciência da Informação. *Informação & Sociedade*, 18(3), 13-27. Recuperado em 28 janeiro, 2021, de <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/1775/2269>
- Oliveira, L. P., Grácio, M. C. C., & Martínez-Ávila, D. (2020). Instrumentos, processos e produtos do tratamento temático da informação: um estudo nos anais da ISKO Brasil (2012-2019). *Scire: Representación y organización del conocimiento*, 26(1), 47-56. Recuperado em 29 janeiro, 2021, de <https://www.iberid.eu/ojs/index.php/scire/article/view/4681>
- Oliveira, L. P., & Martínez-Ávila, D. (2019). Tratamento temático da informação: análise sobre abordagens e enfoques em artigos científicos da área de Ciência da Informação. *Informação em Pauta*, 4(2), 83-100. Recuperado em 27 janeiro, 2021, de <http://www.periodicos.ufc.br/informacaoempauta/article/view/42654/99896>
- Pando, D. A. (2018). *Epistemologia da organização da informação: uma análise de sua cientificidade no contexto brasileiro*. (Tese de Doutorado, Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil). Recuperado em 28 janeiro, 2021, de https://repositorio.unesp.br/bitstream/handle/11449/153395/pando_da_dr_mar.pdf?sequence=3&isAllowed=y
- Rodrigues, M. R., & Cervantes, B. M. N. (2015). Análise de assunto e mapas conceituais: semelhanças nos processos. *Perspectivas em Ciência da Informação*, 20(4), 35-56. Recuperado em 27 janeiro, 2021, de <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2423>
- Sales, R. (2012). *A presença de Kaiser no quadro teórico do tratamento temático da informação (TTI)*. (Tese de doutorado, Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil). Recuperado em 27 janeiro, 2021, de https://repositorio.unesp.br/bitstream/handle/11449/103381/sales_r_dr_mar.pdf?sequence=1&isAllowed=y
- Sales, R. (2017). Diferentes perspectivas nos contextos do GT2 da ANCIB e da ISKO-Brasil. In Encontro Nacional de Pesquisa em Ciência da Informação: *Anais do XVIII Encontro Nacional de Pesquisa em Ciência da Informação*. Marília: UNESP. Recuperado em 18 junho, 2021, de <https://www.brapci.inf.br/index.php/res/download/124907>

Sousa, B. P. (2013). Representação temática da informação documentária e sua contextualização em biblioteca. *Revista Brasileira de Biblioteconomia e Documentação*, 9(2), 132-146. Recuperado em 26 janeiro, 2021, de <https://rbbd.febab.org.br/rbbd/article/view/249/265>

ESTUDO DA PAISAGEM RURAL E SUA REDE DE CONCEITOS: ANÁLISE DE DOMÍNIO A PARTIR DAS PERSPECTIVAS DA COMUNIDADE DISCURSIVA

Vera Lucia Punzi Barcelos Capone¹, Rosa Inês de Novais Cordeiro²

¹Bibliotecária do Instituto Brasileiro de Geografia e Estatística (IBGE, Rio de Janeiro, Brasil). Doutora em Ciência da Informação, Universidade Federal Fluminense (Niterói, Brasil) caponevera@gmail.com
ORCIDiD <https://orcid.org/0000-0001-6384-5201>

²Professora Titular do Departamento de Ciência da Informação da Universidade Federal Fluminense (Niterói, Brasil). Doutora em Comunicação e Cultura, Universidade Federal do Rio de Janeiro (Rio de Janeiro, Brasil) rosanovais@id.uff.br ORCIDiD <https://orcid.org/0000-0003-1871-4995>

Resumo

Pesquisa sobre a análise da paisagem rural que objetiva identificar o conjunto de traços comuns que representam o padrão de comunicação da comunidade discursiva de geógrafos que analisa a paisagem como objeto de estudo, por meio do paradigma da análise de domínio, a fim de estabelecer uma rede de conceitos para constituição de Sistemas de Organização do Conhecimento (SOCs). Utiliza como procedimentos metodológicos as pesquisas bibliográfica e documental. A pesquisa bibliográfica, em bases de dados nacionais e estrangeiras, foi dividida entre os eixos da Ciência da Informação, voltada para a área de Organização e Representação do Conhecimento; e Geografia, direcionada ao estudo da paisagem. Na pesquisa documental examinou-se os Censos Agropecuários de 1970 e 2006 e o Atlas do Espaço Rural Brasileiro, em sua primeira edição de 2011, do Instituto Brasileiro de Geografia e Estatística (IBGE), como também, a identificação da representação da paisagem nos sistemas de classificações bibliográficas, tesouros e ontologias. No resultado, chegou-se à terminologia consensual e padronizada da paisagem por meio da conceituação das categorias que a representam, oriundas do domínio da Geografia, agregadas às informações provenientes dos censos agropecuários. Essas categorias foram sistematizadas e representam a estrutura classificatória da rede consensual identificada.

Palavras-chave: Paisagem rural, Censos Agropecuários, Atlas do espaço rural brasileiro (2011), Sistemas de Organização do Conhecimento (SOCs), Organização e representação do conhecimento (Geografia).

1 Introdução

Comunicação que apresenta um recorte do resultado de pesquisa de doutorado concluída na área de Organização e Representação do Conhecimento Geográfico registrado de diversas naturezas, no contexto dos Sistemas de Organização do Conhecimento (SOCs ou *Knowledge Organization Systems-KOS*). Objetiva identificar o conjunto de traços comuns que representa o padrão de comunicação da comunidade discursiva de geógrafos que analisa a paisagem como objeto de estudo, a fim de estabelecer uma rede de conceitos para constituição de SOCs. Procura-se apresentar a dimensão ontológica da paisagem.

A paisagem está, na maioria das vezes, ligada a quem a observa, sendo por isso sua percepção arbitrária. No entanto, a conceituação de paisagem pode se situar para além de descrições subjetivas e nos direciona à trajetória do conhecimento que extrapola os lugares-comuns, residindo nos centros de cálculos, categoria defendida por Latour (2000), para quem “há um certo risco em limitar a ecologia dos lugares de saber aos signos ou à simples matéria do escrito”.

O estudo do conceito de paisagem na geografia pós-moderna e sua compreensão como fonte documental direcionada aos SOCs é o tema desta pesquisa, que tem como referência empírica da pesquisa documental os Censos Agropecuários de 1970 e 2006¹ e o *Atlas do Espaço Rural Brasileiro*, em sua primeira edição de 2011, doravante identificado por *Atlas*, ambos sob a responsabilidade do Instituto Brasileiro de Geografia e Estatística (IBGE). Soma-se a isso, o exame da representação da paisagem nos sistemas de classificações bibliográficas, tesouros e ontologias.

A escolha dos censos agropecuários como objeto específico de análise deve-se à sua privilegiada condição de repertório consensual das áreas de Geografia e Estatística e ao seu potencial informativo como fonte de pesquisa singular que retrata a realidade do Brasil agrário e possibilita extrair deles informações acerca da paisagem rural. É oportuno observar que os censos agropecuários não analisam paisagens, e muito menos se reportam às variáveis incluídas como paisagens. Ao direcionarmos nosso olhar para as variáveis que investigam a realidade do Brasil agrário, percebemos que elas evocavam a paisagem rural, por serem portadoras de uma acentuada carga icônica.

O emprego do *Atlas* atua na condição de referencial de análise e se apresenta como a interpretação geográfica das informações agropecuárias obtidas pelos censos agropecuários por meio dos mapas, que são, segundo apontamentos teóricos formulados por Bertin (1988); Kozel (2013); Gil e Barleta (2015); e MacEachren (2004), uma representação gráfica que nos auxilia a compreender espacialmente objetos, conceitos, condições, processos e fatos sociais que se vinculam à percepção visual como assimilação básica da representação do conhecimento, substituindo o código escrito.

Entre os vários recortes para examinar a paisagem em outras áreas do conheci-

¹ Para este estudo não foi considerado o Censo agropecuário 2017, tendo em vista que o projeto de pesquisa foi apresentado em 2016, quando os resultados do Censo agropecuário 2017 ainda não tinham sido divulgados.

mento ou contextos, a exemplo das Engenharias agrônoma e florestal, Biologia, Ecologia, Arquitetura, Arqueologia, História, Economia, Turismo e Geografia, conforme destacado por Konkoly-Gyuró (2018) em seu trabalho sobre a conceitualização e percepção da paisagem e suas mudanças numa área transfronteiriça entre França e Alemanha, optamos nesta pesquisa pela abordagem teórico-metodológica sob o ponto de vista da Geografia para análise da paisagem rural.

Adotamos a análise de domínio postulada por Hjörland e Albrechtsen (1995) e Hjörland (2002, 2017) como paradigma analítico no desenvolvimento desta pesquisa, que assume protagonismo ao fornecer uma perspectiva integrada da área de Geografia e encontra suas várias concepções, facilitando explorar seus fundamentos teóricos ao funcionar como uma ferramenta orientadora e legitimadora em relação ao domínio analisado (Zins et al., 2007). Conforme pontua Guimarães (2014), a análise de domínio deriva da aplicação de fundamentos referentes à organização do conhecimento, ou seja, a categorização, que assim possibilita reunir objetos semelhantes e separar aqueles diferentes e dessa forma compreender os limites, fundamentos e relacionamentos que norteiam a cadeia comunicacional na esfera da informação geográfica no que diz respeito à paisagem.

A temática da análise de domínio é examinada à luz da Ciência da Informação (CI) na condição de dispositivo de mapeamento da área de Geografia para conhecermos como a paisagem se estrutura teoricamente enquanto objeto de estudo da comunidade discursiva de geógrafos. Prieto-Diaz (1990) define análise de domínio como um processo, no qual as informações aplicadas no desenvolvimento de sistemas de *software* são identificadas, capturadas e organizadas com o objetivo de se tornarem reutilizáveis na criação de outros sistemas. Por sua vez, Hjörland e Albrechtsen (1995) abordam análise de domínio sob a ótica de paradigma direcionado à identificação, descrição, organização e comunicação de fontes informacionais a fim de atender objetivos específicos e contribuir com o estudo de uma área de conhecimento.

López-Huertas (2006) compreende o domínio sob a perspectiva de modelos disciplinares que se constituem de fontes terminológicas e estruturais que demarcam a especialização do domínio, delimitando-lhe as fronteiras e seus limites.

Nossa opção pela análise de domínio se deu no intuito de conhecer, compreender e organizar o conjunto de conceitos e o universo vocabular relativos à paisagem numa estrutura lógica e coerente, cujo recorte refletisse o objeto de estudo da comunidade discursiva de geógrafos.

O conceito de comunidade de discurso foi formulado por John Swales, em 1990, e se estende à relevância do propósito comunicativo de um grupo de pessoas e seu discurso operando de acordo com as convenções definidas pela comunidade, seja na área acadêmica ou profissional. Para Swales (1990) o discurso é um meio de manter e ampliar o conhecimento do grupo e iniciar novos membros. A definição de comunidade discursiva se fundamenta, portanto, na atividade desenvolvida por seus integrantes e requer uma rede de comunicação e objetivos comuns, mesmo que haja uma distância geográfica entre os participantes. Entre os princípios fundamentais que orientam uma comunidade de discurso, destacam-se, segundo Swales (1990), objetivos comuns com o fito de fortalecer procedimentos de interlocução entre os participantes com vistas ao aprimoramento na troca de informações entre os mem-

bros da comunidade, que compartilham um léxico específico e à associação de outros membros com nível de *expertise* e discurso relevantes.

Guimarães (2014), Hjörland e Albrechtsen (1995) se apropriam da expressão comunidade de discurso com uma nova perspectiva de abordagem ao propor à CI estudar os domínios do conhecimento como categorias discursivas, alicerçada numa relação dialética entre uma comunidade e seus membros, compartilhando conhecimento e questionamentos em torno de disciplina específica, cuja ênfase recai no contexto sociológico e menos no indivíduo. Essa rede de conceitos proveniente tanto dos censos agropecuários como da comunidade discursiva dos geógrafos traduz-se em configurações da paisagem e torna-se informação quando são atribuídos sentido e significado ao conjunto de objetos informacionais fornecidos por meio do mapeamento da literatura da área de Geografia, ou seja, o arranjo que combina consensualmente categorias, conceitos e terminologia padronizada, resultante da organização do espaço geográfico que se manifesta nas Regiões Centro-Oeste e Sul do Brasil.

Entre o conjunto de onze proposições para estudar um domínio específico, selecionamos a segunda abordagem proposta por Hjörland (2002, 2017), que versa sobre a construção de terminologia especializada, portais, esquemas classificatórios, taxonomias e ontologias como possíveis aplicações de SOCs, partindo do conhecimento gerado pela comunidade discursiva dos geógrafos no âmbito da Geografia. A organização da informação e do conhecimento em SOCs se justifica pela necessidade de se conceituar e especificar um universo vocabular, tendo em vista que a importância dos SOCs é medida pelo grau de organização que lhe é atribuído (Gnoli, 2013).

A visão analítica do domínio da Geografia com ênfase na paisagem, por meio da ferramenta do paradigma de análise, lança luz sobre sua terminologia e proporciona visibilidade ao *corpus* do domínio, favorecendo o atendimento às demandas específicas dos usuários deste domínio. De acordo com Hjörland (2002), os domínios do conhecimento não devem ser vistos como semelhantes, mas examinados sob a perspectiva de diferentes comunidades discursivas.

Entre as condições consensualmente aceitas pela CI que qualificam a análise de domínio se situam as garantias literária e semântica por concentrarem as ideias nucleares do domínio e são fontes para extração e validação da terminologia de uma área de conhecimento a ser integrada a um sistema de classificação ou qualquer outro sistema de organização do conhecimento, que se sustenta por meio da documentação científica e especializada e provem da análise de domínio.

Desse modo, nos pautamos nos trabalhos de Barité et al. (2010), Beghtol (1986) e Guedes & Moura (2016), que tratam das garantias literária e semântica por proporcionarem a identificação das ideias nucleares do domínio geográfico e dos censos agropecuários, no tocante à paisagem rural, e assim exercerem a função de fontes para extração e validação da terminologia a ser integrada ao Painel conceitual da paisagem (Figura 1) e à Estrutura classificatória da paisagem (Figura 2), que resultam da análise de domínio e são assegurados pelas garantias literária e semântica.

Visando conhecer o padrão de comunicação da comunidade discursiva de geógrafos, o campo teórico sobre a dimensão ontológica da paisagem se fundamenta no estudo do conceito de paisagem e é referendado por textos acadêmicos sob a chancela de Bertrand (2004), entre outros, que se refere ao conceito de paisagem como

um sistema de conceitos; Santos (2014) que vê na paisagem uma sucessão de forças naturais e artificiais; Saquet (2015) que a depreende como meio de representação dos símbolos da vida cotidiana; Sauer (1998) percebe na paisagem a associação de forças físicas e culturais e partindo dessa compreensão introduz sua classificação em natural e cultural.

Os elementos perceptíveis da paisagem originados de atividades econômicas, como é o caso das práticas agrícolas; áreas ocupadas por estabelecimentos agropecuários, tais como estábulos, fazendas, armazéns, caminhos, entre outros; utilização das terras, abrangendo lavouras, pastagens, matas e florestas; e recursos hídricos, suscitam e englobam a paisagem rural. Essa paisagem nos vem dos censos agropecuários, pois assumem a condição de informações icônicas porque evocam imagens da paisagem rural e possui fortes implicações visuais; corresponde às variáveis investigadas pelos censos agropecuários selecionadas para esta pesquisa, estando seu conceito intimamente ligado às ações antrópicas, por um lado (Bertrand, 2004; Manual, 2013; Santos, 2014), ou seja, deve ser analisada pelas diferentes formas de uso da terra como recurso econômico, e por recursos naturais, por outro lado, sendo impactados pelas atividades humanas, como argumentam Veronezzi e Fajardo (2015, p. 214-215): “As atividades socioeconômicas são elementos cruciais e permitem entender os processos que produzem as paisagens rurais [...]. Entender a paisagem rural é compreender tanto a dinâmica econômica, quanto a dos indivíduos”.

A paisagem rural está intimamente ligada ao uso da terra e à sua cobertura. Conceitualmente, o uso da terra se associa às atividades exercidas pelo homem numa extensão de terra ou a um ecossistema com o objetivo de extrair produtos e benefícios por meio dos recursos da terra. Quanto à cobertura da terra, seu conceito é atribuído aos elementos da natureza, entre os quais a vegetação: natural (matas, florestas), plantada (agricultura), e também pela água, gelo, rocha e areia, segundo o Manual (2013) com base nos estudos de Bie, Leeuwen e Zuidema (1996).

Lepczyk, Lortie e Anderson (2008), autores de um de dois estudos sobre ontologia da paisagem, recuperados na plataforma ScienceDirect, descrevem a paisagem por meio de um padrão (sua composição e característica), processo (fluxos bióticos) e faz parte de uma ecorregião. Para estes autores, a paisagem é definida por sua vegetação, cobertura e uso da terra e seu ecossistema. O outro estudo, de Konkoly-Gyuró (2018), a paisagem é o resultado de várias interações e pode ser representada pelo padrão de cobertura e uso da terra, sendo este último um processo contínuo a fim de atender às várias necessidades humanas. Além dos fatores naturais, o uso da terra sofre influência de condições socioeconômicas, entre outros, conjugados pela situação geopolítica. A paisagem é parte da identidade regional e nacional.

Sendo um mundo de diversidades, a paisagem rural combina fatores de produção devidos à agricultura e diversidade de climas que influenciam o plantio e a colheita, a rotatividade dos solos, as técnicas, estruturas e sistemas de produção, que vão das mais elementares como as queimadas e alternância dos solos às técnicas mais sofisticadas em relação aos contextos econômicos (Cavaco, 2005).

O quadro conceitual da paisagem é formado por diferentes abordagens teóricas e seu entendimento é mediado pela sinergia que se estabelece da depreensão de seus sistemas naturais e antrópicos, que agrupados revelam a dimensão de suas características e, por conseguinte, as categorias que atuarão para formar sua classificação

no contexto da organização do conhecimento registrado e no domínio da Geografia. O mapeamento de diferentes enfoques teóricos sobre a paisagem, no âmbito da CI, nos fornece um conjunto de conceitos atualizados da área de Geografia e enfatiza a noção classificatória e categórica da paisagem.

2 Metodologia

O problema de pesquisa com o qual nos deparamos é: como tornar os tópicos teóricos relacionados à paisagem numa concepção unificada e sistemática que possa fornecer seu padrão de organização quanto a conceitos, categorias e significados? Nossa hipótese é de que o mapeamento da área de Geografia por meio de seu *corpus* teórico oferece um instantâneo da área e, igualmente, fundamentos coerentes baseados no consenso acadêmico que viabilizam a sistematização do conhecimento produzido pela comunidade discursiva dos geógrafos que configura o padrão de organização relativo à paisagem. Um corpo de informações organizadas sistematicamente qualifica o conhecimento.

Os procedimentos metodológicos adotados nesta pesquisa com vistas a fornecer e organizar o conjunto de conceitos e o universo vocabular relativos à paisagem, cujo recorte refle o objeto de estudo da comunidade discursiva de geógrafos foram divididos em pesquisa bibliográfica e pesquisa documental. A pesquisa bibliográfica foi sistematizada em dois eixos, a saber: no âmbito da CI, contemplando a área de representação da informação e organização do conhecimento; e na área de Geografia, direcionada ao estudo da paisagem.

A pesquisa documental teve por escopo examinar como o assunto paisagem é organizado e representado nos sistemas de Classificação Decimal Universal (CDU) e Classificação Decimal de Dewey (CDD), bem como em tesouros (Agrovoc; Thesaurus das Nações Unidas; Everglades Online Thesaurus; Agricultural Thesaurus; e Geoethno Thesaurus) e ontologias. A pesquisa documental nos forneceu o retrospecto dos censos agropecuários e suas variáveis, tendo como ponto de partida o primeiro censo agropecuário mundial patrocinado pelo Instituto Internacional de Agricultura (IIA). Os Censos Agropecuários de 1970 e 2006 e o *Atlas* são as fontes documentais que deram início ao desenvolvimento dessa investigação em sua dimensão empírica.

Do censo agropecuário de 1970, selecionamos as seguintes características que se associam como variáveis relacionadas à hipótese de pesquisa: estabelecimento e utilização das terras. No censo agropecuário de 2006, nossas opções recaíram sobre: estabelecimento agropecuário, práticas agrícolas, recursos hídricos, terras irrigadas e método de irrigação, e utilização das terras.

A sistematização do levantamento bibliográfico no âmbito da CI contemplou a área de representação da informação e organização do conhecimento, destacando-se os seguintes assuntos: análise de domínio, terminologia, organização do conhecimento, representação da informação, sistemas de organização do conhecimento, teoria da linguagem, teoria do conceito, análise de gênero, comunidade discursiva, as garantias literária e semântica e os conceitos de informação e de documento, partindo da década de 1970 até dezembro de 2020. Esse levantamento teve como

suporte teórico os dados obtidos nas bases de dados Scientific Electronic Library Online (SciELO), Library and Information Science Abstracts (LISA), no *website* da International Society for Knowledge Organization (ISKO), no Repositório Institucional da UFF e Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (Brapci) e pesquisa livre na *Web*.

Em seguida, pesquisamos na literatura da área de Geografia os seguintes assuntos: paisagem, paisagem rural, pensamento geográfico, teoria geográfica, espaço, território, região, configuração espacial, categorias geográficas e ontologia, combinando com o termo paisagem. Incluímos, também, neste rol os seguintes termos de busca: modernidade, pós-modernidade e cultura.

Dessa forma, optamos por um *corpus* que fosse constituído por fontes primárias, composto por artigos de periódicos e teses e fontes secundárias, entre as quais livros, dicionários, tesouros, ontologias e taxonomias, nacionais e internacionais, publicados a partir da década de 1960. Os artigos de periódicos foram pesquisados na plataforma ScienceDirect; e nas bases de dados SciELO e LISA, acessadas por meio do Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Para as fontes secundárias, efetuamos nossa busca na Rede Sirius da Universidade Estadual do Rio de Janeiro (UERJ), no Repositório Institucional da UFF, na Rede de Bibliotecas do IBGE e pesquisa livre na *Web* contemplando termos relativos à Geografia estudados anteriormente em pesquisa de mestrado, a saber: paisagem, pensamento geográfico, teoria geográfica, espaço, território, região, configuração espacial, categorias geográficas, modernidade, pós-modernidade, cultura, taxonomia e ontologia. Os termos ontologia, taxonomia, cultura, modernidade e pós-modernidade foram combinados em pesquisa booleana com os termos relacionados ao domínio geográfico, pois era de nosso interesse conhecer como o tema paisagem se estrutura no campo da Geografia.

O estabelecimento da análise de domínio sobre a paisagem, mediante sua comunidade discursiva, foi assegurado com a realização das pesquisas bibliográfica e documental, evidenciado pelas garantias literária e semântica. Para isso, consideramos a garantia literária, a partir de Barité et al. (2010), como elemento aglutinador dos tópicos que estabelecem a literatura de uma área e proporciona as condições necessárias na representação de sua estrutura conceitual. Com relação à garantia semântica, seu conceito é aplicado aos sistemas de classificação bibliográfica, na visão de Beghtol (1986). Para a autora a garantia semântica não só governa as classes que devem compor um sistema de classificação, como também valida os relacionamentos entre os conceitos que são incorporados à estrutura classificatória. O mapeamento indicou autores que revelaram a garantia literária e forneceram elementos semânticos substanciais sobre a paisagem com vistas à organização de uma estrutura, que possibilitou a delimitação de um eixo comum e nuclear, que caracteriza o domínio, bem como sua comunidade discursiva.

3 Resultados

Os resultados alcançados nesse estudo levam em conta o objetivo de pesquisa em traçar um painel conceitual da paisagem para conhecer sua classificação, características e componentes à luz da análise de domínio da Geografia, no âmbito da CI e na área de organização e representação do conhecimento, no contexto dos censos agropecuários de 1970 e 2006.

O mapeamento da área de Geografia, por meio do *corpus* teórico selecionado, quando aplicada a análise de domínio, foi fundamental por ter exercido a função de ferramenta orientadora ao nos proporcionar um instantâneo e perspectiva integrada da área, levando-nos a identificar os conceitos-chave atribuídos à paisagem, tendo em vista as prerrogativas da garantia literária como justificação e validação da terminologia, que personaliza o padrão de comunicação da comunidade discursiva dos geógrafos.

O Painel conceitual da paisagem (Figura 1) revelou sua noção classificatória e categórica (Figura 2), tendo nos fornecido um conjunto de conceitos atualizados da área de Geografia e especificidade quando agregadas as informações extraídas dos censos agropecuários. Esse resultado reforça a importância da análise de domínio a fim de atender ao objetivo específico de se conhecer, identificar, descrever e organizar o conhecimento e estruturas de conhecimento resultantes de fontes informacionais para fins de prospecção de elementos sógnicos de um domínio, contribuindo, desse modo, com o estudo de uma área de conhecimento, que se consolida por informações especializadas produzidas pelos membros que compõem as comunidades de discurso e por meio das quais a disciplina e o ambiente são as unidades que constituem e alicerçam essas comunidades.

Figura 1 – Painel conceitual da paisagem

Termo	Conceito
Paisagem	Conjunto de elementos visíveis de uma região formado por objetos naturais e antrópicos apreendidos de símbolos sociais e naturais que refletem o espaço vivido pela sociedade e seus modos de produção
Paisagem Natural	Conjunto de formas compreendido pela cobertura da terra, águas (rios, pântanos, cursos de água, alagados), vegetação (estruturas campestres e florestais), e áreas de mineração
Paisagem cultural	Conjunto de elementos que resultam de atividades econômicas e sociais exercidas por uma sociedade que utiliza técnicas agrícolas ou industriais, redes de comunicação e outras técnicas que englobam tanto o campo como a cidade
Áreas urbanizadas	Parte da paisagem cultural, são áreas correspondentes às cidades (sede municipais), vilas (sedes distritais) e complexos industriais
Áreas agrícolas	Parte da paisagem cultural, são áreas utilizadas para a produção de alimentos
Estabelecimento agropecuário	Parte das áreas agrícolas, é a unidade dedicada às atividades agropecuárias: fazendas e hortos
Práticas agrícolas	Dão formas à paisagem cultural e resultam de técnicas quanto ao plantio, limpeza de terrenos ou proteção de encostas: plantio em nível; pouso ou descanso de solos; proteção e/ou conservação de encostas; queimada; rotação de culturas; uso de lavouras para recuperação de pastagens; e terraços

Recursos hídricos	Parte das áreas agrícolas, pode estar ou não em estabelecimento agropecuário: poços comuns, poços artesianos e cisternas
Terras irrigadas	Parte das áreas agrícolas, resultam da prática e método de aplicação de água, que não da chuva, diretamente à superfície do solo cultivado
Utilização das terras	Parte das áreas agrícolas, compreendem áreas plantadas ou em preparo para o plantio: lavoura permanente (cultura de longa duração, tais como: café, laranja, cacau, banana, uva, entre outros, que após a colheita não necessitam de novo plantio); lavoura temporária (inclusive horticultura e culturas de curta duração e só produzem uma vez); forrageiras para corte (uso na alimentação de animais); matas e/ou florestas naturais; florestas plantadas com essências florestais (nativas e exóticas); pastagem natural (campos naturais); pastagem plantada; construções, benfeitorias e caminhos; terras degradadas (erodidas, desertificadas, salinizadas); e terras inaproveitáveis (áreas inadequadas para plantio ou pastos: encostas íngremes, pedreiras, pântanos)

Fonte: Dados da pesquisa.

Figura 2 – Estrutura classificatória da paisagem

Classe	Subclasse	Subsubclasse	Subconjuntos
Paisagem natural	Cobertura da terra	Águas	Alagados; Cursos de água; Pântanos; e rios
		Áreas de mineração	
		Vegetação	Campos; matas; e florestas
Paisagem cultural	Áreas urbanizadas	Cidades, vilas e complexos industriais	
		Áreas agrícolas	Estabelecimento agropecuário
		Práticas agrícolas	Plantio em nível; pousio ou descanso de solos; proteção e/ou conservação de encostas; queimada; rotação de culturas; lavouras para reforma e/ou renovação e/ou recuperação de pastagens; e terraços
		Recursos hídricos	Poços artesianos; poços comuns; e cisternas
		Terras irrigadas	Gotejamento; Inundação; Pivô central; Sulcos
		Utilização das terras	Construções; benfeitorias e caminhos; florestas plantadas com essências florestais; forrageiras para corte; lavoura permanente; lavoura temporária; matas e florestas naturais; pastagem natural; pastagem plantada; terras degradadas; e terras inaproveitáveis

Fonte: Dados da pesquisa.

Essa rede discursiva proveniente tanto dos censos agropecuários como da comunidade dos geógrafos, por meio da literatura pesquisada, se traduziu em conceitos, categorias e comprova a importância que a terminologia padronizada assume na sistematização da informação e do conhecimento produzido pelas comunidades discursivas, que por sua vez são precedidas pelas garantias literária e semântica, preconizadas, entre outros, por Barité et al. (2010) e Beghtol (1986). Como dispositivos de justificação, essas garantias se confirmam como princípios de qualificação e aplicabilidade para a representação da paisagem, tanto no que diz respeito ao *corpus* terminológico que a define, quanto em sua estrutura classificatória, identificados no desenvolvimento de nossa pesquisa.

Considerando-se o contexto dinâmico da informação em ambientes como a *web*, a representação da paisagem por meio de sua conceitualização (Figura 1), ou seja, sua representação formal e uniforme, promove a qualificação tanto dos termos, quanto de sua estrutura classificatória (Figura 2), em vista de sua utilidade em responder perguntas formuladas por usuários sobre essa temática e, também, de poder promover a troca de informações e seu reuso por outras comunidades, como profissionais da área de CI e interessados em estudos voltados à organização e representação dessa área de assunto, tendo em vista que os esquemas de classificação, tesouros e ontologias disponíveis mostraram-se limitados na representação da paisagem na área de especialidade da Geografia.

4 Conclusão

Nossos objetivos, tanto geral quanto específicos, foram alcançados em decorrência da metodologia adotada. Foi muito significativo, em termos de método de pesquisa, o levantamento bibliográfico e documental.

A pesquisa bibliográfica seletiva apontou uma numerosa bibliografia sobre paisagem abrangida por outras áreas do conhecimento, fato que reforçou o valor das bases de dados consultadas para refinamento dos resultados pretendidos.

A análise da área de Geografia, por meio da bibliografia selecionada, nos lançou ao desafio de detectar sua rede de termos chancelada pela garantia literária. A terminologia identificada e conceituada nos forneceu subsídios que respalda a estrutura classificatória da paisagem e se consolida na ontologia da paisagem no âmbito dos SOCs. No entanto, a construção de SOCs, dos mais simples ao mais elaborados, deve seguir critérios teóricos e metodológicos, pautando-se na definição e terminologia padronizadas úteis com o objetivo de recuperação de informações dessa natureza. Em vista do exposto, partindo dessa ferramenta de representação da paisagem esperamos poder contribuir com outros profissionais da área de CI no compartilhamento dessas informações e seu reuso, tendo em vista que ainda são poucas as iniciativas nessa esfera voltadas à paisagem.

O uso do arsenal teórico-metodológico dos saberes próprios da organização do conhecimento se comprova por meio do mapeamento das construções discursivas da comunidade dos geógrafos ao demonstrar sua importância na especificação das unidades de conhecimento do universo terminológico que define seu padrão de organização, no que diz respeito à conceitualização da paisagem e, por conseguinte, em sua classificação.

Partindo da avaliação dos censos agropecuários, observamos que estávamos diante de outra comunidade discursiva composta por geógrafos, estatísticos e economistas, partilhando conhecimento e questionamentos de forma consensual, sinalizando uma identidade em torno de um objetivo específico, com base num conjunto de elementos linguísticos, semânticos e pragmáticos verificáveis no elenco das variáveis investigadas pelos censos, que proporciona um caráter de totalidade comunicativa a essa outra comunidade.

Determinado o discurso de especialidade da área de Geografia, o exame da representação do termo paisagem nas classificações bibliográficas, tesouros e ontologias apontou que a paisagem está relacionada, em sua maioria, na área de informações agrícolas, conservação do meio ambiente, arquitetura da paisagem, engenharia ambiental, antropologia, e retratado muito timidamente na área de Geografia. Embora sejam instrumentos criados a partir de um conjunto de regras e princípios de contextualização, não refletem a paisagem especificamente no domínio da Geografia e, portanto, não atuam na sua representação efetiva como recurso informacional voltado à análise e síntese de registros informacionais com ênfase na paisagem cultural no contexto das áreas agrícolas.

Referências

- Atlas do espaço rural brasileiro. (2011). IBGE. <http://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=263372>
- Barité, M. et al. (2010, maio/ago.). Garantia literária: elementos para uma revisão crítica após um século. *Transinformação*, 22(2), 123-138.
- Bawden, D., & Robinson, L. (2013). *Introduction to information science*. Neal-Schuman.
- Beghtol, C. Semantic validity: concepts of warrant in bibliographic classification systems. (1986, April/June). *Library Resources & Technical Services*, 109-125.
- Bertin, Jacques. (1988). Ver ou ler: um novo olhar sobre a cartografia. Em Andrade, M. M. de, & Vasconcellos, R. *Cartografia temática*. AGB.
- Bertrand, G. (2004). "Paisagem e geografia física global: esboço metodológico". *Revista RêGa*, 8, 141-152.
- Bie, C. A. J. M., Leeuwen, J. A. van, & Zuidema, P. A. (1996). *The land use database: a knowledge-based software program for structured storage and retrieval of userdefined land use data sets: user's reference manual. Version 1.04 for MS-DOS*. International Institute for Geo-Information Science and Earth Observation, Food and Agriculture Organization, United Nations Environment Programme, & Wageningen University (Eds).
- Cavaco, C. (2005). "As paisagens rurais: do 'determinismo natural ao determinismo político'?" *Finisterra*, 40(79).
- Censo agropecuário. (1970). IBGE, 3. <http://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?acervo=periodico&campo=titulo&opeqry=&texto=censo%20agropecuario&digital=false&fraseexata=>
- Censo agropecuário: Brasil, grandes regiões e unidades da federação: segunda apuração. (2006). IBGE. <http://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?acervo=periodico&campo=titulo&opeqry=&texto=censo%20agropecuario&digital=false&fraseexata=>
- Gil, T. L.; Barleta, L. (2015). Formas alternativas de visualização de dados na área de História: algumas notas de pesquisa. *Revista de História*, 73, 427-455.
- Gnoli, C. (2013). Knowledge Organization Systems (KOSs). Em Bawden, D., & Robinson, L. *Introduction to information science*. Neal-Schuman.

- Guedes, R. de M. & Moura, M. A. (2016). *O princípio da garantia semântica e os estudos da linguagem*. Encontro Nacional de Pesquisa em Pós-Graduação em Ciência da Informação, Salvador.
- Guimarães, J. A. C. (2014, jan./abr.). Análise de domínio como perspectiva metodológica em organização da informação. *Ciência da Informação*, 41(1), 13-21.
- Hjörland, B. (2017). Domain analysis. *Knowledge Organization*, 44(6). <https://www.nomos-elibrary.de/10.5771/0943-7444-2017-6/ko-knowledge-organization-jahrgang-44-2017-heft-6>
- Hjörland, B. (2002). Domain analysis in information science: eleven approaches: traditional as well as innovative. *Journal of Documentation*, 58(4), 422-462.
- Hjörland, B., & Albrechtsen, H. (1995). Toward a new horizon in information science: domain analysis. *Journal of the American Society for Information Science*, 46(6), 400-425.
- Instituto Internacional de Agricultura. (1939, vol. 1). *The first world agricultural census (1930)*.
- Konkoly-Gyuró, E. (2018). “Conceptualisation and perception of the landscape and its changes in a transboundary area: a case study of the southern german-french borderland”. *Land Use Policy*, 79, 556-574.
- KozeL, S. (2013). Comunicando e representando: mapas como construções socioculturais. *Geograficidade*, 3(número especial), 58-70.
- Latour, B. Redes que a razão desconhece: laboratórios, bibliotecas, coleções. (2008). Em Baratin, M., & Jacob, C. (Org.). *O poder das bibliotecas: a memória dos livros no Ocidente* (3.ª ed). Ed. UFRJ.
- Lepczyk, C. A., Lortie, C. J., & Anderson, L. J. (2008). “An ontology for landscapes”. *Ecological Complexity*, 5, 272-279.
- López-Huertas, M. J. (2006). Análisis del dominio interdisciplinario para la representación y organización del conocimiento. Em González de Gómez, M. N., & Orrico, E. G. G. D. (Org.). *Políticas de memória e informação: reflexos na organização do conhecimento*. Ed. UFRN. 209-235.
- López-Huertas, M. J. (2015). Domain analysis of interdisciplinary knowledge domains. *Knowledge Organization*, 42, 570-580.
- MacEachren, A. (2004). *M. How maps work: representation, visualization, and design*. Guilford Press.
- Manual técnico de uso da terra. (2013). IBGE, Coordenação de Recursos Naturais e Estudos Ambientais (3.ª ed.). (Manuais técnico em geociências, 7).
- Nascimento, D. M., & Marteleto, R. M. (2008). Social field, domains of knowledge and informational practice. *Journal of Documentation*, 64(3), 397-412.
- Prieto-Díaz, R. (1990, April). Domain analysis: an introduction. *Software Engineering Notes*, 15(2), 47-54.
- Santos, Milton. (2014). *Metamorfoses do espaço habitado*. Edusp.
- Saquet, M. A. (2015). Território e paisagem: da construção material à representação? Em Saquet, M. A. *Abordagens e concepções de território* (4.ª ed). Outras Expressões.
- Sauer, Carl. O. (1998). A morfologia da paisagem. Em Roberto Lobato Corrêa, & Zeny Rosendahl (Org.). *Paisagem, tempo e cultura*. Ed. UERJ.
- Smiraglia, R. P. (2012). Epistemology of domain analysis. Em Lee, H. L., & Smiraglia, R. P. (Ed.). *Cultural frames of knowledge*. Ergon, 111-124.
- Swales, J. M. (1990). *Research genres: English in academic and researching settings*. Cambridge University Press.
- Tennis, J. T. (2012, Jan./Jun.). What does a domain analysis look like in form, function, and genre? *Brazilian Journal of Information Science*, 6(1), 3-14.

- Veronezzi, F., & Fajardo, S. (2015). “A paisagem na análise geográfica: considerações sobre uma paisagem rural em Guarapuava – PR”. *GEOgraphia*, 17(34).
- Zins, Chaim, *et al.* (2007). “Mapa do conhecimento da Ciência da Informação: implicações para o futuro da área”. *Brazilian Journal of Information Science*, 1(1). <http://www2.marilia.unesp.br/revistas/index.php/bjis/issue/view/11.doi.org/10.36311/1981-1640.2007.v1n1.02.p3>.

(Página deixada propositadamente em branco)

MAPA TEÓRICO CONCEITUAL PARA A CONSTRUÇÃO DE CATEGORIAS EM PROJETOS TRANSDISCIPLINARES¹

Andréa Doyle¹, Vera Dodebei²

¹Doutora pelo PPGCI IBICT/UFRJ, Email: andrea@hibrida.art.br,
ORCID iD 0000-0002-2387-5438

²Professora titular do PPGMS Uni-Rio, Email: dodebei@gmail.com,
ORCID iD 0000-0003-3401-6257

Resumo

Em 1997, Vera Dodebei desenvolve um método de organização teórico-conceitual para construir o conceito de documento para a Memória Social. Em 2021, Andréa Doyle recupera esta proposta para delinear a ideia de crítica feminista de sua tese. O presente trabalho tem o objetivo de apresentar o método original, sua adaptação e demonstrar a funcionalidade do Mapa Teórico-Conceitual (MTC). Foi verificada a consistência teórico-metodológica do MTC para construção de quaisquer conceitos em situações transdisciplinares, sendo a transdisciplinaridade entendida a partir de Guattari (1992). Conclui que o MTC é uma ferramenta útil para organizar referências heterogêneas como a construção de categorias de análise em trabalhos transdisciplinares.

Palavras-chave: Mapa Teórico-Conceitual. Transdisciplinaridade. Ciência da Informação. Memória Social. Teorias feministas.

Introdução

Duas teses de doutorado em Ciência da Informação, com 24 anos entre elas, dão origem a este artigo. Vera Dodebei (1997) desenvolveu um método de organização teórico-conceitual que ajudou a construir o conceito de documento para a área da Memória Social. Andréa Doyle (2021) recuperou este método para delinear a ideia de crítica feminista e construir as categorias com que analisou a literatura sobre práticas de ensino da informação.

As autoras (mãe e filha) desenvolvem a presente reflexão, cujo objetivo é apresentar os processos e resultados dessa reciclagem científica. Pretendemos demonstrar a funcionalidade do Mapa Teórico Conceitual para a organização do conhecimento e como recurso metodológico para a construção de categorias de análise em projetos transdisciplinares.

¹ O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) via Bolsa de Produtividade em Pesquisa 2.

Vale destacar que a transdisciplinaridade é entendida, aqui, a partir da visão de Félix Guattari (1992), como uma necessidade de cooperação para a solução de problemas complexos, não só entre as disciplinas científicas mas, principalmente, entre ciência, tecnologia e sociedade. Para o autor, a pesquisa social só faz sentido se conseguir envolver redes de agentes que atuam naquela realidade para, coletivamente, solucionar os problemas estudados. Guattari (1992) indica duas frentes de ampliação da ideia de transdisciplinaridade: a inclusão de visões consideradas periféricas como a de mulheres e países em desenvolvimento e o estabelecimento de uma ecologia de pesquisa que leve em consideração o planeta e as localidades.

A transdisciplinaridade, como movimento interno de transformação das ciências, aberta para o social, o estético e o ético, não nascerá espontaneamente. A vida científica internacional fica, frequentemente, presa a rituais formais, numa interdisciplinaridade de fachada. Seu aprofundamento implica uma permanente “pesquisa sobre a pesquisa”, uma experimentação de novas vias de constituição de agrupamentos coletivos de enunciação. (GUATTARI, 1992, p. 26)

É exatamente nesse movimento de experimentação de ‘pesquisa sobre a pesquisa’ que apresentamos o Mapa Teórico Conceitual (MTC) concebido por Dodebei (1997) e seu reuso feito por Doyle (2021) para definir e relacionar conceitos, contextos, categorias, discursos, teorias, pensadoras e pensadores. Por fim, refletimos sobre sua utilidade futura em campos transdisciplinares.

O mapa teórico conceitual e seu reuso

Dodebei (1997), em sua tese de doutorado, constrói um conceito de documento para a memória social, trazendo teorias e argumentos de diversos campos tangenciais à Ciência da Informação. Uma das principais originalidades da pesquisa, para além de uma nova definição de documento, foi o método de investigação desenvolvido para dar conta de seu projeto. A autora produziu um Mapa Teórico Conceitual em que conceitos foram extraídos de teorias (das mais variadas origens, incluindo, filosofia, linguística, sociologia e disciplinas da informação) e agrupados e analisados de acordo com o ciclo de vida da informação.

À época, o ciclo se apresentava em quatro fases: produção, seleção, organização e uso. Mais tarde, a autora aprimorou a compreensão do ciclo e, recentemente, produziu uma atualização dessa proposta, que inclui os objetos digitais, conforme explicitado na Figura 1, acima.

A motivação para a investigação da tese foram as trocas com colegas nos departamentos de Biblioteconomia, Arquivologia e Museologia e sua inquietação diante do fato que nenhum daqueles entendimentos de ‘documento’ era amplo o suficiente para contemplar o outro. Foi o questionamento do próprio campo que permitiu não só a contribuição para sua solidificação teórica, como abrisse a possibilidade de desenvolvimento de um outro campo de estudos, também interdisciplinar, e com interseções com a Ciência da Informação – a Memória Social.

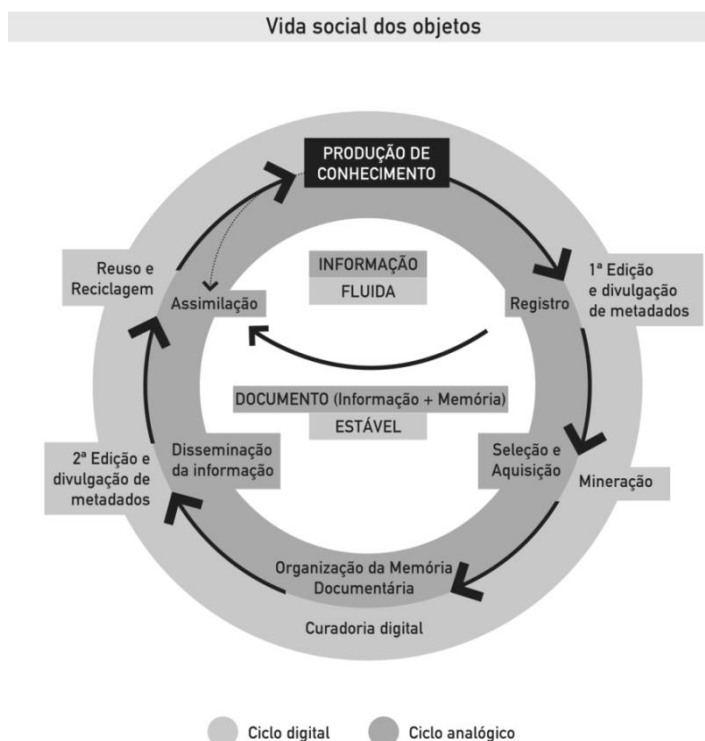


Figura 1 – A vida social dos objetos analógicos e digitais

Fonte: Dodebei (2021)

A classificação simbolizada no quadro a seguir não exclui as interfaces de pensamento entre os discursos, isolando-os em categorias estéreis e imutáveis. Pelo contrário, a correlação dos conceitos gerais com os instrumentos teóricos e seus discursos, mostrou-se como uma tentativa de uma ordenação cognitiva de construção de uma tela onde se entrecruzam os fios ideológicos em que as interseções podem ser pinçadas para a construção de novos conceitos. (DODEBEI, 1997, p. 28).

Quando descreve as características e defende sua proposta teórico-metodológica, a autora chama a atenção para algumas de suas qualidades. Trata-se de um esquema que, longe de ser uma classificação rígida ou excludente, se propõe a reunir e organizar ideias para utilizá-las na pesquisa, conforme Figura 2 abaixo.

MAPA TEÓRICO – CONCEITUAL			
Conceitos gerais		Instrumentos teórico/ metodológicos	Discurso teórico (rupturas e continuidades)
P R O D U Ç Ã O	CULTURA	Teorias da Cultura	Bosi, Maffesoli, Muniz Sodré, Geertz, Cohen
	MEMÓRIA	Teorias da Memória	Jeady, LeGoff, Namer, Nora
	OBJETOS	Teoria dos Objetos	Moles, Baudrillard, Boudon
S E L E Ç Ã O	REPRESENTA ÇÃO	Teorias Sociológicas	Becker, Gomez
		Teorias da Comunicação	Eco,
		Teoria da Linguagem	Spirkin, Gorsky, Pinto, Virilio
		Teoria Analítica dos Conceitos	Dahlberg
O R G A N I Z A Ç Ã O	DOCUMENTO	Teoria da Memória Do- cumentária	Otlet, La Fontaine, Malclès
U S O	INSTITUIÇÕES DE MEMÓRIA	Teorias Arquivísticas	Burke, Miller, Posner
		Teorias Museológicas	Chagas, Russio, Mensch
		Teorias Biblioteconômicas	Estivals, Foucault, Mostafa

Figura 2 – Mapa Teórico Conceitual da tese de Dodebei

Fonte: Dodebei (1997)

A proposta de Doyle (2021) para concluir o capítulo sobre o campo transdisciplinar dos Estudos de Gênero, foi executar um exercício similar. O exemplo de Dodebei (1997) foi utilizado para extrair noções e questionamentos das teorias e pedagogias feministas com vistas a construir um entendimento do que seria uma prática crítica a partir do olhar feminista. Esse delineamento do conceito de crítica guiou a seleção e análise de práticas de ensino-aprendizado de competências em informação, mídias e

tecnologias digitais, objetos de pesquisa da tese. A partir dele, a autora pode selecionar, dentre o conjunto de práticas críticas analisadas, aquelas que se propõem a lutar contra o sexismo, promovendo a desconstrução dos estereótipos de gênero.

Mas de onde surgiu a ideia desse reuso? Durante a escrita da tese, a autora acompanhou um curso intitulado *Pensadoras Negras Brasileiras* (Machado, 2020). Muitas das aulas ressaltavam a importância da valorização da ancestralidade no feminismo, para nosso próprio fortalecimento e desenvolvimento. Trata-se de nos reconhecermos nas lutas das mulheres que vieram antes de nós, identificar suas conquistas e limitações e assumir a responsabilidade de continuá-las.

Anielle Franco (2020), irmã e fundadora do instituto Marielle Franco, pediu ao público, no início de sua fala, para se perguntar e responder: De onde eu venho? Onde estou e o que estou fazendo? Pra onde quero ir? Essa aula representou o início de uma sequência de soluções que foram se encadeando. A autora decidiu ler a tese da mãe, sua ancestral física e acadêmica e considerou organizar suas leituras feministas a partir do Mapa Teórico Conceitual apresentado (Figura 2).

Quando Dodebei (1997), concebe essa sistematização de influências interdisciplinares para definir seu conceito de documento, ela apresenta uma matriz em que as seções horizontais são organizadas de acordo com o ciclo de vida dos objetos informacionais/memoriais e as colunas representam conceitos, teorias e intelectuais. A leitura de Doyle (2021) foi a de que a estrutura da matriz era possível para sua tese, mas o ciclo de vida da informação não tinha utilidade para a nova pesquisa. Lembrando as perguntas que iniciaram todo esse processo, a autora esboça uma organização dos elementos da ideia de crítica a partir de:

- a) Origem/Início: autocrítica, questionamento de si, de suas premissas, seus fazeres e seus objetivos. É o lugar de onde viemos, nosso ponto de partida, as ancestralidades, o lugar de fala. Inclui a corporeidade, a parcialidade, o saber localizado, a experiência, o elemento circunstancial e específico de cada grupo/situação. Inclui também: interesses, desejos, vieses e limitações: tudo o que motiva ou desmotiva o fazer.
- b) Processo/Meio: trata-se da consciência da construção coletiva do conhecimento. Inclui uma nova compreensão da ideia de autoridade, colaboração e cooperação. Se traduz em horizontalidade, na escuta, na inclusão pela valorização de cada pessoa com suas marcas sociais e identitárias diversas; na escolha de temas de trabalho pautados na realidade de vida de cada grupo social; no estudo e na compreensão de mecanismos da construção social da educação, da ciência, da tecnologia, da informação, da comunicação e também dos estereótipos e preconceitos.
- c) Objetivo/Fim: o propósito do processo educativo. Que cada participante se perceba sujeito da sua história, compreenda melhor sua sociedade e suas desigualdades, que encontre sua voz e seu lugar no mundo, e que trabalhe para a redução de injustiças em geral. Para o feminismo, isso representa o fim do patriarcado e o bem viver.

Tabela 1 – Mapa Teórico Conceitual da ‘crítica’ a partir do olhar feminista

Lugar /Momento	Conceitos gerais	Discursos teóricos
ORIGEM / INÍCIO	Autocrítica, reflexividade	Gayatri Spivak, bell hooks, Márcia Tiburi, Estelle Ferrarese
	Ancestralidade	Anielle Franco, Constança Duarte, Heloísa Buarque de Hollanda, Beatriz Nascimento, Lorena Cabnal
	Lugar de fala, contexto, corporeidade, sexualidade	Márcia Amaral, Djamilá Ribeiro, Guacira Louro, bell hooks
	Saber localizado, parcialidade, ponto de vista	Donna Haraway, Gayatri Spivak, Grada Kilomba, Simone de Beauvoir, Iris M. Young
PROCESSO / MEIO	Interseção de marcadores sociais, experiências, diversidade	Lélia Gonzales, bell hooks, Chimamanda Adichie
	Horizontalidade, escuta, troca, comunidade, coletividade	bell hooks, Grada Kilomba, Julieta Paredes, Lorena Cabnal
	Desconstrução de sistemas de opressão – patriarcado – (hierarquia, neutralidade, autoridade, violência)	Todas
OBJETIVO / FIM	Conscientização, busca por equidade	Todas
	Emancipação, autonomia, entusiasmo, empoderamento, auto-estima	Todas
	Bem viver, prazer, diversão, plenitude	Todas
	Resistência contínua	Todas

Fonte: Doyle (2021)

O quadro não inclui a coluna ‘Instrumentos teórico-metodológicos’, pois todos os conceitos usados vêm de teorias feministas. Existem diversas vertentes feministas, como o feminismo negro, decolonial, comunitário cujas teorias são desenvolvidas em diversos campos acadêmicos como Filosofia, Pedagogia, História ou Direito, mas nenhum desses recortes apresentou interesse para a pesquisa. Assim, optou-se por eliminar a coluna.

O MTC mostra as teóricas que informaram a compreensão das diversas dimensões do conceito de crítica feminista. Percebe-se que algumas aparecem em apenas um aspecto, como Márcia Amaral no ‘lugar de fala’, visto que a tese só dialogou com um trabalho específico da autora. Por outro lado, quando se trata do objetivo da

ação feminista, independentemente das origens ou dos métodos preconizados por essa ou aquela corrente feminista, todas são voltadas tanto para a desconstrução do patriarcado quanto para os objetivos de conscientização, de emancipação, de resistência contínua e de bem viver.

Esse quadro foi complementado com o capítulo teórico sobre competências em informação, mídias e tecnologias digitais (AMI) e deu origem a seis categorias de análise: Ponto de Vista e Autocrítica (Origem); Colaboração e Desconstrução (Processo); e Empoderamento e Bem Viver (Objetivo). Essas categorias foram fundamentais para a identificação e análise das práticas de ensino de AMI que contribuem para a desconstrução de estereótipos de gênero.

Conclusões

O experimento de Doyle (2021), a saber, retornar à proposta metodológica de Dodebei (1997) em outro contexto, tanto temporal quanto temático, foi bem sucedido. O método se mostrou adaptável e foi muito útil para a organização de conceitos do campo das teorias feministas. Assim, foi verificada a consistência teórico-metodológica do MTC para construção de quaisquer conceitos em situações transdisciplinares.

O processo todo foi também um exercício ecológico transdisciplinar visto que promoveu o diálogo entre a Memória Social, a Ciência da Informação e os Estudos de Gênero, todos campos interdisciplinares de conhecimento. Ademais, as teorias feministas incarnam bem a transdisciplinaridade guattariana, visto que nascem dos movimentos sociais, se fixam na luta política por direitos, estabelecem uma estética própria e objetivam soluções duráveis de equidade e bem viver para todas as pessoas.

Concluimos que o Mapa Teórico Conceitual é uma ferramenta útil para a construção de categorias de análise para reunir, sistematizar e interpretar ideias, conceitos e teorias heterogêneas em projetos transdisciplinares. O eixo horizontal, entendido como o elemento contextual de cada pesquisa, possibilita o uso dessa matriz em diversos campos e situações de pesquisa. Encarada como um convite à reflexão, a matriz “arruma a cabeça”, ou seja, organiza o conhecimento.

Referências

- Dodebei, V. (1997). *O sentido e o significado de documento para a Memória Social*. Tese (Doutorado) – Universidade Federal do Rio de Janeiro. Disponível em: <http://objdig.ufrj.br/30/teses/VeraLuciaDoyleDodebei.pdf> Acesso em: Fev. 2021.
- Dodebei, V. (2021). *A vida social dos objetos analógicos e digitais*. Atualização do Ciclo da Informação publicado em Dodebei, v. (2014). *Tesouro: linguagem de representação da memória documentária*. Intertexto/Interciência.
- Doyle, A. (2021). *Competências em informação, mídia e tecnologias digitais e a desconstrução de estereótipos de gênero: um mapeamento sistemático de literatura sobre de práticas de ensino críticas*. Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia.

- Franco, A. (2020) *Marielle Franco: pensadora, ativista e política negra 1979-2018*. (Curso Pensadoras Negras 1. ed.). Slides de Aula.
- Guattari, F. (1992). Fundamentos ético-políticos da interdisciplinaridade. *Tempo Brasileiro*, 108, 19-26.
- Machado, R. (2020). Curso Pensadoras Negras Brasileiras. *Rede Brasileira de Mulheres Filósofas*. Disponível em: <https://www.filosofas.org/post/pensadorasnegras> Acesso em: Ago. 2020.

EPISTEMOLOGÍA FEMINISTA Y ORGANIZACIÓN DEL CONOCIMIENTO EN EL CONTEXTO DE ISKO IBÉRICO

Carlos Cândido de Almeida¹, Rosa San Segundo Manuel²,
Daniel Martínez-Ávila³

¹Universidad Estatal Paulista – UNESP,
Instituto de Estudios de Género – Universidad Carlos III de Madrid,
carlos.c.almeida@unesp.br, ORCiD iD 0000-0002-8552-1029

²Universidad Complutense de Madrid, danima27@ucm.es,
ORCiD iD 0000-0002-1449-8175

³Universidad Carlos III de Madrid, dmartine@bib.uc3m.es,
ORCiD iD 0000-0003-2236-553X

Resumen

Pese a que los estudios críticos, cuyo interés es fomentar el debate epistemológico en el seno de la ciencia, se han convertido en un tema de interés general para la organización del conocimiento, la epistemología feminista no ha llamado la atención de los expertos de la misma manera. En este sentido, es importante analizar los avances del capítulo de ISKO España e Ibérico hacia una epistemología feminista de la organización del conocimiento y extraer inferencias útiles para la reflexión. En el presente trabajo, se realizó una revisión de todas las actas publicadas hasta la fecha de los congresos de los capítulos de ISKO España e Ibérico (14 ediciones, de 1993 a 2019). Para identificar aquellos estudios que tratan sobre temas feministas, se analizaron los títulos de las ponencias en busca de las palabras clave “feminismo”, “género” y “mujer” en castellano, inglés y portugués. Del total de 782 trabajos publicados solamente 7 fueron identificados según estos criterios, lo que representa un 0,89% de los trabajos. A partir del análisis se concluyen los siguientes puntos: los estudios no suelen utilizar la expresión epistemología feminista, aunque incluyen premisas y categorías como la de género y androcentrismo; el significado del concepto no es una manifestación genérica de la representación neutral de la realidad humana; los movimientos para la renovación del lenguaje y su uso no sexista deben recibir el apoyo de los expertos y profesionales de la organización del conocimiento; el movimiento feminista ha sido tratado en la bibliografía de la documentación de manera muy general, lo que puede impedir su aceptación como aportación teórica; el conocimiento adecuado sobre los grupos minoritarios en las sociedades complejas depende de una perspectiva de género, no como simple punto de vista o visión distinta de la estándar; la organización del conocimiento, bajo una epistemología feminista que cuestione la formación de la representación del conocimiento en los sistemas, debe asumir el compromiso de rever la construcción de la memoria a través del examen de los términos, clases y categorías.

Palabras clave: Epistemología feminista, Feminismo, Teoría de la Organización del Conocimiento, Estudios Críticos, ISKO Ibérico.

No es fácil tener conversaciones sobre género. Incomodan a la gente y a veces la irritan. Tanto hombres como mujeres se resisten a hablar de género o bien tienden a restar importancia rápidamente a los problemas de género. Porque siempre incomoda pensar en cambiar el estado de las cosas. (Chimamanda Ngozi Adiche, 2020, p. 47)

Introducción

El 25 de septiembre de 2015 se aprobó en la Asamblea General de las Naciones Unidas la Agenda 2030 para un desarrollo sostenible en la que se propone como uno de sus objetivos principales lograr la igualdad de género y empoderar mujeres y niñas. Este acontecimiento supone una conquista del movimiento feminista y un paso adelante hacia la comprensión específica de los problemas asociados a la discriminación de las mujeres en la sociedad. Por otra parte, en el ámbito científico, en las interpretaciones científicas de determinados fenómenos sociales, la lucha teórica de las mujeres continua viva y es objeto de varios campos del conocimiento. La biblioteconomía y documentación, la ciencia de la información y la organización del conocimiento tienen la responsabilidad también de contribuir con propuestas teóricas en ese sentido.

Las contribuciones epistemológicas en la organización del conocimiento y en el marco de ISKO España e Ibérico han sido muy fructíferas en las últimas décadas. Han sido publicados diversos trabajos de foco interdisciplinar incluyendo aspectos de la lingüística, semiótica, lógica, filosofía del lenguaje, teoría del concepto, sociología, estadística y computación. Existe también una corriente de estudios que han evaluado el desarrollo del pensamiento del capítulo ISKO (García Marco, 1996, García Marco 1997, Moneda Corrochano, López Huertas & Jiménez Contreras, 2012).

No obstante, aunque los estudios críticos, cuyo interés es fomentar el debate epistemológico en el seno de la ciencia, se han convertido en tema de interés general para la organización del conocimiento, la epistemología feminista no ha atraído el interés de los expertos de la misma manera. De hecho, se considera aquí que la propuesta feminista podría reunir, bajo el mismo paraguas, otras teorías usadas en el contexto de los estudios críticos que pueden analizar el ámbito cultural y temáticas específicas como la representación del conocimiento de grupos aislados del poder y marginalizados por la estructura dominante. Dicha teoría podría conjugar otras explicaciones de matiz crítico en la teoría de la organización del conocimiento y hacer incidencia en el convencimiento de que hay una línea coherente de estudios socioculturales en el marco de la organización del conocimiento.

La epistemología feminista, según Serret Bravo (2020, p. 85), señala “los fallos y las incongruencias de la epistemología tradicional”. En este sentido, se recurre a algunas ideas fundamentales para comprender la teoría feminista del conocimiento, a saber: a) la teoría tradicional no permite conocer la realidad humana ya que oculta la visión de la mitad de la población; b) el lenguaje, los conceptos y las perspectivas invisibilizan a la mayoría de los seres humanos, empezando por las mujeres; c) la idea de ser humano se centra en los varones, blancos y heterosexuales y, d) por

último, la visión feminista pretende acercarse de manera más correcta de la realidad humana (Serret Bravo, 2000, pp. 85-87). Históricamente las mujeres tampoco han tenido espacio en la academia ni han ocupado puestos privilegiados entre los científicos. El problema para ellas comienza ya desde que son niñas y se enfrentan a todo tipo de dificultades para recibir una educación de calidad y sin prejuicios.

Para reivindicar la centralidad del tema e impulsar la teoría de los comportamientos sociales, el movimiento feminista, ya sea académico o político, ha desarrollado conceptos claves para comprender la realidad social, tales como: patriarcado, género, androcentrismo y sexismo (Varela Menéndez, 2019, p. 175). Explicado de forma breve, la idea general es que estamos en un sistema y una forma de gobierno que subyuga las prácticas económicas, políticas, religiosas y culturales bajo la autoridad del varón. Aunque el patriarcado sea complejo y tenga unos niveles más centralizados o explícitos y otros más sutiles y livianos, el núcleo de la representación de la realidad está ubicada en el hombre. Él es el estándar del ser humano que construye las expresiones de la realidad cultural y conforma lo que se ha llamado androcentrismo. En este sentido, es importante para el feminismo que las ciencias sociales se emancipen del androcentrismo en el que se encuentran subyugadas. Por otra parte, el concepto de género es definido como “la construcción social y cultural de la feminidad y la masculinidad” (Rodríguez Magda, 2020, p. 119). Se trata de un hecho sociológico y no biológico. El patriarcado, para mantener el sistema de reproducción de los roles sociales y de las ventajas de la estructura de poder, recurre a medios, estrategias y métodos. Esta es la forma de poner en práctica la idea general de subordinación de las mujeres y así mantener la desigualdad. Para Varela Menéndez (2019, p. 180), “machismo es un piropo mientras que sexismo es la división de la educación por sexos”, es lo que establece los deberes del hombre y de las mujeres como forma de evitar la igualdad en las obligaciones.

La epistemología feminista puede ser comprendida en términos de niveles de crítica a la ciencia. Se puede decir que hay un nivel más bajo en el que se trata de cuestionar la poca presencia de mujeres científicas, mientras que en los niveles elevados se busca pensar en otras formas menos machistas y masculinizadas de hacer ciencia, respetando el por ejemplo el medioambiente, y donde la premisa de la dominación del otro no sea la más importante. Según Janack (2021), “El común a las epistemologías feministas es un énfasis sobre la importancia epistémica del género y el uso de él como una categoría analítica para las discusiones, críticas y reconstrucciones de prácticas epistémicas, normas e ideales.”

Para Harding (2016, p. 11) “La postura feminista radical sostiene que las epistemologías, metafísicas, éticas y políticas de las formas dominantes de la ciencia son androcéntricas”. Para las preguntas “¿cómo puede incrementar la objetividad de la investigación una indagación tan politizada?” y “¿sobre qué fundamento podrían justificarse tales afirmaciones feministas?”, Harding (2016, p. 23) sugiere tres soluciones epistemológicas: el *empirismo feminista*, el *punto de vista feminista* y *post-modernismo feminista*. Podría entenderse como puntos clave del debate epistemológico feminista.

La primera solución epistemológica sostiene que el sexismo y el androcentrismo son sesgos sociales corregibles utilizando las normas metodológicas de la propia ciencia. El hecho de tener más mujeres en la ciencia permitiría percibir mejor los

sesgos androcéntricos. Así pues, resultaría fácil resolver el problema, aunque las normas de la ciencia seguirían intactas. La idea es que el método científico, cuando aplicado por más mujeres, corregiría los sesgos. “Se presume que el método es capaz de eliminar los sesgos debidos al hecho de que los investigadores concretos sean blancos o negros, chinos o franceses, hombres o mujeres”(Harding, 2016, p. 24).

No obstante, para el empirismo feminista, las investigadoras pueden obtener resultados no sesgados más que los investigadores hombres, sean ellos feministas o no. El problema, sostiene Harding (2016), es que no es la ciencia la que se corrige a si misma al paso de los años con el uso correcto del método científico de forma que se garantiza la objetividad, sino que son los movimientos de liberación social los que han aumentado la objetividad científica. La idea del empirismo feminista, como solución a la objetividad, puede que sea muy ingenua ya que propone que el cambio depende de la estructura interna de la ciencia.

Por otra parte, el punto de vista feminista de la ciencia considera que el conocimiento del hombre es parcial y perverso en la medida que subyuga a las mujeres, así como ocurre en todos los aspectos de la vida social. Hay una vinculación estrecha de este punto de vista con el marxismo, pero añade que considerar el punto de vista de las mujeres abre posibilidades de conocimiento antes no posibles por la comprensión masculina que, además, según ella, explota a las mujeres.

“Aunque esta tentativa de solución de la paradoja epistemológica evita los problemas que plantea el empirismo feminista, también genera sus propias tensiones. En primer lugar, quienes están apegados al empirismo se mostrarán muy reacios a aceptar la idea de que la identidad social del observador puede ser una variable importante con respecto a la objetividad potencial de los resultados de la investigación. Desde un punto de vista estratégico, esta explicación de la mayor adecuación de las proposiciones feministas es menos convincente, salvo para quienes ya están convencidos; es muy improbable que los científicos naturales o los partidarios de la ciencia natural de ambos géneros acepten esta postura” (Harding, 2016, p. 25).

Harding plantea incluso aún más cuestiones sobre la teoría del punto de vista feminista, por ejemplo, cómo tener “un” solo punto de vista sin considerar las posiciones sociales construidas. Cómo sería posible tener un punto de vista sin considerar las diferentes visiones: feministas, no feministas, de clases sociales, de raza o de culturas distintas. En ese sentido, habría que respetar los varios puntos de vista feministas relacionados con la clase, raza, profesión y demás determinantes sociológicos y culturales.

Por su parte, la corriente epistemológica feminista postmoderna niega los presupuestos anteriores del empirismo y del punto de vista feminista. La corriente, según expone Harding (2016, p. 26), reúne a pensadores como Nietzsche, Derrida, Foucault, Lacan, Rorty, Cavell, Feyerabend, Gadamer, Wittgenstein o Unger. También incluye disciplinas y teorías como la semiótica, la deconstrucción, el psicoanálisis, el estructuralismo, la arqueología, la genealogía y el nihilismo. De acuerdo con Flax (citado en Harding, 2016, p. 26), las feministas “comparten un profundo escepticismo respecto a los enunciados universales (o universalizadores) sobre la existencia, la naturaleza y las fuerzas de la razón, el progreso, la ciencia, el lenguaje y el ‘sujeto/yo’”.

De hecho, el rechazo a todo lo establecido puede parecer más una estrategia política que una aportación de origen teórico, pues se plantean interpretaciones que esperan que sean universales o, por lo menos, se entienden como más explicativas o comprensivas que las anteriores. Harding (2016) afirma que estas son las tensiones y contradicciones de la propia epistemología del postmodernismo feminista, así como de otras posiciones epistemológicas feministas.

Sin embargo, hay algo fundamental en el debate epistemológico feminista que no es siempre conocido, y es que éste cuestiona el relativismo. Por más increíble o contradictorio que parezca, algunas líneas de la epistemología feminista no defienden el relativismo. Para Harding (2016, p. 25), no se sostienen la equivalencia entre anunciados sexistas y no sexistas, o si la condición de la mujer es biológica o social como propuestas igualmente aceptables. Los valores no son los mismos, la propuesta feminista no considera que su posición agnóstica en la aceptación del carácter hipotético de muchas afirmaciones científicas tenga que ver con un tipo de relativismo.

En concreto, el relativismo trae más problemas para la argumentación de una epistemología feminista que soluciones. Uno de los problemas relativistas es la visión de que mundo de las mujeres es una postura más que se dispone en la sociedad, así como también se tiene la de los varones. En el feminismo se defiende que hay que rechazar el relativismo pues este permite que se pueda hablar de todo y explicar con supuesto valor científico cualquier cosa, sin la debida necesidad de contrastar los hechos.

Cuando el discurso feminista protesta en contra de la violencia machista, lo hace con datos de la policía, encuestas, estudios, etc. y eso no tiene nada que ver con la defensa del relativismo. Las investigaciones sobre el tiempo que las mujeres dedican a los cuidados del hogar, a los niños y a los mayores, comparado al de los varones, no tiene nada de relativo, no es solamente un punto de vista. Si la idea es poner de manifiesto los problemas de la ciencia y las supuestas verdades emitidas por el discurso científico que se consideran machistas, lo importante es plantear nuevas hipótesis que corrijan estas consideraciones hacia una nueva representación adecuada de la realidad.

En ese sentido, aunque no lo parezca en el discurso postmoderno de la epistemología feminista, la intención es fundar una perspectiva basada en el realismo más que en el relativismo. Por lo tanto, comprender como se ha investigado el feminismo en cuanto respuesta epistemológica puede servir para cambiar las formas de pensar en los objetos de la organización del conocimiento. En ese sentido, se justifica la relevancia de analizar los avances del capítulo ISKO España e Ibérico hacia una epistemología feminista de la organización del conocimiento y sacar inferencias útiles para la reflexión.

Metodología

Se realizó una revisión de todas las actas de los congresos de los capítulos de ISKO España e Ibérico publicadas hasta la fecha (14 ediciones, de 1993 a 2019). Para identificar aquellos estudios que tratan sobre temas feministas, se analizaron los

títulos de las ponencias en busca de las palabras clave “feminismo”, “género” y “mujer” en castellano, inglés y portugués.

Del total de 782 trabajos publicados, únicamente 7 fueron identificados según los criterios seguidos, lo que representa un 0,89% de los trabajos. Por supuesto, existen otros trabajos presentados en las ediciones del Capítulo España e Ibérico que son de gran interés para la discusión de la epistemología feminista, sobre todo los que respectan a estudios críticos o de fundamentos de la organización del conocimiento, pero sin embargo estos no fueron incluidos en el estudio ya que en este primer momento lo que se buscó fue obtener conclusiones cualitativas de las ideas relacionadas con la teoría feminista.

Resultados

El análisis de las actas revela que existen numerosos trabajos que trataron sobre el uso de las tecnologías de la información y la inteligencia artificial para representar el conocimiento. En las democracias desarrolladas, es común que los asuntos que afectan a un progreso material de las personas sean vistos como de mayor interés para la ciencia, mientras que los debates sociológicos y críticos sean percibidos como resultados de cuestiones ideológicas y políticas.

En el caso del feminismo parece haber ocurrido algo semejante: existe una inhibición de los debates académicos en este sentido en el campo de la documentación por entenderse que son expresiones de movimientos de lucha por la justicia social, contra la discriminación, querellas políticas, etc., las cuales no siempre tienen cabida en el contexto científico y por lo que su potencial para la teoría de la organización del conocimiento ha sido subestimado. Entre las excepciones explícitas en el contexto de ISKO Ibérico se encuentran los trabajos señalados a continuación.

El trabajo de Martín-Santos, Modejar-Medina y Santos Sierra (1997) sobre la Clasificación Decimal Universal (CDU), la perspectiva de género y la discriminación de minorías, inauguró las preocupaciones con la temática. El trabajo analizó distintas discriminaciones en el sistema, de género, sexuales, sociales, de minorías e imperialistas, y señaló que en varias clases de la CDU se manifiesta un enfoque cristiano-católico que impide una clasificación objetiva. Los aspectos morales están presentes en muchos apartados de forma innecesaria, por lo que temas relacionados con las mujeres están muchas veces acompañados de un sesgo moralista y machista y, en concreto, poco igualitario. Pese a que en el estudio son comentados otros casos, lo importante para nuestro estudio es la proposición de una discusión sobre el rechazo a los esquemas universales de clasificación y el promover el debate sobre la importancia de la perspectiva de género en el área. Se entiende que la revisión de los sistemas es el primer paso para promover una representación más justa de la realidad.

En 2001, Rodríguez Bravo y Morán Suárez reflexionaron sobre la imagen de la mujer en la CDU. Se asumió la idea general de que las clasificaciones son creaciones sociales compuestas por conceptos con una intención de universalidad y neutralidad. Estos sistemas son elaborados según los discursos dominantes en la cultura.

“Estas nuevas corrientes historicistas-constructivistas chocan frontalmente con las perspectivas tradicionales positivistas y empíricas, bajo las cuales se han creado las principales clasificaciones, incluida la CDU (Clasificación Decimal Universal), objeto de nuestra investigación, fiel reflejo de la estructura del conocimiento existente a finales del siglo XIX y representación de las divisiones científicas, culturales y sociales de los distintos campos del conocimiento” (Rodríguez Bravo; Morán Suárez, 2001, p. 2).

En ese sentido, las autoras afirman que la CDU pone de manifiesto el androcen-trismo, pues el hombre es usado como ser genérico que representa a la humanidad, es el estándar. El objetivo fue estudiar los contextos en que aparecen las mujeres en los sistemas para dar a conocer las características y situaciones que denotan la imagen de la mujer. Para ello, las autoras compararon en las ediciones de 1995 y 2000 de la CDU las tablas principales donde aparece la mujer (aun sin mención expresa como en el caso del aborto) y la experiencia femenina en el contexto doméstico. También se buscaron las cuestiones de género en las tablas de auxiliares especiales comunes de personas y de forma y en el índice alfabético.

La intención del estudio fue entender los avances y las limitaciones aún presentes en el sistema. Las autoras identifican que, por primera vez, el feminismo es tratado entre los sistemas filosóficos o puntos de vista (141.72), pues representa una visión de la realidad. No obstante, la desaparición de los temas de la mujer como materia les pareció preocupante. Por ejemplo, no están presentes los temas emancipación, igualdad, mujer y derecho, educación de la mujer, trabajo de las mujeres, situación de la mujer en la familia, mujer y hogar o la mujer y la política.

Como resultado concluyen que sigue habiendo una dominación desde el punto de vista androcéntrico que se manifiesta en la invisibilidad de las mujeres con el uso del masculino genérico, aunque sea común justificarlo por la llamada tesis de la economía del lenguaje. “Percibimos que la presencia de lo femenino en esta clasificación es escasa, inapropiada y asimétrica con respecto a la masculina. Se nombra poco a las mujeres y, cuando se hace, se las asocia con escasas parcelas de la realidad” (Rodríguez Bravo; Morán Suárez, 2001, p. 3).

Las críticas de Rodríguez Bravo e Morán Suárez demostraron que los sistemas no estaban plenamente familiarizados con las demandas del feminismo y que es un logro muy reciente pasar a considera el feminismo como teoría sociológica y filosófica.

En la edición de 2003, Izquierdo Alonso propuso el análisis de género como metodología para la organización del conocimiento y Hope Olson discutió la des-construcción en la organización del conocimiento desde una perspectiva postcolo-nial.

El trabajo de Izquierdo Alonso (2003), aunque tenga indicado en el título la expresión análisis de género, puso énfasis en la vertiente pragmática en el estudio del contenido documental. La autora presentó el análisis de género de Swales como herramienta lingüística que relaciona la forma de contenido y la función textual. Tal como indica la autora: “El análisis de género se aplica al estudio de las regularidades lingüísticas y estructurales de los tipos textuales y al papel que juegan éstos dentro de una comunidad discursiva. Este análisis pone su énfasis en el objeto comunicati-

vo y en las convenciones que comparte la comunidad discursiva que lo utiliza respecto al contenido, la estructura y el estilo” (Izquierdo Alonso, 2003, p. 748).

Como se ha indicado, la idea de género social no es específicamente lo que se trata en la teoría de género, en la que se entiende género desde un punto de vista lingüístico. Por otra parte, lo que sí parece interesante es el uso de los estudios pragmáticos, enmarcados en el análisis del discurso. Estos estarían incluidos bajo la lingüística documental para estudiar las prácticas comunicativas de las comunidades discursivas.

Por su parte, Olson (2003) tuvo como objetivo proponer la deconstrucción de Derrida como método adaptado a dos líneas que considera transgresivas: la feminista y la postcolonial. Ambas líneas han señalado las limitaciones de la organización del conocimiento y deberían figurar entre las contribuciones del área. Según indica Olson (2003, p. 732), Derrida propone discutir el concepto de diferencia, es decir, el entendimiento de que los conceptos no son estables y las oposiciones conceptuales binarias presumen la dominación de un concepto sobre el otro (unidad-diversidad, identidad-diferencia, centro-periferia, macho-hembra). De esa forma el concepto dominante se define por no ser subordinado. El proceso de deconstrucción se centra en las oposiciones binarias y sostiene que el otro, el dominado, es un concepto construido, no innato o natural.

La deconstrucción es una vertiente postestructural que rechaza las verdades universales, “la deconstrucción revela que nuestras presunciones de principios universales son solo interpretaciones posibles más que fundamentos absolutos” (Olson, 2003, p. 732). Los datos utilizados como objeto de la deconstrucción en el estudio son textos, en este caso de Dewey, Ranganathan, tesauros, metadatos, etc. De acuerdo con Derrida (1988, p. 85 citado por Olson, 2003), la deconstrucción no es una teoría, filosofía, discurso, acto, práctica o método, sino un evento. La idea sería revelar lo que está oculto en las presunciones fundamentales. La deconstrucción no es destrucción como la suelen acusar sus críticos.

Por una parte, las feministas han utilizado Derrida para negar la esencia o la verdad sobre la identidad de la mujer, pues la deconstrucción feminista revela el orden patriarcal en las oposiciones binarias para reconocer a menudo las voces dominadas y deconstruir las ideas de sujeto activo y sujeto pasivo. Por otra parte, el postcolonialismo, como un tipo de deconstrucción transgresiva, presupone cuestionar las diferencias entre los conceptos binarios colonizador-colonizado, occidente-oriente, norte-sur, etc. Uno cuestiona la faceta sexual, mientras el otro la faceta cultural.

En ese sentido, ¿cuál sería el papel de la deconstrucción transgresiva en la investigación del campo de la organización del conocimiento? Según la autora, “la deconstrucción tiene el potencial para explorar nuestras ideas sobre temas de la organización del conocimiento como la garantía” (Olson, 2003, p. 739) y entender que nuestras estructuras no están inscritas en los astros sino que reflejan la tradición de Bacon o de actuales disciplinas académicas. Lo más importante es entender que para la deconstrucción, las categorías construidas en prácticas hegemónicas y las universalidades son fluidas.

Para concluir, Olson (2003, p. 739) afirma que las deconstrucciones postcoloniales ofrecen la oportunidad de revisar las oposiciones binarias de universalidad-diversidad local impuestas. Por lo tanto, deben analizarse los sistemas CDD,

OCLC, CDU, etc. y entender que “el más importante trabajo de deconstrucción en el campo de la organización de conocimiento es revelar la especificidad de nuestra ‘universalidad’”(Olson, 2003, p. 739). Tal como menciona la autora, “la deconstrucción transgresiva nos mantiene culturalmente honestos”. Hull y Broady-Preston, en 2005, discutieron las influencias de la clase social y el género en la propuesta de competencias en el marco de un programa de alfabetización informacional. Aunque no sea una aportación encuadrada totalmente en la organización del conocimiento, nos trae la cuestión de la durabilidad y profundidad de la enseñanza de habilidades informacionales a estudiantes de grado en dos universidades de Gran Bretaña, la Universidad de Teesside y la Universidad de Wales (Aberystwyth).

En las dos primeras fases de la investigación emplearon encuestas y en la tercera fase entrevistas semiestructuradas, entre seis y doce meses siguientes de la conclusión del entrenamiento. Las pruebas estaban basadas en ejercicios para ordenar signatura de un listado de materias. En los resultados no hubo diferencias sustanciales entre hombres (68%) y mujeres (65%) en términos de éxito en la ejecución de las prácticas. Sin embargo, en la muestra utilizada en Teesside los hombres tuvieron más éxito, en un 74% frente un 52% de mujeres, mientras que en Aberystwyth un 79% de mujeres tuvieron más aciertos frente a un 63% de los varones. Las autoras justifican el hecho en que Teesside es una zona de antiguas industrias químicas y de ingeniería pesada donde muchos trabajadores manuales utilizaban las matemáticas diariamente. En términos de uso de internet, el 100% de los estudiantes prioriza la web para cuestiones de búsqueda, mientras que los encuestados de las clases sociales más bajas ven internet como más una fuente interesante que útil.

Concluyen las autoras que no se puede hacer generalizaciones sobre los alumnos cuando llegan a la universidad ya que hay brechas de conocimiento incluso entre los futuros profesionales de documentación. Éstos tal vez piensan que el haber escogido la carrera específica de documentación ya les otorga automáticamente los prerrequisitos para actuar en el área de manera adecuada. Otra brecha también se refiere al nivel de inglés de los estudiantes, pues en la Universidad de Wales el inglés no es la primera lengua de muchos estudiantes, por lo que tienen dificultades para realizar tareas simples, aunque eso no interfiera en el estudio de la carrera.

El género volvió a ser el foco de una ponencia en la edición de 2011. El género neutro en la práctica de la organización del conocimiento fue estudiado por Alfaya Lamas, revisando extensamente la obra de Hope Olson. La autora no solamente hace alusión al problema del género, sino discute cuestiones más generales, como hasta qué punto los sistemas de organización del conocimiento manifiestan subjetividad y en ellos se reflejan las necesidades de unos grupos específicos de usuarios o de todos. En concreto, Alfaya Lamas (2011) defiende que los sistemas de organización del conocimiento no son neutrales. La idea sería aportar nuevas teorías y prácticas a la organización del conocimiento, pero no sin antes tener conciencia de la noción de género.

La organización del conocimiento, como señala la autora, está volcada en las actividades de descripción, indización y clasificación de documentos y da atención tanto a los procesos como a los sistemas. Para justificar esta idea sobre la organización de conocimiento, la autora recurre al referente Hjørland (2008) para sistematizar los enfoques teóricos: tradicional, analítico de facetas, recuperación de la información, orientado al usuario, bibliométrico y analítico de dominio.

No obstante, pone énfasis en el enfoque analítico de dominio ya que lo considera el más sociológico y epistemológico, además de examinar los problemas de la subjetividad de la organización del conocimiento, el lenguaje y los factores culturales. Discutir la subjetividad es poner de manifiesto los problemas que afectan a las materias/representaciones relacionadas con las mujeres en los sistemas, especialmente los sistemas tradicionales. Alfaya Lamas (2011, p. 256) combina estas teorías para demostrar que “los sistemas de organización del conocimiento siguen siendo subjetivos y asumen una visión androcéntrica del mundo”. Así, un sistema no será nunca una expresión transparente de la realidad del conocimiento (Beghtol, 2001 citado por Alfaya Lamas, 2011).

Las principales aportaciones que los estudios críticos de las prácticas de organización del conocimiento pueden ser así resumidas de acuerdo con el trabajo de Alfaya Lamas (2011):

1 – Los significados de las palabras están influidos por el discurso androcéntrico.

2 – El foco en el dominio rompe con la reproducción del poder en los sistemas de organización porque promueve una reflexión más general y proporciona cambios importantes a los sistemas de información bibliotecarios.

3 – Cada sistema de organización del conocimiento, tal como afirma Hjørland (2008), está sesgado según un punto de vista, cualquiera que sea, y por lo tanto no hay neutralidad. El enfoque analítico de dominio puede indicar los conocimientos son necesarios para los sistemas de información bibliotecarios.

4 – Como modelos a seguir, Alfaya Lamas menciona tres ejemplos. En primer lugar, está el estudio de Hope Olson para la elaboración de términos utilizando un vocabulario usado por feministas (A Women’s Thesaurus de 1987). Se parte de la base de que los sistemas universales no ejercen bien la tarea de representar esos términos porque no tienen en cuenta a las mujeres y sólo representan el discurso dominante masculino. Adicionalmente, no permiten incorporar las voces de grupos marginados, excluidos, colonizados y explotados. En segundo lugar, está el trabajo de la escritora y bibliotecaria Nynne Koch (Anna Lise Koch, Copenhague, 1915-2001) en la Biblioteca KVINFO, un referente danés. Esta autora desarrolló un sistema de clasificación para los “estudios de la mujer”, además de otras actividades como base de datos de mujeres danesas y expertas. Finalmente, la autora menciona los estudios de Hope Olson y Rose Schlegl sobre los encabezamientos de materias de la Library of Congress, pues crean barreras para el acceso y promueven la opresión produciendo omisiones y posiciones racistas, sexistas y xenófobas.

5 – Se debe buscar formas de no asumir la visión androcéntrica, ya que cuando se hace el sistema solo es válido para la mitad de la población.

Para concluir la comunicación, Alfaya Lamas (2011) lanza dos preguntas fundamentales: “¿hay subjetividad en los sistemas de organización del conocimiento?” y “¿se reflejan las necesidades de un determinado grupo de usuarios/as o de todos/as?”. La respuesta a la primera pregunta, por supuesto, es sí y a la segunda es un rotundo no, porque los intereses de todos no siempre se ven contemplados en los sistemas. “Los elementos que no se expresan en una clasificación como la relevancia del feminismo para los hombres, por ejemplo, los aspectos políticos del feminismo,

o que se trata de un movimiento filosófico, son formas de opresión, marginalización e invisibilización de la mujer” (Alfaya Lamas, 2011, p. 261). Finalmente recomienda a los profesionales reivindicar su responsabilidad frente al problema y colaborar para la crítica y producción de sistemas de organización del conocimiento más igualitarios (Alfaya Lamas, 2011, p. 261).

Sin lugar a duda, el trabajo de Alfaya Lamas ha aportado muchas cuestiones y ha reivindicado algunas actitudes de los profesionales. En cierta medida, hace eco a trabajos más prácticos que, desde otros ámbitos, han demostrado la importancia para la sociedad de esas acciones. El problema de la mala representación se vuelve más sensible cuando se nombran los crímenes contra las mujeres. En muchos casos, la mala representación convierte las agresiones en hechos invisibles.

Tras algunas ediciones sin trabajos recuperados siguiendo los criterios utilizados en la presente investigación, en 2019, Fioravanti, Nascimento y Sabbag presentaron los resultados de una investigación realizada en Brasil sobre el uso de términos de indización para representar los crímenes contra la mujer en el Estado de Sao Paulo. Estos crímenes al no haber recibido la atención merecida no fueron representados con la clase/materia feminicidio, cuya clasificación jurídica en Brasil es muy reciente. Como conclusión afirman que “la gran mayoría de los crímenes que resultaron en la muerte de la mujer han sido clasificados e indizados como homicidio cualificado y no como feminicidio, evitando así la visibilidad que el proceso debería tener, hecho que dispersa los datos y dificulta la recuperación de la información” (Fioravanti, Nascimento & Sabbag, 2020, pp. 490-491).

El estudio realza la importancia de revisar los términos y los sistemas de clasificación existentes, no solo de bibliotecas, para sensibilizar y visibilizar los crímenes contra las mujeres de manera general. Los temas relacionados con la prevención de todos los tipos de violencia contra las mujeres podrían formar parte de las preocupaciones de una epistemología feminista aplicada al contexto de la organización del conocimiento.

La mala representación produce consecuencias, desvía la mirada y desorienta el pensamiento ante la realidad social. Esa premisa se hace presente en varias aportaciones lanzadas a lo largo de los años en el marco de ISKO Ibérico.

Conclusiones

Con el análisis de este limitado conjunto de ponencias, se concluye que en el contexto de ISKO España e Ibérico no suele utilizar la expresión epistemología feminista en los estudios. No obstante, las premisas epistemológicas están presentes en los trabajos, por ejemplo, en el uso de las categorías género y androcentrismo. Dichas categorías forman parte de los fundamentos de las diversas posiciones de la epistemología feminista, desde la empirista hasta la postmoderna.

Sin embargo, se debe impulsar la unificación de las colocaciones publicadas para cohesionar los argumentos y aportaciones en consonancia con lo que defiende la agenda de la Naciones Unidas. Se hacen necesarias más investigaciones sobre la temática desde un punto de vista más integrado que no solamente estudie empíricamente los esquemas, las tablas de clasificación y el uso de términos de indización, sino que se reconozca como una discusión más amplia.

Una de las principales contribuciones de la teoría feminista es la epistemología del conocimiento, la cual cuestiona los fundamentos de muchos de los problemas actuales de la sociedad. Una aportación teórica de la organización del conocimiento que responda a los retos de la Agenda 2030 y de las necesidades de cambio social podría fundamentarse sobre todo en la epistemología feminista. A continuación se presentan algunas ideas que se pueden extraer de los estudios mencionados y de la epistemología feminista como aportación a la organización del conocimiento:

1 – Los estudios exponen los principales problemas de la excesiva visión androcéntrica de los sistemas de organización del conocimiento, las cual propicia sesgos no solamente contra las mujeres, sino que también afecta a las minorías políticas. Las contribuciones indican que esto ocurre con varios sistemas y en muchos procesos de representación (clasificación, indización, análisis documental y catalogación). La línea epistemológica empirista del feminismo, como se ha visto, sugiere que más mujeres y representantes de las minorías deben formar parte de las comisiones que deciden sobre la elaboración de los sistemas y consorcios bibliotecarios. El hecho de incluir también otras minorías respondería a limitaciones de la epistemología del punto de vista feminista, garantizando la participación de muchas voces. Además, teniendo en cuenta la propuesta epistemológica postmoderna, se podría utilizar la deconstrucción como método para comenzar el trabajo de análisis de un sistema determinado.

2 – La epistemología feminista puede ayudarnos a cuestionar el significado de un concepto como una manifestación genérica y neutral de la representación de la realidad humana. Los conceptos, aunque fundamentales para ordenar los sistemas de organización del conocimiento, no pueden entenderse de manera neutra, especialmente en los temas sociales y en las humanidades. La formación y la ordenación de conceptos en la organización de la sociedad, fundamentales para construir sistemas de organización del conocimiento, siguen todavía una perspectiva patriarcal. Se entiende que la deconstrucción epistemológica postmoderna sirve solamente para comenzar el trabajo, el punto de vista feminista propone tener más voces que participan el proceso, y el empirismo epistemológico feminista apoyaría la construcción y la defensa de mejores métodos como formas de corregir los sesgos sin caer en el rechazo de todo que sea método. Un entendimiento en esta línea podría mejorar la comprensión del concepto y su importancia para los sistemas.

3 – Los movimientos para la renovación del lenguaje y su uso no sexista deben recibir el apoyo de los expertos y profesionales de la organización del conocimiento. Alfaya Lamas (2011) ha reivindicado la responsabilidad de los profesionales. Por supuesto, un cambio total no será posible sin el compromiso y colaboración internacional entre universidades y organizaciones para propagar la necesidad de actualización de los términos. Por tanto, la teoría feminista, que no tiene nada de restrictiva y no afecta únicamente a los intereses de las mujeres, puede ofrecer una gran contribución para el área.

4 – Una hipótesis que hay que considerar para profundizar en el diálogo con la epistemología feminista es que el movimiento feminista ha sido tratado en la bibliografía del área de la documentación de manera muy general, como simple movimiento político o solamente como una minoría más. Incluso los expertos en el área

reproducen la idea de aislamiento o infravaloración del feminismo bajo el rótulo de un simple movimiento reivindicativo. Por supuesto, las reivindicaciones son fundamentales para la rama política del movimiento, pero no representan la totalidad de las aportaciones. Lo mismo no ocurre con corrientes marxistas, que suelen ser caracterizadas de críticas y de expresión general de la reflexión de vanguardia. En general, las reivindicaciones políticas suelen no aludir a la teoría epistemológica feminista como propuesta de cambio de la teoría del conocimiento.

5 – El conocimiento adecuado sobre los grupos en las sociedades complejas depende de una perspectiva de género, no como simple punto de vista o mirada distinta de la estándar, sino como la principal forma de abordar las causas de los problemas sociales. Estas causas afectan la producción del conocimiento, especialmente en lo que respecta a los saberes no científicos (los prejuicios de género, clase, raza, de etnia, de origen, etc.). Un pequeño ejemplo son los prejuicios de clase y de origen que impiden que se utilicen conocimientos científicos producidos en países pobres y las pocas citas que autores ajenos al sistema académico europeo o norteamericano reciben en las publicaciones. La variable interseccionalidad muy discutida en la teoría feminista, incluso en la ciencia, podría mejorar nuestra visión de la producción y acceso al conocimiento.

6 – La organización del conocimiento, bajo una epistemología feminista que cuestione la formación de la representación del conocimiento en el lenguaje y principalmente en las ciencias sociales y otros sistemas, deberá tener el compromiso de rever la construcción de la memoria a través del examen de los términos, clases y categorías. Un ejemplo de este punto sería la citada clasificación de los crímenes. El tema y clasificación de un documento no es solo una cuestión que afecta a la recuperación del contenido, sino que también representa un reconocimiento de la historia, se trata de una acción de justicia social.

Así, la teoría sugiere que hay responsabilidades sociales y que la organización del conocimiento, como otras actividades que afectan el ámbito público, tiene que rendir cuentas. La representación justa de los conocimientos es sin duda uno de los retos que afectará al campo de la información de cara a la Agenda 2030.

Referencias

- Adiche, Chimamanda Ngozi (2020). *Todos deberíamos ser feministas*. Barcelona: Penguin Random House Grupo Editorial.
- Alfaya Lamas, E. (2011). La asunción del género neutro en la teoría y práctica de la organización del conocimiento. In M. C. Pérez Pais & M. A. González Bonome (2012). *20 años del Capítulo Español de ISKO: actas del X Congreso ISKO-España* (pp. 249-263). Ferrol, 30 de junio. La Coruña: Universidade da Coruña, Servicio de Publicaciones.
- Bouveresse, Jacques (2020). *Nietzsche contra Foucault: sobre la verdad, el conocimiento y el poder*. Barcelona: Ediciones del Subsuelo.
- Cobo Bedia, Rosa & Ranea Triviño, Beatriz (2020). *Breve diccionario de feminismo*. Madrid: Editorial Catarata.

- Fioravanti, D. C. B.; Nascimento, F. A. & Sabbag, D. M. A (2020). A contribuição da organização de conhecimento no procedimento de classificação e indexação e nos processos crime com perspectiva de genero: interpelações acerca dos feminicídios no Estado de São Paulo – Brasil. In J. Tramullas, P. Garrido-Picazo & G. Marco-Cuenca (2020). *Actas del IV Congreso ISKO España-Portugal 2019, XIV Congreso ISKO España* (pp. 485-492). Zaragoza: Capítulo Ibérico.
- García Marco, Francisco Javier (1996). Avances en Organización del Conocimiento en España: los II encuentros sobre organización del conocimiento en los sistemas de información y documentación. In F. J. García Marco (1996). *Organización del conocimiento en sistemas de información y documentación: actas del II encuentro de ISKO-España* (pp. 7-12). Getafe (Madrid), 16 y 17 de noviembre de 1996. Zaragoza: Universidad de Zaragoza.
- García Marco, Francisco Javier (1997). Casi una década de la Sociedad Internacional para la Organización del Conocimiento en España. In F. J. García Marco (1997). *Organización del conocimiento en sistemas de información y documentación: actas del III encuentro de ISKO-España* (pp. 7-21). Getafe (Madrid), 19 a 21 de noviembre de 1997. Zaragoza: Universidad de Zaragoza.
- Harding, S. (2016). *Ciencia y feminismo*. San Sebastián de los Reyes: Ediciones Morata.
- Hull, B. & Broady-Preston, J. (2005). Social class and gender as predictors of information literacy skills: report of research in progress. In J. Gascón; F. Burguillos Martínez & A. Pons i Serra (2005). *VII Congreso del Capítulo Español de ISKO. La dimensió humana de l'organització del coneixement* (pp. 309-325). Barcelona: Universitat de Barcelona, Facultat de Biblioteconomia i Documentació.
- Izquierdo Alonso, M. (2003). El análisis de género como metodología para la organización y representación del conocimiento. In C. Travieso Rodríguez & J. A. Frías Montoya (2003). *Tendencias de investigación en organización del conocimiento* (pp. 747-754). Actas del VI Congreso del Capítulo Español de ISKO, Salamanca 2003. Salamanca: Ediciones Universidad de Salamanca.
- Janack, Marianne (2021). Feminist Epistemology. In: Internet Encyclopedia of Philosophy. ISSN 2161-0002. Disponible en: <https://iep.utm.edu/fem-epis/>. Acesado en:16/02/2021.
- Madruga Bajo, Marta (2020). *Feminismo e Ilustración*. Madrid: Ediciones Cátedra.
- Mondéjar Madina, A.; Marín García, T.& Santos Serra, R.(1997). CDU: perspectivas de género y discriminación de minorías. In F. J. García Marco (1997). *Organización del conocimiento en sistemas de información y documentación: actas del III encuentro de ISKO-España* (pp. 103-118). Getafe (Madrid), 19 a 21 de noviembre de 1997. Zaragoza: Universidad de Zaragoza.
- Moneda Corrochano, M.; López Huertas Pérez, M. J. & Jiménez Contreras, E. (2012). La investigación sobre organización del conocimiento en España (2002-2010). In M. C. Pérez Pais & M. A. González Bonome (2012). *20 años del Capítulo Español de ISKO: actas del X Congreso ISKO-España* (pp. 91-106). Ferrol, 30 de junio. La Coruña: Universidade da Coruña, Servicio de Publicaciones.
- Naciones Unidas (2015). *Transformar nuestro mundo: la Agenda 2030 para el Desarrollo Sostenible*. (Resolución aprobada por la Asamblea General el 25 de septiembre de 2015).https://www.agenda2030.gob.es/recursos/docs/APROBACION_AGENDA_2030.pdf.
- Olson, H. (2003). Transgressive deconstructions: feminist/postcolonial methodology for research in knowledge organization. In C. Travieso Rodríguez & J. A. Frías Montoya (2003). *Tendencias de investigación en organización del conocimiento* (pp. 731-740). Actas del VI Congreso del Capítulo Español de ISKO, Salamanca 2003. Universidad de Salamanca, Ediciones Universidad de Salamanca.

- Rodríguez Bravo, B. & Morán Suárez, M. A. (2001). La imagen de la mujer en la Clasificación Decimal Universal (CDU). In A. I. Extremeño Placer (2001). *La representación y organización del conocimiento: metodologías, modelos y aplicaciones: actas del V Congreso ISKO-España*, 25-27 de abril de 2001 (pp. 1-12), Alcalá de Henares, Madrid: Universidad de Alcalá.
- Rodríguez Magda, Rosa María (2020). Género. In R. Cobo Bedia & B. Ranea Triviño (Ed.). *Breve diccionario de feminismo* (pp. 119-121). Madrid: Editorial Catarata.
- Serret Bravo, Estela (2020). Epistemología feminista. In R. Cobo Bedia & B. Ranea Triviño (Ed.). *Breve diccionario de feminismo* (pp. 85-87). Madrid: Editorial Catarata.
- Varela Menéndez, Nuria (2019). *Feminismo para principiantes*. Barcelona: Penguin Random House Grupo Editorial.

Agradecimiento

El presente trabajo se ha realizado con el apoyo de la Coordinación de la Formación de Personal de Nivel Superior – Brasil (CAPES) – Código de la Inversión 001 (Proceso: 88887.571329/2020-00).

(Página deixada propositadamente em branco)

SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO E O CONCEITO DE REGIME DE INFORMAÇÃO: RELAÇÕES POSSÍVEIS

Ana Flávia Dias Zammataro¹, Ana Cristina de Albuquerque²

¹Universidade Estadual de Londrina (UEL), afzammataro@gmail.com,
ORCID iD <https://orcid.org/0000-0001-6520-3229>

²Universidade Estadual de Londrina (UEL), albuanti@uel.br,
ORCID <https://orcid.org/0000-0003-3506-0479>

Resumo

Os Sistemas de Organização do Conhecimento são definidos como instrumentos que servem para sistematizar e demonstrar as relações dos conceitos que representam um domínio, permitindo assim que informações sejam recuperadas. Considerando como Sistemas de Organização do Conhecimento os sistemas de classificação, as taxonomias, os tesouros, as ontologias, entre outros, esse trabalho tem como objetivo problematizar relações possíveis entre os Sistemas de Organização do Conhecimento e o conceito de regime de informação. Assim, além de se trazer para discussão a literatura que versa sobre os Sistemas de Organização do Conhecimento no contexto da Ciência da Informação, discorre-se sobre o conceito de regime de informação, compreendendo sua capacidade de exercer influências sobre o meio no qual opera. Essas discussões respaldam a resposta ao seguinte problema de pesquisa: é possível estabelecer relações entre os Sistemas de Organização do Conhecimento e o conceito de regime de informação? Os conceitos norteadores do trabalho, bem como seu objetivo, problematização e resultados, são abordados por meio da metodologia de revisão bibliográfica, com base em análise qualitativa dos trabalhos referenciados, que discutem os elementos de elaboração dos Sistemas de Organização do Conhecimento, assim como as definições de regime de informação. O levantamento da literatura foi feito em bases de dados e em repositórios de dissertações e teses da área de Ciência da Informação, além de trabalhos publicados no âmbito da ISKO. Posteriormente, seguiu-se nas análises desses trabalhos, buscando identificar como compreendem os Sistemas de Organização do Conhecimento e o conceito de regime de informação, a fim de se estabelecer relações entre os conceitos. Os resultados permitem estabelecer relações entre os Sistemas de Organização do Conhecimento enquanto “modelos de realidade” e os regimes de informação, posto que se tratam de sistemas elencados e selecionados com base em escolhas (e, portanto, exclusões), relações de poder e necessidades dos agentes envolvidos. O cenário colocado pelos resultados torna possível corroborar a relação existente entre os Sistemas de Organização do Conhecimento na perspectiva da área de Organização do Conhecimento e o conceito de regime de informação na Ciência da Informação, além de permitir refutar

argumentos que defendem os Sistemas de Organização do Conhecimento como sistemas ou dispositivos imparciais ou neutros, contribuindo para pensar os processos e práticas da Organização do Conhecimento em uma perspectiva mais pragmática.

Palavras-chave: Organização do Conhecimento; Sistemas de Organização do Conhecimento; Regime de informação.

Introdução

Os Sistemas de Organização do Conhecimento são definidos como instrumentos que, a partir dos diversos níveis em que são estruturados, servem para sistematizar e demonstrar as relações dos conceitos que representam um domínio e permitem que informações sejam recuperadas, levando em consideração um “modelo da realidade”. Os SOC podem se caracterizar como sistemas de classificação, taxonomias, tesouros, ontologias, entre outros estudados na literatura da área. São estruturas essenciais no contexto da Ciência da Informação, assim como elementares para a organização de domínios do conhecimento e para a representação e recuperação da informação registrada, sendo que, suas funções serão determinadas pelo grau de complexidade e também pelo objetivo a que foi criado e será utilizado.

De acordo com Moreira (2018, p. 107), terminologicamente os SOC possuem o compromisso da garantia literária, ou seja: “[...] um SOC deve refletir analítica e criticamente o conjunto de conceitos e termos usualmente expressos no interior de um determinado domínio ou discurso.” O autor complementa que usualmente os SOC são produtos institucionalizados, onde termos e conceitos são adequados a certa comunidade e instituição e refletem as relações conceituais determinadas (Moreira, 2018).

Pode-se inferir, a partir dos argumentos do autor, que os SOC são desenvolvidos por meio de uma ação, no caso, institucional, que servirá de fundamento relacional de conceitos que representarão uma comunidade específica, assim se configurando em sua elaboração, o envolvimento a um regime de informação.

A partir do exposto, questiona-se: é possível estabelecer relações entre os Sistemas de Organização do Conhecimento (SOC) e o conceito de regime de informação? Atentando que o regime de informação se caracteriza como uma: “[...] formação social conjunta de elementos em rede – como atores sociais (sujeitos, dispositivos e tecnologias), regras de poder, a organização e a gestão política da informação que se operacionalizam em práticas sociais com produtos e serviços.” (Bezerra et al, 2016, p. 61), pretende-se problematizar nesse artigo essas possíveis relações, de forma teórica e conceitual, a fim de visualizar diferentes compreensões quanto a elaboração, definição, características e usabilidade dos SOC.

Para compor a discussão, utilizou-se a pesquisa bibliográfica com abordagem qualitativa, por meio de leituras que se dedicam a discutir os Sistemas de Organização do Conhecimento, assim como autores que tecem considerações sobre o conceito de regime de informação, como Bernd Frohmann (1995), Sandra Braman (2004) e Maria Nélide González de Gómez (2012a e 2012b). Os levantamentos foram realizados em bases de dados e em repositórios de dissertações e teses da área de Ciência da Informação, além de trabalhos publicados no âmbito da ISKO.

Os resultados demonstram que há relações entre os SOC e o conceito de regime de informação, posto que, o estabelecimento e definição dos sistemas estão imbricados de relações de poder, escolhas e silenciamentos, atendendo assim à demandas institucionais, bem como às necessidades dos usuários desses sistemas.

Metodologia

Essa pesquisa identifica-se como bibliográfica, de caráter qualitativo, com base em referencial teórico que aborda a temática trabalhada. Conforme escreve Almeida Júnior (2007, p. 100): “a pesquisa bibliográfica é a atividade de localização e consulta a fontes diversas de informação escrita, para coletar dados gerais ou específicos a respeito de determinado tema.”. O autor esclarece ainda que esse tipo de pesquisa se mostra, assim, como um eficiente método de estudo na obtenção de informações.

A pesquisa bibliográfica ganhou corpo por meio das seguintes etapas:

1. Primeiramente, definiu-se o tema da pesquisa, considerando suas relações com o eixo 4, da ISKO, ou seja: Teoria e Epistemologia da Organização do Conhecimento. Assim, decidiu-se trabalhar com os Sistemas de Organização do Conhecimento na perspectiva da área de OC pensados sob o conceito de regime de informação.

2. A partir de então, iniciou-se o levantamento bibliográfico de trabalhos que desenvolveram suas discussões a partir da temática do Knowledge Organization Systems (Sistemas de Organização do Conhecimento) na área de Organização do Conhecimento e também a partir do conceito de regime de informação. Essas pesquisas foram feitas em bases de dados como a Brapci e em repositórios de dissertações e teses da área de Ciência da Informação, além de trabalhos publicados nos periódicos da ISKO. Para realizar essa pesquisa, foram feitas buscas de palavras-chave como: organização do conhecimento, sistema de organização do conhecimento, regime de informação, ciência da informação.

3. A partir da identificação dos trabalhos que abordam os conceitos desenvolvidos neste artigo, foram feitas leituras daqueles que se dedicam a estudar os SOC na OC e na CI como um todo, passando por trabalhos que fazem uma revisão bibliográfica para identificar como os autores do campo da CI entendem os SOC até aqueles que se debruçam sobre princípios de elaboração de taxonomias, tesouros, sistemas de classificação, entre outros. Procedeu-se da mesma maneira com os trabalhos que discutem o conceito de regime de informação na perspectiva da Ciência da Informação.

4. Posteriormente à leitura da bibliografia identificada, foi realizado um cruzamento de semelhanças e diferenças entre as abordagens das referidas temáticas (conceito de SOC, princípios para elaboração de SOC, regime de informação na CI), possibilitando fundamentar, então, a discussão proposta nesse trabalho. A partir disso, foi possível estabelecer relações entre os SOC e o conceito de regime de informação.

As análises de natureza qualitativa voltaram-se para o estudo desses artigos, buscando compreender de que maneira abordaram os SOC e o conceito de regime de informação. Nesse aspecto, é relevante destacar o caráter subjetivo da pesquisa ao se contextualizar essas abordagens a fim de responder ao problema colocado. Conforme Richardson (1999): “[...] Os estudos que empregam uma metodologia qualitativa podem descrever a complexidade de determinado problema, analisar a interação de certas variáveis, compreender e classificar processos dinâmicos vividos por grupos sociais, contribuir no processo de mudança de determinado grupo e possibilitar, em maior nível de profundidade, o entendimento das particularidades do comportamento dos indivíduos.” (Richardson, 1999, p. 80).

Por meio da abordagem qualitativa foi possível traçar algumas relações entre os SOC e o conceito de regime de informação, pela possibilidade interpretativa e subjetiva do método, sempre considerando o rigor científico ao se trabalhar com os aportes teóricos escolhidos para desenvolver essa pesquisa.

Os Sistemas de Organização do Conhecimento

Os Sistemas de Organização do Conhecimento se caracterizam como tipos de instrumentos, que traduzem conteúdos de documentos originais, estruturando-se de maneira sistemática e têm a finalidade de organizar a informação e o conhecimento, facilitando assim a recuperação informacional (Carlan & Medeiros, 2011).

Entre os SOC existentes podemos citar os tesouros, as taxonomias, as classificações, as ontologias, cujos conceitos apresentam-se semanticamente estruturados, facilitando assim a indexação e também orientando os usuários desses respectivos sistemas.

Na perspectiva de Lara (2015), além de facilitar as atividades de armazenamento e recuperação da informação, os Sistemas de Organização do Conhecimento promovem também a gestão da informação e do conhecimento: “[Os Sistemas de Organização do Conhecimento] compreendem uma série de instrumentos que tratam do ‘conteúdo’ dos documentos para fins de armazenamento e recuperação ou para promover sua gestão.” (Lara, 2015. p. 92).

Lara (2015), esclarece ainda que o termo Knowledge Organization Systems (KOS) ou Sistemas de Organização do Conhecimento (SOC) foi utilizado pela primeira vez por Hodge (2000), para se referir a esquemas que serviam para representar, organizar e gerir a informação e o conhecimento, envolvendo assim, desde sistemas de classificação até dicionários, tesouros e ontologias.

Pensados a partir das perspectivas da área de Organização do Conhecimento, os SOC são estudados quanto à sua natureza e qualidade, desempenhando a função de ajudar em questões de padronização terminológica, a fim de organizar e recuperar documentos (Silva, 2017).

Os Sistemas de Organização do Conhecimento podem ser entendidos, ainda, como métodos de Organização do Conhecimento, na perspectiva de Birger Hjørland (2003). Para o autor, tanto os métodos de construção de Sistemas de Organização do Conhecimento (como os sistemas de classificação e os tesouros, por exemplo), quanto os processos de indexação e classificação, podem ser considerados métodos de OC (Hjørland, 2003).

Nesse sentido, é possível afirmar que para que um método de construção de um SOC seja estruturado, é importante pensar diversos pontos norteadores, como as necessidades dos usuários desses sistemas, refletindo assim sobre o tipo de sistema mais adequado a determinada comunidade, além do tipo de *hardware* ou *software* que comportaria esse sistema (Carlan & Medeiros, 2011). Assim, concorda-se com as autoras quando afirmam: “Um SOC pode ser significativo e vantajoso para uma cultura, uma coleção ou um domínio e para outros pode não ser.” (Carlan & Medeiros, 2011, p. 56).

Ao pensar-se nos objetivos e usabilidade dos Sistemas de Organização do Conhecimento, concorda-se com Moreira (2018), quando afirma que os SOC visam atender às necessidades de uma organização específica. Nesse sentido, cabe aos elaboradores desses sistemas compreender as demandas dos cenários, seus usuários, as características da organização, ou seja, considerar questões que influenciam diretamente no modelo de SOC a ser usado, determinando assim, seus atributos e propriedades. Moreira (2018, p. 107-108) esclarece: “Nesses casos, é preciso também identificar e respeitar as relações de significação entre termos e conceitos bem como as relações conceituais praticadas por uma determinada organização e por uma determinada comunidade usuária.”

Afirma-se, a partir do exposto, que os contextos influenciam diretamente nas características dos Sistemas de Organização do Conhecimento, determinando o tipo de sistema a ser usado, sua estrutura, seu modo de funcionamento, a relação entre os termos e conceitos desses sistemas, bem como as necessidades de seus respectivos usuários, o que leva a inferir sobre possíveis relações com o conceito de regime de informação.

Discutindo o conceito de regime de informação

Inaugurando estudos sobre o referido conceito no campo da Ciência da Informação, Bernd Frohman (1995) expõe que descrever um regime de informação significa fazer um mapeamento dos interesses, discursos, artefatos científicos e tecnológicos, que podem resultar em conflitos ou estabilizações dentro de um determinado grupo social. Frohman (1995), nesse sentido, está preocupado com o regime de informação que fundamenta os processos pelos quais os artefatos estão construídos (Bezerra, 2017).

Pensado então como um modo informacional em uma determinada instituição social, o regime de informação aproxima-se da perspectiva estabelecida por Maria Néida González de Gómez (2012a), que explica: “[...] um regime de informação seria o modo informacional dominante em uma formação social, o qual define quem são os sujeitos, as organizações, as regras e as autoridades informacionais e quais os meios e os recursos preferenciais de informação, [...] enquanto vigentes em certo tempo, lugar e circunstância.” (Gómez, 2012a, p. 43).

Como um modo informacional dominante, um regime de informação relaciona-se à distribuição de poder entre diversos sujeitos dentro de uma comunidade específica, seu modo de funcionamento e suas regras, que irão refletir diretamente na formação dos SOC para atender às necessidades de informação. Na formação desses dispositi-

vos ou sistemas, existem então “recursos de visibilidade informacional” (Gómez, 2012b, p. 28), e ao mesmo tempo o silenciamento de outras informações, mobilizando assim determinados valores de informação disponíveis nesses SOC (Gómez, 2012b) em detrimento de outros valores, que nesse processo, são apagados e silenciados.

Pensando ainda a partir da perspectiva bramaniana de que um regime de informação se edifica de maneira semelhante ao que ocorre com o processo de codificação do conhecimento, têm-se esses regimes informacionais permeados nas práticas sociais, tornando-os aceitos por meio de um consenso socialmente estabelecido dentro das instituições e comunidades, e podendo ser expressos, por exemplo, por meio das tecnologias (Braman, 2004).

Nas discussões propostas por Sandra Braman (2004) acerca do conceito, a autora argumenta que existiria um único regime de informação na atualidade, que além de global, seria também emergente, envolvendo o poder do Estado e os atores não-estatais. Nessa perspectiva, a informação é entendida como insumo de poder, sendo capaz de controlar e manipular regras, valores e comportamentos por meio das bases informacionais dos contextos estatais e não-estatais.

Acerca do exposto sobre o conceito de regime de informação, é possível inferir sobre sua capacidade de exercer influências sobre o meio no qual opera, e por extensão, sobre a própria identidade de uma comunidade ou da sociedade. Nesse sentido, essas influências tornam-se materiais pelas demandas impostas pelo regime de informação e isso pode acontecer, por exemplo, por meio dos Sistemas de Organização do Conhecimento.

Problematizando a relação entre os sistemas de organização do conhecimento e os regimes de informação

Discutindo as relações de poder envolvidas nos sistemas de organização e representação da informação e do conhecimento, Moura (2018) compreende tais elementos como dispositivos, por meio dos quais: “estruturam-se sistemas conceituais que articulam os termos, as definições, os relacionamentos e as propriedades dos conceitos em um domínio específico.” (Moura, 2018, p. 12).

Denominados como dispositivos ou na acepção mais pura da área de Organização do Conhecimento como SOC, a autora destaca os critérios de escolha para que este ou aquele termo represente determinado conteúdo, colocando em evidência a não-neutralidade desses dispositivos ou sistemas.

Reconhecendo os SOC como dispositivos ou sistemas não-neutros, é possível relacioná-los ao conceito de regime de informação. Quando Frohmann (1995) esclarece que os regimes de informação, quando mapeados, identificam interesses, discursos e artefatos científicos e tecnológicos, ele está também preocupado com o regime de informação, que fundamenta os processos pelos quais esses artefatos estão construídos (Bezerra, 2017). Assim, os processos que determinam a elaboração dos Sistemas de Organização do Conhecimento podem ser relacionados ao viés institucional aos quais esses sistemas atendem, consolidando práticas institucionais, científicas e sociais, ou seja, os regimes de informação o que também permite relacionar

os processos que determinam a elaboração dos dispositivos ou Sistemas de Organização do Conhecimento ao viés institucional aos quais esses sistemas atendem, consolidando práticas institucionais, científicas e sociais.

Relacionando o conceito de regime de informação de Braman (2004) aos SOC, torna-se possível perceber esses instrumentos como carregados de valores, regras e práticas de sujeitos diversos – sejam eles ligados ou não ao Estado – ao determinar quais são os termos mais adequados para representar os conceitos, ao se ater a determinados discursos em detrimento de outros, e ao fazer escolhas sobre os tipos de sistemas e a maneira como aplicá-los em seus respectivos contextos. Com as explicações da referida autora, é possível inferir que as tecnologias – compreendidas aqui como dispositivos, instrumentos, sistemas – se destacam também como os Sistemas de Organização do Conhecimento e os discursos a eles relacionados, além das normas, padrões e valores sociais, ou seja, os próprios regimes de informação.

O conceito de regime de informação está relacionado, portanto, à estruturação de um SOC, posto que, constituídos por “termos, conceitos e relações semânticas” (Carlan & Medeiros, 2011, p. 64), que se referem ao próprio funcionamento do sistema e suas características. Tanto esses elementos, quanto as características dos contextos nos quais esses SOC se inserem são constituintes do conceito de regime de informação, pois dizem respeito a um “modo informacional dominante”, na perspectiva de Gómez (2012a, p. 43).

O esquema a seguir representa a relação entre os SOC e o conceito de regime de informação.

Figura 1. Regime de Informação e Sistemas de Organização do Conhecimento: relações possíveis



Referência: Elaborado pelas autoras

O esquema é representativo dos resultados identificados no trabalho, evidenciando a relação entre os SOC e o conceito de regime de informação, a partir da percepção de que, a elaboração de dispositivos e de SOC, é determinada pela distribuição de poder entre atores e organizações, pelas necessidades dos usuários, pelos termos e conceitos selecionados, assim como, pela escolha dos *softwares* ou *hardwares*. Tais escolhas operacionalizam dinâmicas sociais e culturais em regulamentos, que são então estabelecidos pelos respectivos regimes de informação.

Um SOC é desenvolvido com o objetivo de representar um determinado domínio do conhecimento e suas relações semânticas, funcionando como instrumentos fun-

damentais de interpretação, de forma organizada, das estruturas do conhecimento. Assim, relacionados aos regimes de informação vigentes no ato de sua construção ou envolvidos na função que terão, permitem também direcionamentos quanto aos caminhos do controle terminológico, norteamento sobre uso ou não de determinados conceitos, entendimento do domínio a partir de interlocuções da comunidade discursiva e uma comunicação ou delimitação dos conceitos que conseqüentemente implicarão em menores ambigüidades, assim como farão um “retrato” do domínio representado.

O cenário desses diferentes contextos, ou seja, as próprias organizações e suas necessidades, seus usuários, etc., compõem demandas e relações de poder que podem ser identificadas como regimes de informação, estabelecendo-se assim uma possível relação entre os SOC e o conceito de regime de informação.

Conclusões

Com base nos resultados identificados neste artigo, torna-se possível corroborar a relação existente entre os Sistemas de Organização do Conhecimento na perspectiva da área de OC e o conceito de regime de informação na Ciência da Informação.

As pesquisas científicas envolvendo os SOC e os regimes de informação são importantes para refutar argumentos que defendem os SOC como sistemas ou dispositivos imparciais ou neutros, contribuindo para pensar os processos e práticas da Organização do Conhecimento em uma perspectiva mais pragmática, considerando as influências sociais nos processos e sistemas de OC e “a intermediação cultural na relação das pessoas com a informação” (Hjørland, 2003, p. 99).

Nessas pesquisas, é também importante que se ressalte o alcance dos SOC para não viabilizar a circulação de informações falsas e/ou enviesadas, evitando assim, a difusão de discursos autoritários e também as assimetrias de informação (Moura, 2018).

Referências

- Almeida Júnior, J. B. de (2007). O estudo como forma de pesquisa. In: Carvalho, M. C. M. de (Org.). *Construindo o saber – Metodologia científica: fundamentos e técnicas*. 18a ed. Campinas: Papirus.
- Bezerra, E. P., Guimarães, I. J. B, Silva, Z. C. G. da, Souza, E. D. de. (2016). Regime de informação: abordagens conceituais e aplicações práticas. *Em Questão*, Porto Alegre, 22(2), 60-86. Recuperado de <http://dx.doi.org/10.19132/1808-5245222.60-86>
- Bezerra, A. C. (2017). Vigilância e cultura algorítmica no novo regime global de mediação da informação. *Perspectivas em Ciência da Informação*, Belo Horizonte, 22(4), 68-81. Recuperado de <http://dx.doi.org/10.1590/1981-5344/2936>
- Braman, S. (2004). The emergent global information policy regime. In: Braman, S. (Ed.). *The emergent global information policy regime*. Houndsmills, UK: Palgrave Macmillan.
- Carlan, E., Medeiros, M. B. N. (2011). Sistemas de Organização do Conhecimento na visão da Ciência da Informação. *Revista Ibero-Americana De Ciência Da Informação*, 4(2), 53-73. Recuperado de <https://periodicos.unb.br/index.php/RICI/article/view/1675>

- Frohmann, B. (1995). Taking information policy beyond information science: applying the actor network theory. Annual Conference Canadian Association For Information Science. Anais do Annual Conference Canadian Association For Information Science, Edmonton, Alberta, 23. Recuperado de <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=5E0EF0CB76AEA2AD7B0D9FDE0EDEFCA6?doi=10.1.1.521.6657&rep=rep1&type=pdf>
- González de Gómez, M. N. (2012a). Regime de informação: construção de um conceito. *Informação & Sociedade*, 22(3), 43-60. Recuperado de <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/14376/8576>
- González de Gómez, M. N. (2012b). As Ciências Sociais e as questões da informação. *Morpheus – Revista Eletrônica em Ciências Humanas*, 9(14), 18-37. Recuperado de <http://www.seer.unirio.br/index.php/morpheus/article/view/4832/4322>
- Hjørland, B. (2003). Fundamentals of Knowledge Organization. *Knowledge Organization. Knowl. Org.*, 30(2), 87-111. Recuperado de <http://ppggoc.eci.ufmg.br/downloads/bibliografia/Hjorland2003.pdf>
- Lara, M. L. G. (2015). Propostas de tipologias de KOS: uma análise das referências de formas dominantes na organização do conhecimento. *Encontros Bibli*, 20(1), 89-107. Recuperado de <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2015v20nesp1p89>
- Moreira, W. (2018). *Sistemas de organização do conhecimento: aspectos teóricos, conceituais e metodológicos*. (Tese de Doutorado). Universidade Estadual Paulista – Unesp – Campus de Marília, SP, Brasil. Recuperado de <http://hdl.handle.net/11449/190878>
- Moura, M. A. (2018). Organização social do conhecimento e performatividade de gênero: dispositivos, regimes de saber e relações de poder. *Liinc em Revista*, Rio de Janeiro, 14(2), 118-135. Recuperado de <http://dx.doi.org/10.18617/liinc.v14i2.4472>
- Richardson, R. J. (1999). *Pesquisa social: métodos e técnicas*. São Paulo: Atlas.

(Página deixada propositadamente em branco)

SEMIÓTICA DOCUMENTAL: RETOMANDO UM DIÁLOGO

Alexandre Robson Martines¹, Carlos Cândido de Almeida²

¹Unesp (FFC) Marília-SP, alexandre.martines@unesp.br,
ORCID iD 0000-0002-4524-0978

²Unesp (FFC) Marília-SP, carlos.c.almeida@unesp.br,
ORCID iD 0000-0002-8552-1029

Resumo

A Semiótica Documental é uma disciplina preocupada com o tratamento documental e com a semiose documental. Assim, o objetivo desta pesquisa foi revisitar as propostas teóricas apresentadas por Izquierdo Arroyo explorando a fundamentação teórica de C. S. Peirce e C. W. Morris, além de ampliar a discussão sobre o tratamento documental e a semiose documental, inserindo contribuições da semiótica francesa, sob o viés teórico de A. J. Greimas e J. Fontanille. Aplicou-se metodologia de natureza qualitativa, de caráter exploratório, método Análise de Conteúdo através das etapas: a) pré-análise; b) exploração do material; c) tratamento dos resultados, a inferência e a interpretação. A Semiótica Documental baseia-se na teoria do pragmatismo para efetuar o tratamento documental, em busca do melhor estágio do signo para a representação, além de debater sobre o processo cognitivo, como o hábito mental para efetuar a representação. Além disso, as práticas semióticas e a análise da práxis enunciativa podem ampliar a atuação da Semiótica Documental. Frente a isso, espera-se ampliar as discussões das teorias semióticas já aplicadas à disciplina, assim como debater acerca das contribuições da semiótica francesa.

Palavras-chave: Semiótica Documental, Documentação, Izquierdo Arroyo, Tratamento Documental, Organização do conhecimento.

Introdução

O documento é o principal objeto de estudo de áreas como a Arquivística, a Diplomática, Ciência da Documentação e, em destaque nesta pesquisa, a Organização do Conhecimento, justamente por ser o suporte e veículo da informação. O tratamento documental, conjunto de procedimentos e operações técnicas que incluem a análise, avaliação, extração, classificação, descrição, representação e indexação de conceitos presentes em um documento, revolucionou a forma das sociedades sistematizarem e organizarem o conhecimento.

Elaborar e estabelecer a linguagem especializada, aplicada para efetuar a representação do conhecimento é normalizar um mecanismo de poder. Um poder que

legítima e institucionaliza os documentos e toda a produção de conhecimento, o desenvolvimento de técnicas e os fundamentos de organização como: construção, descrição, catalogação, ordenação, classificação, gestão, uso, evolução de sistemas de descrição e comunicação para que se possa testemunhar, conservar, transmitir o conhecimento, configurar os atos da humanidade.

Ademais, permite explorar a informatividade do documento tratando de seu conteúdo, inclusive possibilitando a mediação da informação produzida e de sua veiculação nos âmbitos sociais, por sistemas informacionais, que validam ou refutam, nos processos discursivos, as evidências dos aspectos semânticos de sua aceitação como modelização de conhecimento, por conseguinte de hábitos, de valores, de comportamentos e de condutas. Além de as práticas de tratamento documental analisar os conceitos apresentados para que sejam evidenciados através de linguagem especializada, é importante dialogar com a linguagem natural para garantir a atualização na representação, por conseguinte inovando e produzindo novos conhecimentos (García Gutiérrez, 2002).

Diante disso, é válido destacar a linguagem como elemento mediador da representação da realidade, da significação e da organização do conhecimento, assim seu estudo é marcado por teorias, práticas e estratégias bastante complexas. Nessa linha, entende-se que a Semiótica – ciência responsável pelos estudos dos signos, assim como pela elaboração de um sistema de significação, como ainda responsável por compreender os processos de semiose (Greimas & Courtés, 2016), também denominada de teoria geral dos signos (Peirce, 2017), é “um campo de estudo que tem por objetivo quaisquer tipos de signos” (Santaella, 2020).

Ainda sobre Semiótica, pode-se definir que “é uma ciência rigorosa, construtora de formas ideais, pelas quais por via dedutiva, e por tanto, a modo de necessidade, demonstrará suas conclusões” (Silveira, 2007) – é uma área que propõe não só estudos acerca da significação e aplicação do signo para estabelecer a representação, mas também é capaz de traçar procedimentos estratégicos e metodológicos para balizar as práticas, as estratégias, as valências que envolvem o tratamento documental a fim de evidenciar as marcas textuais e discursivas, que refletem a rede de conceitos, em dialogismo, na composição do documento.

Sendo assim, o objetivo desta pesquisa foi revisitar as propostas teóricas apresentadas por Izquierdo Arroyo acerca da Semiótica Documental – professor na Universidade de Múrcia, alinhado aos cursos de Biblioteconomia e Documentação. Responsável pelo marco acadêmico, em que inseriu na grade de graduação a disciplina Linguística Documental, da qual desenvolveu pesquisas para organizar os conceitos sobre a Semiótica Documental, em 1990 –, além de ampliar a discussão sobre o tratamento documental e a semiose documental e aproximar essa discussão das teorias do discurso alinhadas à Semiótica Francesa, na linha desenvolvida por J. Fontanille e A. J. Greimas, denominada de semiótica discursiva.

Frente a isso, espera-se contribuir para o tratamento e análise documental na identificação dos conceitos que evidenciam o conhecimento manifestado pelo documento, de modo a legitimar seus aspectos semânticos, os quais sejam propícios para representar o conhecimento, não apenas para organizá-lo em um determinado domínio, mas oferecendo a condição de veridicção e inquirição frente aos aspectos culturais, sociais, históricos, políticos, ideológicos, de representatividade e cientifi-

cos. Outrossim, projeta-se alcançar a inserção da Semiótica Francesa como teoria de análise e interpretação, bem como as especificações de texto e discurso aplicados aos procedimentos de tratamento documental, a fim de ampliar as discussões sobre o trabalho realizado por Izquierdo Arroyo.

Metodologia

A metodologia aplicada possuiu natureza qualitativa, já que os objetivos apresentam o caráter epistemológico, visto que visa à análise de teorias da semiótica (Greimas, 1975, 2014; Greimas & Fontanille, 1993; Fontanille, 2008, 2019; Silveira, 2007; Ibri 2015, 2020; Peirce, 2017; Santaella 2017, 2020), a fim de explicar suas contribuições para o tratamento documental, decorrentes dos estudos acerca da Linguística Documental (García Gutiérrez, 1984, 1989, 1990) e da Semiótica Documental (Izquierdo Arroyo, 1990, 1994, 1995).

Para organização dos dados, aplicou-se a Análise de Conteúdo “um conjunto de instrumentos metodológicos [...] sutis em constante aperfeiçoamento, que se aplicam a discursos (conteúdos e continentes) extremamente diversificados” (Bardin, 2016, p. 15). Caracteriza-se como um método eficiente, que atente às necessidades desta pesquisa, pois é inerente à técnica de investigação, assim como a descrição sistematizada e objetiva dos discursos.

A análise ocorreu à base de investigação de estudos que exploraram a semiótica como recurso aplicado para a organização e representação do conhecimento. Foram consultadas as bases de dados: Brapci, Dialnet, Google Scholar, recuperando trabalhos publicados nos idiomas português, espanhol, inglês e francês, sem demarcação específica de período de publicação. Diante do material investigado, destacam-se: Thellefsen & Thellefsen (2004); Lara (1993, 2006, 2009); Moura (2003, 2006, 2011); Almeida (2016); Ziller & Moura (2020), além dos estudos já apontados de Izquierdo Arroyo (1990, 1994, 1995); García Gutiérrez (1984, 1989, 1990).

Dessa forma, a análise evidenciou os conceitos: tratamento documental, linguagem especializada, técnicas e procedimentos para a representação aplicados ao estudo sobre o Documentação; bem como conceitos Linguística Documental, linguagem natural, disciplina, linguagem lógica, técnicas de catalogação e semiose documental acerca da Semiótica Documental. Para os estudos sobre a obra peirceana, recuperaram-se os conceitos: semiótica, categorias fenomenológicas, experiência colateral, signos lógicos; por fim, sobre a Semiótica discursiva e prática semiótica, apontaram-se os conceitos: práxis enunciativa, práxis discursiva, percurso gerativo da significação, níveis de pertinência, níveis de imanência, semiótica discursiva e práticas semióticas. Sendo assim, o estudo se caracteriza pela interpretação desses conceitos, bem como sua aplicabilidade na ampliação das reflexões sobre as práticas e estratégias que envolvem a atuação e estudo da Semiótica Documental, decorrentes da construção de inferência a partir do cruzamento das informações, resultado de um trabalho de análise interpretação interdisciplinar.

Resultados

O desenvolvimento das técnicas acerca do tratamento do documento na Espanha é fruto de muito estudo e dedicação de especialistas como Lasso de la Vega (1949), López Yepes (1978), García Gutiérrez (1984, 1989, 1990), Izquierdo Arroyo (1990, 1994, 1995) entre outros.

A aplicação da linguagem e a necessidade de tratamento especializado para garantir a representação adequada do documento, através de uma terminologia especializada, ocorre por técnicas e por procedimentos adequados devido à necessidade de precisão e de normalização. Nessa linha, García Gutiérrez (1984) propõe uma obra denominada *Linguística Documental*, em que aponta a necessidade de alinhar duas áreas do conhecimento: Linguística e Documentação a fim de constituir uma disciplina capaz de garantir subsídios teóricos para tratamento documental à base da linguagem natural.

O autor espanhol aponta preocupações com o tratamento e com a análise documental e como relacionar essas práticas à aplicação da linguagem, já que “ao conceitualizar e descrever os documentos, cria-se um terreno propício para estabelecer aqueles tipos de documentos pertinentes à elaboração de uma terminologia documental normalizada no campo da comunicação social a partir de considerações linguístico-documental e estatística” (García Gutiérrez, 1984, p. 75, tradução nossa). No entanto, o autor defende que para a realização da análise documental era preciso normalizar léxicos científicos para estabilizar a representação, o que seria possível com a normalização da linguagem documental, garantindo objetividade e unicidade, ao mesmo tempo, em que acompanha as atualizações semânticas da linguagem natural.

Contudo, há outro ponto que merece destaque no desenvolvimento das discussões: a fiabilidade dos resultados, decorrentes do processo de interpretação, pois “o analista pode indexar de modo diferente em distintos períodos dependendo de certos fatores que influenciam sua mente”, ou ainda, é possível reconhecer que “um documento indexado por dois indivíduos tem-se como resultado um produto similar, porém não é idêntico” (García Gutiérrez, 1984, p. 89, tradução nossa). Tal fator é decorrente de que “a interpretação nem sempre tem uma causa cultural ou psicológica, já que o nível de profundidade requerida nas análises, conscientemente, provoca alterações nos resultados, sendo este um motivo objetivo”, assim o problema básico para realização da análise “reside na eleição dos temas ou conceitos apropriados”, além disso “existem também os perigos de sinonímia e homonímia, acidentes estáveis graças ao uso de linguagens coordenadas” (García Gutiérrez, 1984, p. 89, tradução nossa). Diante disso, entende-se que a análise documental necessita de normalização para alcançar seus objetivos, todavia há a complexidade nas descrições dos conteúdos, fator que o autor acredita poder se ajustar devido ao “arbítrio das linguagens documentais” (García Gutiérrez, 1984, p. 89, tradução nossa).

Sendo assim, García Gutiérrez (1984, p. 135, tradução nossa) aponta a linguagem documental como “controle dos documentos”, isto é, “meios de expressão criados por e para o controle dos documentos e mais concretamente de seus conteúdos, como pontes que possibilitam a comunicação entre homens, intermediários e sistema

de documentação acumulada”. Dessa forma, a linguística, por se tratar de uma ciência que estuda sistemas de signos e como atuam para configurar atividades humanas relacionadas à produção de mensagens, pode, alinhada à ciência da Documentação e à aplicação da lógica, tratar da normalização da linguagem documental. Nessa linha, a Linguística Documental exerce sua função decorrente da “aplicação técnica e concreta dependente do marco geral da Ciência da Documentação” (García Gutiérrez, 1984, p. 137, tradução nossa). Portanto, sua aplicação está direcionada à linguagem dos documentos, por conseguinte às mensagens documentais, situação que a caracteriza como “processos informativo-documentais (científico-informativo e profissional) que tem por objetivo o estabelecimento de um efetivo controle documental mediante a utilização de mecanismos léxicos” (García Gutiérrez, 1984, p. 138, tradução nossa).

Izquierdo Arroyo, seguindo essa linha de pensamento, lança um livro: *Esquemas de Linguística Documental* (1990), motivado pelos trabalhos produzidos por López Yepes e García Gutiérrez, em que propõe a elaboração de um projeto acadêmico, que consiste na elaboração de um currículo a fim de introduzir a Linguística Documental como matéria da disciplina *Análise e Linguística Documentais*, no curso de Biblioteconomia e Documentação na *Universidad* de Murcia, como indicava a resolução 536/1989 de 5 de dezembro de 1988.

Essa disciplina apresentava como objetivo a “exposição das técnicas de catalogação de todo tipo de documentos desde um enfoque convencional ao computacional. Teoria da classificação e construção de tesouros. Além de análises de conteúdo dos documentos e metodologia do resumo científico” (Izquierdo Arroyo, 1990, p. 3, tradução nossa).

Uma das preocupações principais do professor espanhol era o processo de leitura-interpretação, que transpassava os problemas de catalogação e classificação, as quais dependiam das descrições externas do documento, como ainda as circunstâncias internas decorrentes da interpretação, perfazendo os atributos centrais da análise documental. Sendo assim, acrescia às competências a serem desenvolvidas pela disciplina de Linguística Documental, direcionadas à elaboração da linguagem documental e sua aplicação nas representações dos documentos, as preocupações acerca das “técnicas do trabalho intelectual”, ou seja, era preciso proporcionar experiências de tratamento, que se fundamentavam por exercícios cognitivos (Izquierdo Arroyo, 1990, tradução nossa).

As bases para o desenvolvimento da disciplina são aplicadas a partir da estruturação dos conhecimentos e das atividades, além da aplicação do chamado “poder efetivo”, que corresponde à “qualquer forma particular de estruturar um campo de conhecimento que se relaciona com o valor gerador de uma série de proposições apreendidas”, assim como há o poder gerativo, por sua vez, é de aplicação lógica, inerente a um determinado campo, “associando ou combinando os conceitos (a clássica associação das ideias)” (Izquierdo Arroyo, 1990, p. 29, tradução nossa).

Frente a isso, a concepção de um currículo disciplinar estabeleceu que a Linguística Documental se caracterizasse por ser uma “disciplina teórico-prática que se ocupa de um problema de armazenamento racional e ulterior recuperação de conteúdo analítico de qualquer documento” (Izquierdo Arroyo, 1990, p. 36, tradução nossa). Nessa perspectiva, a linguagem documental, denominada de “meios semióti-

cos” pode resolver esses problemas de armazenamento, através da interação com “agentes qualificados e especializados”, que garantem uma sistematização corporativa e institucional (Izquierdo Arroyo, 1990, p. 36, tradução nossa).

Assim, a elaboração da linguagem documental se realiza por um processo semiótico porque trata das características e das substâncias dos conteúdos dos documentos, isto é, “o objeto mediador (imediato ou processual) é o tratamento semiótico (ou lógico-linguístico) e controle sistemático e racional do referido conteúdo mediado”, que faz da Linguística Documental uma técnica, uma “Arte” de tratar o conteúdo documental (Izquierdo Arroyo, 1990, p. 39, tradução nossa).

Outro ponto que merece destaque nas construções da Semiótica Documental é a aproximação dos elementos que constituem a informação para C. Peirce – extensividade e profundidade – na construção de uma proposição capaz de agrupar e esclarecer termos fundamentais para a constituição da linguagem documental, ou seja, uma proposição que se perfaz pelo signo dicente da teoria peirceana, que contempla um argumento – interpretante lógico – que se estabelece como terminologia especializada para efetuar a representação e a indexação, pois carrega em sua forma conceitual, o conhecimento representado.

Tal processo é denominado por Izquierdo Arroyo de “linguagem mista”, já que sua estruturação é dada a partir da linguagem natural, extraída do documento e, na relação sujeito e atributos, elementos desprovidos da linguagem de catalogação, efetiva-se a linguagem documental, isto é, “o sujeito se designa mediante uma expressão – que pertence ao que é chamado de linguagem de catalogação, enquanto o predicado se expressa com a linguagem documental” (Izquierdo Arroyo, 1990, p. 47, tradução nossa). Ademais, os componentes básicos de uma linguagem científica são os enunciados, isto é, a sequenciação de linguagens de catalogação envolvidas à linguagem de documentação permite a extensividade dos conceitos manifestados através da informação veiculada no enunciado dos documentos.

Outrossim, o processo de transposição das informações do documento para a representação ocorre através de uma tradução, porém não uma simples tradução de pensamento, mas sim por uma semiose. Izquierdo Arroyo, indica a importância do hábito mental, isto é, da cognição e do pragmatismo tratado por C. S. Peirce e C. W. Morris. Sendo assim, o autor espanhol defende a ideia de uma semiose documental, pois o processo de representação envolve o reconhecimento dos termos da informação na análise documental, assim como a prática de interpretação, resultado de um processo cognitivo que se constitui com a atualização do hábito mental.

Sendo assim, estabeleceram-se cinco características fundamentais das estruturas de aplicação semiótica no desenvolvimento da disciplina, conforme aponta Izquierdo Arroyo (1990, p. 17, tradução nossa): a) comunicação e linguagem; b) a linguagem formal e a lógica; c) cálculo lógico: lógica proposicional; d) cálculo lógico: lógica de classes; e) metodologia do conhecimento científico. A Semiótica Documental se aplica com a finalidade de estabelecer uma espécie de “controle semiótico dos conteúdos documentais”, fato que efetiva o tratamento semiótico em busca da normalização da terminologia aplicada, ou seja, efetiva-se “a Normatização como se encontra na Ética, no Direito e na Gramática especializada”, tendo como exemplo o dicionário terminológico (Izquierdo Arroyo, 1990, p. 40, tradução nossa).

Ao aplicar a concepção de semiótica ao tratamento documental, evidencia-se que a coordenação dos descritores é orientada pela distinção de “signos lógicos e descritivos”, sendo passível a definição precisa de “signos sinônimos e proposições equipolentes” decorrentes da extensividade, caracterizada pelos “aspectos pragmáticos, semânticos e sintáticos” no tratamento semiótico, além de serem aplicados a fim de “caracterizar o conteúdo de uma proposição, ocupar-se dos paradoxos lógicos, classificar certos tipos de expressões e clarificar as expressões da necessidade, possibilidade e impossibilidade” (Morris, 1976, p. 18).

Ademais, a propósito de esclarecer as técnicas ou métodos e instrumentos que permitam tratar os conteúdos dos documentos, é preciso destacar cada uma das funções até a garantia da recuperação da informação, pois devem estar “presentes nas funções de representação e reconhecimento dos conteúdos documentais: A) Análise (comparar, distinguir, deduzir); B) Síntese (transformar, reconstruir, condensar); C) Avaliação (julgar, selecionar e definir valor” (Izquierdo Arroyo, 1990, p. 78, tradução nossa).

Na perspectiva de expansão das ações referentes à Semiótica Documental, proposta originalmente por Izquierdo Arroyo, entende-se que a linguagem é fundamental para a veiculação do conhecimento, para tanto é preciso procedimentos que analisem os níveis de significação no plano textual, discursivo e cognitivo. Nesta dinâmica, o eixo de teorias da linguagem denominado de Semiótica apresenta um rol de possibilidades para o tratamento da informação, por conseguinte instrumentos para normalização terminológica, assim como a confirmação dos aspectos que validam o conhecimento.

A Semiótica proposta por C. S. Peirce possui um arcabouço teórico muito rico e sua atuação corrobora muitas circunstâncias relacionadas à significação. Nessa análise, é importante destacar as categorias fenomenológicas (Primeiridade, Secundidade, Terceiridade) devido à aplicação da linguagem para representar o objeto e aos fatores de tratamento do fenômeno e a compreensão se sua existência, resistência e generalização por uma mente capaz de aprender com a experiência. Sendo assim, os aspectos inerentes à experiência colateral são elementos importantes na concepção de conhecimento, assim como sua aquisição (Silveira, 2007).

Sobre as categorias fenomenológicas, a primeiridade seria “algo que é aquilo que é sem referência a qualquer outra coisa dentro dele, ou fora dele, independentemente de toda força e de toda razão” (Peirce, 2017, p. 24), assim está “relacionada com as ideias de acaso, oriência, originalidade, presentidade, imediatividade, frescor, espontaneidade, qualidade, sentimento, impressão” (Santaella, 2004, p. 30), portanto “onde quer que haja um fenômeno, há uma qualidade, de modo que deveria quase parecer que não há nada mais no fenômeno. As qualidades misturam-se umas às outras. Elas não têm identidades perfeitas, mas apenas semelhanças, ou identidades parciais” (Santaella, 2020, p. 36).

A secundidade é “obsistência”, que sugere “obviar, objeto, obstinado, obstáculo, insistência, resistência”, ou seja, “é aquele elemento que, tomado em conexão com a originalidade, faz de uma coisa aquilo que uma outra a obriga a ser” (Peirce, 2017, p. 27), dessa maneira entendida como “ideias de ação e reação, esforço e resistência, conflito, surpresa, luta, aqui e agora” (Santaella, 2004, p. 31), que “engloba os fatos

atuais. As qualidades, por serem gerais, são, de algum modo vagas e potenciais. Mas uma ocorrência é perfeitamente individual” (Santaella, 2020, p. 37).

A terceiridade é mediação, “transuasão” (Peirce, 2017, p. 27), dessa forma constituindo as ideias de “generalidade, continuidade, crescimento, aprendizagem, tempo, evolução”, portanto deve ser denominada com lei (Santaella, 2004, p. 31).

Nessa concepção, a transuasão assume o caráter de signo, que “é tudo aquilo que está relacionado a uma segunda coisa, seu objeto, com respeito a uma qualidade de modo a trazer uma terceira coisa, seu interpretante, para uma relação com o mesmo objeto” (Peirce, 2017, p. 28), logo “todo pensamento se dá em signos. Qualquer coisa que esteja presente à mente é signo”, fator que ratifica que “são os signos que dão materialidade ao pensamento” (Santaella, 2020, p. 10).

O signo é composto por uma tricotomia, seu primeiro correlato é denominado de *representamen*, o qual permite a possibilidade de representação, visto que “é aquilo que, sob certo aspecto ou modo, representa algo para alguém”, é a representação de uma ideia, uma potencialidade, a força de um pensamento, “alguma qualidade de signo”, captada somente por via de abstração, assim ficará no lugar do objeto (Silveira, 2007, p. 30-31). Sua tricotomia é caracterizada por mera qualidade (qualisigno), existência concreta (sinsigno), ou lei geral (legissigno).

O segundo correlato é o objeto, o qual, de fato, é a realidade, assim o signo não o representa inteiramente, porém o signo de um objeto, embora “a realidade é sempre mais vasta e complexa do que o campo restrito a que o signo se reporta”, estabelece uma “relação de determinação do objeto para o signo, e uma relação de representação do signo para o objeto”, desse modo a relação entre signo e objeto é “a fonte da semiose” (Santaella, 2020, p. 15). Sua tricotomia é caracterizada por traços qualitativos do objeto (ícone), aplicativos (índices) e representacionais (símbolo). Frente a isso, entende-se que a relação legissigno e símbolo estabelece a linguagem de catalogação, a qual é réplica da linguagem natural, um sinsigno pela réplica da lei (termo) e indicial por relacionar a um valor semântico estabelecido na linguagem natural. Nesses dois correlatos se concentram duas preocupações da Semiótica Documental: a habilidade de representação (linguagem de catalogação) e a atualização da substância da palavra (semiose).

O terceiro correlato é o interpretante, é um signo de razão, sua tricotomia é constituída de possibilidade qualitativa (rema), de existência (dicente), de lei, representação do objeto e seus caracteres (argumento) (Peirce, 2017, p. 53), também podendo ser identificados como emocional (sentimento), energético (relação envolvendo o mundo exterior ao sujeito de semiose) e lógico (a natureza de um conceito) (Silveira, 2007, p. 52-53). É a junção dos três correlatos que se pode efetuar as considerações acerca da Semiótica Documental, pois o signo dicente é constituído por um sujeito (índice), o qual é decorrente da linguagem de catalogação e um predicado (ícone), que atualiza as atribuições do objeto, de tal forma que seu resultado é a linguagem documental, decorrente de uma investigação, uma inquirição das correlações de suas qualidades, possibilitando a extensividade da informação, tal resultado, portanto, constitui-se em um argumento, assim atualizando o hábito mental aplicado à representação do conhecimento.

Por outro lado, para a semiótica europeia, o signo se caracteriza de forma diádica, abrange os tipos textual-descritiva, práticas semióticas, níveis de pertinência, níveis

de imanência, semiótica enunciativa, semiótica discursiva, semiótica da cultura, assim a semiose é a constante transformação da substância, é a contínua revitalização da significação (Fontanille, 2008). Dessa maneira, reconhece-se nessa teoria elementos para a Semiótica Documental, já que a atualização da significação se realiza na construção do discurso e nos procedimentos de sua enunciação. Tal contribuição destaca os estudos franceses de Greimas (1966, 1975, 2014) e Fontanille (2008a, 2008b, 2019), já que oferecem recursos para o tratamento documental em níveis do enunciado, da enunciação e do discurso, resultantes de intenções manifestantes, simulacros e práxis enunciativas.

Dessa forma, analisar o impacto desses estudos no processo de tratamento e análise documental sob o viés que caracteriza o percurso gerativo da significação, bem como os níveis de pertinência semiótica a fim de averiguar o impacto da cognição pela interação com os níveis de imanência, que efetuam “a própria práxis semiótica (a enunciação em ato)” (Fontanille, 2008b, p. 18) e com formas de vida no processo de representação, podendo assim oferecer ao tratamento documental modulações além do nível do signo e do texto como já é praticado, mas incrementar com elementos modais e axiomáticos a significação inerente ao objeto, práticas, estratégias e formas de vida (Fontanille, 2008a). Além das construções tensivas que evidenciam os simulacros e as valências na constituição da isotopia (Greimas & Fontanille, 1993), estabilizando a significação para efetuar a representação, fatores que evidenciam as manifestações semânticas no documento, sendo analisadas como semiótica-objeto, o que contribui para a representação do conhecimento, já que destaca os conceitos capturados por um processo lógico-cognitivo de veridicção que confronta hábitos estabilizados e abre-se caminho para a aplicação do pragmatismo.

Além de propor uma semiótica discursiva, procurando evidenciar a significação além do nível textual, tendo como base o enunciado para verificar a isotopia, reconhecendo uma práxis enunciativa que estabelece em ação, ou seja, o discurso é a linguagem em ação, portanto uma práxis discursiva, constituída pelas ações sociais, pelas paixões e pela lógica, responsáveis por proporcionar semioses.

Assim, dá-se início às Práticas Semióticas (Fontanille, 2008a) em que se destacam as cenas práticas, ou seja, o percurso gerativo de significação. Nessa linha, pode-se atribuir à análise documental, que é resultado de uma prática de investigação da semântica, a aplicação de uma linguagem documental, ou seja, a significação não está pronta no documento, se constrói no enunciado, estabilizando a isotopia, verificada nos níveis de pertinência: a) figuras-signos, responsáveis pela identificação dos signos, dos conceitos veiculados no documento; b) textos-enunciados, níveis em que ocorre a interpretação, isto é, a identificação da rede semântica que estabelece a malha semântica do documento; c) objetos, a corporeidade do documento, sua natureza como suporte; d) cenas predicativas, a materialidade das descrições conceituais, o conjunto informacional que configura o conhecimento articulado no texto; e) as estratégias, a construção de conjunturas, seja na organização do documento por parte do produtor, do profissional da representação, como ainda o usuário; f) as formas de vida, ponto de identificação da forma informacional que revela comportamentos, condutas, ethos e hábitos, podendo ser decorrentes das descrições ou da representação (Fontanille, 2008a, 2008b).

A relação entre os níveis de pertinência e níveis de imanência revela a construção dos simulacros que se manifestam no documento a fim de mediar a isotopia, através da figuratividade do pensamento em relação ao objeto; o plano da expressão é essa conexão, todavia “o sentido não existe para se pegar (como seria um tesouro quando se cava a superfície”, o sentido “em todos os casos, deve ser construído: compreender é fazer, é operar, é construir” (Landowski, 2001, p. 29). Dessa forma, entende-se que é no reconhecimento da práxis enunciativa, constituída por simulacros e valências que se efetiva no envolvimento de temas e figuras a isotopia conceitual que se revela na interpretação do documento sob aspectos semióticos, os quais devem nortear a organização e a representação da informação e do conhecimento, assim evitando aspectos heurísticos e subjetivos no processo de representação, pois “não somente o sentido dos textos a serem lidos se constrói em ato, mas o próprio ato de leitura realizado em situação assume valor de ato” (Landowski, 2001, p. 31), é nesse confronto, praticamente uma secundidade, que a prática revela no enunciado as marcas discursivas de uma terceiridade.

Trata-se, portanto, de conjuntos teóricos, que alinhados à semiótica peircena, possibilitam uma compreensão dos conceitos veiculados, explorando os aspectos estéticos, éticos e lógicos, evidenciados por interpretantes emotivos, energéticos e lógicos, assim garantindo que o tratamento documental marcado pela cognição, que age no devir à base de ações deliberativas, constituídas por hábitos mentais atualizados pela inquirição e pela veridicção dos simulacros que constituem a práxis enunciativa e discursiva, confrontando conceitos e valências, apresentados através da investigação dos níveis de imanência decorrentes dos níveis de pertinência que compõem o objeto para garantir a interpretação adequada para que além de reconhecimento da isotopia traduzida em termos, apresente o cuidado com a evidência dos aspectos semióticos responsáveis pela interdiscursividade, permitindo uma maior abrangência da representação acerca de conhecimentos produzidos pelo documento.

Conclusões

A Semiótica Documental é objeto de anos de pesquisa de vários estudiosos já mencionados, diante dos quais esta pesquisa se fundamentou nos resultados obtidos por García Gutiérrez (1984) e Izquierdo Arroyo (1990), pois tratam da materialização de um projeto disciplinar que envolve tratamento documental e evidências de um processo cognitivo fundamental para esse processo ter sucesso, pois demonstra que a representação vai além de seleção de termos, mas é realizada pela tradução de um texto, de um discurso em um instrumento responsável por ampliar a aplicação das informações em um sistema informacional, através do apontamento, seleção e organização de conceitos, que possibilita a organização e a representação do conhecimento e da informação presentes no documento.

Somado a isso, os resultados desta pesquisa apresentam avanços para o estudo da Semiótica Documental e para a Organização do Conhecimento em:

- a) a compreensão das teorias matrizes da Semiótica Documental;
- b) aplicação do pragmatismo para a organização do conhecimento, norteador a investigação dos valores dos conceitos e a atualização do hábito mental para o

aperfeiçoamento autocontrolado e autoconsciente para efetuar o tratamento documental adequado;

c) incluir os conceitos inerentes à Semiótica discursiva e práticas semióticas na contribuição do tratamento documental ao nível textual, mas também a nível discursivo;

d) incluir o percurso gerativo da significação e as práticas semióticas no processo de tratamento documental proposto pela Semiótica Documental, bem como aplicar os níveis de imanência para inquirição dos objetos a fim de estabilizar a isotopia, assim possibilitando a aplicação do pragmatismo, da semiose documental e da atualização do hábito mental, fundamental para que o processo cognitivo presente no tratamento documental esteja afinado aos discursos, aos domínios e à linguagem documental;

e) a realização de forma mais exequível o tratamento do enunciado dos documentos, indicando o conteúdo dos níveis de imanência chegando às cenas práticas de produção de texto a fim de evidenciar as marcas de diversos domínios que possam se interessar por aquele conhecimento;

f) além de destacar as formas de vida presentes no documento e, assim, garantir a representatividade no processo de análise documental.

Portanto, além de a Semiótica Documental, fundamentada na semiótica peirceana, demonstrar que o documento é um instrumento humano para traduzir a realidade e exerce a função de registro e prova da cognição se efetuando, logo é instrumento com grande potencial para explorar o processo cognitivo. Somado a isso, os resultados obtidos apontam que o tratamento documental é marcado pelas práticas e estratégias de interpretação, por isso a semiótica discursiva apresenta ferramentas metodológicas para estabilizar a isotopia do documento e efetuar a representação.

Referências

- Almeida, C. C. (2016). A semiótica na ciência da informação brasileira: ideias e tendências. *Tendências da Pesquisa Brasileira em Ciência da Informação*, 9(2).
- Bardin, L. (2016). *Análise de conteúdo*. Trad. Luís Antero Reto; Augusto Pinheiro. São Paulo: Edições 70.
- Fontanille, J. (2019). *Semiótica do discurso*. Trad. Jean Cristtus Portela. 2.^a ed. São Paulo: Contexto.
- Fontanille, J. (2008a). *Prácticas semióticas*. Trad. Desiderio Blanco. Coleção Biblioteca Universidad de Lima, Livro digital. Lima: Fondo Editorial.
- Fontanille, J. (2008b). Práticas Semióticas: Imanência e pertinência, eficiência e otimização. In. *Semiótica e mídia: textos, práticas, estratégias*. Maria Lúcia Vissotto Paiva Diniz; Jean Cristtus Portela (Org.). Bauru: UNESP/FAAC.
- García Gutiérrez, A. L. (1984). *Linguística Documental: Aplicacion a la Documentacion de La Comunicación Social*. Mitre: Barcelona, Espanha.
- García Gutiérrez, A. L. (1989). Teoria da la inidización: nuevas parámetros de investigación. *Transinformação*, 1(2). Disponível em: <http://hdl.handle.net/20.500.11959/brapci/22889>. Acesso em: 13 abr. 2021.

- García Gutiérrez, A. L. (1990). *Estructura lingüística de la documemación: teoría y método*. Murcia: Universidad de Murcia.
- Greimas, A. J. (1966). *Sémantique Structurale*. Paris: Librairie Larousse.
- Greimas, A. J. (1975). *Sobre o sentido: ensaios semióticos*. Trad. Ana Cristina Cruz Cezar et al. Petrópolis: Vozes.
- Greimas, A. J. (2014). *Sobre o sentido II: ensaios semióticos*. Trad. Dilson Ferreira Cruz. São Paulo: Nankin: Edusp.
- Greimas, A. J. & Fontanille, J. (1993) *Semiótica das paixões: dos estados de coisas aos estados de alma*. Trad. Maria José Rodrigues Coracini. São Paulo: Ática.
- Greimas, A. J. & Courtés, J. (2016). *Dicionário de semiótica*. São Paulo: Contexto.
- Ibri, I. A. (2015). *Kósmos noetós: a arquitetura metafísica de Charles S. Peirce*. São Paulo: Paulus.
- Ibri, I. A. (2020). *Semiótica e pragmatismo: interfaces teóricas: vol 1*. São Paulo: Cultura Acadêmica.
- Izquierdo Arroyo, J. M. (1990). *Esquemas de lingüística documental*. Barcelona: Promociones y Publicaciones Universitarias. Tomo I.
- Izquierdo Arroyo, J. M. & Moreno Fernandez, L. M. (1994). Listas de encabezamientos de matéria y Thesauri em perspectiva comparada. *Documentación de las Ciencias de la Información*, (17), 287-310.
- Izquierdo Arroyo, J. M. (1995). Estructuras conceptuales para la representación documental. *Organización del conocimiento en sistemas de información y documentación: actas del I Encuentro de ISKO-España*, Madrid, 4 y 5 de noviembre de 1993, 27-50.
- Izquierdo Arroyo, J. M. & Moreno Fernandez, L. M. (1995). Problemas de terminología metalingüística em los lenguajes documentales de estrutura combinatoria. *Organización del conocimiento en sistemas de información y documentación: actas del I Encuentro de ISKO-España*, Madrid, 4 y 5 de noviembre de 1993, 51-64.
- Landowski, E. (2001). *O olhar comprometido*. Galáxia, São Paulo: PUC, n. 2, p. 19-56.
- Lara, M. L. L. G. (1993). Algumas contribuições da semiologia e da semiótica para a análise das linguagens documentárias. *Ciência da Informação*, 22(3).
- Lara, M. L. L. G. (2006). é possível falar em signo e semiose documentária? 10.5007/1518-2924.2006v11nesp3p18. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, n. esp., p. 18-29.
- Lara, M. L. L. G. (2009). *Linguística Documentária: seleção de conceitos*. 2009. Tese (Livre Docência em Análise Documentária) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo.
- López Yepes, J. & Garcia, J. R. (1993) *¿Qué es Documentación? Teoría e História del concepto em España*. SÍNTESIS: Madrid, Espanha.
- López Yepes, J. (1995) *La Documentacion como disciplina: Teoria e História*. 2.^a ed. EUNSA: Navarra, Espanha.
- Morris, C. W. (1976). *Fundamentos da teoria dos signos*. São Paulo Livraria Eldorado Tijuca; Edusp.
- Moura, M. A. (2003). Semiótica e mediações digitais: o processo de criação e recepção de hipermídias. *Perspectivas em Ciência da Informação*, 8(2).
- Moura, M. A. (2006). Ciência da informação e semiótica: conexão de saberes 10.5007/1518-2924.2006v11nesp3p1. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, n. esp., p. 1-17.
- Moura, M. A. (2011). Interoperabilidade semântica e ontologia semiótica: a construção e o compartilhamento de conceitos científicos em ambientes colaborativos online. *Informação & Informação*, 16(2), p. 165-179.

- Peirce, C. S. (2017). *Semiótica*. Trad. José Teixeira Coelho Neto. 4. ed. São Paulo: Perspectiva.
- Silveira, L. F. B. (2007). *Curso de Semiótica geral*. São Paulo: Quartier Latin.
- Santaella, L. (2004). *O método anticartesiano de C. S. Peirce*. São Paulo: Editora Unesp.
- Santaella, L. (2008). *O que é semiótica*. São Paulo: Brasiliense.
- Santaella, L. (2020). *Charles Sanders Peirce: excertos*. São Paulo: Paulus.
- Santaella, L. & Nöth, W. (2017). *Introdução à semiótica: passo a passo para compreender os signos e a significação*. São Paulo: Paulus.
- Thellefsen, T. L. & Thellefsen, M. M. (2004). Pragmatic semiotics and knowledge organization. *Knowledge Organization*, v. 31, n. 3, p. 177-187.
- Ziller, J. & Moura, M. A. (2010). Semiose e fluxos informacionais: os agenciamentos coletivos e a condição de usuário em ambientes digitais. *Liinc em revista*, 6(2).

(Página deixada propositadamente em branco)

PRAGMATISMO E CATEGORIAS FENOMENOLÓGICAS DE PEIRCE: UM DEBATE SOBRE A NOÇÃO DE CONCEITO

Alexandre Robson Martines¹, Valdirene Aparecida Pascoal²,
Carlos Cândido de Almeida³

¹Unesp (FFC) Marília-SP, alexandre.martines@unesp.br, ORCID iD 0000-0002-4524-0978

²Unesp (FFC) Marília-SP, valdirene.pascoal@unesp.br, ORCID iD 0000-0002-3695-6560

³Unesp (FFC) Marília-SP, carlos.c.almeida@unesp.br, ORCID iD 0000-0002-8552-1029

Resumo

A Organização do Conhecimento tem como um dos seus objetivos e atividades estabelecer a teoria conceitual, logo preocupa-se com classificação e indexação e representação do conhecimento. Nessa linha, o objetivo desta pesquisa é aplicar as teorias desenvolvidas por C. S. Peirce acerca das categorias fenomenológicas e do pragmatismo para compreender a constituição do conceito a partir da perspectiva da experiência e da construção de valores mediante a inquirição. Para tanto, aplicou-se metodologia qualitativa, de caráter epistemológico, exploratório, sob o método de Análise do Conteúdo. Como resultado, o conceito pode ser estabelecido a partir da experiência e compreendido através da generalidade decorrente das categorias fenomenológicas e do pragmatismo por não ser mera vontade da razão, mas o conceito é fortalecido através da dúvida, decorrente da sua veracidade em ato. Portanto, o pragmatismo e as categorias fenomenológicas apresentam fundamentos teóricos capazes de enriquecer o tratamento do conceito.

Palavras-chave: Organização do Conhecimento; Conceitos; Categorias Fenomenológicas; Pragmatismo; Semiótica peirceana

Introdução

A linguagem é recurso simbólico que promove as dinâmicas de significação. Nessa linha, a produção de documentos se utiliza da linguagem natural, pois é a linguagem humana, de símbolos convencionalizados pela prática, pelo uso, pelo hábito e pela história, que evidencia sua construção e evolução ao longo das eras, no entanto, ao se tratar de comunicação científica, a linguagem a ser aplicada é a linguagem especializada, de natureza técnica, a qual é decorrente dos estudos de documentos e denominada de linguagem documental.

O processo de tratamento e análise documental é constituído pela identificação de conceitos evocados na construção dos enunciados dos documentos a fim de estabilizar o conhecimento proposto. Sendo assim, essa prática ou técnica tem como

objetivo a tradução da linguagem natural em linguagem especializada, com a orientação de sistemas de organização do conhecimento (SOCs). Por conseguinte, esse processo de identificação e extração dos conceitos de um documento e tradução para uma linguagem técnica, cuja função é estabilizar a semântica, é uma das preocupações centrais da Organização do Conhecimento.

Diante disso, o estudo da linguagem configura-se ser importante para o tratamento do conceito, pois se realiza como “processo de modelagem do conhecimento que visa à construção de representações do conhecimento. Esse processo tem por base a análise do conceito e de suas características para o estabelecimento da posição que cada conceito ocupa num determinado domínio” (Brascher & Café, 2008, p. 8), já que a Organização do Conhecimento “analisa os termos e conceitos associados, pois são a forma de veicular os conteúdos documentais” (Barité, 2015, p. 13, tradução nossa).

A Organização do Conhecimento estuda “as leis, os princípios e os procedimentos pelos quais se estruturam o conhecimento especializado em qualquer disciplina, com a finalidade de representar tematicamente e recuperar a informação contida em documentos de qualquer natureza” (Barité, 2015, p. 120, tradução nossa), além de ser uma “estrutura conceitual que representa modelos de mundo”, visa “à construção de modelos de mundo que se constituem em abstrações da realidade” (Brascher & Café, 2008, p. 6).

Nessa linha, a Organização do Conhecimento tem como um dos seus objetivos e atividades estabelecer a teoria conceitual, logo preocupa-se com a linguagem adequada para realizar a classificação, a indexação e a representação do conhecimento, pois “não entendemos apenas a estrutura lógica da representação conceitual, mas também todas as questões de nomear conceitos pelos termos adequados” (Dahlberg, 2006, p. 11, tradução nossa).

A relação interdisciplinar entre a Organização do Conhecimento com teorias da linguagem e da representação é marcada por buscas de técnicas, práticas e estratégias para o entendimento e tratamento dos conceitos, já que estes estabelecem o todo organizado do fluxo informacional presente nos documentos. Assim, as teorias da Organização do Conhecimento devem centralizar seus esforços em “conceitos, critérios para a inclusão de categorias, significado, indexação, relações semânticas, assuntos e pontos de acesso do sujeito” (Hjørland, 2003, p. 88, tradução nossa).

Entende-se que os conceitos possam ser estudados decorrente a sua relação com a experiência, a qual interage com os fenômenos e, a partir de signos, estabelecem categorias para representar a realidade, sendo estas capazes de organizar o conhecimento humano, através de estágios de cognição.

Devido a isso, o objetivo desta pesquisa é aplicar as teorias desenvolvidas por C. S. Peirce acerca das categorias fenomenológicas e do pragmatismo para compreender a constituição do conceito a partir da perspectiva da experiência e da construção de valores mediante a inquirição.

Como resultados, estabelecem-se caminhos para aplicar as categorias fenomenológicas e o pragmatismo peirceano no tratamento dos conceitos e, assim, correlaciona-se a sua existência à experiência, confirmando a veracidade em ato. Dessa forma, é possível contribuir para a área da Organização do Conhecimento, já que o pragmatismo se trata de método de inquirição do valor de verdade, o qual pode ser aplicado aos conceitos manifestados no documento.

Metodologia

A metodologia aplicada possui natureza qualitativa, pois os objetivos apresentam caráter epistemológico, exploratório, no direcionamento de evidenciar as constituições teóricas desenvolvidas pelas áreas apontadas, com o fito de inter-relacioná-las em uma investigação interdisciplinar, por isso se efetivando em uma pesquisa bibliográfica. Aplicou-se o método de Análise do Conteúdo, efetuado em três etapas: pré-análise, em que ocorreu o levantamento de materiais a partir de bases de dados como Brapci, Dialnet, Google Scholar, recuperando pesquisas desenvolvidas em idiomas português, espanhol e inglês, sendo que não se estabeleceram datas para as publicações recuperadas, pois o foco da pesquisa está na temática. Dentre os materiais consultados, destacam-se: Dahlberg (1978, 2006), Koselleck (1992, 2020) e Hjørland (2002, 2009) para debater fundamentos sobre a concepção de conceito, assim como as teorias peirceanas sobre as categorias fenomenológicas e o pragmatismo: Peirce (1993, 1994, 2000, 2017), Silveira (2007), De Waal (2007) e Ibri (2015). Alinhado a isso, também realizou-se levantamento de trabalhos que tenham debatido a relação do conceito com a linguagem ou com a semiótica no cenário da Organização do Conhecimento. Destacam-se Izquierdo Arroyo (1990, 1993); Almeida (2012, 2013, 2015); Moron, Almeida & Moreira (2013).

A análise evidenciou os principais conceitos aplicados ao desenvolvimento da pesquisa. São eles: conceito, predicação, uniformidade, enunciados verdadeiros, sentido comum, associações semânticas, teorias epistemológicas e abstração relacionados à teoria dos conceitos e à Organização do Conhecimento. Já os conceitos extraídos da análise da filosofia peirceana são: cognição, objeto, dúvida, vontade da razão, categorias fenomenológicas, pragmatismo, consequências práticas de aplicação, conduta racional, hábito mental, realidade, extralinguístico, experiência colateral e relação lógica.

A interpretação foi conduzida pelos cruzamentos das informações e pelas inferências para evidenciar as contribuições peirceanas para o tratamento do conceito, pois visa à análise das categorias fenomenológicas e do pragmatismo a fim de explicar o tratamento do conceito à base da experiência.

Resultados

A teoria do conceito entende que a formação dos conceitos se confirma como “a reunião e compilação de enunciados verdadeiros a respeito de um determinado objeto, fixados por um símbolo linguístico” (Dahlberg, 1978, p. 102). Para a formação do conceito, há uma relação de signos por associações, as quais “pressupõem um mínimo de sentido comum, uma pré-aceitação de que trata de palavras importantes e significativas” (Koselleck, 1992, p. 135). Além disso, para a Organização do Conhecimento, entende-se “o conceito como uma entidade que pode representar a realidade e o conhecimento”, enquanto para “a comunicação e a mediação da informação, condicionadas a um relativismo, definem conceito de uma maneira mais flexível, e não tão próspera para representar o real e o conhecimento”, pois salientam que “o

conhecimento avança continuamente, por isso mesmo inadequada a qualquer forma de acomodá-lo semanticamente” (Almeida, 2012, p. 50).

Alinhado a isso, pode-se apontar ainda que “a construção do conceito depende basicamente de um referente, da emissão de juízos sobre o referente, de uma forma verbal (um termo ou um nome) e de uma maneira de usar essa forma verbal em um universo discursivo” (Francelin & Kobashi, 2011, p. 210). Ademais, pode ser definido como “qualquer unidade de pensamento. Noção selecionada para reter como unidade de análise semântica”, por conseguinte é um “elemento do pensamento expresso, em geral, por um termo ou por um símbolo literal ou outro” (Faria & Pericão, 2008, p. 188), o que possibilita a comunicação científica e o diálogo entre os campos científicos.

Todo conceito é não apenas “efetivo enquanto fenômeno linguístico; ele é também imediatamente indicativo de algo que se situa para além da língua” (Koselleck, 1992, p. 136), isto é “conceito pode ser definido como “abstração ou noção que se refere a uma unidade de conhecimento, independente, de sua expressão linguística e compreende o conjunto de seus recursos essenciais” (Barité, 2015, p. 50, tradução nossa). Assim, o conceito é mecanismo responsável pela compreensão do sistema informacional, já que “um conceito relaciona-se sempre àquilo que se quer compreender, sendo portanto a relação entre o conceito e o conteúdo a ser compreendido, ou tomado inteligível, uma relação necessariamente tensa” (Koselleck, 1992, p. 136).

Dentre as teorias epistemológicas apresentadas por Hjørland na sistematização da teoria do conceito, destaca-se o pragmatismo. Sendo assim, tem-se a definição de que o pragmatismo “é o ideal de basear o conhecimento na análise de objetivos, propósitos, valores e consequências” (Hjørland, 2009, p. 1526, tradução nossa).

Além disso, o pragmatismo tenta ser “explícito sobre o propósito da pesquisa e da cognição”, por isso “é baseado no pressuposto de que o conhecimento não pode ser neutro (devido à sua natureza teleológica) e, portanto, é importante descobrir os valores e consequências inerentes a qualquer afirmação de conhecimento, em qualquer concepção e em qualquer classificação” (Hjørland, 2009, p. 1526, tradução nossa). Por conseguinte, “o pragmatismo entende os conceitos como uma forma de fixar partes da realidade no pensamento, na linguagem e em outros sistemas simbólicos” (Hjørland, 2009, p. 1526, tradução nossa).

Hjørland adota a teoria do pragmatismo de Willian James, que compreende que ações são feitas para um fim, ou seja, “para atingir clareza perfeita em nossos pensamentos de um objeto precisamos somente considerar quais efeitos de uma espécie concebivelmente prática o objeto pode envolver – quais sensações devemos esperar dele, e quais reações devemos preparar” (De Waal, p. 52). Porém, de fato, o pragmatismo é originalmente apresentado por Peirce, em 1878 em *Como tornar nossas ideias claras e reformulado em 1905, em O que é o Pragmatismo*.

Por sua vez, esta pesquisa aborda as teorias apresentadas por C. S. Peirce para debater sobre a constituição do signo, caracterizado como um elemento fundamental na compreensão da representação dos fenômenos do mundo. Em linhas gerais, na conceituação peirceana, o pragmatismo é o método para a definição dos significados. A máxima pragmática peirceana sustenta que o significado de um conceito está contido nas consequências práticas de sua aplicação. Peirce, por não aprovar a

conceituação que Willian James deu à palavra pragmatismo, propõe o termo *pragmaticismo* para referenciar sua teoria.

Em 1905, Peirce reformula sua máxima: “todo propósito intelectual de qualquer símbolo consiste na totalidade dos modos gerais de conduta racional, que na dependência de todas as possíveis e diversas circunstâncias e desejos, assegurariam a aceitação do símbolo” (Peirce, 2000b [1905], p. 287). Em vista disso, a ênfase da máxima pragmática peirceana antes dada “às consequências práticas” se transforma em “conduta racional”. Nas palavras do filósofo, *pragmaticismo*:

[...] o significado racional de uma palavra ou alguma outra expressão se encontra, exclusivamente, na concepção da sua influência concebível sobre a conduta da vida [...] uma vez que é óbvio que algo que não resulte de um experimento possa exercer qualquer influência direta sobre a conduta, se alguém for capaz de definir, com exatidão, todos os experimentos fenomênicos concebíveis que a afirmação ou negação de um conceito possa implicar, teria, então, a definição completa do conceito. (PEIRCE, EP 2. 332, 1992-1998).

Dessa forma, a ampliação da máxima pragmática peirceana ocorre no entendimento de que a conduta está alicerçada na fundamentação de um método que alia teoria e prática. De acordo com Almeida (2014), o significado de um conceito, nessa perspectiva *pragmaticista*, inicia-se pela esquematização a partir da imaginação, das consequências práticas de sua representação. O significado só pode ser atribuído com a confirmação de sua existência no mundo a partir da experiência. Em outras palavras, o conceito só se confirma a partir de sua aplicabilidade, é contínuo de possibilidades e não pode ser confundido com instanciações.

Almeida (2014) aprofunda a explicação e afirma que a capacidade de expressar relações de ordem dos objetos entre objetos ocorre porque o conceito estabelece a relação entre os fatos brutos do mundo. A relação ali exposta configura a conduta deste objeto. A conduta do objeto direciona a conduta do intérprete do objeto. Desse modo, as relações sgnicas compõem o desenvolvimento de um hábito de conduta, ou seja, tal hábito é o resultado do significado interpretado de um conceito. Assim, o signo atua na mediação do objeto frente ao universo fenomenológico, dessa maneira “é necessário observá-lo onde quer que se manifeste e tal manifestação, distinguindo da pura produção da razão”, tal fator retrata, portanto, que jamais “a Lógica o terá exposto à sua explicação, mas precisará buscá-lo no universo da experiência e construir sobre ele hipóteses que o expliquem” (Silveira, 2007, p. 21).

Nessa perspectiva, a experiência “é aquilo que diretamente é revelado pela arte observacional da ciência” (Ibri, 2015, p. 23). Tal processo está “conectado e assimilado ao conhecimento já possuído e derivado de outro modo”, visto que a “experiência é o inteiro resultado cognitivo do viver”, ou seja, “a experiência é o curso da vida” (Ibri, 2015, p. 23). Nessa medida, à base de um resultado cognitivo proposto pela experiência, é possível “semear conceitos que moldam a conduta humana”.

Peirce desenvolveu sua teoria das categorias fenomenológicas universais ao longo de trinta anos, de 1867 a 1897, a qual foi apresentada no ensaio *Sobre uma nova lista de categorias*. No entanto, apenas, em 1902, Peirce aplica as categorias fenomenológicas como base geral para sua doutrina lógica (cf. Braga, 1999). Em seu

escrito *Metafísica tem de considerar todo o universo do ser* (CP. 6.214.), por *ser*, Peirce entende três modos (CP 1.23): “O ser de possibilidade qualitativa positiva, o ser do fato atual, e o ser da lei que governa os fatos no futuro” (Peirce, CP 1.23, 1903), caracterizadas, segundo o filósofo, em três modos de ser: o ser da possibilidade, o ser do fato e o ser da lei, que governará os fatos no futuro. Destarte, as categorias fenomenológicas se caracterizam em primeiridade, secundidade e terceiridade.

Primeiridade é a categoria dos fenômenos em si mesmos, existentes, independente de qualquer fenômeno ou coisa. Representa a casualidade, aquilo que é espontâneo. Os fenômenos descritos como primeiridade não são determinados por nada além deles mesmos, não passam de possibilidades. Relacionam-se à primeiridade as ideias de frescor, liberdade, algo que é livre em si mesmo sem que haja outro determinando suas ações (Peirce, CP 1.303, 1903). Para Peirce (CP 8.328, 1903): “Primeiridade é o modo de ser daquilo que é tal como é positivamente e sem referência a qualquer outra coisa”. Santaella (1983, p. 46) reitera que a primeiridade não pode ser articulada conceitualmente, pois no momento em que há determinação e significado, perde toda a característica de primeiridade e passa a ser secundidade. A categoria da primeiridade é a categoria do sentimento sem reflexão, da espontaneidade pura, da sensação sem nome, do imediato (Nöth; Santaella, 2017). Alguns exemplos que auxiliam na compreensão dos fenômenos de primeiridade são fornecidos por Peirce: “[...] uma sensação vaga de vermelho, ainda não objetivada, do sabor de sal, apenas o sabor nele mesmo, uma dor ou tristeza vaga, o puro sentimento em si de alegria ou de uma nota musical prolongada” (Peirce, CP 1303, 1894, apud Nöth; Santaella, 2017, p. 37).

A secundidade está no nível da experiência, da coisa ou do evento, é aquilo que existe, e impõe resistência, é a categoria da ação e da reação. O imediatismo do presente está contido na segunda categoria. A secundidade é compreendida como o fato em si, sendo brutalmente imposto, é a alteridade das relações entre um e outro, é a vida acontecendo no mundo existencial e material. É aquilo que existe e para existir é necessário a constituição do tempo e do espaço. Peirce (CP 2.237, 1903) define a secundidade da seguinte maneira: “O Segundo Correlato é um dos três que é considerado de complexidade média, de modo que se existir dois que são da mesma natureza, podem ser meras possibilidades, existências reais, ou leis, assim o Segundo Correlato é uma existência atual”. Peirce exemplifica a secundidade como: “(a secundidade) nos aparece em fatos tais como o outro, a relação, compulsão, efeito, dependência, independência, negação, ocorrência, realidade, resultado”. (CP 1.358, apud Nöth; Santaella, 2017, p. 38).

A relação entre o primeiro e o segundo fenômeno possibilita o terceiro modo de ser. A terceiridade é a categoria da continuidade e da mediação por um terceiro entre um primeiro e um segundo. É a categoria da semiose e dos signos. Nas palavras de Peirce: “Terceiridade nada é senão o caráter de um objeto que incorpora a Mediação nas suas formas mais simples e rudimentares; e eu a uso como o nome daquele elemento do fenômeno que é predominante onde quer que a Mediação seja predominante, e que encontra sua plenitude na representação” (Peirce, CP 5.104, 1878). É imprescindível enfatizar que as três categorias são indissociáveis e estão contidas, em algum grau, nos fenômenos da experiência. O que possibilita a classificação do

fenômeno em primeiridade, secundidade e terceiridade é a predominância de elementos da categoria em determinado fenômeno.

Em decorrência da experiência, da lógica e da concepção da verdade, o pragmatismo se configura como uma ciência geral, “comum a todas as inteligências capazes de aprender com base na experiência” (Silveira, 2007, p. 182). Portanto, é “a investigação de como procederiam os sentidos e suas impressões mais imediatas confeririam instrumentos para analisar nossas complexas representações e esclarecê-las” (Silveira, 2007, p. 183).

Do mesmo modo que as categorias da experiência se fundamentam no mundo, o signo se caracteriza por meio de relações triádicas que se estabelecem entre o signo (*representamen*), objeto e interpretante. A construção dessas relações triádicas se dá à medida que o signo produz na mente do intérprete outro signo. A partir de tal relação, Ibri (2020) afirma que estudiosos do pragmatismo peirceano estabelecem vínculos entre o termo prático e a conduta. No entanto, a conduta não pode ser reduzida à mera instância da ação. Seu significado deve ser geral, repleto de terceiridade e dependente de observação espaço-temporal, ou seja, situada na secundidade. Dessa forma, Ibri (2020) compreende que quando a conduta se situa como a generalização da ação ou modo de agir, ela se situa como terceira categoria, bem como o conceito.

Sendo o conceito estabelecido a partir da experiência e compreendido através da generalidade decorrente das categorias fenomenológicas, pode atuar como a representação do conhecimento, não por uma ideia abstrata definida anteriormente, ou por mera vontade da razão, mas essa crença se constitui e se fortalece através da dúvida, princípio posto em prova pelo pragmatismo decorrente da sua veracidade em ato, gerando a conduta racional.

Desse modo, “o conceito é visto como a unidade indivisível de conhecimento, aquela que representa os objetos da realidade e representa o teor de significado e sentido de um termo científico e especializado”, “o conceito enquanto signo refere-se a um fundamento. O signo compartilha com o objeto um aspecto, uma qualidade ou uma característica”, contudo “o conceito não é formado apenas por qualidades. Na tipologia dos signos, Peirce dá um espaço importante aos chamados signos de convenção, sendo assim, “o conceito é um símbolo, resultante da união entre signo e objeto, ou da segunda tricotomia, que estabelece também o ícone e o índice” (Almeida, 2012, p. 51).

No entanto, o conceito é uma unidade de pensamento, e, por isso, pode ser um signo de terceiridade, pois está relacionado a um objeto, a um fenômeno, o que o evidencia na representação como uma convencionalidade simbólica, todavia também é compreendido decorrente a atividades cognitivas que são constituídas pela interação da experiência; ele estabelece um hábito de conduta, porque é resultado de semiose. Em sua aplicação frente a uma sociedade científica ou cultural é um hábito mental estabelecido, convencional, um símbolo, todavia um conceito é atualizável conforme os aspectos semióticos geram a semiose, atualização e crescimento dos signos, na tricotomia do interpretante, na categoria discente, e, por isso, há um processo pragmático na inquirição acerca da sua atualização.

Nessa linha, o conceito convencionalizado, já estabilizado na Organização do Conhecimento, aplicado a categorias e classificações, trata-se de um conceito sim-

bólico, devido à sua relação com o objeto, todavia a interação, a contínua descoberta, a ampliação da experiência proporcionam a esse conceito uma ampliação em sua significação, por isso ele pode ser visto como resultado de debates, conflitos e construções cognitivas, e é neste ponto em que o pragmatismo se torna fundamental como método para inquirição dos valores lógicos que se configuram como novo argumento, pois proporciona a atualização no hábito de conduta.

Não obstante, o conceito é atualizável, pois carrega em si fatores de primeiridade em relação ao seu objeto, já que a espontaneidade, a causalidade e possibilidade perfazem sua relação com a mera percepção, o conjunto de qualidades que são meros sentimentos. Assim, em confronto, em relação a um outro, o conceito em secundidade toma caminhos de existência, decorrente das diversas possibilidades construídas no âmbito social e cultural. Dessa forma, entre ser uma sensação a ser compreendida e o conflito de possibilidades, as áreas científicas se encarregam de definir as atribuições e os atributos do conceito e determinar sua função na modelação do conhecimento. Estando os seres humanos em pensamento, é o conceito, em seu estágio de terceiridade, que efetiva o construto da realidade.

O conceito é moldado decorrente às interações sociais e culturais, as quais são demarcadas e orientadas pela interação com fenômenos de secundidade e elementos extralinguísticos, a qual evolui e atualiza sua significação devido às necessidades, às descobertas, aos conflitos ideológicos, aos confrontos de perspectivas e superação de preconceitos, sistematizados em uma representação da realidade ao devir, por conseguinte “sem formações e sem os conceitos com os quais tais formações – de modo reflexivo ou autorreflexivo – tentam definir e vencer os desafios que se lhe apresentam, não existe história” (Koselleck, 2020, p. 19). Sendo assim, as investigações estão sempre se confrontando com novas descobertas em relação ao seu objeto de pesquisa. Dessa maneira, a relação objeto e conceito constrói influência sobre a conduta da vida social, cultural e científica, através de raciocínios e métodos norreados pela lógica, e é como consequência da dúvida e da experiência que a compreensão sobre a realidade se renova.

Ademais, o esclarecimento sobre o objeto ou fenômeno é decorrente de descrição de suas ideias, conseqüentemente sobre a organização e a representação da informação que se apresenta para regularizar as evidências sobre o objeto, para, assim, ser testado por um processo de inquirição na sua legitimação como conceito, o qual aciona processos cognitivos. Dessa forma, o conceito se estabiliza diante de fatores sociais e históricos, não apenas como objeto de convenções, mas também na interação com a realidade e o extralinguístico, consolidando comportamentos, condutas e valores demarcados por signos, contudo na ausência desses fatores, a história “não pode ser objeto nem de experiência nem de interpretação, nem de representação nem de narração” (Koselleck, 2020, p. 19), porém há o perigo de se realizar decorrente da vontade da razão, por isso é necessário a veracidade a fim de garantir que a atualização do conceito esteja consoante à atualização dos objetos e fenômenos, além de mensurar as conseqüências do falibilismo.

Ainda nessa linha, há uma interação constante entre a história social e a história dos conceitos, “a história dos conceitos tematiza um vínculo entre eventos sincrônicos e estruturas diacrônicas” (Koselleck, 2020, p. 28). Desse modo, um hábito pode sofrer atualizações conseqüentes à novas estruturas e proposições, entretanto na

aplicação do conceito, há fundamentos e características do estágio anterior do hábito que permanecem, pois “aquilo que ocorre pode ser singular e novo, mas nunca é tão novo a ponto de não ter sido viabilizado por condições sociais preestabelecidas no longo prazo” (Koselleck, 2020, p. 28).

É preciso apontar que há implicações individuais e fatos históricos que possam interferir ou contribuir na constituição do conceito, todavia é na relação com o outro, no confronto da representação da realidade que suas particularidades, seus atributos se configuram ou se renovam, isto é, o conceito é um construto de pensamento e engrenagem de conhecimento que se constrói e se efetiva na presença do outro, efetivando uma conduta racional. Sendo assim, analisar a constituição de um conceito sob os aspectos teóricos das categorias fenomenológicas possibilita a compreensão de seu processo de representação, desde sua gênese, na relação com as qualidades, percepção dos seus atributos, resistência aos seus critérios de existência, sua adequação à regularidade de interações sociais, culturais e históricas, das quais seu valor representacional e convencional é investigado pelas áreas interessadas, até sua normalização como material de pensamento, como elemento lógico.

Neste aspecto, o pragmatismo aplica-se como uma espécie de método de inquirição a fim de testar as possibilidades significativas, assim como os recursos e parâmetros aplicados na configuração do conceito, garantido que este seja constituído à base de fundamentos lógicos, organizados por um processo imerso em raciocínio a fim de garantir que o valor de verdade estabilizado no procedimento de organização e representação do conhecimento seja consoante aos aspectos da realidade, resultante dos processos de semiose, evidenciados em hábitos de conduta.

Conclusões

A teoria do conceito se caracteriza por ser o um dos objetos centrais da Organização do Conhecimento, como visto em Dahlberg (1978, 2006); Hjørland (2002, 2009) e sua preocupação é inerente ao tratamento dos documentos, como visto em Izquierdo Arroyo (1990, 1993); Almeida (2013); Almeida & García Marco (2015).

A teoria do conceito proposta por Hjørland evidencia a importância das teorias epistemológicas para as discussões acerca do conceito, porém a filosofia apresenta um posicionamento sobre o conceito para a representação da realidade a partir da experiência. Dessa forma, o pragmatismo e as categorias fenomenológicas ganham importância nos estudos atuais. Portanto, os resultados obtidos demonstram que:

- O pragmatismo e as categorias fenomenológicas propostas por Charles S. Peirce apresentam fundamentos teóricos capazes de enriquecer o tratamento do conceito;

- As categorias fenomenológicas se relacionam com o conceito em seu processo de constituição de hipóteses, de existência, de confronto até a confirmação de suas características lógicas, assim o conceito, por ser de terceiridade, é construto de pensamento e engrenagem de conhecimento;

- A atualização do conceito é decorrente das interações sociais frente às descobertas e à ampliação das experiências decorrentes de fatores extralinguísticos, assim o signo assume uma nova relação com o objeto e gera um novo valor cognitivo, desse modo a semiose atualiza o hábito, o qual, embora possa propor uma nova

significação, mantém característica e fundamentos do antigo estágio, visto que este também foi construído através de métodos lógicos;

– O objeto não é estático, nem exclusivamente externo, real, é também imaginário, fictício e interno. Desse modo, na representação, o signo é a orientação do objeto, e o conceito como signo indica o conhecimento. Ademais, sempre que um intérprete adquire um novo conhecimento, surge um novo objeto do signo quando um novo pensamento se refere a esse conhecimento.

– O conceito é uma unidade de pensamento, com isso pode evidenciar que a conduta se situa como a orientação cognitiva da ação ou modo de agir, ela se situa como terceira categoria;

– O conceito atua como mediação entre o real bruto e a experiência, esta é capaz de se atualizar conforme o confronto com os limites da percepção, fato que evidencia o falibilismo;

– O conceito não será compreendido apenas como uma abstração decorrente de uma ideia estabelecida anteriormente, ou decorrente da vontade da razão, mas sim como uma soma de predicativos que efetuam a generalidade decorrente da soma de consequências de experiência em ato e a compreensão de sua atualização, bem como sua disposição para o devir;

– Com a aplicação do falibilismo, o conceito caracteriza o objeto por um todo de concepção decorrente da experiência e está em constante atualização, assim não se limitando a ser pertencente a um único domínio.

Referências

- Almeida, C. C. (2012). Conceito como signo: elemento semiótico para análise e mediação da informação. // *Scire*. 18:2. 49-55. <https://www.ibersid.edu/ojs/index.php/scire/article/download/3961/3706/>
- Almeida, C. C. (2013). Notas sobre os conceitos semióticos fundamentais à organização do conhecimento. II |*Congresso ISKO España y Portugal/ XII Congreso ISKO España*. ISBN 978-989-8648-10-5, págs. 68-81.
- Almeida, C. C. & Marco, F. J. G. (2015). Aportaciones semióticas de la documentación en España a la Organización del Conocimiento: un análisis preliminar. II *Congreso ISKO España-Portugal, Organización del conocimiento: sistemas de información abiertos*. ISBN 978-84-608-3558-5, págs. 498-507.
- Almeida, R. V. (2014). Pragmatismo e Pragmaticismo – O Embate Peirceano. *Revista do Instituto Humanistas Unisinos*, v. 457, ed. Ano XIV, págs. 56-64, out. 2014.
- Bardin, L. (2016). *Análise de conteúdo*. Trad. Luís Antero Reto; Augusto Pinheiro. São Paulo: Edições 70.
- Barité, M. et al. (2015). *Diccionario de Organización del Conocimiento: Clasificación, Indización, Terminología*. 6.a ed. corregida y aumentada. Montevideo: Csic.
- Bräscher, M. & Café, L. (2008). Organização da informação ou organização do conhecimento? In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. *Anais eletrônicos*[...] São Paulo: USP, ANCIB.
- Dahlberg, I. (1978b). Fundamentos teórico-conceituais da classificação. *Revista de Biblioteconomia de Brasília*. Brasília, 6(1), p. 9-21, jan./jul.
- Dahlberg, I. (2006). Knowledge organization: A new science?. *Knowledge Organization*, 33, p. 11-19.

- De Wall, C. (2007). *Sobre Pragmatismo*. Trad. Cassiano Terra Rodrigues. São Paulo: Edições Loyola.
- Faria, M. I.; Pericão, M. G. (2008). *Dicionário do Livro: da escrita ao livro electrónico*. São Paulo: Almedina.
- Francelin, M. M. & Kobashi, (2012). N. Y. Concepções sobre o conceito na organização da informação e do conhecimento. *Ciência da Informação, [S. l.]*, 40(2). Disponível em: <http://revista.ibict.br/ciinf/article/view/1311>. Acesso em: 21 ago. 2021.
- Hjørland, B. (2002). Domain analysis in information science: Eleven approaches—traditional as well as innovative. *Journal of Documentation*, 58(4), p. 422-462, ago. 2002.
- Hjørland, B. (2009). Concept Theory. *Journal of the American Society for Information Science and Technology*, 60(8), p. 1519-1536. DOI: <https://doi.org/10.1002/asi.21082>.
- Izquierdo Arroyo, J. M. (1993). De la semiótica del discurso a la semiótica documental. En: Moreno González, J. A. (1993). *Aplicación de las ciencias del texto al resumen documental*. Madrid: Universidad Carlos III. 199-216.
- Izquierdo Arroyo, J. M. (1990). *Esquemas de lingüística documental*. Barcelona: Promociones y Publicaciones Universitarias. Tomo I.
- Ibri, I. A. (2015). *Kósmos Noetós: a arquitetura metafísica de Charles S. Peirce*. São Paulo: Paulus.
- Ibri, I. A. (2020). Introdução. In: *Semiótica e pragmatismo: interfaces teóricas*: vol. I [online]. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica; FiloCzar, p. 23-27, 2020.
- Koselleck, R. (2020). *História dos conceitos: estudos sobre a semântica e a pragmática da linguagem política e social*. Rio de Janeiro: Contraponto.
- Koselleck, R. (1992). Uma história dos conceitos: problemas teóricos e práticos. Trad. Manoel Luís Salgado Guimarães. *Estudos Históricos*, Rio de Janeiro, 5(10), p. 134-146.
- Moron, R. A, Almeida, C., & Moreira, W. (2013). *Diretrizes para análise conceitual. Como perspectivas de Hjørland, Dahlberg e Lakoff*.
- Nöth, W.; Santaella, L. (2017). *Introdução à Semiótica: passo a passo para compreender os signos e a significação*. São Paulo: Paulus, 2017.
- Peirce, C. S. (2017). *Semiótica*. Trad. José Teixeira Coelho Neto. 4.^a ed. São Paulo: Perspectiva.
- Peirce, C. S. (2000b) *O que é o pragmatismo*. Tradução J. T. C. Neto. In: Guinsburg, J. (Ed.). *Semiótica*. São Paulo: Perspectiva, 2000b [1905]. p. 283-99. (Coleção Estudos, 46).
- Peirce, C. S. (1994). *The collected papers of Charles Sanders Peirce*. Ed. C. Hartshorne and P. Weiss. Edição eletrônica reproduzindo os seis primeiros volumes. Cambridge: Harvard University Press, (1866-1913). (citados aqui como CP, seguido pelos números referentes a volume e parágrafo).
- Peirce, C. S. (1993). *Semiótica e Filosofia*. Trad. Octanny Silveira da Mota e Leonidas Hegenberg. São Paulo: Editora Cultrix.
- Santaella, L. (1983). *O que é semiótica*. São Paulo: Brasiliense.
- Silveira, L. F. B. (2007). *Curso de Semiótica geral*. São Paulo: Quartier Latin.

(Página deixada propositadamente em branco)

CUNHO EVIDENCIAL E INDICIAL DA FOTOGRAFIA NA ORGANIZAÇÃO DO CONHECIMENTO ARQUIVÍSTICO

Bruno Henrique Machado¹, Rafael Semidão²,
Telma Campanha de Carvalho Madio³

¹ Universidade Estadual Paulista (UNESP), Faculdade de Filosofia e Ciências, Marília, ORCID iD <https://orcid.org/0000-0002-3034-3122>, machadobrunohenrique@gmail.com

² Universidade Federal do Rio Grande (FURG), Rio Grande.

ORCID iD <https://orcid.org/0000-0002-9715-3702>, rafaelsemidao@gmail.com

³ Universidade Estadual Paulista (UNESP), Faculdade de Filosofia e Ciências, Marília, ORCID iD <https://orcid.org/0000-0002-7031-2371>, telma.madio@unesp.br

Resumo

O trabalho figura como uma reflexão de natureza qualitativa e exploratória que tem por objetivo apresentar uma abordagem dual a respeito da representação das fotografias institucionais, enquanto documentos de arquivo, considerando-as como evidência contextual decorrente das funções e atividades de uma instituição e as imagens que veiculam como indícios informativos a constarem nos instrumentos de representação, a partir de valores de credibilidade. Dessa maneira, apresentamos a fotografia como imbuída de capacidade comunicativa na relação entre o emissor e o receptor, assegurando que a comunicação e a circulação da informação orgânica registrada no contexto institucional sejam efetivadas a partir de medidas de representação e organização do conhecimento arquivístico. Acerca da abordagem do método, foi realizado um levantamento bibliográfico sobre tópicos que incidam sobre a temática desenvolvida. Em seguida foi desenvolvida uma reflexão qualitativa de natureza exploratória que buscou, à guisa de ensaio, contribuir para novas reflexões a respeito da fotografia compreendida como documento de arquivo. Dessa maneira, refletimos sobre uma abordagem dual, em que a razão institucional inicial da produção da fotografia é representada através da manutenção do contexto de produção dos documentos nos instrumentos de representação e organização, respeitando-se a sua proveniência, e a utilização da imagem em si, veiculada pelo registro fotográfico, figura na condição de indício a complementar informativamente a manutenção e compreensão do referido contexto. De modo concreto, propomos que ao ser representada e organizada como documento de arquivo da esfera institucional, a fotografia receba, por exemplo nos instrumentos de pesquisa, como catálogos, uma representação dual em que a descrição do processo de evidenciação da mesma a partir do desempenho de funções e atividades seja dotada de credibilidade como sendo a razão inicial que levou a produção da fotografia; e que a imagem em si, veiculada no registro fotográfico, seja representada como dotada da credibilidade de um indício que complementa informativamente a representação do processo de evidenciação. Compreendemos que nessa abordagem dual que coaduna a representação do pro-

cesso de evidenciação e da imagem, a partir de uma distinção de níveis de credibilidade, a representação e organização da fotografia estariam amparadas em uma acepção ética que procura oferecer ao usuário a consciência do real motivo que levou à produção da fotografia e que dá contexto à sua imagem em uma época em que imagens e demais elementos informacionais estão à mercê de distorções e mal-entendidos a todo instante.

Palavras-chave: *fotografia, evidência, documento de arquivo, organização do conhecimento arquivístico.*

Introdução

A fotografia, enquanto um registro documental e desde a sua divulgação oficial em 1839 com a técnica viabilizada pelo daguerreótipo, tem despertado diversas expectativas quanto à sua utilização na evidenciação documental de fenômenos das esferas científica, natural, social e institucional. Diante disso, nos limitaremos, neste momento, a aprofundar nossa pesquisa na compreensão da fotografia em um ambiente institucional, sabendo o quanto é difícil limitá-la a apenas um propósito, visto que é um documento que possui diversas nuances, sobretudo visuais. Conforme os ideais de Paul Otlet, citados por Guerra e Pinheiro (2009), a fotografia é o documento que melhor representa o conhecimento humano, por ser o mais próximo da realidade. Neste contexto, atribuímos à fotografia um valor de documento, criado por um processo de trabalho referente a uma instituição no desempenho de suas funções e atividades, ou seja, estamos analisando a fotografia como documento de arquivo, produzido e acumulado pela instituição em suas rotinas administrativas, definidas e desencadeadas pela missão institucional. (Machado; Semidão, Madio, 2017).

Tais documentos orgânicos fotográficos são apresentados sob a perspectiva da *Archival knowledge organization* (Guimarães; Tognoli, 2015) que é compreendida como todo o processamento teórico e técnico aplicado aos documentos de arquivo, desde a sua criação até sua destinação final, ou seja, desde uma concepção de processo documental de evidenciação que contextualiza a fotografia junto à gama de funções e atividades da instituição produtora/acumuladora, e que compreende a imagem registrada em si como indício que agrega e complementa esse processo documental.

A perspectiva apresentada foi construída com base em uma reflexão qualitativa, de natureza exploratória, que propõe uma abordagem dual de valor de credibilidade dos elementos de representação das fotografias, informando a autenticidade a partir da consideração da evidência contextual e considerando o fator indicial veiculado na imagem como aspectos informativos a constarem nos instrumentos de representação. Assim sendo, neste texto não iremos apresentar as abordagens metodológicas aplicadas para leitura de fotografias com o objetivo de indexação, como as encontradas em Smit (1997); Smit (1998); Manini (2002); Boccato e Fujita (2006); entre outros, mas sim, abordaremos questões pertinentes ao tema, justificando o quanto é oportuno a compreensão das características singulares desse documento no ambiente da Organização do Conhecimento Arquivístico.

Metodologia

Em termos de caracterização metodológica, o trabalho se apresenta como uma reflexão qualitativa dotada de uma natureza exploratória (Prodanov; Freitas, 2013, p. 51-52), que visa, à guisa de ensaio, estabelecer um nexos entre o atributo de evidencição contextual e do índice da imagem fotográfica, enquanto registro documental – documento de arquivo no âmbito do tratamento técnico dispensado a este documento na Organização do Conhecimento Arquivístico proposta por Guimarães e Tognoli (2015). Dessa maneira, o texto não tem a pretensão de encerrar as discussões acerca do tema, mas pretende contribuir para novos aportes reflexivos a respeito da fotografia como documento de arquivo no que tangencia a compreensão e do entendimento deste documento dentro de qualquer instituição públicas ou privadas ou pessoais.

A fotografia na organização conhecimento arquivístico: uma abordagem evidencial e indicial

Iniciamos pelas palavras de Sontag (2004, p. 172) para quem a fotografia é, “de várias maneiras, uma aquisição”. Atualmente, a variedade de movimentos de circulação das imagens fotográficas, e conseqüente variedade de usos, são incalculáveis. Baudelaire já afirmava, acerca dos usos da fotografia, no “Salão de 1859”, que o “verdadeiro dever” da fotografia consistia em ser serva das artes e das ciências (Baudelaire, 1972). Temos, nesse sentido, uma fotografia como forma de lembrança vicária de uma pessoa, de um momento ou de um objeto querido e, além disso, somos consumidores das imagens fotográficas pelo seu poder de suposta objetividade e reprodutibilidade que proporciona a apropriação de informações e conhecimentos na maioria das vezes dissociados da experiência. Assim, em termos de capilarização e diversidade de usos, a “fotografia tem pouco ou nenhum valor como propriedade, porque não traz em si o valor da raridade” (Berger, 2017, p. 37).

A técnica da fotografia foi sendo aprimorada com o passar dos anos, desde o seu anúncio e, com efeito, a sua concepção, no imaginário geral, passou a ter um “[...] predomínio [da ideia] de cópia do real e seu caráter de representação, com códigos próprios de construção da imagem.” (Madio, 2016, p. 27).

A fotografia correlacionada ao desenvolvimento industrial, do comércio, a descoberta de agentes químicos e as mudanças nos padrões de produção e consumo, estava pronta para uma nova fase de expansão, seja na publicidade, no jornalismo ou em qualquer outra atividade, conforme pontuou Sontag (2004), Tagg (2005), e Rouillé (2009). Dessa maneira, sua objetividade confere, de forma unânime, um papel determinante como documento desde o seu advento” (Guerra; Pinheiro, 2009, p. 2).

A partir da fotografia, as linguagens começaram a ser produzidas pela mediação de aparelhos considerados instrumentos inteligentes e portadores de uma nova inteligência visual. Depois da fotografia, é o cérebro, no seu papel de fonte produtora de linguagens, que começa a ser extra-somatizada. “O cérebro começa, digamos assim, a crescer fora do corpo.” (Santaella, 1996, p. 319).

Assim, a tecnologia da fotografia superou todas as outras maneiras de representação em imagens, ela retirou o estigma da presença do criador de imagem, desse modo, a fotografia não é dependente do criador, com caráter subjetivo, pois é a máquina que executa a captura da imagem, anteriormente com processo óptico-químico, atualmente no meio eletrônico. “A fotografia diferencia-se dos outros tipos de documentos pelo fato de fornecer um testemunho, um recorte do mundo, mais do que evidenciar uma expressão de caráter artístico ou ser propriamente uma interpretação da realidade”. (Sontag, 2004, p. 14-15).

Por sua vez, Barthes (1984, p. 15) observa que a fotografia figura como uma emanção do referente, ou seja, “a fotografia traz sempre consigo seu referente”. Nestes termos, há uma necessidade da conexão de uma existência, “um sentimento de realidade incontornável” (Dubois, 2012, p. 26). Esta afirmação foi justificada por Santaella (2005) ao comentar as três abordagens de argumento proposta por Peirce, sendo elas: primeiridade, secundidade e terceiridade. A autora postula que a fotografia é tipicamente da secundidade, pois “tem uma vocação referencial, o que a categoriza como signo indicial” (Santaella, 2005, p. 15). Assim, “o poder do índice para funcionar [...] vem da conexão com o objeto que o afeta e determina, a função característica do índice” (Santaella, 2005, p. 197).

Nesse sentido, o aspecto indicial, concepção peirceana relacionada à fotografia, em determinadas situações faz com que ocorra a introdução de diversos significados, sendo capaz de abarcar em si uma evidência comunicativa, porém, a imagem fotográfica não deverá ser analisada de maneira análoga a sua apresentação, pois a mesma é composta por “generalidades e discursos desenvolvidos, fruto de um desconhecimento evidente, e, algumas vezes, entusiasticamente reivindicados” (Rouillé, 2009, p. 16).

Dessa maneira, ao analisar a proposta do objeto à imagem, segundo o autor, o movimento anti representativo não sacrifica a imagem fotográfica a função do referente, mas reconhece que a fotografia inventa mundos. (Rouillé, 2009). Dito isto, a fotografia, “como produto humano, cria também com esses dados luminosos de uma realidade que não existe fora dela, nem antes dela, mas precisamente nela”. (Machado, 2015, p. 45).

Assim, é necessário compreender que a fotografia é uma construção humana, e que desde a sua produção, que inclui a imagem fixada na veiculação do documento, “é preciso dizer, de novo, que a fotografia-documento não garante relação direta – nem mesmo relações reduzidas ou transparentes – entre imagens e as coisas. É isso que é acentuado pelas imagens das imagens, as representações, as falsificações (efeitos especiais), etc.” (Rouillé, 2009, p. 158). Nesse sentido, Burke (2001, p. 13) pondera que “para não sermos enganados por fotografias, sejam fixas ou móveis, precisamos assim como no caso dos textos prestar atenção à mensagem e ao remetente, perguntando quem está tentando nos dizer o quê e por que motivos.” Por isso é justificada a preocupação de Kossoy (1993, p. 14) ao afirmar que “as fotografias não podem ser aceitas imediatamente como espelhos fiéis dos fatos. Assim como os demais documentos elas são plenas de ambiguidades, portadoras de significados não explícitos e de omissões pensadas, calculadas, que aguardam pela competente decifração”

Em termos de variedade de usos possíveis, é importante afirmar que “[...] a fotografia se define e se desdobra em múltiplas funções e expressões, seja como signo,

estética da fotografia e/ou documento fotográfico, tanto no sentido da prova, seja ela real ou fabricada” (Lima; Murguia, 2008, p. 4), revelando com isso, a diversidade quase incalculável de usos, dependendo da razão inicial que conduziu o registro.

Nestes termos, a imagem fotográfica, em sua base tecnológica, é considerada a essência do registro, o qual, em seu contexto social-ideológico, remete a um discurso a ser instaurado e contextualizado em uma perspectiva que procure, de certo modo, sustentar o entendimento global da fotografia na instituição. Neste contexto, aprofundamos no entendimento da constituição da imagem fotográfica institucional sob dois aspectos, da indicial da imagem fotográfica e a evidenciação contextualizada.

Diante dessas considerações acerca do caráter alusivo e indicial da fotografia, gostaríamos de perspectiva-lá desde o ponto de vista da Organização do Conhecimento Arquivístico, refletindo sobre a possibilidade de uma compreensão dual de tratamento documental que reconheça e utilize o fator indicial da imagem e o processo de evidenciação documental da fotografia enquanto resultante do exercício das funções e atividades de uma instituição, isto é, da sua contextualização a partir de sua proveniência.

Dessa maneira, sob a perspectiva da avaliação da informação, Rockembach (2015) aborda as noções de indício, evidência e prova, enquanto aspectos apresentados pela informação a ser tratada. Acerca do indício, o referido autor afirma que

São levados em consideração pistas, sinais ou vestígios, que auxiliam na reconstrução do passado. O indício também possui características não-intencionais, pois estes vestígios não são produzidos com a intencionalidade comum ao registro/documento, são inferências realizadas a partir de uma análise contextual do ambiente da produção desta informação (Rockembach, 2015, p. 94).

Já quanto à evidência, Rockembach (2015) a conecta ao conhecimento, à ação de conhecer algo pela captação de linguagem pelos sentidos, sendo os que indícios, em contrapartida, permitiriam apenas a inferência. E sobre a prova, o autor postula que a mesma “[...] nasce da necessidade de legitimação de determinada informação ou conjunto de evidências, e isto não está livre da intencionalidade do sujeito” (Rockembach, 2015, p. 94). Desse modo, o autor compreende a relação complementar entre as três noções como estando posta na forma de círculos concêntricos, na reunião dos quais o indício ocupa o círculo mais externo e tem como atributos informacionais as ações cognitivas de suspeitar, inferir e pressupor; enquanto a evidência apresenta os atributos de mostrar e ostentar; e a prova, por sua vez, os atributos de contar e ilustrar (Rockembach, 2015, p. 95).

Desta forma, tomando por base as reflexões do autor supracitado, compreendemos o indício como característica de um documento, ou de parte de um documento, que conduz à inferência daquilo (fenômeno, objeto e, sobretudo, ação) que nele foi registrado. Entendemos a evidência como característica documental que leva ao conhecimento do que foi registrado e de seu contexto e, por fim, entendemos a prova como característica, baseada em indícios e evidências, que levam ao reconhecimento do que foi consignado no documento. No contexto deste trabalho, nos ateremos às características de evidência para pensá-las em sua integridade nos processos de

representação do conhecimento arquivístico, por exemplo, como a descrição. Como visto em parágrafos anteriores, a imagem formada pela fotografia não tem a chancela para constar como expressão direta e fiel de uma realidade concreta. Essa situação pode ser, hoje em dia, facilmente verificada pela abundância de casos de manipulação e de *fake news* envolvendo fotografias. Contudo, é importante afirmar que desde os primeiros usos da fotografia a manipulação da imagem se fazia presente.

A respeito da compreensão da fotografia enquanto documento de arquivo, Rezende e Lopez (2014, p.03, tradução nossa) afirmam que a “entrada tardia da fotografia no universo das administrações públicas, como forma de registro de funções e atividades, fez com que fossem: a) produzidas sem configurações definidas legal ou administrativamente; b) utilizadas de forma autônoma e menos controlada do que os documentos tradicionais; c) acumuladas à parte dos demais documentos, tanto em termos físicos como lógicos. Além disso, as imagens fotográficas, ao contrário de muitos documentos textuais, geralmente não apresentam dados explícitos sobre seu contexto de produção”. Essa última situação mencionada pelos autores vem de encontro à proposta aqui aventada, de forma que o contexto de produção, nos procedimentos de representação, necessita ser evidenciado e isso somente ocorre, no caso da fotografia, quando o processo de evidenciação é retomado desde o registro, com a constituição da imagem, até às funções e atividades que deram início ao mesmo. Já a imagem em si, também ao participar dos procedimentos de representação, pode apenas complementar, à guisa de índice, a retomada do processo de evidenciação.

Em termos de representação em instrumentos de pesquisa descritivos, Rezende e Lopez (2014, p.07), com base na Norma Geral Internacional de Descrição Arquivística – ISAD-G, propõem elementos de descrição específicos para a fotografia. Dentre esses elementos, os referidos autores apontam, junto à área de conteúdo e estrutura, os seguintes elementos: “âmbito e conteúdo”, “marcas e/ou anotações”, “lugar fotografado”, e “gênero fotográficos” (Rezende & Lopez 2014, p.07).

Todos esses elementos propostos contribuem com um melhor esclarecimento acerca do contexto da imagem registrada. Em nossa perspectiva, caberia, nessa área da norma, também relacionar esse contexto da imagem com a vinculação da fotografia ao seu contexto de produção em um elemento de descrição que informasse a proveniência da mesma e, ao mesmo tempo, declarasse a credibilidade indicial do contexto da imagem, credibilidade que complementar, em termos informativos e de esclarecimento, a credibilidade evidencial do contexto de produção representado em outras áreas da norma, nos instrumentos de pesquisa.

Retomando o já mencionado caráter alusivo e indicial da fotografia e o fato de ela, em sua imagem, ser sempre uma criação humana que envolve uma miríade de aspectos, sobretudo linguístico-discursivos, nossa proposta de dupla credibilidade na representação da fotografia institucional como documento de arquivo envolveria um tipo de intervenção que abrange fatores éticos ao explicitar nos instrumentos de pesquisa, baseados na ISAD-G, como no exemplo dado acima, ou em outros padrões, que a representação com valor evidencial é aquela que informa o contexto de produção no processo de evidenciação e que a representação da imagem em si, por mais contextualizada que seja, tem o valor de índice que pode, complementarmente, conduzir ao contexto de produção.

ÁREA DE CONTENIDO Y ESTRUCTURA		
Alcance y contenido	Alcance y contenido	Apuntar informaciones relevantes sobre el contexto registrado en la imagen. Puede incluir nombres de personas, lugares (cuando sean diferente(s) del/los registrado(s) en "lugar fotografiado"), además de otras características significativas del contenido.
	Marcas y/o anotaciones	Registrar cualquier información apuntada (o fijada por otros medios) en el documento original, sea en el anverso, en el verso o en otro soporte de fijación de la fotografía (tarjeta, álbum, etc.); en el caso de los documentos digitales, se deben registrar las informaciones colocadas para la identificación del fichero y/o en sus metadatos. La información debe ser integralmente transcrita, en el idioma original. También se pueden describir otros aspectos relevantes, por ejemplo: información manuscrita o impresa, tipo y color de la tinta, ubicación en el documento original, etc.
	Sitio fotografiado	Identificar el espacio físico retratado.
	Género fotográfico	Atribuir un género que describa la configuración de los elementos retratados por la imagen fotográfica (ejemplos: (fotografía documental, retrato, retrato con paisaje, paisaje, etc.).

Figura 1. Recorte de Rezende & Lopez (2014, p. 07).

Em última análise, a proposta aqui apresentada consiste em observar o princípio da proveniência como “meio pragmático de lutar contra injustiças e *fake news* no contexto da organização do conhecimento arquivístico” (Machado et. al. 2020, p. 557, tradução nossa), oferecendo aos usuários dos instrumentos de representação de fotografias as condições para saber a motivação inicial que levou ao seu registro dentro do seu contexto de produção e uma margem de uso da imagem em si dentro de um domínio de significado restringido pelo caráter evidencial do contexto de produção e complementado pelo fator indicial da imagem veiculada.

Conclusões

Nesse sentido, nossa proposta de tratamento arquivístico na representação do conhecimento de documentos fotográficos institucionais consiste, em um primeiro momento, em uma compreensão ética. Essa compreensão assenta suas bases sobre o que optamos por chamar de valor de credibilidade contextual na representação do conhecimento arquivístico de fotografias.

Propomos que, ao representar/descrever fotografias, haja, nos instrumentos de representação, uma explícita divisão entre os elementos que descrevem a fotografia em seu contexto de produção e elementos que a descrevem de forma indicial, a partir da imagem em si. Os elementos que descrevem os meandros do contexto de produção da fotografia, de acordo com a proveniência, teriam um valor evidencial de credibilidade contextual que, ademais, também informa a autenticidade dos documentos relacionados e posicionados no contexto, e os elementos que descrevem apenas a imagem teriam um valor de indício que apenas poderia levar à inferência.

O ponto ético da perspectiva estaria em informar ao usuário quais os pesos de credibilidade que cada elemento de descrição teria, deixando a ele a opção de livre uso da representação feita. A evidência alcançada refere-se apenas ao contexto, enquanto que a representação da imagem figura como indício, sendo complementar. Complementar ao conhecimento obtido pela representação da evidenciação do contexto de produção, uma vez que a imagem não poderia ser tomada como expressão direta e comprovação de uma realidade; mas poderia, com base em interpretações, gerar informações que nutram o conhecimento do contexto. Como exemplo, além do mencionado, poderíamos pensar em uma descrição multinível em que junto aos elementos que representam o contexto, fossem inseridos elementos que representem a imagem em si mesma, mas com o aviso de que aquele elemento tem uma credibilidade indicial e não uma credibilidade evidencial, pois não pode ser relacionado organicamente às funções e atividades da instituição produtora. Uma opção mais específica seria, conforme visto, duplicar os elementos que descrevem o âmbito e conteúdo da fotografia, de modo que uma parte informaria o conteúdo relacionado ao contexto de produção e outra a imagem com seus indícios complementares ao referido contexto.

Referências

- Barthes, R. (1984) *A câmara clara*, nota sobre fotografia. Rio de Janeiro: Nova Fronteira.
- Baudelaire, C. (1972) *Selected writings on arts and artists*. Cambridge: Cambridge University Press.
- Berger, J. (2017) *Para entender uma fotografia*. São Paulo: Companhia das Letras.
- Boccatto, V. R. C.; Fujita, M. S. L. (2006) Discutindo a análise documental de fotografias: uma síntese bibliográfica. *Cadernos BAD (Portugual)*, n. 2, 2006. <http://hdl.handle.net/20.500.11959/brapci/82351>.
- Burke, P. (2001) Como confiar em fotografias. *Folha de São Paulo*, São Paulo, 04 de fevereiro de 2001, Disponível em: <https://www1.folha.uol.com.br/fsp/mais/fs0402200105.htm>
- Guerra, C. B. Pinheiro, L. V. R. (2009) *A imagem fotográfica como documento: desideratos de Oilet*. Encontro Nacional de Pesquisa em Ciência da Informação, 10., 2009, João Pessoa. Anais do X ENANCIB, João Pessoa: UFPB, 2009. <https://ridi.ibict.br/handle/123456789/69>
- Prodanov, C. C.; Freitas, E. C. de. (2013) *Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico*. Novo Hamburgo: Feevale.
- Dubois, P. (2012) *O ato fotográfico e outros ensaios*. 14. ed. Campinas: Papirus.
- Guimarães, J. A. C.; Tognoli, N. B. (2015) Provenance as a Domain Analysis Approach in Archival Knowledge Organization. *Knowledge Organization*, v. 42, n. 8, p. 562-569. <https://repositorio.unesp.br/handle/11449/164830>
- Kossov, B. (1993) Estética, memória e ideologia fotográficas: decifrando a realidade interior das imagens do passado. *Acervo*, Rio de Janeiro, v. 6, n.1/2, p. 13-24.
- Lima, M. L.; Murguia, E. I. (2008) *Fotografia e Informação*. Encontro Nacional de Pesquisa em Ciência da Informação, 9, 2008. São Paulo, Anais do IX ENANCIB, São Paulo, USP, 2008.
- Machado, A. (2015) *A ilusão especular: uma teoria da fotografia*. São Paulo: Editora Gustavo Gilli, 2015.
- Machado, B. H.; Semidão, R.; Madio, T. C. C. (2017) Reflexões sobre uma organização do conhecimento fotográfico segundo seu contexto de produção documental. *Tendências atuais e perspectivas futuras em organização do conhecimento: atas do III Congresso*

- ISKO Espanha e Portugal – XIII Congresso ISKO Espanha.* (coord.) Simões, Maria da Graça Melo, Borges, Maria Manuel, Coimbra, Universidade de Coimbra, 2017. http://sci.uc.pt/eventos/atas/comunicacoes/isko2017/isko2017_1020_1028.pdf
- Machado, B. H.; Semidao, R.; Madio, T. C. C.; Martinez-Avila, D. (2020). Provenance as an Ethical Measure for the Archival Knowledge Organization of Photographs. In: Knowledge Organization at the Interface. (Org.). *Advances in Knowledge Organization*, 1.^a ed.: Ergon, v. 17, p. 557-558.
- Madio, T. C. de C. (2016) *Documento de Arquivo: fotografia*. Tese (Livre-Docência) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Filosofia e Ciências.
- Manini, M. P. (2002) *Análise documental de fotografias: um referencial de leitura de imagens fotográficas para fins documentários*. Tese (Doutorado) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo.
- Santaella, L. (1996) *Produção de linguagem e ideologia*. 2 ed. rev. ampl. São Paulo: Cortez.
- Santaella, L. (2005) *Matrizes da linguagem e pensamento: sonora visual verbal: aplicações na hipermídia*. São Paulo: Iluminuras.
- Sontag, S. (2004). *Sobre fotografia*. São Paulo: Companhia das Letras.
- Smit, J. W. (1996) A representação da imagem. *Informare*, Rio de Janeiro, v. 2, n. 2, p. 28-36.
- Smit, J. W. (1997) *Propostas para a indexação de informação iconográfica*, (Mimeo).
- Rezende, D. S.; Lopez, A. P. A. (2014). Adecuación de la descripción archivística de documentos fotográficos a los estándares internacionales. 2.^a *Conferencia Anual de Archivos: Archivos e Industrias Culturales*, 2014, Girona. Papers. Girona: Ajuntament de Girona. <https://www.girona.cat/sgdap/docs/qo4xhr0id164.pdf>
- Rockembach, M. (2015) Evidência da informação no contexto dos arquivos digitais. *Ponto de Acesso*, Salvador, v. 9, n. 2, p. 50-64. <https://periodicos.ufba.br/index.php/revistaici/article/view/12258/9840>
- Rouillé, A. (2009). *A fotografia: entre documento e arte contemporânea*. São Paulo: Editora Senac São Paulo.

(Página deixada propositadamente em branco)

**AFIRMANDO PEREIRA:
A INTEGRATIVE LEVELS CLASSIFICATION
NA REPRESENTAÇÃO DO ASSUNTO FICCIONAL**

Patrícia de Almeida¹, Claudio Gnoli²

¹Integrative Levels Classification research project, Itália, mebpatria@gmail.com,
ORCID iD 0000-0002-8061-8622

²Integrative Levels Classification research project, Itália, claudio.gnoli@unipv.it,
ORCID iD 0000-0002-4721-7448

Resumo

As classificações bibliográficas do conhecimento com base no assunto facilitam não só a organização dos serviços, mas também o acesso e a recuperação da informação aos seus utilizadores. No entanto, no que toca à Literatura (ficção), o acesso por assunto ao documento encontra-se comprometido, uma vez que, no caso dos documentos de cariz ficcional, a representação privilegia a forma (género literário), em detrimento do conteúdo temático e intelectual. No final do século XX e início do século XXI, a representação do assunto ficcional despertou o interesse de diversos investigadores e tem sido estudada na Organização do Conhecimento, sem que, genericamente, a prática tradicional classificatória se altere.

A *Integrative Levels Classification (ILC)* é um sistema classificatório facetado assente em fenómenos, que, sem o espartilho das tradicionais áreas do conhecimento, poderá propiciar uma melhor representação do assunto literário. Neste contexto, pretende-se atestar o carácter distintivo e as potencialidades da ILC, no âmbito da representação do assunto da ficção. Para tal, realiza-se um estudo exploratório, comparativo em distintos sistemas de Organização do Conhecimento (*Dewey Decimal Classification*, *Universal Decimal Classification* e *Library of Congress Classification / Library of Congress Subject Headings*), relativo ao romance italiano *Sostiene Pereira* [Afirma Pereira] do escritor Antonio Tabucchi.

Os resultados mostram que, nos sistemas de organização do conhecimento em análise, esta obra literária se encontra classificada numa perspetiva tradicional e que a ILC permite não só uma melhor representação do seu conteúdo intelectual, mas também a expressão de relações entre documentos que abordem as mesmas temáticas. Paralelamente, e no que toca aos elementos formais de assunto, verifica-se que não ocorre perda significativa de informação. Assim, conclui-se que a ILC se destaca pelo seu maior potencial informativo na representação do assunto literário e ficcional, sobressaindo uma maior versatilidade e um aprofundamento da relação entre os conceitos. Neste sentido, considera-se que a representação proporcionada pela ILC facilita o acesso e a recuperação da informação de cariz ficcional aos utilizadores deste sistema, bem como reforça a exequibilidade da representação do conteúdo temático literário.

Palavras-chave: Classificação de ficção, representação de assunto, assunto ficcional, *Integrative Levels Classification (ILC)*.

Introdução

Em qualquer unidade de informação, as classificações bibliográficas do conhecimento com base no assunto facilitam não só a organização dos serviços, mas também o acesso e a recuperação da informação aos seus utilizadores. Todavia, na tradição da organização do conhecimento, a representação de documentos de cariz literário e ficcional segue princípios distintos daqueles utilizados para os documentos de cariz mais técnico-científico. No que toca à Literatura, o acesso por assunto ao documento encontra-se comprometido, como se constata nas principais estruturas classificatórias e como tem sido apontado na literatura científica.

Muito embora sejam consideradas a primeira e mais antiga linguagem de recuperação da informação (Barbosa, 1972), a verdade é que, nesta área específica do conhecimento, os sistemas classificatórios não se alteraram de forma significativa, mantendo-se a organização da ficção por idioma, género e tempo (forma), ao invés de se socorrerem de tópicos de assunto (conteúdo intelectual). No final do século XX e início do século XXI, a representação do assunto ficcional (re)despertou o interesse e tem sido estudada na organização do conhecimento por diversos investigadores, sobressaindo as consequências e, principalmente, as alternativas à prática tradicional. A classificação de ficção é profundamente estudada no final do século XX, em particular por Annelise Mark Pejtersen (1979) e por Clare Beghtol (1994). No entanto, até à atualidade e no que toca à classificação bibliográfica da Literatura, pouco parece ter-se alterado nos esquemas tradicionais.

Em relação à representação do assunto, as estruturas classificatórias de natureza facetada mostram-se potencialmente mais expressivas. Nesta linha, surge a *Integrative Levels Classification (ILC)*, um sistema que assenta em fenómenos e não nas tradicionais áreas do conhecimento, com potencial para propiciar uma melhor representação do assunto literário. É neste contexto teórico que se desenrola a presente investigação, que inicia com uma breve revisão sobre a representação dos documentos literários e ficcionais e sobre a *Integrative Levels Classification (ILC)*. A fim de se atestar o caráter distintivo e as potencialidades deste sistema de organização do conhecimento, realiza-se um estudo de caso, exploratório e de índole comparativa, tendo como amostra o romance italiano *Sostiene Pereira* [*Afirma Pereira*] de Antonio Tabucchi, e discutem-se os resultados. O trabalho termina com algumas considerações sobre o tópico em análise, que se desejam como mais um contributo para a alteração do paradigma tradicional da representação do assunto literário e ficcional.

Representação do assunto literário e ficcional

A literatura científica reconhece que os sistemas de classificação bibliográfica tradicionais se desenvolvem especialmente para os documentos de natureza técnico-científica e que não refletem a diversidade e a riqueza dos documentos de ficção (Hypén, 2014). Segundo Beghtol (1989), isto sucede, porque a ciência e a tecnologia monopolizam a atenção dos classificacionistas, tanto a nível teórico como a nível do desenvolvimento dos próprios sistemas. Efetivamente, esta atração pela ciência

nasce de uma moderna preocupação com o pensamento científico e de uma forte percepção sobre a necessidade de acesso aos documentos científicos, também referida em García-Marco, Moraes, García-Marco e Guimarães (2010). Beghtol (1994) considera que as classificações bibliográficas para as áreas relativas às Humanidades não se mostram tão desenvolvidas como para as áreas mais científico-técnicas e que, em particular, os sistemas não se encontram adaptados para a representação do assunto literário e ficcional. Este paradigma não parece ter sofrido alterações nos últimos anos. Ainda recentemente, Ward e Saarti (2018) referem que a classificação de documentos literários ficcionais parece causar alguma consternação ou estranheza, quando comparada com a classificação bibliográfica da não ficção.

De uma forma genérica, constata-se que os sistemas classificatórios de tipo enumerativo não representam o assunto dos documentos literários, de uma forma suficientemente específica e equitativa face aos documentos não-literários. Tradicionalmente, estes sistemas colocam o género literário enquanto unidade distintiva base da classificação, o que, por um lado, nada diz sobre o conteúdo intelectual da obra (assunto) e, por outro, colide com potenciais problemas teóricos relativos à identificação e definição de géneros literários, a atestar nos Estudos Literários. Em Rafferty (2013), confirma-se que o género literário constitui a base epistemológica para os sistemas de recuperação da informação ficcional e serve o princípio de categorização de Literatura nos tradicionais sistemas de organização do conhecimento, obtendo, por isso, um valor funcional. Embora se valorize o carácter pragmático e operativo desta opção, concorda-se com a autora e admite-se que o conceito de *género* é instável e discutível. Sendo reconhecido que o assunto literário se mostra complexo ou mais especializado, considera-se provável a inadequação da atual base teórica de representação. Com efeito, poderá afirmar-se que as classificações tradicionais nem sempre se mostram adequadas às necessidades, tendo como principais inconvenientes a dificuldade de atualização e de expressão de assuntos complexos, a rigidez e a obrigatoriedade de encaminhamento de assuntos por aproximação (Chaumier, 1988).

Desta feita, considera-se que *autor, período, língua e nacionalidade* constituem informações relevantes, contudo a um nível pouco profundo e insuficientemente distintivo. Os comuns qualificadores proporcionam, de facto, uma noção metatópica de identidade autoral e cultural, porém demasiado genérica, para que se consiga a devida representação e futura recuperação documental. Por exemplo, um assunto representado como *Romance português – século XX – José Saramago* funcionará como um metatópico de alcance informativo muito limitado, por não permitir a representação distintiva entre Romances do mesmo autor e por nada dizer sobre o conteúdo intelectual de uma determinada obra literária. Pensando na disposição física em estantes, verifica-se que dois documentos de natureza ficcional surgem lado a lado e nada têm em comum no que toca ao seu assunto, o que leva a questionar até que ponto esta organização seja proveitosa para os utilizadores. Já relativamente ao acesso à informação e em situação de pesquisa nos catálogos, para um assunto como *ficção francesa ou romance português*, obter-se-ão centenas de documentos, o que impossibilita uma busca e recuperação da informação fácil e rápida. Recorde-se que, desde há muito, se estabeleceu a lei que determina “poupar tempo ao leitor” (Ranganathan, 1963).

A pouca utilidade deste tipo de representação do assunto para os utilizadores dos serviços encontra registo na literatura científica, nomeadamente em Eriksson (2005), que exemplifica:

If a reader wants, for example, a French novel written in the first half of the 19th century, the class number for those novels will probably include way too many titles to be much of a help, and anyway the reader is seldom looking for a book described in so general terms. And if, more commonly, the reader do not care where or when the book was written but just wants a novel about a certain subject matter, for example the Napoleonic wars, including certain themes, for example love and death, and prefers a certain literary style or genre, the classification has almost nothing to offer [sic].

Na atualidade, esta visão distintiva da Literatura parece já não se justificar nem mesmo do ponto de vista económico, como explica Rafferty (2001, p. 184):

Reading novels for pleasure is no longer considered morally dangerous or time-wasting in a capitalist society which encourages the commodification of leisure and leisure products. Fiction has economic and ideological value in contemporary capitalism. Librarians are increasingly interested in retrieval systems which facilitate access and encourage fiction reading, perhaps influenced by the sophisticated retrieval tools created and implemented by Amazon.com, a commercial site which sees the value of investing in information (bibliographic information) and interpretation (e.g. readers' reviews) for material and economic ends.

DeZelar-Tiedman (1996, p. 204) defende que a representação de ficção merece receber o mesmo nível de atenção que a representação de não ficção, questionando: “how does a library measure the cost-effectiveness of such a practice? [...] Now that fiction has a well-respected, and expected, place on library shelves, isn't it time that cataloging practice allowed the same type of subject access to these works that has long been available for nonfiction?”.

Ainda que externos ao conteúdo, reconhece-se que os tópicos tradicionais constituem pontos de acesso úteis, isto é, através dos quais um documento pode ser efetivamente recuperado (Lancaster, 1993). Com efeito, estes permitem uma localização física e alguma recuperação de informação, bem como constituem elementos objetivos, em acordo com a tradição histórico-linguística a que está sujeita a classificação nos domínios das Línguas e das Literaturas. Por conseguinte, as abordagens a aspetos externos poderão ser de alguma utilidade para uma classificação cujo propósito seja o da organização física de documentos em estantes, mas não tanto àquela que procure descrever os tópicos internos do assunto, tendo em vista um futuro acesso e recuperação da informação. De uma forma geral, aos tradicionais pontos de acesso falta *granulosidade* (especificidade) e, principalmente, a perspetiva de conectividade entre os documentos e os leitores, suas necessidades e desejos (García-Marco et al., 2010). Não surpreende, pois, que tenham surgido os *Library of Congress Subject Headings e Guidelines on Subject Access to Individual Works of Fiction, Drama, Etc* (American Library Association, 1990) e sistemas como *Kirjasampo* (ver

<https://www.kirjasampo.fi/>) e o *WorldCat Fiction Finder* (ver <http://experimental.worldcat.org/xfinder/fictionfinder.html>).

Broughton (2004, p. 289) especifica e destaca algumas particularidades dos sistemas enumerativos tradicionais:

In general terms UDC has the best provision for scientific subjects, whereas LCC contains very considerable detail in the arts and humanities, particularly in respect of naming individual writers, artists, persons in history, events and primary texts of all kinds of and I all subjects. DDC probably has the best balance between subjects, with less detail but fewer obvious shortcomings.

Contudo, Campos (1975) afirma que, por mais que cresçam e se desenvolvam, as tradicionais classificações não conseguirão acompanhar o ritmo de desenvolvimento e a complexidade da moderna documentação. Na atualidade, tal convicção parece correta, face à cada vez maior heterogeneidade dos documentos e à complexidade dos seus assuntos. Como já apontado por Sapp (1986, p. 488), considera-se que a ficção não se ajusta aos típicos esquemas taxonómicos e hierárquicos da não-ficção por possuir uma lógica interna própria e multidimensional – “the topical diversity of fiction defies monolithic classification”. Também Tristão, Fachin e Alarcon (2004) dão conta do caráter limitativo destas classificações e da dificuldade de inserção de novos termos de assunto. Poderá afirmar-se que as classificações enumerativas refletem uma visão filosófica e não uma visão pragmática da documentação (Costa, 2010), mais próxima das reais necessidades dos utilizadores. Talvez esta seja a razão pela qual os princípios de uma teoria facetada têm vindo a ser discutidos como uma possível solução para problemáticas relativas à organização do conhecimento, muito devido às suas potencialidades em acompanhar as constantes mudanças e a evolução dos assuntos.

Integrative Levels Classification (ILC)

A generalidade da literatura científica elogia as capacidades das classificações facetadas para lidar com assuntos complexos. Por comparação e no que toca ao processo de análise do assunto, estes sistemas mostram limites menos estreitos e menor rigidez (Broughton, 2004). No entanto, para a representação dos documentos de ficção literária, verifica-se que também os tradicionais sistemas facetados não trazem vantagens observáveis ou maiores desenvolvimentos, por manterem como categorias de análise características externas, deixando o assunto (conteúdo intelectual) fora do sistema classificatório. Pogorelec e Šauperl (2006) salientam que os bibliotecários e os utilizadores dos serviços das bibliotecas necessitam de coleções de ficção organizadas, de forma que os documentos possam ser recuperados, tanto no catálogo como nas estantes. Por conseguinte, as questões que estes autores colocam são apenas duas: quais os critérios de classificação de documentos ficcionais; e qual o sistema classificatório a utilizar, abrindo como possibilidade a criação de uma nova proposta.

A Classificação de Níveis Integrativos (*Integrative Levels Classification – ILC*) é um sistema facetado de organização do conhecimento, de caráter experimental e em

desenvolvimento desde 2004 (Gnoli, 2020). Como características distintivas, destaca-se o facto de esta estrutura classificatória ser completamente facetada e de assentar em fenómenos, por oposição às disciplinas tradicionais do conhecimento. Aqui, os fenómenos do mundo encontram-se listados, de acordo com a sequência natural dos níveis integrativos – a tabela da ILC encontra-se disponível no sítio *web* do projeto (<http://www.iskoi.org/ilc/2/ilc.php>). Desta forma, cada conceito pode ser livremente combinado com outros por relacionamentos significativos (facetas), sem o espartilho das tradicionais áreas do conhecimento.

O projeto de investigação da ILC fundamenta-se, teórica e tecnicamente, no trabalho desenvolvido na década de 1960 pelo Classification Research Group. A ideia básica deste sistema consiste em enumerar as classes de fenómenos, por exemplo, a censura ou a coragem, em vez das tradicionais disciplinas. Já os níveis integrativos constituem um princípio útil para estabelecer uma determinada sequência entre todos os tipos de fenómenos existentes. Apesar de se inspirar nestes princípios, o sistema ILC desenvolve-se de raiz, sem qualquer correspondência de notação com as estruturas classificatórias pré-existentes. Desde 2004, a ILC encontra-se em constante evolução, o que torna necessário fixar as diferentes edições. Para este estudo, utiliza-se a segunda edição (ILC2), no entanto encontra-se já em desenvolvimento uma nova edição.

Na atualidade, o sistema é desenvolvido por uma equipa internacional de investigadores (pesquisadores, bibliotecários, cientistas da computação e filósofos), liderada por Claudio Gnoli. O intuito deste grupo de trabalho é testar um sistema de organização do conhecimento baseado em fenómenos, numa abordagem ontológica e em acordo com as tendências de interdisciplinaridade e inter-relação entre diversas áreas, em conformidade com o declarado no Manifesto de León (León Manifesto, 2007). É neste contexto que se pretende atestar o caráter distintivo e as potencialidades da ILC, no âmbito da representação do assunto literário e ficcional.

Metodologia

Para cumprimento deste objetivo, justifica-se um estudo exploratório de índole comparativa, designadamente um estudo de caso qualitativo, tendo como amostra por conveniência um romance do escritor italiano Antonio Tabucchi – *Sostiene Pereira* [*Afirma Pereira*] (1994). Seleciona-se esta obra literária por ser bastante conhecida e por já ter sido lida pelos dois autores desta investigação, algo fundamental para se verificar o grau de profundidade da representação obtida. Especificamente, pretende-se comparar a representação do assunto do romance obtido na ILC com distintas classificações bibliográficas, em particular as tipo enumerativo, por se constituírem como hegemónicas na generalidade dos serviços de informação, a saber, *Dewey Decimal Classification*, *Universal Decimal Classification* e *Library of Congress Classification / Library of Congress Subject Headings*, em acordo com as *Guidelines on Subject Access to Individual Works of Fiction, Drama, Etc* (American Library Association, 1990).

Após a consulta das respetivas tabelas classificatórias, confrontam-se resultados com a informação recolhida em catálogos *online* de serviços de informação; a título

de exemplo, a recolha de dados é realizada nos catálogos da Biblioteca Nacional de Portugal, das Bibliotecas do Sistema Provincial de Turim e da Biblioteca do Congresso. A informação da amostra é alvo de análise de conteúdo e apuram-se as categorias/facetes de assunto, relativas aos tópicos internos e externos que poderão constituir o assunto literário e ficcional. Os resultados são apresentados num quadro, do qual se fará a devida leitura comparativa-interpretativa e consequente discussão.

Resultados e discussão

Sistema Classificatório	Notação obtida	Representação obtida	Categorias/Facetes representadas (tópicos internos)	Categorias/Facetes representadas (tópicos externos)
ILC	ragi5plw42yacc2ttez p19pxrwi00xlt009c0 025qvUI	“bravery, through clear awareness, facing censorship, in Portugal, in August 1938, as attested in novels, in written texts, in Italian” [coragem, pela consciência lúcida, perante a censura, em Portugal, em agosto de 1938, atestada em romances, em textos escritos, em italiano]	Personagem Evento Espaço Tempo (diegese)	Género literário Tipo de texto Língua
UDC	821.131.1- -31”19/20”	Literatura em italiano, romance, autor do século XX-XXI	—————	Tipo de documento Língua Género literário Tempo (autor)
DDC	853.9	Literatura italiana, narrativa, século XX	—————	Tipo de documento Língua Modo literário Tempo (autor)
LCC/LCSH	<i>LC classification:</i> PQ4880.A24 S6613 1995 <i>LC Subjects:</i> Portugal--History-- -1910-1974-- -Fiction.	Literaturas românicas, italiana, autor individual 1968-2000, autor, obra, ano de edição Portugal–História–1910-1974– Ficção	Espaço Tempo (diegese)	Tipo de documento Língua Tempo (autor) Autor Título Tempo (edição) Subgénero literário

Figura 1. Quadro comparativo da representação de “Sostiene Pereira”

(Elaboração própria)

A figura 1 contém os dados recolhidos em: (1) Classificação atribuída pelos autores; (2) Catálogo da Biblioteca Nacional de Portugal – <http://catalogo.bn.pt>

gal.gov.pt/ipac20/ipac.jsp?session=U629391F5260Y.14398&profile=bn&source=~! bnp&view=subscriptionssummary&uri=full=3100024~!525200~!1&ri=1&aspect=su btab11&menu=search&ipp=20&spp=20&staffonly=&term=sostiene+pereira&index =.GW&uindex=&aspect=subtab11&menu=search&ri=1; (3) Catálogo das Bibliote- cas do Sistema Provincial de Turim [https://sbam.erasmo.it/Opac/Esame Tito- lo.aspx?IDTIT=n8Q4nora5i4=](https://sbam.erasmo.it/Opac/Esame Titolo.aspx?IDTIT=n8Q4nora5i4=); e (4) Catálogo da Biblioteca do Congresso – <https://catalog.loc.gov/vwebv/holdingsInfo?searchId=13787&recCount=25&recPoin ter=0&bibId=1145265>.

No geral, constata-se a tradicional predominância da representação dos tópicos externos, como seria expectável. Neste âmbito, o número de categorias/facetas é variável de três (ILC) a sete (LCC), sendo estas relativas ao documento (tipo, géne- ro, língua), autor (nome, tempo) e obra (título, edição). O quadro revela que não se encontram informações sobre tópicos internos na UDC e na DDC, sendo que a primeira se mostra um pouco mais específica no género literário – de facto, o romance constitui um género do modo narrativo. A similitude entre os dois sistemas é reconhecida historicamente, visto que o primeiro parte da 5.^a edição do segundo. Por sua vez, a ILC e a LCC mostram uma representação de tópicos internos dentro das categorias/facetas reconhecidas na literatura (Almeida, 2019), pelo menos desde o *Experimental Fiction Analysis System* de Clare Beghtol (1994): personagem, evento, espaço e tempo (da diegese).

De acordo com Langridge (1989), existem duas questões que devem ser respon- didas no tratamento documental: “What is it?”, que corresponderá aos tópicos exter- nos, e “What is it about?”, refletindo tópicos internos. Num eventual documento técnico *Manual de indexação*, o primeiro elemento responderia a o que é e o segun- do aquilo que se trata ou é sobre – tratar-se-ia, efetivamente, de um documento sobre indexação, na forma de um manual (livro de instruções). Contudo, este raciocínio não poderia ser efetuado com o Romance *Manual dos Inquisidores* do escritor português António Lobo Antunes – não se trata de um livro com indicações para inquisidores ou sobre inquisição. Logo, no que toca à Literatura, o quadro confirma que as classificações bibliográficas respondem à primeira pergunta de Langridge, mas não à segunda. Portanto, acompanha-se Saarti (2019) quando afirma que as classificações bibliográficas deixam intocado o assunto da ficção literária: “these classifications leave the idea of describing the subject content of fiction – what the fiction is about – untouched”. Conclui-se, pois, que UDC e DDC apenas respondem à primeira questão e que ILC e LCC respondem às duas questões de Langridge. No entanto, a ILC revela-se mais exaustiva, uma vez que vai um pouco mais além da localização espaço-temporal da ação, revelando características das personagens (coragem, lucidez) e eventos (tomar consciência, enfrentar a censura). Paralelamente, verifica-se maior especificidade dos descritores, por exemplo na referência ao tempo da diegese (agosto de 1938).

Estes resultados mostram que, nos sistemas de organização do conhecimento tra- dicionais, a obra literária de Tabucchi se encontra associada a outros romances italianos do século XX do mesmo autor, sem que seja possível estabelecer uma ligação ao conjunto de outras obras que tratam a censura ou cuja ação decorre durante o regime ditatorial de Salazar. Estas poderão constituir ficção ou não ficção, muito embora o sistema permita a devida distinção. Ao dar prioridade a fenómenos,

a ILC permite interligar *Sostiene Pereira* a documentos que também abordem, por exemplo, a censura em Portugal, tanto de cariz ficcional como de cariz mais técnico-científico.

Neste sistema classificatório, as relações semelhantes entre a forma e o conteúdo intelectual podem ser expressas por uma linguagem verbal de indexação, representando-se distintos elementos diegéticos temáticos de obras de ficção. Desta feita, o papel dos elementos formais de *assunto* (género, forma física, meio...) pode ser mais bem esclarecido na ILC, verificando-se como uma mesma obra (tomada no sentido de *Functional Requirements for Bibliographic Records*) se pode representar nas suas diferentes expressões. Na ILC, por exemplo, é possível associar-se o romance original de Antonio Tabucchi a:

- uma tradução em língua portuguesa – ragi5plw42yacc2ttezp19pxrwi00xlt009c005qvUT0053qvUI (“bravery, through clear awareness, facing censorship, in Portugal, in August 1938, as attested in novels, in written texts, in Portuguese, translated from Italian”);
- uma leitura áudio transmitida em rádios italianas (audiolivro) – ragi5plw42yacc2ttezp19pxrwi00xlt009n0025qvUI (“bravery, through clear awareness, facing censorship, in Portugal, in August 1938, as attested in novels, in sound records, in Italian”);
- uma adaptação a filme do realizador Roberto Faenza – ragi5plw42yacc2ttezp19pxrwi00xsi(31xlt)0025qvUI (“bravery, through clear awareness, facing censorship, in Portugal, in August 1938, as attested in feature films (evolved from novels), in Italian”);
- e uma adaptação a romance gráfico em banda desenhada da autoria de Magliani e D’Aponte – ragi5plw42yacc2ttezp19pxrwi00xgo(31xlt)0025qvUI (“bravery, through clear awareness, facing censorship, in Portugal, in August 1938, as attested in comics (evolved from novels), in Italian”).

É certo que as notações de classe produzidas por este sistema se revelam um pouco longas, tornando-as especialmente úteis para a recuperação em catálogos digitais, quando comparadas com as utilizadas na arrumação física de documentos, onde algumas políticas de encurtamento podem ser adequadas.

Em suma, conclui-se que, quando comparada com os sistemas tradicionais de organização do conhecimento, a ILC se destaca pelo seu carácter distintivo e por apresentar um maior potencial informativo na representação do assunto ficcional, sobressaindo uma maior versatilidade e relação entre conceitos. Neste sentido, considera-se que a representação proporcionada pela ILC é mais facilitadora do acesso e da recuperação da informação ficcional aos utilizadores do sistema. Em sentido inverso, verifica-se que a ILC não representa tão exaustivamente tópicos externos, não se registando dados sobre autor, sub-género ou edição, encontrados nos sistemas pré-existentes. Para os tópicos internos, e no que respeita à localização da diegese, constata-se que a atual tabela da ILC poderá mostrar-se mais específica, pela inclusão de referências à região do país onde se desenrolam os acontecimentos (“Lisbon district” em vez de “Portugal”).

Considerações finais

Com efeito, o virar do século trouxe um novo interesse pela Literatura, sendo-lhe reconhecida relevância na generalidade das bibliotecas, principalmente públicas e escolares. Não obstante esta mudança, verifica-se que as classificações se mantêm como instrumentos de representação do conhecimento pouco adequados à Literatura, por não refletirem algumas das características mais importantes da ficção literária (Eriksson, 2010). No que toca à representação dos documentos literários, a expressividade das classificações bibliográficas é limitada e enviesada, na medida em que os metatópicos não refletem o assunto. Um género, um autor ou um período constituem a moldura do documento, porém não formam a sua identidade e não asseguram a representação do seu conteúdo intelectual. Na prática, as bibliotecas continuam a não representar o assunto literário, parecendo confiar essa tarefa a editores e livreiros, nomeadamente em ambiente *web* (Ward & Saarti, 2018).

É possível uma representação mais detalhada do assunto ficcional, através da abordagem baseada em fenómenos da ILC. Na verdade, embora a área do conhecimento dos documentos seja sempre “Literatura”, a classificação por fenómenos permite representar tanto os tópicos internos como os tópicos externos mais relevantes. A sintaxe facetada permite combinar todos os elementos numa ordem de citação útil. Por si só, a sintaxe facetada não se mostra suficiente quando o sistema é baseado em disciplinas, uma vez que isso apenas resultará numa classificação por facetadas de tópicos externos. A verdadeira novidade da ILC é apresentar-se como uma classificação baseada em fenómenos, desta forma atestando o seu carácter distintivo em potencial e, no que respeita à Literatura, afirmando a representação do assunto ficcional.

Este estudo encontra limitações na amostra, pelo que não se estabelecem generalizações, no entanto apresenta-se como uma das várias investigações em curso pelo grupo de investigadores da ILC e nelas encontra o seu contexto e alcance. Paralelamente, este trabalho assume-se como mais um onde se demonstra a exequibilidade da representação do conteúdo temático ficcional, em linha com as tendências das últimas décadas em organização do conhecimento.

Referências bibliográficas

- Almeida, P. (2019). Tendencias europeas en la indización por asunto de ficción narrativa. *Anales De Documentación*, 22(1), pp. 1-10. Disponível em <https://doi.org/10.6018/analesdoc.22.1.339791>
- American Library Association – ALA. (1990). *Guidelines on Subject Access to Individual Works of Fiction, Drama, etc.* Chicago: American Library Association.
- Barbosa, A. P. (1972). Classificações Facetadas. *Ciência da Informação*, 1(2), pp. 73-81. Disponível em <http://revista.ibict.br/ciinf/article/view/10>
- Beghtol, C. (1989). Access to fiction: a problem in classification Theory and practice. Part I. *Int.Classif.*, 16(3), pp. 134-140. Disponível em <https://www.nomos-elibrary.de/10.5771/0943-7444-1989-3-134/access-to-fiction-a-problem-in-classification-theory-and-practice-part-i-volume-16-1989-issue-3>

- Beghtol, C. (1994). *The Classification of Fiction: The Development of a System Based on Theoretical Principles*. Metuchen: Scarecrow Press.
- Broughton, V. (2004). *Essential Classification*. London: Facet Publishing.
- Campos, A. (1975). A teoria das classificações analítico-sintéticas, ou facetadas, e a sua influência sobre a reforma da Classificação Decimal Universal (CDU). *Revista Biblioteconomia de Brasília*, 3(1), pp. 23-36. Disponível em <https://periodicos.unb.br/index.php/rbbsb/article/view/29079>
- Chaumier, J. (1988). Indexação: conceito, etapas e instrumentos. *Revista Brasileira de Biblioteconomia e Documentação*, 21, n. ½, pp. 63-79.
- Costa, L. (2010). Aproximações teórico-conceitual entre as categorias de Ranganathan, o discurso retórico e a narrativa literária. *Pesquisa Brasileira em Ciência da Informação*, 3(1), pp. 169-184. Disponível em <https://brapci.inf.br/index.php/res/v/119336>
- DeZelar-Tiedman, C. (1996). Subject Access to Fiction: An Application of the Guidelines. *Library Resources and Technical Services*, 40(3), pp. 203-210. Disponível em <https://doi.org/10.5860/lrts.40n3.203>
- Eriksson, R. (2005). The classification and indexing of imaginative literature. *American Society for Information Science and Technology's 16th Workshop for the Special Interest Group on Classification Research*, Charlotte, NC, October 29, 2005. Disponível em <https://doi.org/10.1002/bult.1720310403>
- Eriksson, R. (2010). Klassifikation og indeksering af skønlitteratur: Et teoretisk og historisk perspektiv. [The classification and indexing of fiction – a theoretical and historical perspective]. Tese de doutoramento. Danmarks Biblioteksskole. Disponível em [https://research.ku.dk/search/result/?pure=en/publications/klassifikation-og-indeksering-af-skoenlitteratur\(1857cd42-4a4d-4ed4-8d13-a032a7cc48a9\).html](https://research.ku.dk/search/result/?pure=en/publications/klassifikation-og-indeksering-af-skoenlitteratur(1857cd42-4a4d-4ed4-8d13-a032a7cc48a9).html)
- García-Marco, F. J., Moraes, J. B. E., García-Marco, L. F., & Guimarães, J. A. C. (2010). Knowledge organization of fiction and narrative documents: A challenge in the age of the multimedia revolution. In *Advances in Knowledge Organization, Paradigms and conceptual systems in knowledge organization Proceedings of the Eleventh International ISKO Conference 23-26 February 2010 Rome, Italy* (Vol. 12, pp. 262-268). Würzburg: Ergon Verlag.
- Gnoli, C. (2020). Integrative Levels Classification. In *ISKO Encyclopedia of Knowledge Organization*. Disponível em <https://www.isko.org/cyclo/ilc>
- Hypén, K. (2014). Kirjasampo: Rethinking metadata. *Cataloging and Classification Quarterly*, 52(2), pp. 156 -180. Disponível em <https://doi.org/10.1080/01639374.2013.848389>
- Lancaster, F. W. (1993). *Indexação e resumos: teoria e prática*. Brasília: Briquet de Lemos.
- Langridge, D.W. (1989). *Subject analysis: principles and procedures*. London: Bowker-Saur.
- Léon Manifesto (2007). *Knowledge organization*, 34, pp. 6-8. Disponível em <http://www.iskoi.org/ilc/leon.php>
- Pejtersen, A. M. (1979). The meaning of 'about' in fiction indexing and retrieval. *Aslib Proceedings*, 31(5), pp. 251-257. Disponível em <https://www.emerald.com/insight/content/doi/10.1108/eb050682/full/html>
- Pogorelec, A., & Šaupperl, A. (2006). The Alternative Model Of Classification Of Belles-Lettres In Libraries. *Knowledge Organization*, 33(4), pp. 204-214. Disponível em <https://www.nomos-elibrary.de/10.5771/0943-7444-2006-4-204/the-alternativemodel-of-classification-of-belles-lettres-in-libraries-volume-33-2006-issue-4>
- Rafferty, P. (2001). The Representation of Knowledge in Library Classification Schemes. *Knowledge Organization*, 28(4), pp. 180-192. Disponível em https://www.ergon-verlag.de/isko_ko/downloads/ko2820014d.pdf

- Rafferty, P. (2013). Epistemology, literary genre and knowledge organisation systems. *Actas Del X Congreso de ISKO-España. Ferrol 20 de Junio-1 de Julio de 2011*, 553–565. Disponível em http://ruc.udc.es/dspace/bitstream/handle/2183/11640/CC_132_art_39.pdf%0Ahttp://ruc.udc.es/dspace/handle/2183/11640
- Ranganthan, S. R. (1963). *The five laws of library science* (2.^a edição revista). Bombay: Asia Publishing House.
- Saarti, J. (2019). Fictional literature classification and indexing. in ISKO Encyclopedia of Knowledge Organization, eds. Birger Hjørland and Claudio Gnoli. Disponível em <http://www.isko.org/cyclo/fictional>
- Sapp, G. (1986). The Levels of Access: Subject Approaches to Fiction. *RQ*, 25(4), pp. 488-497. Disponível em <https://www.jstor.org/stable/25827716?seq=1>
- Tabucchi, A. (1994). *Sostiene Pereira*. Milano: Feltrinelli.
- Tristão, A. M. D., Fachin, G. R. B., & Alarcon, O. E. (2004). Sistema de classificação facetada e tesouros: instrumentos para organização do conhecimento. *Ciência da Informação*, 33(2), pp. 161-171. Disponível em <https://doi.org/10.1590/S0100-19652004000200017>
- Ward, M., & Saarti, J. (2018). Reviewing, rebutting, and reimagining fiction classification. *Cataloging and Classification Quarterly*, 56(4), pp. 317-329. Disponível em <https://doi.org/10.1080/01639374.2017.1411414>

AS TAXONOMIAS NAVEGACIONAIS FACETADAS E A PRODUÇÃO CIENTÍFICA DA CIÊNCIA DA INFORMAÇÃO: TENDÊNCIAS TEMÁTICA E DIACRÓNICA (2011-2020)

Ana Gouveia Coelho¹, Gercina Ângela de Lima², Maria Manuel Borges³

¹Universidade de Coimbra, gouveiacolho@student.fl.uc.pt, ORCID-ID 0000-0001-8520-2522

²Universidade Federal de Minas Gerais, glima@eci.ufmg.br,

ORCID ID 0000-0003-0735-3856

³Universidade de Coimbra, CEIS20, mmb@fl.uc.pt, ORCID ID 0000-0002-7755-6168

Resumo

As taxonomias navegacionais facetadas (TNF) são um sistema de organização do conhecimento (SOC) que surge da necessidade de estruturar conteúdos organizacionais no âmbito da Web, especialmente de sites de comércio eletrónico e constituem, atualmente, o SOC mais utilizado nesse âmbito. É objetivo deste estudo identificar as tendências temática e diacrónica da produção científica da ciência da informação (CI), entre 2011 e 2020, acerca das TNF, bem como inferir da respetiva tendência de produção científica futura. Usou-se um método qualitativo de tipo exploratório, incluindo uma análise textual de conteúdo com software. Recolheu-se um corpus de artigos científicos (2011-2020) sobre TNF na WOS, SCOPUS, Library & Information Science Source. As palavras-chave dos autores foram transformadas em linguagem controlada tendo como base o plano de classificação constante do Tesouro Brasileiro em Ciência da Informação. A análise textual de conteúdo foi feita com base nas palavras-chave e com auxílio do software R para cálculo do grafo de coocorrências e comunidades lexicais, o que permitiu inferir acerca da tendência temática da produção científica acerca de TNF. A linha de tendência diacrónica da produção científica foi calculada através do MSEXCEL®. Os resultados indicam que a produção científica da CI acerca das TNF tem vindo a diminuir, prevendo-se que assim continue para lá de 2020, o que poderá ser explicado pelos temas predominantemente escolhidos pelos autores, para o enquadramento da produção científica sobre TNF.

Palavras-chave: Taxonomias, Taxonomias Navegacionais Facetadas, Sistemas de Organização do Conhecimento

Introdução

O desígnio de contextualizar a informação traduz uma necessidade humana e constitui um objeto da ciência da informação. Segundo Brandt e Medeiros (2010), pode-se entender que a organização do conhecimento “está relacionada com um

processo de análise conceitual de um domínio do conhecimento e a partir daí, sua estruturação, gerando uma representação do conhecimento de tal domínio”. Ou seja, a informação necessita ser categorizada, classificada e relacionada (Simões & Freitas, 2011). Também Maculan e Lima (2011), destacam a análise do domínio como passo fundamental para a organização do conhecimento, enfatizando a metodologia teórico-prática e social do processo, interpretando os fluxos de comunicação no âmbito de uma comunidade onde esse discurso é formulado, bem como a relação dessa comunidade com a sociedade onde se insere.

Para Barité (2015), o domínio constitui um universo autónomo mais ou menos autossuficiente, que se apoia numa estrutura relacional definida pelos interesses de uma comunidade de utilizadores, por exemplo, a comunidade da ciência da informação. Por isso, é algo mais abrangente que apenas uma disciplina ou ciência, pois tem um enfoque comunitário, constituído pelos membros que dela fazem parte, mas também, constituído pelas respetivas crenças e atitudes, que afetam o modo de processamento conhecimento científico, através dum filtro social. Partindo dessa premissa, os SOC são o reflexo da mente coletiva dos autores que os edificaram e naturalmente, a tipologia dos sistemas de organização do conhecimento varia de acordo com os objetivos a atingir e do contexto – para quem e quando – em que são criados. Ainda, segundo Lima (2013) “os conceitos existem desde que o homem teve a capacidade de se abstrair” A partir do momento em que o homem desenvolveu socialmente a capacidade em distinguir diferenças e semelhanças, estabeleceu, concomitantemente, relacionamentos entre os objetos que o rodeiam e deu o primeiro passo para a organização do conhecimento de forma sistémica. Segundo Hjørland e Albrechtsen (1995), apesar dos conceitos poderem ser comumente vistos como independentes nunca poderão ser verdadeiramente compreendidos sem o entendimento do contexto global onde se inserem, contexto esse que constitui a matéria-prima de qualquer sistema de informação. Segundo Chiavenato (1993) um sistema constitui-se por um conjunto de elementos unidos por alguma forma de interação ou interdependência. Qualquer conjunto de partes unidas entre si pode ser considerado como um sistema, se as relações entre as partes e o comportamento do todo for o ponto principal abordado. De acordo com Bertalanffy (1968), o criador da Teoria Geral dos Sistemas e das leis que os regulam um sistema é o conjunto de unidades em inter-relações mútuas. Para Morin (1977), um sistema constitui-se como uma inter-relação de elementos que formam uma entidade ou unidade global. Das três definições prevalece a ideia de unidade dinâmica, de fluxos de relacionamentos constituídos em torno de um objetivo, ou conceito central. O sistema detém um equilíbrio dinâmico. Por sua vez, o principal problema de um sistema de informação é refletir o domínio. Porque “domínio” e “sistema de informação” são conceitos interligados, qualquer sistema de informação encerra, intrinsecamente, a vertente dinâmica inseparável da mudança, ou seja, o aspeto temporal, que, aliás, subjaz a tudo.” A compreensão de como a informação e os conceitos fluem através do tempo promove também a criação e transferência de novo conhecimento e a tomada de decisão, especialmente na presença de grandes volumes de documentos, que tornam mais difícil, se não impossível, sem um tratamento metodológico adequado, a intuição de como esta decorre” (Chen, *et al*, 2017). A visualização dos

domínios decorre numa estrutura designada por Sistemas de Organização do Conhecimento (SOC). Os SOC destinam-se a organizar a informação armazenada, sendo, por essa razão, fundamentais para a sua recuperação com vista à geração de um novo ciclo de conhecimento.

Como referem Borner, *et al* (2003), o principal problema da constituição de um sistema de informação é a profundidade da análise. Atualmente, avanços na técnica de visualização de domínios, através de uso de software adequado aplicado à mineração automática de dados textuais, auxiliam a compreensão do todo. Tais técnicas, não visam a compreensão do documento (parte do sistema), mas sim, do conjunto (*corpus* em análise), cuja interpretação faz surgir nova informação (meta-informação), que se converte, num processo social, em conhecimento. Fundamentos derivados da linguística são habitualmente aplicados, juntamente com métodos de índole estatístico, à compreensão de grandes volumes de documentos sendo comumente aplicados em diversas ciências sociais, como a ciência de informação. “A nova informação que surge é também dinâmica e possui a sua própria linha de evolução/difusão, sofrendo momentos de retração, de expansão, pontos de viragem e também, desaparecimento”. (Chen, 2006). Todos os conceitos se exprimem em palavras ou conjuntos de palavras, as quais traduzem a evolução e/ou difusão dos conceitos afetos a um domínio. Se no início algumas palavras surgem como “palavras de topo”, passado algum tempo podem perder esse estatuto, o que acontece quando se ligam a outros conceitos. Alguns autores (Chen, *et al* (2017); Williams & Balaz, (2014), estabeleceram um paralelismo entre este fenómeno e a demografia concernente às próprias migrações humanas. “Ao conhecer as palavras podemos conhecer os conceitos e vice-versa, pois a palavra, uma vez migrada, representará uma ideia diferente, ao ligar-se a um outro contexto.” (Iles *et al.*, 2004). Também segundo Holme e Saramäki (2012), para que seja possível entender um sistema complexo é necessário remover a atenção dos detalhes e olhar para o todo. Uma maneira simples de concretizar este desiderato é transformar o sistema num grafo. No início do século XXI, estudos desenvolvidos por Newman (2001), demonstraram a total aplicabilidade da teoria dos grafos para a visualização de redes de informação social e por tal, a sua aplicabilidade no seu estudo. Neste caso, os “vértices denotam entidades e as ligações, a comunicação entre as entidades, num misto de modelização e representação gráfica” (Chen, 2006); (Ahmed & Chen, 2016).

No que toca à Ciência da Informação (CI), a deteção de tendências (expansão, retração, desaparecimento, aparecimento). “É um campo de estudo que tem interesse especialmente à bibliometria, em estudos de citações de artigos científicos, cujas flutuações permitem a deteção de interesses prestes a surgir ou a desaparecer no campo científico e que para o efeito, utilizam grafos de evolução temporal” (Erten, *et al*, 2003). Existem vários estudos desta natureza no campo da análise bibliográfica, nomeadamente, na deteção de tendências emergentes de investigação científica, por exemplo, o estudo desenvolvido por Klavans & Boyack (2007) no desenvolvimento de um mapa geral da ciência por intermédio da rede de citações.

As taxonomias são o SOC formal mais antigo de que há conhecimento. Já na Grécia Antiga havia a noção de taxonomia, quando Aristóteles (384-322 a.C.) concebeu a sua famosa classificação das ciências. A taxonomias tiveram o seu apogeu

na publicação da obra *Systema Naturae* em 1735, em que Lineu (1707-1778), publicou a sua famosa classificação biológica. As taxonomias detêm ainda a particularidade de serem a estrutura de todos ou outros SOC. Tal acontece porque, segundo Zeng (2008), refletem o entendimento básico acerca das coisas e é a esse nível básico que repousa todo o edifício da nossa sabedoria.

Indiferentes à sua antiguidade, as taxonomias permaneceram até aos dias de hoje, expandindo-se na Web. “Somaram aos tradicionais relacionamentos hierárquicos, relacionamento associativos, que funcionam como elos semânticos de diferentes cadeias hierárquicas e transformaram-se no SOC atualmente mais utilizado para modelar e mediar conteúdos especialmente em sites de comércio eletrónico” (Cavalcante & Brasher, 2014). Apesar de terem surgido no âmbito da Ciência da Computação como resposta a uma necessidade, nelas se espelha um clássico da CI, a teoria da classificação facetada de Ranghanatam (1967). Se as Taxonomias Navegacionais Facetadas (TNF) forem bem construídas, em termos da construção das facetas e se seguirem o princípio, especialmente importante na pesquisa on-line, de *userfriendliness* (Hjørland, 2012), constituem o SOC mais adaptado na modelação de conteúdos na rede tal como hoje a conhecemos. Considerando a importância e omnipresença das TNF (quem nunca fez uma compra on-line?) faz sentido analisá-las sob o ponto de vista único da ciência da informação.

Nesse âmbito, escolheu-se traçar o retrato, ainda que tendencial, de como a comunidade científica da CI tem olhado estes SOC, nas vertentes temática e diacrónica, refletidas na respetiva produção científica. Esta, constituiu a nossa problemática e consubstanciou a pergunta de investigação: “Quais as tendências temáticas e diacrónica da produção científica da Ciência da informação (2011-2020) acerca das taxonomias navegacionais facetadas?”

Objetivos

Conhecer as tendências temática e diacrónica da produção científica da CI sobre TNF no período 2011-2020, constitui o objetivo geral. Como objetivo específico conjugar os resultados obtidos e inferir como a ciência da informação encara as TNF na atualidade e como deve encarar prospectivamente.

Método

Foi utilizado um método qualitativo exploratório que incluiu uma análise textual de conteúdo com recurso a software.

Foram recolhidos documentos produzidos sobre TNF¹, no âmbito da CI selecionando o período 2011-2020, nas bases de dados, nomeadamente, Web of Science, Library and Information Science Source, Scopus. Utilizou-se o termo *faceted taxonomy* (sem aspas), nos campos Título, o Resumo e Palavras-Chave, para recuperar outros equivalentes e que dessem maior abrangência ao *corpus*. A pesqui-

¹ Em apêndice

sa foi restringida ao domínio da Ciência da Informação. A recolha resultou em 39 documentos e 193 palavras-chave.

As palavras-chave foram o insumo para o estudo temático. Estas seguidamente foram classificadas de acordo com o plano de classificação constante do Tesouro Brasileiro em Ciência da Informação (Pinheiro & Ferrez, 2014)²

As palavras-chave, através da classificação, foram transformadas em linguagem controlada, constituindo o *corpus*. Este foi operacionalizado com o software gratuito de análise de dados R, interface IRaMuteq, selecionando o cálculo de coocorrências e clusters de proximidade, visando a demarcação da principal tendência temática do conjunto. Segundo Barabasi (2016), a emergência da ciência das redes no início do século XXI e a sua aplicação às ciências sociais faz todo o sentido, pois cada sistema de informação social complexo encontra-se firmado numa teia de ligações que explicitam as conexões entre os componentes do sistema e o tornam tangível se representadas graficamente. A análise textual de conteúdo por máquina é ainda, segundo Cervi (2018), menos subjetiva do que aquela produzida pela mente do leitor, nomeadamente, através de outros métodos de índole qualitativa. Aliás, constitui um método híbrido, pois tem uma vertente quantitativa, firmada na análise estatística.

A linha de tendência diacrónica da produção científica foi realizada com recurso ao programa EXCEL® e servirá para aferir da evolução diacrónica da produção científica da CI acerca das TNF, também para além de 2020.

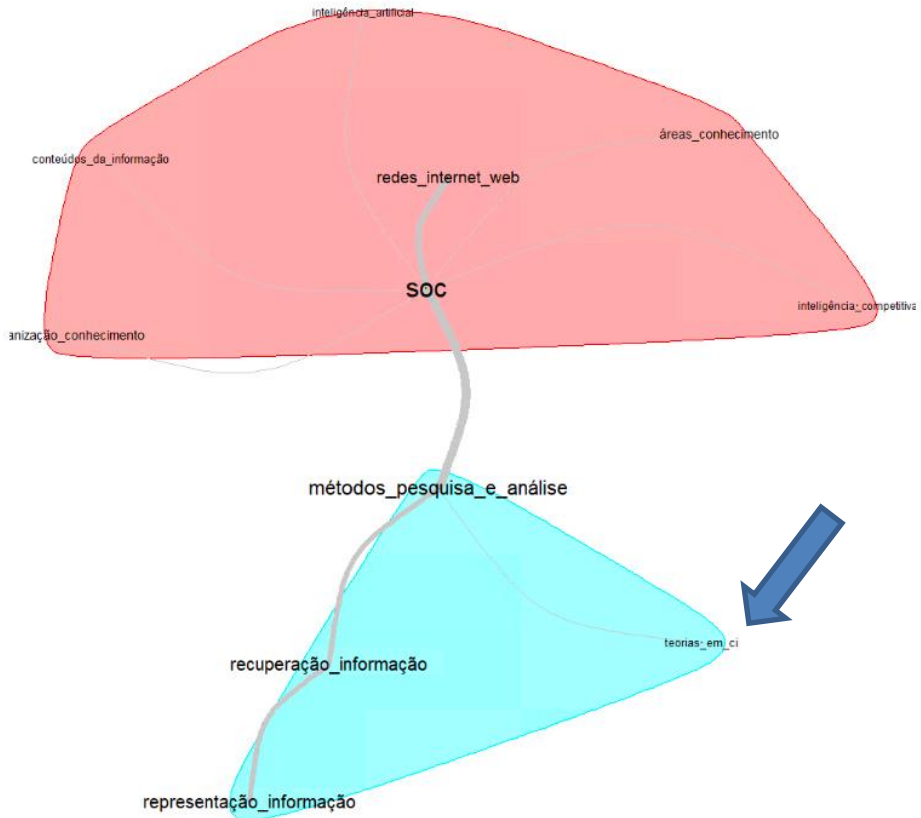
Resultados

1 – Tendência Temática da Produção Científica

A operacionalização das palavras-chave por intermédio do software R resultou no seguinte grafo de coocorrências, que reflete os principais temas, do domínio da CI, (em conformidade com o Tesouro), no âmbito dos quais a produção científica das TNF (2010-2020), tem sido tendencialmente enquadrada:

² Em apêndice

Grafo n.º 1 – As principais temáticas que enquadram a produção científica 2010-2020 sobre TNF



Fonte: Output R. Coocorrências entre os artigos do *corpus* através das respetivas palavras-chave (em linguagem controlada). Os artigos são a variável de factorização. Layout de Freutchman-Reingold com representação das Comunidades Lexicais e respetivo halo, para melhor leitura.

O nó principal do grafo é gerado a partir do termo *SOC*. Tal não constitui surpresa. As TNF são um SOC, sendo, portanto, um conceito transversal a todo o corpus lexical. Este termo insere-se numa comunidade de proximidade com outros termos, mas, principalmente, com “*redes internet e web*”, com o qual apresenta uma forte relação. As TNF são um SOC que existe no seio da WWW. Ou seja, a ligação “*SOC*” – “*Redes, internet e web*”, representa as TNF.

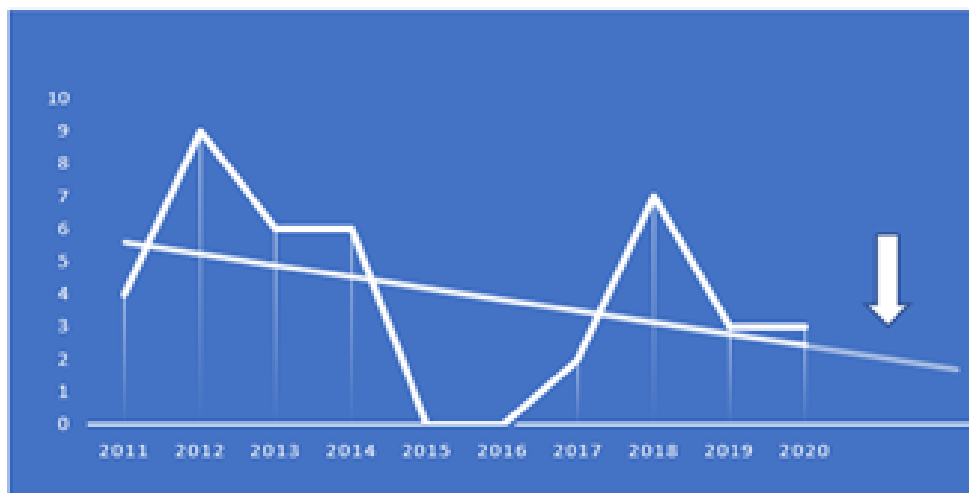
O principal fluxo do grafo é gerado entre “*SOC*” e “*Métodos de pesquisa e análise*”. Este último consubstancia o tipo de estudos realizados, por exemplo, estudos de caso, descritivos, exploratórios ou outros. Independentemente do tipo de método aplicado, este tem sido dirigido ao objetivo de “*recuperação da informação*” e “*representação da informação*”, ambos subalíneas do ponto 2-*Organização do conhecimento e recuperação da informação* do Tesouro utilizado (ver apêndice)

Um fluxo de menor intensidade é traduzido entre “*Métodos de pesquisa e Análise*” e “*Teorias em CI*”, que nos remete para estudos relacionados com a principal característica epistemológica das TNF, a análise facetada.

Em termos globais, infere-se que os artigos têm sido tendencialmente elaborados no âmbito de construção de TNF para modelação de conteúdos de domínios específicos.

2 – Tendência diacrónica da produção científica da CI sobre TNF

Gráfico n.º 1 – Previsão linear da produção científica da CI sobre TNF (2020-)



O eixo dos YY indica o n.º de artigos e o eixo dos XX os anos. O declive da reta, para lá de 2020, continua negativo

Fonte: Output do MSEXCEL

O cálculo da previsão linear mostrou um declive negativo, mesmo depois de 2020. Ilustra também uma contínua diminuição do interesse da comunidade científica da CI sobre TNF.

Conclusões

Se o e-commerce tem crescido ao longo da década, por que razão o estudo, por parte da CI, dos SOC preferidos para estruturar os conteúdos de sites, tem diminuído? Uma possível explicação reside na perspetiva temática dada ao estudo das TNF. A CI debruçou-se tendencialmente, no período, na perspetiva de construção de TNF para modelação de conteúdos. Tal é corroborado pela revisão de literatura do *corpus* pois dos trinta e nove artigos recolhidos, vinte e um dizem respeito a estudos dessa natureza, o que está em consonância com os resultados automáticos veiculados pelo *software*.

O aumento de TNF consequência do crescimento do e-comercio, o que foi corroborado com a pandemia, é um bom indicador que a CI deveria prestar mais atenção a estes SOC. Estes continuam e deverão continuar, por uns bons anos, na ordem do dia. Mas o foco temático da CI deveria mudar e olhar para o objeto mais sobre a sua única perspectiva. A avaliação das TNF é pouco difundida, mormente, no que toca à correta constituição das facetas. Daí poderiam advir ganhos em termos económicos, consubstanciados na melhoria funcional das TNF, mormente em sites de e-comercio e constituindo uma mais-valia para empresas e utilizadores. Seria mais uma oportunidade para a afirmação da CI.

Implicações teóricas e práticas para a ciência da informação

O estudo traduz um retrato de uma década acerca da produção da CI sobre TNF. Através do cálculo da linha de tendência indica como esta pode se desenrolar prospectivamente. Reflete uma perspectiva pouco explorada, no âmbito da CI, acerca das TNF.

Recomendações

As TNF acompanham a expansão do e-comercio, que não está a abrandar. A avaliação das TNF é um campo ainda pouco difundido, o que consubstancia uma oportunidade para mais estudos no âmbito da CI.

Limitações

A análise textual de conteúdo realizada, só abrangeu artigos publicados nas bases de dados referidas, não incluindo teses ou dissertações ou outro tipo de documentos.

Referências bibliográficas

- Ahmed, N. M., & Chen, L. (2016). An efficient algorithm for link prediction in temporal uncertain social networks. *Information Sciences*, 331, 120-136. <https://doi.org/10.1016/j.ins.2015.10.036>
- Aquino, I. J., Carlan, E. & Medeiros, M. (2009). Princípios classificatórios para a construção de taxonomias. *Ponto de Acesso*, 3(3), 196-215.
- Barabási, A.L. (2016). *Network Science*. Cambridge University Pres. <http://networksciencebook.com/>
- Barité, M (2015) *Diccionario de Organización del Conocimiento: Clasificación, Indización, Terminología*. Montevideo
- Bertalanffy, L. V. (2006). *Teoria Geral dos Sistemas: Fundamentos, desenvolvimentos e Aplicações* (2.^a, Vol. 1-1). Vozes.
- Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37(1), 179-255. <https://doi.org/10.1002/aris.1440370106>

- Campos, C. (2004). Método de análise de conteúdo: ferramenta para a análise de dados qualitativos no campo da saúde. *Revista Brasileira de Enfermagem*, 7(5), p. 611-4.
- Carlan E. & Medeiros, M. (2011). Sistemas de Organização do Conhecimento na visão da Ciência da Informação. *Revista Ibero-Americana de Ciência Da Informação*, 4(2), 53-73. <https://doi.org/10.26512/rici.v4.n2.2011.1675>
- Cavalcante, R. & Brasher, M. (2014). Taxonomias navegacionais em sítios de comércio eletrônico: critérios para avaliação. *Transinformação*, 26(2), 191-201. <https://doi.org/10.1590/0103-37862014000200008>
- Cervi, E. (2018). Análise de conteúdo automatizada para conversações em redes sociais online: uma proposta metodológica. [Paper presentation]. 42.^o Encontro ANPOQS, Mato Grosso. <https://bit.ly/38UkYtz>
- Chen, C. (2006). *Information visualization: Beyond the horizon* Springer International Publishing; Scopus. <https://doi.org/10.1007/1-84628-579-8>
- Chen, B., Tsutsui, S., Ding, Y., & Ma, F. (2017). Understanding the topic evolution in a scientific domain: An exploratory study for the field of information retrieval. *Journal of Informetrics*, 11(4), 1175–1189. Scopus. <https://doi.org/10.1016/j.joi.2017.10.003>
- Chiavenato, I. (1993) *Introdução à teoria geral da administração*. Makron Books,
- Erten, C., Harding, P. J., Kobourov, S. G., Wampler, K., & Yee, G. (2004). *Exploring the computing literature using temporal graph visualization. Visualization and Data Analysis*, 45–56. <https://doi.org/10.1117/12.539245>
- Freitas, M. & Simões, M. (2011). Perspetivas metodológicas inovadoras para a Ciência da informação: dois estudos desenvolvidos no âmbito do programa de doutoramento da Universidade de Salamanca. *Límites, Fronteras y Espacios Comunes: Encuentros y Desencuentros*. 438–448. <https://bit.ly/35CEKsD>
- Holme, P., & Saramäki, J. (2012). Temporal networks. *Physics Reports*, 519(3), 97-125. <https://doi.org/10.1016/j.physrep.2012.03.001>
- Hjorland, B., & Albrechtsen, H. (1995). Toward a New Horizon in Information-Science-Domain-Analysis. *Journal of the American Society for Information Science*, 46(6), 400-425. [https://doi.org/10.1002/\(SICI\)1097-4571\(199507\)46:6<400::AID-ASI2>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASI2>3.0.CO;2-Y)
- Hjorland, B. (2012). User-based and Cognitive Approaches to Knowledge Organization: A Theoretical Analysis of the Research Literature. *Knowledge Organization*, 40, 11-27
- Iles, P., Ramguttty-Wong, A., & Yolles, M. (2004). HRM and knowledge migration across cultures: Issues, limitations, and Mauritian specificities. *Employee Relations*, 26(6), 643-662. <https://doi.org/10.1108/01425450410562227>
- Klavans, R., & Boyack, K. W. (2007). Is there a convergent structure of science? A comparison of maps using the ISI and scopus databases. [Paper presentation]. *Proceedings of ISSI*. 437–448
- Lima, G. (2004). O modelo simplificado para análise facetada de Spiteri a partir de Ranganathan e do classification research group. *Information, Cultura y Sociedad*. 11. pp. 57-72
- Lima, V. (2013). A organização do conhecimento no domínio da Ciência da informação: o mapa conceitual e terminológico como instrumento referencial para o ensino e a pesquisa. InCID: *Revista de Ciência da informação e Documentação*, 4(1), pp. 26-48. <https://doi.org/10.11606/issn.2178-2075.v4i1p26-48>
- Lopes, P., Aganette, E. & Maculan, B. (2020). Taxonomia corporativa e taxonomia facetada: usos e aplicações da ciência de Informação no Brasil. *Investigación Bibliotecologica*, 34(82).
- Maculan, B. & Lima, G. (2011). Modelo para análise conceitual de teses e dissertações com vistas à criação de taxonomia facetada. *Informação & Sociedade: Estudos*.

- Medeiros, J (2013). *Taxonomia Navegacional Facetada: Análise à luz dos princípios teóricos da classificação facetada*. [dissertação de mestrado]. Universidade Federal de Santa Catarina
- Nascimento, A. (2015). *Análise dos catálogos da Estante Virtual e Cultura a partir dos Princípios Teóricos da Classificação Facetada* [dissertação de mestrado]. Universidade de Brasília. <https://bit.ly/3soBb3a>
- Newman, M. (2001). The structure of scientific collaboration networks. *PNAS*, 98(2). <https://www.pnas.org/content/98/2/404>
- Pinheiro, L. V., & Ferrez, H. (2014). *Tesouro Brasileiro de Ciência da Informação*. Instituto Brasileiro de Informação em Ciência e Tecnologia.
- Ranganatham, S. (1967). *The five laws of library science*. Mandras University.
- Spiteri, L. (1998). Simplified model for facet analysis: Ranganathan 101. *Canadian Journal of Information and Library Science*. 23, 1-30
- Williams, A., & Balaz, V. (2013). Tourism, Risk Tolerance and Competences: Travel Organization and Tourism Hazards. *Tourism Management*. <https://doi.org/10.1016/j.tourman.2012.07.006>

APÊNDICES

Tabela 1

Plano de Classificação constante do Tesauro Brasileiro de Ciência da Informação
8-Áreas do Conhecimento
7-Documento e informação como componente
7.1-Tipos de documento
7.2-Suportes da informação
7.3-Conteúdos da informação
6- Comunicação e acesso à informação
6.1-Comunicação científica
6.1.1-Produtividade, métricas
6.1.2- Publicações científicas
6.2- Transferência e acesso à comunicação
6.2.1-Propriedade intelectual
6.3-Indústria de informação
6.4-Sociedade da Informação
6.4.1-Literacia da informação
5-TIC
5.1-Equipamentos de computador
5.2-Programas de computador
5.3- Bases de dados e extração da informação
5.4- Redes, internet, web
5.5-Gestão das TIC
5.6-Inteligência artificial e engenharia do conhecimento
4-Informação e Conhecimentos Estratégicos nas organizações
4.1-Inteligência competitiva
4.1.1- Métodos de análise na inteligência competitiva
4.2-Gestão do conhecimento
3-Gestão da Informação

3.1-Gestão de bibliotecas e Recursos de Informação
3.1.1.-Serviços de Biblioteca
3.1.2- Desenvolvimento de Coleções
3.1.3-Preservação de documentos
3.2- Utilizadores e usos da informação
3.3-Serviços de informação
2-Organização do Conhecimento e Recuperação da Informação
2.1-Organização do Conhecimento
2.1.1-Representação da Informação
2.1.2-SOC
2.2-Recuperação da Informação
2.2.1-Avaliação de sistemas de RI
1-Epistemologia da CI
1.1-História da CI
1.2-Teorias da CI
1.3-Interdisciplinaridade
1.4-Métodos de pesquisa e análise
1.4.1-Métricas em CI
1.5-Pesquisa e Ensino em CI
1.6- Profissão e mercado de trabalho

**Tabela 2 – Documentos da CI acerca TNF (2011-2020)
recolhidos nas bases de dados**

Autoress	Título	Palavras-chave do autor	Ano de publicação
Smiraglia,RP; Henry,J; Milonas, E; Marchese,C; Zherebchevsky, S	A Formal Taxonomy of Knowledge Organization: Meta-Analysis and Facet Analysis	Knowledge organization, domain analysis, terms, taxonomy, classification	2020
Teixeira, RD; de Sousa, BP	The use of figures of speech in the lesbianism domain in Lesbian Herstory Archives’ photographic collection: a taxonomy proposal	Figures of speech, thematic representation, taxonomy, Lesbians, Lesbian herstory Archives	2020
Lopes, PTD; Aganette, EC; Maculan, BCMS	Corporate and faceted taxonomies: analysis of its uses and applications in the information science of Brazil	Corporate Taxonomy; Faceted Taxonomy; Uses and Applications; Information Science	2020
Rosenzweig, JW; Thill, M; Lambert, F	Student Constructions of Authority in the Framework Era: A Bibliometric Pilot Study Using a Faceted Taxonomy	Bibliometric study, Authority Faceted taxonomy	2019
Maculan, BCMS; Lima, GABO	Faceted taxonomy for accessing digital libraries	Faceted taxonomy; Facet analysis theory; Domain analysis theory; Content analysis; Categorical thematic analysis technique; Scientific communication; Information retrieval	2014
Pontes, FV; Lima, GABD	Knowledge organization in digital environments: faceted classification theory applied	Knowledge organization; Theory of faceted classification; Digital libraries; Faceted taxonomy	2012
Leeder, C; Markey, K; Yakel, E	A Faceted Taxonomy for Rating Student Bibliographies in an Online Information Literacy Game	Faceted taxonomy, literacy, online	2012
Maculan, BCMD; Lima, GABD	Faceted Taxonomy as a Mechanism for Browsing and Accessing Digital Libraries of theses and Dissertations: A Case Study	Faceted taxonomy, digital libraries, case study, theses and dissertations	2012
Maculan, BCMD; Lima, GABD	Model For Conceptual Analysis on Thesis and dissertations for the faceted taxonomy creation	Model Reading Technique; Conceptual Analysis; Document Indexing; Information Retrieval	2011

Rosenzweig, James; Thill, Mary; Lambert, Frank.	A model for analyzing and understanding how novice researchers construct source authority	Information literacy, faceted taxonomy, coding, citation analysis	2018
Silva, M. F	Proposta de modelo de colaboração para catálogo web facetado	Ferramentas de busca na Web, classificação facetada, ciência da informação, Sistemas de recuperação da informação	2013
Maculan, B	Taxonomia facetada navegacional: construção a partir de uma matriz categorial para trabalhos acadêmicos	Recuperação da informação, bibliotecas digitais, ferramentas de busca na web, ciência da informação, organização da informação	2011
Plentz, Samuel Sebben	Taxonomia para técnicas criativas aplicadas ao processo de projeto	Classificação facetada, projeto, técnicas criativas	2011
Cunha, José Humberto da Cruz	Estudo do emprego da taxonomia como instrumento auxiliar para decisões táticas no processo de auditoria	Taxonomia, auditoria, gestão de riscos, decisões	2017
Condotta M	A semiotic model as a conceptual link between conception and detailed design	Cognitive tools; Detailed design; Faceted Taxonomy; Linked Building Data; Semiotic model	2019
Litovkin D., Anikin A., Kultsova M., Sarkisova E	Representation of what-knowledge structures as ontology design patterns	Data description; Data mining; Inverse problems; Ontology; Semantics; Taxonomies; Concept definitions; Concept description; Description logic; Disjoint relation; Family resemblance; Knowledge structures; Scientific research; Semantic	2019
Pittore M., Haas M., Megalooikonomou K.G.	Risk-oriented, bottom-up modeling of building portfolios with faceted taxonomies	Building stock; Exposure; Multi-hazard; Seismic risk; Taxonomy; Vulnerability	2018
Scaturro I	Faceted taxonomies for the performing arts domain: The case of the European collected library of artistic	faceted taxonomy, arts domain, library	2013
Menard, E	Multilingual taxonomy development for ordinary images: Issues and challenges	taxonomy, image indexing, image retrieval, image description, controlled vocabularies,	2012

		bilingual environment, multilingualism, methodological approach, language issues, tagging	
Pontes, FV	A organização do conhecimento em ambientes digitais: aplicação da teoria da classificação facetada	faceted classification, digital libraries, information organization, library science, bibliographic classification, information retrieval, knowledge management, knowledge organization, faceted taxonomy	2012
Papadakos, P	On exploiting static and dynamically mined metadata for exploratory web searching	faceted taxonomies, dynamic taxonomies, information retrieval, results clustering, web searching	2012
Shiri, A	User evaluation of searching and Thesaurus: Multilingual Thesaurus-Enhanced Visual Interfaces for Digital libraries	digital libraries, thesaurus, user evaluation, model, dynamic taxonomies	2013
Gornstein, L	Information and library science, changes that influenced it's new character, direction and research: A bibliometric study, 1985-2006	evolution, bibliometric study	2013
Maculan, B	Faceted taxonomy for accessing digital libraries	digital libraries, algorithms, information storage, retrieval systems, academic dissertations, science publishing, categorical thematic analysis technique, content analysis, domain analysis theory, faceted analysis theory, faceted taxonomy, information retrieval, scientific communication	2014
Papadakos, P	Interactive explorations of multidimensional information spaces with preference support	information retrieval, querying, access to information, metadata, web search engines	2014
Lima, G., Maculan, B, Maia, L.	Taxonomia dos tipos de relações semânticas para a organização e a representação do conhecimento	relações Semânticas, Organização do Conhecimento, Representação do Conhecimento, Taxonomia de Relações Semânticas	2018

Loehrlein, A	The classification of financial products	Information resources management, action research, economics, international agencies, investments, hierarchies, finance, risk management, faceted classification	2014
Bedford, D	Understanding and managing taxonomies as economic goods and services	knowledge management, economics, semantics, subject heading, access to information, costs, economic of information, index language construction, information production, taxonomies	2014
Rosenzweig, J	A model for analyzing and understanding how novice researchers construct source authority	wikipedia, webometrics, information literacy, acquisition of data, citation analysis, education research, coding, faceted taxonomy	2018
Zong, N. Kim, H., Nam, S.	Constructing faceted taxonomy for heterogeneous entities based on object properties in linked data	Ontology learning, taxonomy construction, t-box learning, faceted taxonomy, linked data	2018
Wei, B., Liu, J., Ma, J., Zheng, Q., Zhang, W., & Feng, B.)	DFT-extractor: a system to extract domain-specific faceted taxonomies from wikipedia	information systems, data mining, software	2011
Putkey, T.	Using SKOS to Express Faceted Classification on the Semantic Web	faceted classification, semantic web, SKOS	2013
Morton, J. E., & Scherr, M. H.	Relevance ranked faceted metadata search method	faceted taxonomy, searching, application	2012
Dadić, J., Despotović-Zrakić, M., Barać, D., Paunović, L., & Labus, A.	Managing eGovernment Information Resources Using Faceted Taxonom	egovernment, dynamic taxonomy, information overload, decision making	2012
Zha, H., Shen, J., Li, K., Greiff, W., Vanni, M. T., Han, J., & Yan, X.	FTS: Faceted Taxonomy Construction and Search for Scientific Publications	Phrase Mining; Taxonomy; Text Categorization; Document Search	2018
Trudeau, C., & Guastavino, C.	Classifying soundscapes using a multifaceted taxonomy	information retrieval, faceted taxonomy, classification, soundscapes	2018
GAO, J. Z., & HE, F. J.	Model and Technique of Facet Search: An Overview	faceted search, model, SKOS, comparison	2012
Zolyomi, A.	Challenges of	Domain analysis,	2017

	Constructing a Multiple-Perspective Domain Analysis of Neurodiversity	neurodiversity, autismo	
Cavalcante, R. & Medeiros, M	Taxonomias navegacionais em sítios de comércio eletrônico: critérios para avaliação.	Comércio eletrônico; Navegação WEB; Organização da informação; Sistemas de organização do conhecimento; Taxonomia navegacional. Avaliação	2014
Medeiros, J.	Taxonomia Navegacional Facetada: Análise à luz dos princípios teóricos da classificação facetada	SOC, análise facetada, avaliação	2013

(Página deixada propositadamente em branco)

REPRESENTAÇÃO DA INFORMAÇÃO EM ARQUIVOS DESLOCADOS: UMA METASSÍNTESE DE LITERATURA QUALITATIVA

L. S. Ascensão de Macedo¹, Carlos Guardado da Silva²
e Maria Cristina Vieira de Freitas³

¹Centro de Estudos Interdisciplinares do Século XX, Universidade de Coimbra,
laureano.macedo@madeira.gov.pt, ORCID iD 0000-0001-7251-7314

²Centro de Estudos Clássicos, Faculdade de Letras, Universidade de Lisboa,
carlosguardado@campus.ul.pt, ORCID iD 0000-0003-1490-8709

³Centro de Estudos Interdisciplinares do Século XX, Faculdade de Letras,
Universidade de Coimbra, cristina.freitas@fl.uc.pt, ORCID iD 0000-0002-8849-8792

Resumo

Esta comunicação visa realizar uma síntese qualitativa de literatura relativa à representação da informação em arquivos deslocados. O conceito de “arquivo deslocado” é polissémico e integra diversos termos na literatura arquivística em distintas línguas (u. g., arquivo removido, expatriado, exilado, migrado, confiscado, contencioso arquivístico, etc.). Desta forma, o objetivo deste estudo consiste em realizar uma prospeção de literatura que informe sobre como a comunidade científica se debruçou sobre a representação da informação em arquivos deslocados. Pretende-se responder à seguinte questão de investigação: “Como as entidades arquivísticas têm representado os arquivos deslocados de acordo com a literatura científica produzida entre 1954 e 2019 no domínio da Ciência da Informação?”. Metodologicamente, esta comunicação encontra-se configurada numa metassíntese orientada à construção de teoria, constituindo-se num tipo de estudo de tipo não reativo, de base documental e exploratória, focalizada em artigos e capítulos em livros publicados em inglês entre 1954 a 2019. A recolha de textos é suportada pela estratégia SPICE, aplicada na pesquisa em bases de dados (WoS e EBSCO). Adotamos a análise de conteúdo de acordo com os pressupostos de Charmaz (2006) e de Finfgeld-Connett (2018). Dos 443 registos, que a pesquisa nos devolveu, incluímos 155 textos que responderam aos propósitos do estudo. Predominam estudos de caso, explícitos ou não, simples ou sobrepostos com outros métodos (análises teórico-conceptuais, métodos biográficos, retórica coletiva), análises históricas, análises teórico-conceptuais, métodos etnográficos, auto-etnográficos e investigação naturalística. Os métodos menos frequentes são a análise de conteúdo ou temática, a análise documental, o estudo longitudinal e o relatório de investigação. Os principais fatores incoativos de remoção de arquivos dos seus contextos originários, que emergiram desta metassíntese, derivam de fenómenos ocorridos em contexto de conflito armado, de descolonização, de sucessão de Estados e do tráfico ilícito de bens culturais. Além disto, emergiram novos

casos de arquivos constituídos na diáspora, casos de conflito de custódia em contexto subnacional e a dispersão de arquivos na web/nuvem. Esta diversidade também se manifesta em conjuntos documentais não só com valor histórico e/ou administrativo para as comunidades desapossadas, mas também nas suas múltiplas materialidades (analógico/digital, audiovisual, textual, iconográfico, etc.). Foram constituídos três temas que emergiram da análise de conteúdo em torno do tema suprarreferido: (1) “anarquívismo como (não) representação”: analisámos processos de representação anarquívica praticadas por algumas entidades custodiantes, através da manipulação do conteúdo por motivos políticos e ideológicos no sentido de obscurecer a proveniência ou de descaracterizar a ordem original (u. g. “higienização do passado colonial”, “troféus de guerra”, eufemização); (2) “arquivo do arquivo” (como “meta-arquivo” e “arqui-arquivo”): exploramos como os “arquivos dos arquivos” representam os conjuntos documentais removidos dos seus contextos originários, como forma de empoderar as entidades custodiantes através de mecanismos adotados para o controlo, gestão e determinação de políticas de acesso; e (3) “cânone arquivístico”: verificamos como as representações de arquivos deslocados em instrumentos de acesso visam legitimar cânones arquivísticos institucionais, com ênfase em torno de representações em ambiente digital. Em suma, embora os arquivos deslocados constituam um tema emergente em diversos domínios, importa explorar no âmbito da Organização do Conhecimento.

Palavras-chave: Representação da informação arquivística, arquivos deslocados, síntese qualitativa de literatura, metassíntese geradora de teoria, acesso à informação.

Introdução

Uma das condições necessárias para a identificação de arquivos deslocados consiste na existência de instrumentos de acesso à informação (doravante, IAI). Catálogos, inventários, guias, roteiros, listagens, índices e bases de dados constituem materializações genológicas de “instrumentos de acesso à informação” (Ribeiro, 1998). Para Lowry, “the catalogue is the key” (2017a, p. 8).

Além disto, o conceito de instrumento de acesso à informação encontra-se, também, associado a outro, o de “representação da informação arquivística”. Apesar de não constar na terminografia da especialidade de expressão portuguesa (Silva & Albuquerque, 2018), tem sido explorado no quadro teórico da Organização do Conhecimento (Aguiar & Kobashi, 2013; Barros & Martins, 2015; Silva & Albuquerque, 2018; Tognoli, 2012; Tognoli & Guimarães, 2012; Vital et al., 2017). Para Yakel, a representação arquivística deriva de uma “fluid, evolving, and socially constructed practice” (2003, p. 2), centrado em processos de organização e de descrição de conjuntos documentais com vista à produção de IAI, à produção e controlo de representações substitutas (u. g., micrografias e digitalizações) e de sistemas eletrónicos com vista a facilitar o acesso à informação.

Porém, no que diz respeito aos arquivos deslocados, o conceito também não se manifesta de forma clara na terminografia arquivística, por coexistirem diversas denominações para referir-se, em princípio, ao mesmo fenómeno (Chaterera-Zambuko, 2020; Lowry, 2017a; Macedo, 2019; Ngoepe & Netshakhuma, 2018). A definição mais recente consta no relatório de Lowry, apresentado ao Grupo de Peritos para o Património Arquivístico Partilhado do Conselho Internacional de

Arquivos (vulgo, EGSAH), definido como “archives removed from the place of their creation, where the ownership of the archives is disputed by two or more parties” (Lowry, 2020, p. 5).

Embora Winn (2015) considere que a ausência de IAI como um factor limitador para a identificação de arquivos na condição de deslocados, tais vínculos conceptuais não se manifestam, contudo, de modo evidente na literatura científica. Uma vez que alguns estudos sobre a representação da informação em IAI já evidenciaram que não existem descrições objetivas e imparciais sobre o conteúdo e a estrutura de um fundo ou sistema de informação (Cox, 2008; MacNeil, 2012), tais descrições arquivísticas sustentam-se em abordagens interpretativas que dependem das escolhas adotadas pelas entidades custodiantes, que não são alheias aos contextos políticos, históricos, sócio-culturais e institucionais do meio onde foram produzidos. De acordo com Ketelaar, “archival institutions, unlike libraries, do not publicly display their holdings to offer a panoptic view to their clients. But they do display the knowledge-power of the finding aids, as representation of what the public may not see openly, but may expect to find behind the closed doors of the prison-like repository” (Ketelaar, 2002, p. 147). Estes aspetos resultam ser pertinentes se aplicados na compreensão etiológica do fenómeno dos arquivos deslocados, especialmente como os IAI podem esclarecer ou obscurecer a proveniência deste tipo de recursos informacionais.

De acordo com o mapeamento de literatura realizado por Macedo (2019), verifica-se a existência de um número muito reduzido de artigos que se versaram em torno deste tópico (u. g., Grimsted, 1993a). Como tal, a representação da informação em torno dos arquivos deslocados concita ao seu aprofundamento.

Deste modo, propõe-se realizar uma metassíntese (Finfgeld-Connett, 2018; Grant & Booth, 2009; Sandelowski & Barroso, 2010), trazendo o fenómeno dos arquivos deslocados para o âmbito disciplinar da Organização do Conhecimento, de modo a poder-se explorar que relação existe entre os conflitos pela custódia de arquivos e as suas representações realizadas por entidades arquivísticas em torno dos fundos em disputa.

1. Materiais e métodos

1.1. Problema e objetivos de investigação

Este estudo consiste numa metassíntese, que é um tipo de síntese qualitativa de literatura obtido a partir de estudos de natureza qualitativa (Finfgeld-Connett, 2018; Finfgeld, 2003; Grant & Booth, 2009; Saini & Shlonsky, 2012; Thorne et al., 2004). Estabelecemos como enfoque dois tópicos entre si interrelacionados: (1) representação da informação arquivística e (2) arquivos deslocados.

Para formular a questão de investigação, adotamos a estratégia *SPICE* (Booth, 2006a; Booth et al., 2016), por consistir numa abordagem concebida no âmbito da Ciência da Informação e aplicada a estudos qualitativos. Como tal, a pergunta de investigação deve ser formulada em sentido amplo, podendo ser objeto de reformulação em função dos resultados obtidos (Finfgeld-Connett, 2018; Walsh & Downe, 2005). As especificações da questão de investigação encontram-se a seguir:

- *S(etting)*: instituições arquivísticas (instituições custodiantes);
- *P(erspective)*: representação da informação arquivística;
- *I(ntervention)*: arquivos deslocados;
- *C(omparison)*: IAI;
- *E(valuation)*: literatura científica no domínio da Ciência da Informação (1954-2019).

A questão de investigação pode ser formulada da seguinte forma: “Como as entidades arquivísticas [S] têm representado [P] arquivos deslocados [I] em IAI [C] de acordo com a literatura científica produzida entre 1954 e 2019 no domínio da Ciência da Informação [E]?”.

O objetivo geral deste estudo consiste em perscrutar na literatura científica perspectivas teóricas sobre como os arquivos deslocados têm sido representados, principalmente em IAI. Os objetivos específicos consistem em: (a) levantar a dimensão e as características dos artigos publicados entre 1954 e 2019 em pelo menos duas bases de dados académicas; (b) identificar no conteúdo dos textos incluídos evidências sobre a importância da representação da informação arquivística em IAI para a identificação de arquivos na condição de deslocados; e (c) avaliar questões sobre possibilidades e limitações na identificação de arquivos deslocados em IAI. O resultado que se espera desta metassíntese de literatura consiste em criar bases para a geração de nova teoria, estendendo as perspectivas existentes sobre a representação da informação arquivística aplicadas aos arquivos deslocados.

1.2. Estratégia de pesquisa

Combinando a abordagem *SPICE* (Booth et al., 2016, 2019) com a estratégia de reporte de pesquisa *STARLITE* (Booth, 2006b), a recolha de dados foi realizada de forma iterativa entre 2019-11-19 e 2020-04-21 nas bases de dados *WoS* e *EBSCO* (via *b-on*), por facilitar o uso de truncaturas em campos de título e de tópicos, expostos abaixo no Quadro 1.

Aplicámos descritores utilizados por Macedo (2019), tendo sido acrescentados outros, em inglês, como *colonial archives* (arquivos coloniais), *diasporic archives* (diáspora de arquivos), *seized archives* (confisco de arquivos), *looted archives* (arquivos saqueados) e *expatriated archives* (arquivos expatriados). Aplicámos a estes descritores o algoritmo de *stemming* de Porter para a língua inglesa (Porter, 1980; Porter & Boulton, 2002), para facilitar a geração de truncaturas a aplicar nos motores de pesquisa.

Quadro 1 – Estratégia de reporte de pesquisa STARLITE

Estratégia SPICE	WoS	EBSCO (via b-on)
S + P	Arquivos deslocados em instituições arquivísticas	
I + C	(TS=(archiv* AND (alien* OR captur* OR claim* OR *colon* OR custod* OR diaspor* OR displac* OR fugit* OR fugu* OR inalien* OR loot* OR migrat* OR remov* OR repatri* OR replev* OR restit* OR return* OR seiz*) AND record* AND (alien* OR captur* OR claim* OR *colon* OR custod* OR diaspor* OR displac* OR fugit* OR fugu* OR inalien* OR loot* OR migrat* OR remov* OR repatri* OR replev* OR restit* OR return* OR seiz*)) OR TI=(archiv* AND (alien* OR captur* OR claim* OR *colon* OR custod* OR displac* OR fugit* OR fugu* OR diaspor* OR displac* OR fugit* OR fugu* OR inalien* OR loot* OR migrat* OR remov* OR repatri* OR replev* OR restit* OR return* OR seiz*) AND record* AND (alien* OR captur* OR claim* OR *colon* OR custod* OR diaspor* OR displac* OR fugit* OR fugu* OR inalien* OR loot* OR migrat* OR remov* OR repatri* OR replev* OR restit* OR return* OR seiz*)))	SU (archiv* AND (alien* OR captur* OR claim* OR *colon* OR custod* OR displac* OR fugit* OR fugu* OR inalien* OR loot* OR migrat* OR remov* OR repatri* OR restit* OR return* OR seiz*)) OR SU (record* AND (alien* OR captur* OR claim* OR *colon* OR custod* OR displac* OR fugit* OR fugu* OR inalien* OR loot* OR migrat* OR remov* OR repatri* OR restit* OR return* OR seiz*)) AND TI (archiv* AND (alien* OR captur* OR claim* OR *colon* OR custod* OR displac* OR fugit* OR fugu* OR inalien* OR loot* OR migrat* OR remov* OR repatri* OR restit* OR return* OR seiz*)) OR TI (record* AND (alien* OR captur* OR claim* OR *colon* OR custod* OR displac* OR fugit* OR fugu* OR inalien* OR loot* OR migrat* OR remov* OR repatri* OR restit* OR return* OR seiz*))
E	AND IDIOMA: (English) AND TIPOS DE DOCUMENTO: (Article OR Book Chapter); Refinado por: CATEGORIAS DO WEB OF SCIENCE: (INFORMATION SCIENCE LIBRARY SCIENCE) AND TIPOS DE DOCUMENTO: (ARTICLE OR BOOK CHAPTER); Tempo estipulado: 1954-2019. Índices: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI. (n= 156)	Limitadores – Revistas Científicas (Analisadas pelos Pares); Data de Publicação: 19540101-20191231; Tipo de Publicação: Academic Journal; Tipo de Documento: Article, Book Chapter; Tipo de Publicação: Academic Journal; Tipo de Documento: Article, Book Chapter; Idioma: English; Modos de pesquisa – Booleana/Frase; Interface – EBSCOhost Research Databases; Ecrã e Pesquisa – Pesquisa Avançada; Base de dados – Library & Information Science Source; Library, Information Science & Technology Abstracts (n= 150)

Referência: Elaboração dos autores, modelo adaptado STARLITE (Booth, 2006b).

1.3. Recolha de dados e critérios de elegibilidade

O objetivo deste processo de recolha de dados consiste em reunir uma amostra de textos científicos que perfilhem um paradigma metodológico qualitativo ou mistos e que possibilite a construção de teoria (Finfgeld-Connett, 2018; Finfgeld-Connett &

Johnson, 2013). A construção de teoria é um método a partir do qual “the logic that is used to build the theory is made explicit and accessible to the user of the resulting theory” (Gilliland & McKemish, 2004, p. 155). O processo de recolha de dados parte de uma amostra teoricamente induzida. Três fases foram aplicadas, que a seguir se descrevem:

- **Fase 1: Pesquisa em bases de dados.** Realizamos pesquisas na *WoS* e *EBSCO*, acessíveis por subscrição institucional (Universidade de Coimbra). Cingimos a artigos publicados em revistas científicas e capítulos de monografias, revistos por pares, em inglês. Utilizamos a estratégia de pesquisa constante do Quadro 1. Todas as tipologias documentais que não respondam ao determinado são excluídas. Circunscrevemos as pesquisas no domínio da Ciência da Informação, designadamente em *Library and Information Science (WoS)* e em *Library & Information Science Source (LISS via EBSCO)* e *Library, Information Science & Technology Abstracts (LISTA via EBSCO)*. Os registos obtidos de cada uma das bases de dados foram exportados eletronicamente para o gestor bibliográfico *EndNote™*.
- **Fase 2: Seleção de textos.** Analisamos cuidadosamente os títulos, as palavras-chave (ou tópicos) e resumos em inglês, permitindo relizar uma triagem de textos suscetíveis de serem excluídos, dado serem os mais visíveis e por representarem o teor do seu conteúdo (Haggan, 2004). Além disto, adotamos uma estratégia de pesquisa expansiva recomendada por Fingeld-Connect (2018). Os critérios de seleção foram os seguintes:
 - (a) Artigos: para além dos artigos candidatos à inclusão, escrutinamos as respetivas referências bibliográficas, com base nos mesmos critérios de seleção adotados para esta fase, conferindo manual e iterativamente nas bases de dados *WoS*, *EBSCO* e *Google Scholar*, preferencialmente aqueles textos que se encontrem disponíveis, esgotadas as opções, de forma integral e dentro dos parâmetros base do **Quadro 1**.
 - (b) Capítulos de livro: incluímos os capítulos integrantes de *Displaced Archives* (Lowry, 2017b) e de *Archival returns* (Barwick, Green, & Vaarzon-Morel, 2019).
- **Fase 3: Processamento dos dados.** Uma vez feita esta verificação, os registos foram exportados a partir do *EndNote™* em formatos normalizados (*.ris*) e *outputs* armazenados na plataforma *Open Science Framework* como dados da investigação, acessível em < <https://osf.io/wmz78/>>. Os dados candidatos à análise foram importados para a ferramenta *ATLAS.ti cloud* (ATLAS.ti Scientific Software Development GmbH, 2020), uma plataforma *web* proprietária, que proporciona um fluxo de trabalho automatizado e colaborativo. Cada uma destas fases será representada num fluxo de reporte normalizado adaptado do PRISMA (Moher et al., 2009).

1.4. Critérios para a análise de dados

Para Fingeld-Connett (2018), a extração de dados de estudos qualitativos para o propósito de geração ou refinação de teoria consiste em dois tipos de análise: (a)

enfoque na extração de dados a partir dos resultados qualitativos; e (b) características dos estudos, em termos de objetivos, objeto de estudo e métodos. Para o processo e codificação, partimos da terminologia arquivística e das suas variantes, dado que já se dispõe de uma base conceptual sobre este fenómeno, permitindo transformar e gerar associações e interpretações que respondam à questão principal em estudo (Freitas, 2012). Optámos por uma vertente epistemológica construtivista e pragmática (Charmaz, 2008), pelo facto de ser mais flexível, menos prescritiva e mais aberta à interpretação (Belgrave & Seide, 2019). Sendo assim, aplicámos três etapas incrementais conforme Quadro seguinte:

Quadro 2 – Estratégia de reporte de pesquisa STARLITE

Etapas	Descrição
1 Codificação inicial	<ul style="list-style-type: none"> – leitura cuidada dos textos e codificação em <i>ATLAS.ti cloud</i> (ATLAS.ti Scientific Software Development GmbH, 2020); – sinalização de ocorrências de palavras em segmentos textuais <i>finding aid(s)</i> (IAI e suas variantes genológicas, como catálogos, inventários, etc.), <i>archival representation</i> (representação arquivística), <i>literary warrant</i> (garantias literárias); – produção de <i>memoranda</i> em torno de segmentos textuais que se relacionem com a representação da informação em casos de arquivos deslocados.
2 Codificação focalizada	<ul style="list-style-type: none"> – constituição de <i>clusters</i> de códigos por categorias; – categorização dos <i>memoranda</i> por grandes categorias relacionadas com o fenómeno de interesse.
3 Interpretação e síntese dos resultados	<ul style="list-style-type: none"> – produção de teoria com cunho interpretivista, recorrendo à comparação constante e à saturação na análise dos conceitos e das suas interrelações; – foi constituído <i>a priori</i> um quadro temático para guiar a constituição de categorias e subcategorias, associadas a mecanismos de representação da informação em arquivos deslocados: <ul style="list-style-type: none"> Remoção: códigos que representem mecanismos de (não) representação da informação arquivística independentemente dos fatores incoativos de remoção de arquivos e do destino que foi dado pelas partes intervenientes. Pode incidir no reuso ou na destruição de sistemas de representação da informação existentes ou elaborados para diversos fins, independentemente do estado do conjunto documental (total/parcial) quando do processo de remoção; Recuperação: códigos que representem estratégias de (não) recuperação da informação arquivística. Incide em mecanismos de (não) representação da informação arquivística e das vicissitudes custodiais dos conjuntos documentais. Integra representações sobre o estado/estatuto destes arquivos identificados <i>in situ</i> ou <i>ex situ</i> pelas partes (reivindicantes/custodiantes) e sobre as estratégias/mecanismos de (não) mediação da informação pelas entidades custodiantes; Reconstituição: códigos que representem estratégias de reconstituição da informação arquivística sobre conjuntos documentais removidos, dispersos ou restituídos às comunidades originárias. Incide em processos, técnicas, métodos ou formas de representação de conjuntos documentais com vista à restauração, reparação ou reunificação e pós-restituição.

Referência: Elaboração dos autores.

Para cada um dos temas expostos, constituímos *clusters*, que ajudarão a diferenciar em função dos contextos (conflito armado, descolonização, secessão de Estados, contexto subnacional e diáspora e *sitibus/extra situs*) e usos dos conjuntos documentais no processo de representação, agrupados nos seguintes subtemas iniciais: (a) IAI; (b) substituição de suporte; (c) acessibilidade e custódia; (d) classificação e descrição arquivística; (e) representação da integridade e organicidade; (f) garantias literárias; (g) representação de entidades produtoras/custodiantes. A constituição de categorias finais, como parte do processo de construção de teoria, ampara-se na perspectiva da “critical theory of displaced archives” (Lowry, 2019b, p. 350).

1.5. Validação, triangulação e reporte

Para o processo de triangulação, no sentido de reduzir distorções de interpretação por estarmos perante um estudo conduzido por três investigadores. Um dos membros da equipa realiza o processo de recolha de dados e processamento da codificação inicial. Dois dos membros efetuam a validação dos códigos atribuídos aos textos. Existência de ambiguidades e conflitos é resolvida por consenso entre os autores.

2. Resultados e discussão

2.1. Características gerais dos estudos

Dos 443 registos bibliográficos, 306 registos iniciais foram obtidos de acordo com os critérios de pesquisa do Quadro 1. Além disto, foram integrados manualmente 137 registos bibliográficos a partir da *WoS* e *EBSCO*, com base nas referências citadas dos primeiros resultados. Foram removidos 114 registos, entre duplicados, dados incompletos, tipologias documentais não requeridas e datas fora do âmbito da pesquisa. Foram excluídos mais 164 registos, através da leitura do título, palavras-chaves e resumos, por não responderem ao tópico de estudo, por ter devolvido tipologias documentais não requeridas, por não se dispor de acesso a texto integral e por a leitura integral dos artigos e dos capítulos de livro não terem devolvido os tópicos em estudo, durante o processo de codificação. Assim, a amostra foi reduzida a 155 artigos, produzidos entre 1962 e 2019 e textos em inglês.

A cobertura geográfica de estudos incluídos aponta para uma tendência crescente em torno de casos de arquivos deslocados tanto em contexto internacional como subnacional. A procedência da produção científica indica vários países de diversos continentes. Verifica-se um volume de estudos oriundos do continente americano ($n=80$), Oceania ($n=21$), Europa ($n=40$), África ($n=10$) e Ásia ($n=5$).

Quanto aos autores mais produtivos, refira-se Grimsted ($n=15$), Montgomery ($n=9$), Banton ($n=4$), Cox ($n=4$), Karabinos ($n=4$), Caswell ($n=3$), Kecskeméti ($n=3$), Lowry ($n=3$), Namhila ($n=3$), Punzalan ($n=3$), Anderson ($n=2$), Bastian ($n=2$), Frings-Hessami ($n=2$), Gilliland ($n=2$), Lovering ($n=2$), Phillips ($n=2$), Sela ($n=2$), Shepard ($n=2$). Sobre os restantes autores, apenas obtivemos uma única produção.

De acordo com os critérios de pesquisa utilizados, persiste a ausência de produção científica procedente de países, entre outros, de língua ibero-românica, em específico. Uma parte significativa dos registos recuperados sobre representação da informação arquivística foi excluída por não estar relacionada com o fenómeno de interesse. A Figura 3 sintetiza ao processo de seleção de estudos incluídos e excluídos.

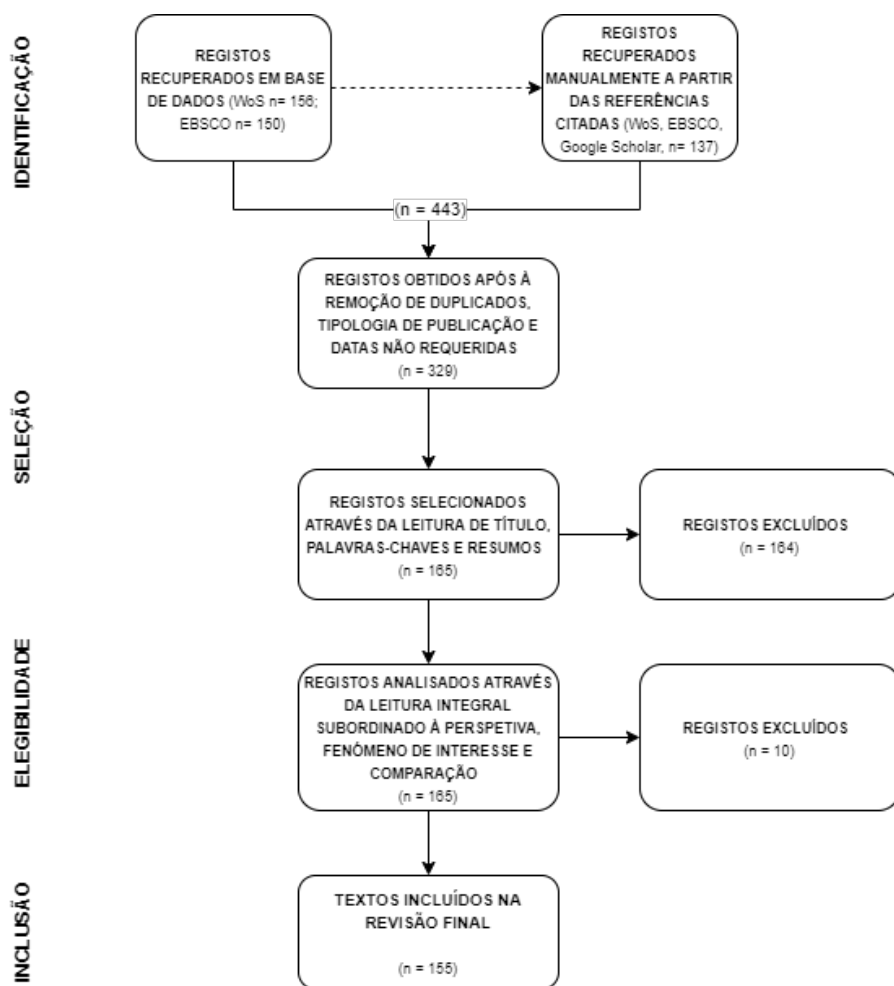


Figura 3. Fluxo de pesquisa e de seleção PRISMA

2.2. Descrição dos estudos incluídos

Foram incluídos 155 textos produzidos entre 1962 e 2019. Entre as décadas de 1960 a 2000, a produção de textos situa-se em 14% ($n= 21$), enquanto, a partir de 2001 a 2019, se assiste a um aumento significativo na ordem de 86% ($n= 134$).

Predominam estudos de caso, explícitos ou não, simples ou sobrepostos com outros métodos (análises teórico-conceituais, métodos biográficos, retórica coletiva)

($n= 50$); análises históricas ($n= 41$); análises teórico-conceituais ($n= 21$); métodos etnográficos, auto-etnográficos e investigação naturalística ($n= 17$); opinião, tomada de posição, ponto de vista ($n= 5$); revisão de literatura ($n= 3$); investigação-ação ($n= 3$); comentário de jurisprudência ($n= 3$); artigo de reação crítica ($n= 2$) e entrevista ($n= 2$). Muitos dos textos incluídos abordam contextos internacionais e subnacionais, como se poderá verificar no anexo <<https://osf.io/p786x/>>.

2.3. Descrição dos temas e apresentação da teoria

O anexo <<https://osf.io/eq7tp/>> categoriza os três eixos temáticos sobre o fenómeno de interesse desta síntese de literatura qualitativa. Estes eixos temáticos proporcionam a criação de categorias metafóricas, que apontam para motivações, estratégias e mecanismos de despossessão física e intelectual através da (não) representação da informação arquivística.

2.3.1. Construto teórico 1: Anarquismo como (não) representação

Constituímos esta categoria tendo como referência o conceito de “anarchive” de Derrida (1995b, 1995a), entendido como “la pulsion de mort, la violence de l’oubli, la *sur-répression*” (1995b, p. 126). Simplesmente, o *anarquismo* consiste em “destructive forces of the archive”, que podem ser representados como “indexes of loss and absence” (Huang, 2020).

A remoção de arquivos dos seus contextos originários não se reduz apenas a uma deslocação física, mas a sua remoção pode ter sido submetida a uma hermenêutica de representação *anarquívica*, especialmente como os traços ou vínculos dos arquivos gerados nos seus contextos originários foram removidos (ou destruídos) através de mecanismos de (não) representação da informação. Constituímos duas subcategorias interrelacionadas, que a seguir se explicam.

Reapropriação/Recetação. Esta subcategoria aponta para mecanismos de (não) representação da informação arquivística por via da remoção dos arquivos à entidade custodiante originária (instituição, comunidade ou indivíduos), tomando-os como se fossem próprios ou adquiridos por vias extra-legais ou ilegais (recetação).

Usamos aqui a reapropriação e a recetação como formas de apropriação de arquivos removidos por via da representação intelectual. Pode ser conferida em Frings-Hessami (2019a, 2019b), que expandiu o modelo de *records continuum* em “modelo de *continuum* de arquivo apropriado”, aplicado ao caso do arquivo do Museu do Genocídio Tuol Sleng dos Khmer Vermelho (Camboja), que foi “politically appropriated by the successor government” (Frings-Hessami, 2019a, p. 260) e reapropriado intelectualmente no sentido de “misrepresent the content of the archive” (Frings-Hessami, 2019a, p. 260). A reapropriação de arquivos em contexto de conflito armado é mais notória com o recurso às tecnologias de reprodução (micrografia/digitalização), não só para suportar a análise de inteligência militar, mas também como mecanismo fungível à repatriação física de arquivos e controlo (ou limitação) de acessos (Auer, 2017; Montgomery, 2001; Pennell, 2019). Quanto aos processos de descolonização ou de secessão/sucessão de Estados, verificam-se

abordagens distintas de reapropriação por via da remoção e da ocultação de arquivos (públicos ou privados) (Agostinho, 2019; Anderson, 2015; Banton, 2013; Karabinos, 2018; Sato, 2017), como o caso dos denominados “migrated archives”, arquivos removidos durante a descolonização britânica (Anderson, 2015; Badger, 2012; Bailkin, 2015; Banton, 2012b, 2012a, 2012c, 2013, 2017; Elkins, 2015; Hampshire, 2013; Hiribarren, 2017; Karabinos, 2018; Lowry, 2019a; Mnjama, 2011; Phillips, 2013, 2016; Sato, 2017).

Reconfiguração. Esta subcategoria encontra-se interrelacionada com a subcategoria anterior. A reconfiguração sobressai, também, como uma dimensão *anarquívica* da (não) representação de arquivos deslocados, especialmente quando o objetivo consiste em obscurecer o conteúdo dos arquivos através de mecanismos diversos de (não) representação da informação arquivística. Vários exemplos sustentam esta ideia, que se enumeram a seguir:

- **Representações eufémicas.** Os arquivos confiscados durante ou após os conflitos armados apoiaram-se em sistemas de classificação eufémicos, concebidos para obscurecer não só a proveniência dos arquivos, mas também o seu conteúdo (Montgomery, 2001; Sela, 2018a). Por exemplo, os arquivos iraquianos do partido Ba’ath, liderado por Saddam Hussein (1937-2006) recorriam a descritores eufemísticos para atenuar atos relacionados com o genocídio da comunidade curda na “Operação Anfal” (1986-1989). Os instrumentos de recuperação de informação iraquianos representavam tais atos como “‘collective measures’, a ‘return to the national ranks’ and the ‘resettlement in the south’” para referir-se a um tipo de “numbing bureaucratic language about ‘liquidations’, ‘expulsions’, and ‘transfers’ of the Kurdish victims” (Montgomery, 2001, pp. 88-89). Outra estratégia de eufemização consiste no tratamento dado aos arquivos capturados aos palestinianos pelas forças militares israelitas, as quais procuraram evitar a reutilização dos sistemas de representação “according to the criteria originally employed by the Palestinians” e, como tal, os arquivos capturados foram “catalogued according to Zionist terminology (...) with no relation to their Palestinian characteristics/cataloging” e “treating them as materials of minor cultural importance” (Sela, 2018b, p. 211). Acresce, além disto, outra estratégia de reconfiguração por eufemização que consiste no recurso à mudança de denominação oficial das instituições custodiantes contestadas como forma obscurecer a natureza dos fundos custodiados (Balcells, 2014; Grimsted, 2010a). Grimsted exemplificou um caso mediático que emergiu em 1990 relativo aos bens culturais confiscados aos nazis pelas forças soviéticas, como as autoridades russas “euphemistically renamed” (Grimsted, 1998, p. 6) ou “euphemistically rebaptized” (Grimsted, 2001, p. 192) o então “Arquivo Especial do Estado Central da URSS” para “Centro de Preservação de Coleções Histórico-Documentais” em 1992. O mesmo caso sucedeu em Espanha com os “Papeles de Salamanca”, que opõe as autoridades públicas da Comunidade Autónoma da Catalunha contra o arquivo nacional localizado em Salamanca, alterando a denominação do então *Archivo General de la Guerra Civil Española* para *Centro Documental de la Memoria Histórica* (Balcells, 2014).

- **Representações ‘contra gentes’.** Esta estratégia *anarquívica* de representação da informação ocorre quando os agentes da remoção se apoiam ou desenvolvem sistemas de representação arquivísticos baseados na discriminação segundo a etnicidade e/ou o credo religioso (Bailkin, 2015; Chilcott, 2019; Mnjama & Lowry, 2017; Montgomery, 2001; Sela, 2018b; Stoler, 2002) ou, ainda, de não representação de determinadas comunidades como forma de exclusão (Caswell, 2012a; Namhila, 2004, 2015). Sela exemplifica como “a longstanding Zionist tradition, a Palestinian is classified as an enemy and a terrorist”, ao ponto de “Palestinian settlements are referred to by Hebrew names to erase their Palestinian identity” (2018b, p. 211). Montgomery revelou como os sistemas de classificação arquivísticos iraquianos representaram os curdos “commonly referred to as ‘traitors’, ‘saboteurs’, ‘criminals’, or ‘human cargo’” (2001, pp. 88–89). Outro mecanismo de representação *contra gentes* foi demonstrado por Namhila, que identificou como arquivos das comunidades “nativas” têm sido preteridas por motivo de “a preconceived attitude that the “Natives” were not important” (2015, p. 4).
- **Representações pelos ‘archiuorum falsarii’.** A alteração deliberada da organicidade/integridade para obscurecer, ocultar ou descaracterizar os arquivos também pode ser verificada com a adulteração dos arquivos de Tuol Sleng dos Khmer Vermelho (Caswell, 2012a; Frings-Hessami, 2019a), onde as forças governamentais atuaram no sentido de expurgar peças potencialmente incriminatórias sobre atos de violação de direitos humanos e, desta forma, “the archive was taken out of its context, and its structure was changed” (Frings-Hessami, 2019a, p. 260). Esta abordagem está relacionada, também, com a instrumentalização da avaliação arquivística como mecanismo de “higienização” do passado colonial e dos crimes de guerra (Anderson, 2015; Elkins, 2015; Grimsted, 1992; Hampshire, 2013; Karabinos, 2018), através de destruições insuficientemente documentadas pelas entidades custodiantes ou baseados em critérios axiológicos conducentes à ocultação e eliminação (Elkins, 2015; Hampshire, 2013; Namhila, 2015; Sato, 2017). Desta forma, os IAI sobre arquivos removidos tendem a ser reconfigurados pelas entidades custodiantes com o intuito de restringir o acesso e possibilidade de identificação (Brower, 1963; Grimsted, 2010a; Montgomery, 2011, 2014, 2015b, 2017; Sela, 2018a).

2.3.2. Construto teórico 2: “Arquivo do arquivo” como “meta-arquivo” e “arqui-arquivo”

Embora os conceitos de *meta-archive* e *archi-archive* se encontrem atestados em Derrida (1995a, pp. 44, 45), estes têm sido escassamente desenvolvidos na literatura arquivística. A intenção epistemológica que nos motiva a adotar estes conceitos consiste na ideia de “arquivo do arquivo”, como base para a recuperação de informações com potencial reconstitutivo de arquivos removidos dos seus contextos de criação e de acumulação originários. Por esta razão, consideramos que “arquivos dos

arquivos” devem ser perspectivados em “how archives can be held accountable for their own actions” (Henry, 2009, p. 6). Constituímos duas categorias:

Restauração. Entre os aspetos mais importantes dizem respeito à existência de IAI (Grimsted, 2001; Winn, 2015). Trata-se de um aspecto fulcral seja em contexto de (pós-)conflito armado, seja (pós-) colonial, seja, ainda, na (pós-)secessão ou sucessão de Estados, contextos subnacionais e comunidades na diáspora. Os instrumentos de acesso permitem não só identificar arquivos que podem estar na condição de deslocados como também possibilitam aferir sobre o grau de dispersão dos fundos removidos dos contextos de produção e/ou de acumulação originários. Além disto, a disponibilidade dos instrumentos de acesso permite assegurar a continuidade de negócio e os direitos patrimoniais das comunidades sucessoras (Gilliland, 2018; Lemmon, 1992). O requisito fundamental destes instrumentos de acesso consiste no grau de atualização e nível de granularidade descritiva que atestem sobre o estado, o estatuto e a proveniência dos arquivos removidos dos seus contextos originários (Grimsted, 1982, 1991b, 1993a, 2017). Estas condições visam garantir a confiabilidade das descrições como instrumentos de *accountability* (Frings-Hessami, 2019a). Verifica-se que não existe absoluto consenso no conjunto de textos analisados quanto à fungibilidade da produção de representações substitutas (micrografias/digitalizações), em alternativa à restituição física dos arquivos (Frick, 2015; Kahunde, 2012; Karabinos, 2013, 2015; Montgomery, 2013; Peterson, 2000; Winn, 2015). Diversos autores confluem na ideia de que a produção de substitutos resolve apenas o problema do acesso ao conteúdo, mas não a custódia física dos arquivos removidos (Ashie-Nikoi, 2019; Banton, 2009; Boserup, 2005; Chebotarev, 2005; Cox, 2017; Danielson, 2013; Frick, 2015; Karabinos, 2017; Lyons & Sands, 2009; Montgomery, 2013; Niles, 2012; Peterson, 2000). Alguns autores sugerem que os “arquivos dos arquivos” são uma base fundamental para a reconstituição do destino dado aos arquivos que se encontram na condição de deslocados (Grimsted, 1992; Hampshire, 2013; Karabinos, 2013, 2017; Mattern, 2016).

Revidação. Esta subcategoria emergiu do conjunto de textos incluídos como consequência da recuperação. A (re)descoberta de conjuntos documentais deslocados junto de entidades custodiantes pode gerar um conjunto de críticas e linhas argumentativas relacionadas com a questão da legitimidade da custódia. O problema do acesso aos arquivos removidos às comunidades originárias levanta questões sobre a “ética da custódia” (Lowry, 2019b, p. 353) e sobre confiabilidade das instituições custodiantes contestadas pelas comunidades reivindicantes (Banton, 2012a; Caswell, 2013). Por exemplo, a inexistência de instrumentos de acesso (O’Neill, 1979) ou a sua imprecisão descritiva (Cox, 2010; Grimsted, 1971, 1991b; Namhila, 2004) ou provisória (Yeo, 2009) pode ter consequências negativas para as comunidades desapossadas, como, por exemplo, a resolução de casos na justiça por motivo de violação de direitos humanos (Anderson, 2011; Badger, 2012; Banton, 2012a; Lowry, 2019a). Arquivos confiscados em contexto de conflito armado tendem a ser representados pelas entidades custodiantes como “troféus de guerra” (Grimsted, 1998, 2001, 2010a, 2010b) enquanto outros revidam com antropomorfismos como “imprisoned photographs” (Sela, 2018a) ou “prisoners of war” (Grimsted, 2017). Às questões de acesso levantam outros imponderáveis relacionados com os arquivos

deslocados: o problema da propriedade e da custódia dos dados no ambiente de nuvem, devido ao modelo de multilocação, e a confiabilidade/autenticidade dos dados de arquivo (Goh, 2014); a obsolescência tecnológica (Treloyn et al., 2016); os custos da custódia quando da produção de substitutos (Grimsted, 1992); as classificações de segurança (Banton, 2017); as questões de propriedade intelectual (Heuman, 2013; Kahunde, 2012); as dificuldades de identificação de entidades sucessoras, para que os arquivos na condição de deslocados não terminem como “órfãos” (Grimsted, 2017), e as dificuldades na recuperação de arquivos na posse de entidades privadas (Dominy, 2013; O’Neill, 1979).

2.3.3. Construto teórico 3: Cânone arquivístico

O conceito de “cânone arquivístico” emergiu da análise do conjunto de textos incluídos como função teleológica da custódia de arquivos. As ocorrências do termo “archival canon” na literatura arquivística surgem por via da transposição de conceitos oriundos dos Estudos Literários. O conceito de cânone tem o sentido metafórico e metonímico de regra, medida, norma (Kolbas, 2001). A acepção que se adota aqui cinge-se ao cânone arquivístico institucional, como *corpus* documental derivado de um processo de escolha realizado por instituições arquivísticas, através de mecanismos de aquisição e de representação da informação, cuja consagração pode gerar reconhecimento ou contestação por parte das comunidades.

Harris acrescenta que “a canon is always in the process of being constructed, and the process expresses prevailing relations of power” (2004, p. 212). Matienzo (2008) estabelece uma relação entre canonização e o conceito de *archivalization* de Ketelaar (1999), com a diferença de a canonização “occur within an institutional context” (Matienzo, 2008). Constituímos duas subcategorias que a seguir se explanam.

Reunificação. A dispersão de arquivos do mesmo produtor entre diversas entidades custodiantes instiga a que se definam estratégias para a sua reunificação. A reunificação não se limita, contudo, à junção de artefactos dispersos entre entidades custodiantes, mas também à reunificação *com* as comunidades reivindicantes. A reunificação física constitui um dos mais complexos processos de resolução, dado que envolve questões patrimoniais que interferem no âmbito jurídico. Apesar de a reunificação física – pela via da restituição, repatriação, retorno ou relocação – constituir uma matéria que motiva a congregação de esforços entre os organismos internacionais e nacionais, Lowry (2017a, 2019b), Cox (2017), Kecskeméti (2017) e Montgomery (2012) realçaram que a via legal tem sido pouco eficaz. A outra via consiste na reunificação virtual (ou digital), que Punzalan define como “the strategy of putting together physically dispersed heritage collections in order to produce a consolidated, digitized representation of scattered artifacts” (2014a, p. 300). Como referido, a existência de IAI que representem arquivos na condição de deslocados constitui um requisito indispensável para o processo de reunificação tanto física como digital (Grimsted, 1987, 1991a, 1993b; Heverly, 2005; Kamba & Mazikana, 1988; Lemmon, 1992; Patkus, 2005). A existência de IAI não é, contudo, um elemento suficiente em si, pois as representações devem ser avaliadas em torno das suas (in)consistências descritivas em matéria de propriedade, proveniência e organi-

cidade (Blouin, 1992; Cox, 2011, 2014; Grimsted, 1991a; Lehane, 2012; Lovering, 2017; Punzalan, 2014b).

Revitalização. Os estudos sobre patrimônio cultural dão especial ênfase à importância da repatriação, da restituição ou do retorno como forma de revitalização das comunidades em situação de vulnerabilidade. A revitalização permite que as comunidades desapossadas recuperem e reinterpretem o(s) seu(s) cânone(s) arquivístico(s). Como apresentado acima, a remoção de arquivos dos seus contextos originários pode gerar contextos disruptivos no ecossistema social e cultural, como a desvitalização da diversidade cultural ou a dificuldade de assegurar o normal funcionamento das instituições (públicas ou privadas) e das comunidades desapossadas. A ausência de instrumentos de acesso sobre arquivos removidos dos seus contextos originários deixou diversas comunidades privadas da possibilidade de recorrerem à justiça para poderem ser ressarcidas de violações de que foram alvo ou de os seus direitos (de identidade, de propriedade, de acesso à informação, etc.) terem permanecido suspensos (Cox, 2014).

Para fazer face a isto, diversos autores apontaram para modelos participados na representação da informação arquivística (Gibson, 2019; Kimberly, 2019), como forma de desconstruir narrativas dominantes que representaram as comunidades desapossadas (Cooke, 2015; Crouch, 2010) e como estratégia para a revitalização de culturas em situação de vulnerabilidade (Barwick, Green, Vaarzon-Morel, et al., 2019; Ellis et al., 2019; Treloyn & Emberly, 2013).

Conclusões e implicações

Esta revisão metassintética possibilitou aferir sobre como o fenómeno dos arquivos deslocados desafia a Arquivística e a Ciência da Informação em termos epistemológicos e metodológicos. Pudemos verificar que qualquer uma das funções da Arquivística – como o acesso, a classificação ou a avaliação – pode obscurecer ou esclarecer a proveniência e a ordem original de arquivos na condição de deslocados.

Dos 155 textos analisados, podemos tecer, a título de síntese, as principais conclusões:

(i) **Remoção associada a fenómenos “anarquivísticos”.** A representação da informação sobre arquivos deslocados tende a torna-los reapropriados/recetados com vista à sua ocultação ou a serem reconfigurados, através da descaraterização da sua organicidade ou do obscurecimento da sua proveniência por meio de denominações eufemísticas ou discriminatórias.

(ii) **Recuperação relacionada com o “arquivo do arquivo”.** A literatura incluída indicou uma pista que aponta para o lugar dos “arquivos dos arquivos” (*arqui-arquivos* e *meta-arquivos*) na forma como estes influem na representação arquivística de arquivos removidos dos seus contextos territoriais e sociais originários. A restauração tende a representar arquivos removidos com recurso a IAI e às novas tecnologias de reprodução, produzidos pelos “arqui-arquivos”. Tal estratégia pode suscitar cenários de revidação pela forma como as entidades custodiantes estabelecem relações de poder sobre as comunidades desapossadas através de IAI.

(iii) **Reconstituição de “cânones arquivísticos”**. A constituição de cânones arquivísticos institucionais a partir dos arquivos removidos dos seus contextos sociais e territoriais originários podem ser reconstituídas dependendo da participação das comunidades desapossadas no processo de representação da informação. No entanto, é possível depreender que a restituição de arquivos pode corresponder a uma “desanonização” não só das representações atribuídas pelas entidades custodiantes, mas também do seu papel como instituições arquivísticas.

Esta metassíntese fornece evidências para a necessidade de se explorarem outras dimensões relacionadas com a representação da informação sobre arquivos deslocados. Considera-se que o conceito arquivo deslocado é ontológica e epistemologicamente pouco clarificador, sugerindo-se exploração conceptual em torno da natureza da contestação.

Entre as implicações práticas desta revisão sistemática, pudemos verificar que a existência de IAI, suportados ou não por garantias literárias para a descrição arquivística, não assegura a identificação inequívoca de arquivos na condição de deslocados.

Referências bibliográficas

Nota: as referências bibliográficas aqui listadas incluem o conjunto de títulos incluídos da pesquisa, acessível em <<https://osf.io/wmz78/>>.

- Agostinho, D. (2019). Archival encounters: rethinking access and care in digital colonial archives. *Archival Science*, 19(2), 141-165. <https://doi.org/10.1007/s10502-019-09312-0>
- Aguiar, F. L., & Kobashi, N. Y. (2013). Organização e representação do conhecimento: perspectivas de interlocação interdisciplinar entre ciência da informação e arquivologia. *XIV Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB 2013)*, *Enancib*. <http://hdl.handle.net/123456789/2421>
- Ahmed, S. (2018). Archives du Maroc? The official and alternative national archives of Morocco. *Archives and Manuscripts*, 46(3), 255-268. <https://doi.org/10.1080/01576895.2018.1558408>
- Amoros, L. G. (2019). The digital return of ILAM’s Zimbabwean recordings: revitalization of the sound archive through postcolonial engagement between ILAM and African universities. *Archives & Records*, 40(3), 281-290. <https://doi.org/10.1080/23257962.2018.1561362>
- Anderson, D. M. (2011). Mau Mau in the High Court and the «Lost» British Empire Archives: Colonial Conspiracy or Bureaucratic Bungle? *The Journal of Imperial and Commonwealth History*, 39(5), 699-716. <https://doi.org/10.1080/03086534.2011.629082>
- Anderson, D. M. (2015). Guilty Secrets: Deceit, Denial, and the Discovery of Kenya’s ‘Migrated Archive’. *History Workshop Journal*, 80, 142-160. <https://doi.org/10.1093/hwj/dbv027>
- Ashie-Nikoi, E. D. (2019). Ghana’s Cultural Records in Diaspora: Perspectives from Papers held at the Schomburg Center for Research in Black Culture, New York. *African Journal of Library, Archives and Information Science*, 29(2), 143-155.
- ATLAS.ti Scientific Software Development GmbH. (2020). *ATLAS.ti | Your Projects*. <https://cloud.atlasti.com/>
- Auer, L. (2017). Displaced Archives in the Wake of Wars. Em J. Lowry (Ed.), *Displaced Archives* (pp. 114-129). Routledge. <https://doi.org/10.4324/9781315577609-9>

- Badger, A. (2012). Historians, a legacy of suspicion and the «migrated archives». *Small Wars & Insurgencies*, 23(4-5, SI), 799-807. <https://doi.org/10.1080/09592318.2012.709761>
- Bailkin, J. (2015). Where Did the Empire Go? Archives and Decolonization in Britain. *The American Historical Review*, 120(3), 884-899. <https://doi.org/10.1093/ahr/120.3.884>
- Balcells, M. (2014). The Knife that Still Divides: The Archive of Salamanca and the Heritage of Spain's Civil War in the 21st century. *Arts and Social Sciences Journal*, 5(2), 1-5. <https://www.hilarispublisher.com/open-access/the-knife-that-still-divides-2151-6200.100086.pdf>
- Banton, M. (2009). «Expatriate» or «migrated» archives: the role of the UK archivist. *Archives*, 34(121), 14-24.
- Banton, M. (2012a). Destroy?'Migrate'? Conceal? British Strategies for the Disposal of Sensitive Records of Colonial Administrations at Independence. *The Journal of Imperial and Commonwealth History*, 40(2), 321-335.
- Banton, M. (2012b). «Lost» and «found»: the concealment and release of the Foreign and Commonwealth Office «migrated archives». *Government Recordkeeping in Sub-Saharan Africa2*, June. <https://doi.org/10.3828/comma.2012.1.04>
- Banton, M. (2012c). 'Lost' and 'found': the concealment and release of the Foreign and Commonwealth Office 'migrated archives'. *Comma*, 2012(1), 33-46. <https://doi.org/10.3828/comma.2012.1.04>
- Banton, M. (2013). Obscured if not hidden: records relating to Africa in the National Archives of the UK (TNA). *African Research & Documentation*, 123(123), 3-13. <https://www.africabib.org/rec.php?RID=391146432>
- Banton, M. (2017). Displaced Archives in The National Archives of the United Kingdom. Em J. Lowry (Ed.), *Displaced Archives* (pp. 41-59). Routledge. <https://doi.org/10.4324/9781315577609-4>
- Barrera, G. (2009). Of condors and judges: archival musings over a judicial investigation. *Archival Science*, 9(3-4), 203-214. <https://doi.org/10.1007/s10502-009-9109-y>
- Barros, T. H. B., & Martins, W. R. (2015). A informação orgânica enquanto um objeto interdisciplinar: as relações entre a Arquivística e a Ciência da Informação no âmbito da representação em arquivos. *Ágora*, 25(51), 132-149.
- Barwick, L., Green, J., Vaarzon-Morel, J., & Zissermann, K. (2019). Conundrums and consequences: Doing digital archival returns in Australia. Em L. Barwick, J. Green, & P. Vaarzon-Morel (Eds.), *Archival returns: Central Australia and beyond* (pp. 1-27). University of Hawai'i Press & Sydney University Press. <http://hdl.handle.net/10125/24875/>
- Barwick, L., Green, J., & Vaarzon-Morel, P. (2019). Archival returns : Central Australia and beyond. Em *LD&C Special Publication*. Sydney University Press. <https://nflrc.hawaii.edu/ldc/sp-18/>
- Bastian, J. (2001). A Question of Custody: The Colonial Archives of the United States Virgin Islands. *The American Archivist*, 64(1), 96-114. <https://doi.org/10.17723/aarc.64.1.h6k872252u2gr377>
- Bastian, J. (2006). Reading Colonial Records Through an Archival Lens: The Provenance of Place, Space and Creation. *Archival Science*, 6(3-4), 267-284. <https://doi.org/10.1007/s10502-006-9019-1>
- Battley, B. (2019). Archives as places, places as archives: doors to privilege, places of connection or haunted sarcophagi of crumbling skeletons? *Archival Science*, 19(1), 1-26. <https://doi.org/10.1007/s10502-019-09300-4>
- Belgrave, L. L., & Seide, K. (2019). Coding for Grounded Theory. Em A. Bryant & K. Charmaz (Eds.), *The SAGE handbook of current developments in grounded theory* (pp. 167-185). SAGE. <https://doi.org/10.4135/9781526485656>

- Belton, T. (2010). Custody, Control, and Confusion: Legal, Historical, and Territorial Aspects of Court Records in Ontario. *Archivaria*, 69, 9–21. [http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502989393\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502989393(=pt-pt&site=ehost-live&scope=site)
- Bernbaum, J. A. (1970). The Captured German Records: A Bibliographical Survey. *The Historian*, 32(4), 564–575. <https://doi.org/10.1111/j.1540-6563.1970.tb00378.x>
- Blouin, F. X. (1992). A Case for Bridging the Gap: The Significance of the Vatican Archives Project for International Archival Information Exchange. *American Archivist*, 55(1), 182–191. <https://doi.org/10.17723/aarc.55.1.j23815ng47ng7144>
- Booth, A. (2006a). Clear and present questions: formulating questions for evidence based practice. *Library Hi Tech*, 24(3), 355–368. <https://doi.org/10.1108/07378830610692127>
- Booth, A. (2006b). «Brimful of STARLITE»: toward standards for reporting. *J Med Lib Assoc*, 94(October), 421–430. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1629442&tool=pmcentrez&rendertype=abstract>
- Booth, A., Noyes, J., Flemming, K., Moore, G., Tunçalp, Ö., & Shakibazadeh, E. (2019). Formulating questions to explore complex interventions within qualitative evidence synthesis. *BMJ Global Health*, 4(Suppl 1), e001107–e001107. <https://doi.org/10.1136/bmjgh-2018-001107>
- Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review*. SAGE.
- Boserup, I. (2005). The Manuscript and the Internet: digital repatriation of cultural heritage. *IFLA Journal*, 31(2), 169+–169+. <https://doi.org/10.1177/0340035205054881>
- Bracknell, C., & Scott, K. (2019). Ever-widening circles: Consolidating and enhancing Wirlomin Noongar archival material in the community. Em L. Barwick, J. Green, & P. Vaarzon-Morel (Eds.), *Archival returns: Central Australia and beyond* (pp. 325–338). University of Hawai'i Press & Sydney University Press. <http://hdl.handle.net/10125/24890>
- Brenda, L. C., Toussaint, S., Meakins, F., & McConvell, P. (2019). «For the children...»: Aboriginal Australia, cultural access, and archival obligation. Em L. Barwick, J. Green, & P. Vaarzon-Morel (Eds.), *Archival returns: Central Australia and beyond* (pp. xi–xvii). University of Hawai'i Press & Sydney University Press. <http://hdl.handle.net/10125/24883>
- Brower, P. P. (1963). The United-States Army's Seizure and Administration of Enemy Records Up to World-War-II. *American Archivist*, 26(2), 191–207.
- Butler, B. (2009). «Othering» the archive-from exile to inclusion and heritage dignity: the case of Palestinian archival memory. *Archival Science*, 9(1-2), 57–69. <https://doi.org/10.1007/s10502-009-9095-0>
- Caswell, M. (2009). Irreparable Damage: Violence, Ownership, and Voice in an Indian Archive The Case of the Bhandarkar Oriental Research Institute and the Sambhaji Brigade. *LIBRI*, 59(1), 1–13. <https://doi.org/10.1515/libr.2009.001>
- Caswell, M. (2012a). Using classification to convict the Khmer Rouge. *Journal of Documentation*, 68(2), 162–184. <https://doi.org/10.1108/00220411211209177>
- Caswell, M. (2012b). Using classification to convict the Khmer Rouge. *Journal of Documentation*, 68(2), 162–184. <https://doi.org/10.1108/00220411211209177>
- Caswell, M. (2013). Rethinking Inalienability: Trusting Nongovernmental Archives in Transitional Societies. *American Archivist*, 76(1), 113–134. [http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=88067414\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=88067414(=pt-pt&site=ehost-live&scope=site)
- Charmaz, K. (2008). Grounded theory as an emergent method. Em S. N. Hesse-Biber & P. Leavy (Eds.), *Handbook of emergent methods* (pp. 155–170). Guilford Press.
- Chaterera-Zambuko, F. (2020, Outubro). *Looking back, looking forward: In quest of closure on issues of displaced archives*. ICABlog. <https://blog-ica.org/2020/10/16/looking-back-looking-forward-in-quest-of-closure-on-issues-of-displaced-archives/>

- Chebotarev, T. (2005). Repatriation of the Bakhmeteff Archive: Russian Dreams and American Reality. *RBM: A Journal of Rare Books, Manuscripts, & Cultural Heritage*, 6(1), 44-51. [http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502953136\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502953136(=pt-pt&site=ehost-live&scope=site)
- Chilcott, A. (2019). Towards protocols for describing racially offensive language in UK public archives. *Archival Science*, 19(4), 359–376.
- Christen, K. (2011). Opening Archives: Respectful Repatriation. *The American Archivist*, 74(1), 185-210. <https://doi.org/10.17723/aarc.74.1.4233nv6nv6428521>
- Coeuré, S. (2017). Cultural Looting and Restitution at the Dawn of the Cold War: The French Recovery Missions in Eastern Europe. *Journal of Contemporary History*, 52(3), 588-606. <https://doi.org/10.1177/0022009416658700>
- Cooke, P. (2015). A Response to Sylvia Nannyonga-Tamusuza and Andrew N. Weintraub's "The Audible Future: Reimagining the Role of Sound Archives and Sound Repatriation in Uganda". *Ethnomusicology*, 59(3), 475-479.
- Cox, D. (2010). Archives and records in armed conflict: international law and the current debate over iraqi records and archives. *Catholic University Law Review*, 59(4), 1001-1056.
- Cox, D. (2011). "Inalienable" archives: Korean royal archives as French property under international law. *International Journal of Cultural Property*, 18(04), 409-423. <https://doi.org/http://dx.doi.org/10.1017/S0940739111000245>
- Cox, D. (2014). The Lost Archives of Noriega: Emancipating Panamanian Human Rights Documents in U.S. Military Custody. *Boston University International Law Journal*. <https://doi.org/10.4337/9781845428297.00102>
- Cox, D. (2017). Revisiting the Law and Politics of Compromise. Em J. Lowry (Ed.), *Displaced Archives* (pp. 196-214). Routledge. <https://doi.org/10.4324/9781315577609-13>
- Cox, R. (2008). Revisiting the Archival Finding Aid. *Journal of Archival Organization*, 5(4), 5-32. <https://doi.org/10.1080/15332740802153245>
- Crouch, M. (2010). Digitization as Repatriation? The National Museum of the American Indian's Fourth Museum Project. *Journal of Information Ethics*, 19(1), 45-56.
- Curran, G. (2019). Incorporating archival cultural heritage materials into contemporary Warlpiri women's yawulyu spaces. Em L. Barwick, J. Green, & P. Vaarzon-Morel (Eds.), *Archival returns: Central Australia and beyond* (pp. 91-110). University of Hawai'i Press & Sydney University Press. <http://hdl.handle.net/10125/24879>
- Danielson, E. S. (2013). Archives and the Ethics of Replevin. *Journal of Information Ethics*, 22(2), 110-136.
- Derrida, J. (1995a). Archive Fever: A Freudian Impression. *Diacritics*, 25(2), 9-63. <http://www.jstor.org/stable/465144>
- Derrida, J. (1995b). *Mal d'archive : une impression freudienne*. Galilée.
- Dominy, G. (2013). Overcoming the apartheid legacy: the special case of the Freedom Charter. *Archival Science*, 13(2-3, SI), 195-205. <https://doi.org/10.1007/s10502-012-9182-5>
- Dritsas, L., & Haig, J. (2014). An archive of identity: the Central African Archives and Southern Rhodesian history. *Archival Science*, 14(1), 35-54. <https://doi.org/10.1007/s10502-013-9204-y>
- Edelstein, J. (2017). Reconnecting with a fugitive collection: a case study of the records of JDC's Warsaw Office, 1945-1949. *Jewish Culture & History*, 18(1), 109-117. [http://search.ebscohost.com/login.aspx?direct=true&db=asn&AN=120931968\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=asn&AN=120931968(=pt-pt&site=ehost-live&scope=site)

- Elkins, C. (2015). Looking beyond Mau Mau: Archiving Violence in the Era of Decolonization. *American Historical Review*, 120(3), 852-868. <https://doi.org/10.1093/ahr/120.3.852>
- Ellis, E. M., Green, J., & Kral, I. (2019). i-Tjuma: The journey of a collection – from documentation to delivery. Em L. Barwick, J. Green, & P. Vaarzon-Morel (Eds.), *Archival returns: Central Australia and beyond* (pp. 303-323). University of Hawai'i Press & Sydney University Press. <http://hdl.handle.net/10125/24889>
- Finfgeld-Connett, D. (2018). *A Guide to Qualitative Meta-Synthesis*. Routledge. <https://doi.org/10.4324/9781351212793>
- Finfgeld-Connett, D., & Johnson, E. D. (2013). Literature search strategies for conducting knowledge-building and theory-generating qualitative systematic reviews. *Journal of Advanced Nursing*, 69(1), 194-204. <https://doi.org/10.1111/j.1365-2648.2012.06037.x>
- Finfgeld, D. L. (2003). Metasynthesis: The state of the art – So far. *Qualitative Health Research*, 13(7), 893-904. <https://doi.org/10.1177/1049732303253462>
- Freitas, M. C. V. de. (2012). Grounded Theory como método de investigação em Arquivologia: subsídios teóricos e práticos. Em M. L. P. Valentim (Ed.), *Estudos avançados em Arquivologia* (pp. 107-134). Oficina Universitária, Cultura Acadêmica. https://www.marilia.unesp.br/Home/Publicacoes/estudos_avancados_arquivologia.pdf
- Frick, C. (2015). Repatriating American film heritage or heritage hoarding? Digital opportunities for traditional film archive policy. *Convergence: The International Journal of Research into New Media Technologies*, 21(1), 116–131. <https://doi.org/10.1177/1354856514560999>
- Frings-Hessami, V. (2019a). Khmer Rouge archives: appropriation, reconstruction, neo-colonial exploitation and their implications for the reuse of the records. *Archival Science*. <https://doi.org/10.1007/s10502-019-09301-3>
- Frings-Hessami, V. (2019b). The flexibility of the records continuum model: a response to Michael Karabinos' "in the shadow of the continuum". *Archival Science*, 1-14. <https://doi.org/10.1007/s10502-019-09316-w>
- Garaba, F. (2011). Provenance, identification, restitution and management of the liberation struggle heritage in the ESARBICA region. *Journal of the South African Society of Archivists*, 44, 26-42. <https://doi.org/10.13140/2.1.4272.7041>
- Gibson, J. (2019). Returning recordings of songs that persist: The Anmatyerr traditions of akiw and anmanty. Em L. Barwick, J. Green, & P. Vaarzon-Morel (Eds.), *Archival returns: Central Australia and beyond* (pp. 65-89). University of Hawai'i Press & Sydney University Press. <http://hdl.handle.net/10125/24878>
- Gibson, J., Angeles, S., & Liddle, J. (2019). Deciphering Arrernte archives: The intermingling of textual and living knowledge. Em L. Barwick, J. Green, & P. Vaarzon-Morel (Eds.), *Archival returns: Central Australia and beyond* (pp. 29-45). University of Hawai'i Press & Sydney University Press. <http://hdl.handle.net/10125/24876>
- Gilliland, A. J. (2017). Networking Records in Their Diaspora: A Reconceptualization of «Displaced Records» in a Postnational World. Em J. Lowry (Ed.), *Displaced archives* (pp. 180-195). Routledge.
- Gilliland, A. J. (2018). To what lengths the 'Physical and Moral Defence of the Record' in times of conflict and exigency? *Archives & Records*, 39(2), 117-138. <https://doi.org/10.1080/23257962.2017.1348940>
- Gilliland, A. J., & McKemmish, S. (2004). Building an infrastructure for archival research. *Archival Science*, 4(3-4), 149-197. <https://doi.org/10.1007/s10502-006-6742-6>
- Glenn, B. (1962). Private records seized by the United-States in wartime – their legal status. *The American Archivist*, 25(4), 399-405.

- Goh, E. (2014). Clear skies or cloudy forecast? Legal challenges in the management and acquisition of audiovisual materials in the cloud. *Records Management Journal*, 24(1), 56-73. <https://doi.org/10.1108/RMJ-01-2014-0001>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91-108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Grimsted, P. K. (1971). Archives in the Soviet Union: Their Organization and the Problem of Access. *The American Archivist*, 34(1), 27-41. <https://doi.org/10.17723/aarc.34.1.01416hv158h31057>
- Grimsted, P. K. (1982). The fate of early records in Lviv archives – documentation from Western Ukraine under Polish rule (15th-century to 1772). *Slavonic and East European Review*, 60(3), 321-346. <https://doi.org/10.2307/4208539>
- Grimsted, P. K. (1987). The archival legacy of Soviet Ukraine: Problems of tracing the documentary records of a divided nation. *Cahiers du Monde Russe*, 28(1), 95-108. https://www.persee.fr/doc/cmr_0008-0160_1987_num_28_1_2103
- Grimsted, P. K. (1991a). The Fate of Ukrainian Cultural Treasures During World War II: The Plunder of Archives, Libraries, and Museums under the Third Reich. *Jahrbücher für Geschichte Osteuropas*, 39(1), 53-80. <https://doi.org/10.2307/41048536>
- Grimsted, P. K. (1991b). «Perestroika» in the Archives? Further Efforts at Soviet Archival Reform. *The American Archivist*, 54(1), 70-95. <https://doi.org/10.17723/aarc.54.1.60h14p374i379289>
- Grimsted, P. K. (1992). Beyond Perestroika – Soviet-area archives after the August Coup. *The American Archivist*, 55(1), 94-124. <https://www.jstor.org/stable/40293629%0A>
- Grimsted, P. K. (1993a). Archival rossica/sovietica abroad – provenance or pertinence, bibliographic and descriptive needs. *Cahiers du Monde Russe*, 34(3), 431-479. https://www.persee.fr/doc/cmr_0008-0160_1993_num_34_3_2364
- Grimsted, P. K. (1993b). Russian Archives in Transition: Caught Between Political Crossfire and Economic Crisis. *The American Archivist*, 56(4), 614-662. <https://doi.org/10.17723/aarc.56.4.40u60mg4238475r6>
- Grimsted, P. K. (1998). «Trophy» archives and non-restitution: Russia's cultural «cold war» with the European community. Em *Problems of Post-Communism* (Vol. 45, Número 3, pp. 3-16). M.E. Sharpe Inc. <https://doi.org/10.1080/10758216.1998.11655785>
- Grimsted, P. K. (2001). Twice plundered or «twice saved»? Identifying Russia's «trophy» archives and the loot of the (Nazi) Reich-Security-Main-Office. *Holocaust and Genocide Studies*, 15(2), 191-244. <https://doi.org/10.1093/hgs/15.2.191>
- Grimsted, P. K. (2005). Roads to Ratibor: Library and archival plunder by the Einsatzstab Reichsleiter Rosenberg. *Holocaust and Genocide Studies*, 19(3), 390-458. <https://doi.org/10.1093/hgs/dci041>
- Grimsted, P. K. (2006). The postwar fate of Einsatzstab Reichsleiter Rosenberg archival and library plunder, and the dispersal of ERR records. *Holocaust and Genocide Studies*, 20(2), 278-308. <https://doi.org/10.1093/hgs/dcl005>
- Grimsted, P. K. (2010a). Legalizing “Compensation” and the Spoils of War: The Russian Law on Displaced Cultural Valuables and the Manipulation of Historical Memory. *International Journal of Cultural Property*, 17(2), 217-255. <https://doi.org/10.1017/S094073911000010X>
- Grimsted, P. K. (2010b). Why do captured archives go home? Restitution achievements under the Russian law. *International Journal of Cultural Property*, 17(02), 291-333. <https://doi.org/10.1017/S0940739110000123>

- Grimsted, P. K. (2017). Pan-European Displaced Archives in the Russian Federation: Still Prisoners of War on the 70th Anniversary of V-E Day. Em J. Lowry (Ed.), *Displaced Archives* (pp. 130-156). Routledge. <https://doi.org/10.4324/9781315577609-10>
- Haggan, M. (2004). Research paper titles in literature, linguistics and science: Dimensions of attraction. *Journal of Pragmatics*, 36(2), 293-317. [https://doi.org/10.1016/S0378-2166\(03\)00090-0](https://doi.org/10.1016/S0378-2166(03)00090-0)
- Halilovich, H. (2014). Reclaiming erased lives: archives, records and memories in post-war Bosnia and the Bosnian diaspora. *ARCHIVAL SCIENCE*, 14(3-4, SI), 231-247. <https://doi.org/10.1007/s10502-014-9227-z>
- Hampshire, E. (2013). 'Apply the Flame More Searingly': The Destruction and Migration of the Archives of British Colonial Administration: A Southeast Asia Case Study. *The Journal of Imperial and Commonwealth History*, 41(2), 334-352. <https://doi.org/10.1080/03086534.2013.799349>
- Harris, V. (2004). Concerned with the writings of others: Archival canons, discourses and voices. *Journal of the Society of Archivists*, 25(2), 211-220. <https://doi.org/10.1080/0037981042000271529>
- Hennerbichler, F., & Montgomery, B. P. (2015). U.S. Restitution of the Iraq Secret Police Files from Saddam Hussein's Regime Regarding the Kurds in Iraq. *Advances in Anthropology*, 5, 31-38.
- Henry, C. M. (2009). *Toward the Archives of Archives: The New Archival History, Accountability and the Documentation of Archival Appraisal*. <https://mspace.lib.umanitoba.ca/xmlui/handle/1993/21532>
- Heuman, J. (2013). Unpacking the gift: negotiating property and custody in archival hardcore blogs. *Popular Music*, 32(2), 177-195. <https://doi.org/10.1017/S026114301300010X>
- Heverly, W. G. (2005). Virtual Repatriation: The Pittsburgh-Konstanz Archival Partnership. *RBM: A Journal of Rare Books, Manuscripts, & Cultural Heritage*, 6(1), 34-43. <http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502953143> (<=pt-pt&site=ehost-live&scope=site
- Hilder, T. R. (2012). Repatriation, Revival and Transmission: The Politics of a Sami Musical Heritage. *Ethnomusicology Forum*, 21(2, SI), 161-179. <https://doi.org/10.1080/17411912.2012.689473>
- Hill, C. G. (2017). Seeds as Ancestors, Seeds as Archives: Seed Sovereignty and the Politics of Repatriation to Native Peoples. *American Indian Culture and Research Journal*, 41(3, SI), 93-112. <https://doi.org/10.17953/aicrj.41.3.hill>
- Hiribarren, V. (2017). Hiding the Colonial Past? A Comparison of European Archival Policies. Em J. Lowry (Ed.), *Displaced Archives* (pp. 74-85). Routledge. <https://doi.org/10.4324/9781315577609-6>
- Huang, C. (2020). Dwelling on the "anarchival": archives as indexes of loss and absence. *Archival Science*, 1-15. <https://doi.org/10.1007/s10502-020-09333-0>
- Iacovino, L. (2010). Rethinking archival, ethical and legal frameworks for records of Indigenous Australian communities: a participant relationship model of rights and responsibilities. *Archival Science*, 10(4), 353-372. <https://doi.org/10.1007/s10502-010-9120-3>
- Irvine, O. U. (2009). The law and ethics of acquisition of expatriate archives: addressing the «lack of guidelines». *Archives*, 34(121), 6-13. <http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=48965709> (<=pt-pt&site=ehost-live&scope=site
- Kahunde, S. (2012). Repatriating Archival Sound Recordings to Revive Traditions: The Role of the Klaus Wachsmann Recordings in the Revival of the Royal Music of Bunyoro-Kitara, Uganda. *Ethnomusicology Forum*, 21(2, SI), 197-219. <https://doi.org/10.1080/17411912.2012.689471>

- Kamba, A. S., & Mazikana, P. C. (1988). Archive repatriation in southern Africa. *records of the Federation of Rhodesia and Nyasaland*, 4, 79-85. <https://doi.org/10.1177/0266666698800400202>
- Karabinos, M. (2013). Displaced Archives, Displaced History: Recovering the Seized Archives of Indonesia. *Bijdragen tot de taal-, land- en volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia*, 169(2-3), 279-294. <https://doi.org/10.1163/22134379-12340027>
- Karabinos, M. (2015). The Djogdja Documenten: the Dutch-Indonesian relationship following independence through an archival lens. *Information & Culture*, 50(3), 372-391. <https://doi.org/10.7560/IC50304>
- Karabinos, M. (2017). Indonesian National Revolution Records in the National Archives of the Netherlands. Em J. Lowry (Ed.), *Displaced Archives* (pp. 60–73). Routledge. <https://doi.org/10.4324/9781315577609-5>
- Karabinos, M. (2018). In the shadows of the continuum: testing the records continuum model through the Foreign and Commonwealth Office ‘Migrated Archives’. *Archival Science*, 18(3), 207-224. <https://doi.org/10.1007/s10502-018-9292-9>
- Kecskeméti, C. (1985). Contested records – The Legal Status of National Archives. *UNESCO COURIER*, 2, 9-11.
- Kecskeméti, C. (1992). Displaced european archives – is it time for a postwar settlement. *The American Archivist*, 55(1), 132-140. <https://doi.org/10.17723/aarc.55.1.g15783174577780t>
- Kecskeméti, C. (2017). Archives Seizures: The Evolution of International Law. Em J. Lowry (Ed.), *Displaced Archives* (pp. 12-20). Routledge. <https://doi.org/10.4324/9781315577609-2>
- Ketelaar, E. (1999). Archivalisation and Archiving. *Archives & Manuscripts*, 27(1), 54–61. <https://doi.org/10.1080/01576895.2018.1410603>
- Ketelaar, E. (2002). Archival temples, archival prisons: Modes of power and protection. *Archival Science*, 2(3-4), 221-238. <https://doi.org/10.1007/BF02435623>
- Kimberly, C. (2019). «The songline is alive in Mukurtu»: Return, reuse, and respect. Em L. Barwick, J. Green, & P. Vaarzon-Morel (Eds.), *Archival returns: Central Australia and beyond* (pp. 153-172). University of Hawai'i Press & Sydney University Press. <http://hdl.handle.net/10125/24882>
- Kolbas, E. D. (2001). *Critical theory and the literary canon*. Westview Press.
- Kratz, J. (2014). «PS: You had better remove the records”». *Prologue: Quarterly of the National Archives and Records Administration*, 46(2), 36-44.
- Lancefield, R. C. (1998). Musical Traces’ Retraceable Paths: The Repatriation of Recorded Sound. *Journal of Folklore Research*, 35(1), 47-68. <https://doi.org/10.2307/3814785>
- Lehane, R. (2012). Documenting sites of creation. *Archives & Manuscripts*, 40(3), 171-180. <https://doi.org/10.1080/01576895.2012.738008>
- Lemmon, A. E. (1992). The Archival Legacy of Spanish Louisiana’s Colonial Records. *American Archivist*, 55(1), 142-155.
- Lovering, T. (2009). Expatriate Archives. *Archives: The Journal of the British Records Association*, XXIV(121), 1-5.
- Lovering, T. (2017). Expatriate Archives Revisited. Em J. Lowry (Ed.), *Displaced Archives* (pp. 86-100). Routledge. <https://doi.org/10.4324/9781315577609-7>
- Lowry, J. (2017a). Introduction: displaced archives. Em J. Lowry (Ed.), *Displaced archives* (pp. 1-9). Routledge. <https://doi.org/10.4324/9781315577609-1>
- Lowry, J. (Ed.). (2017b). *Displaced archives*. Routledge.
- Lowry, J. (2019a). Radical empathy, the imaginary and affect in (post)colonial records: how to break out of international stalemates on displaced archives. *Archival Science*, 19(2, SI), 1-19. <https://doi.org/10.1007/s10502-019-09305-z>

- Lowry, J. (2019b). "Displaced archives": proposing a research agenda. *Archival Science*, 19(4), 349-358. <https://doi.org/10.1007/s10502-019-09326-8>
- Lowry, J. (2020). *Disputed Archival Claims: An International Survey (2018/2019): Report to the International Council on Archives' Expert Group on Shared Archival Heritage*. https://www.ica.org/sites/default/files/disputed_archival_claims_an_international_survey_2018_2019.pdf
- Lyons, B., & Sands, R. M. (2009). A Working Model for Developing and Sustaining Collaborative Relationships Between Archival in the Caribbean and the United States. *IASA Journal*, 32, 26-37. [http://search.ebscohost.com/login.aspx?direct=true&db=lxh&AN=37145297\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=lxh&AN=37145297(=pt-pt&site=ehost-live&scope=site)
- Maaba, B. B. (2008). Challenges to Repatriation and Preservation of Tangible Heritage in South Africa: Black Art and the Experiences of the Ifa Lethu Foundation. *South African Historical Journal*, 60(3), 500-513. <https://doi.org/10.1080/02582470802417557>
- Macedo, L. S. A. de. (2019). Arquivos deslocados: mapeamento de literatura. *Brazilian Journal of Information Science*, 13(4), 5-34. <https://doi.org/10.36311/1981-1640.2019.v13n4.02.p5>
- MacNeil, H. (2012). What finding aids do: archival description as rhetorical genre in traditional and web-based environments. *Archival Science*, 12(4), 485-500. <https://doi.org/10.1007/s10502-012-9175-4>
- Matienzo, M. A. (2008). Canonization, Archivalization, and the «Archival Imaginary» Item Type Meetings and Proceedings. *Archive Fervour/Archive Further: Literature, Archives, and Literary Archives, Aberystwyth, Wales, July 9-11, 2008*. <http://hdl.handle.net/10150/216929>
- Mattern, E. (2016). A six-stage process for recovery of public records: replevin and the state of North Carolina. *Archival Science*, 16(2), 195-212. <https://doi.org/10.1007/s10502-015-9241-9>
- McKemmish, S., Chandler, T., & Faulkhead, S. (2019). Imagine: a living archive of people and place «somewhere beyond custody». *Archival Science*, 19(3, SI), 281-301. <https://doi.org/10.1007/s10502-019-09320-0>
- Melamed, E. (2012). The fate of the archives of the Kiev Institute of Jewish Proletarian Culture: puzzles and discoveries. *East European Jewish Affairs*, 42(2), 99-110. <https://doi.org/10.1080/13501674.2012.699206>
- Mifflin, J. (2009). "Closing the Circle": Native American Writings in Colonial New England, a Documentary Nexus between Acculturation and Cultural Preservation. *American Archivist*, 72(2), 344-382. [http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502988487\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502988487(=pt-pt&site=ehost-live&scope=site)
- Mnjama, N. (2011). Migrated archives revisited. *ESARBICA Journal*, 30, 15.
- Mnjama, N., & Lowry, J. (2017). A Proposal for Action on African Archives in Europe. Em J. Lowry (Ed.), *Displaced Archives* (pp. 101-113). Routledge. <https://doi.org/10.4324/9781315577609-8>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6(7), e1000097–e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Montgomery, B. P. (2001). The Iraqi secret police files: a documentary record of the Anfal genocide. *part of human rights collection at University of Colorado at Boulder*, 52, 69-99. [http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502881704\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502881704(=pt-pt&site=ehost-live&scope=site)
- Montgomery, B. P. (2010). Returning Evidence to the Scene of the Crime: Why the Anfal Files Should be Repatriated to Iraqi Kurdistan. *Part of a special issue on archives and the*

- law, 69, 143-171. [http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502989358\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502989358(=pt-pt&site=ehost-live&scope=site)
- Montgomery, B. P. (2011). Immortality in the Secret Police Files: The Iraq Memory Foundation and the Baath Party Archive. *International Journal of Cultural Property*, 18(3), 309–336. <https://doi.org/10.1017/S094073911100018X>
- Montgomery, B. P. (2012). Saddam Hussein’s Records of Atrocity: Seizure, Removal, and Restitution. *American Archivist*, 75(2), 326–370. [http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=83882367\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=83882367(=pt-pt&site=ehost-live&scope=site)
- Montgomery, B. P. (2013). Rescue or return: the fate of the Iraqi Jewish archive. *International Journal of Cultural Property*, 20(02), 175-200. <https://doi.org/10.1017/S0940739113000040>
- Montgomery, B. P. (2014). US Seizure, Exploitation, and Restitution of Saddam Hussein’s Archive of Atrocity. *Journal of American Studies*, 48(2), 559-593. <https://doi.org/10.1017/S0021875813002004>
- Montgomery, B. P. (2015a). Reconciling the Inalienability Doctrine with the Conventions of War. *American Archivist*, 78(2), 288-316. <https://doi.org/10.17723/0360-9081.78.2.288>
- Montgomery, B. P. (2015b). The Rape of Kuwait’s National Memory. *International Journal of Cultural Property*, 22(1), 61-84. <https://doi.org/http://dx.doi.org/10.1017/S0940739115000053>
- Montgomery, B. P. (2017). Iraq and Kuwait: The Seizure and Destruction of Historical Patrimony. Em J. Lowry (Ed.), *Displaced Archives* (pp. 158-179). Routledge. <https://doi.org/10.4324/9781315577609-11>
- Moustafa, L. H. (2018). Research without Archives? The Making and Remaking of Area Studies Knowledge of the Middle East in a Time of Chronic War. *Archivaria*, 85, 68-95.
- Namhila, E. N. (2004). Filling the Gaps in the Archival Record of the Namibian Struggle for Independence. *IFLA Journal*, 30(3), 224-230. [http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502921006\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502921006(=pt-pt&site=ehost-live&scope=site)
- Namhila, E. N. (2015). *Recordkeeping and Missing “Native Estate” Records in Namibia: An Investigation of Colonial Gaps in a Post-colonial National Archive*. Tampere University.
- Namhila, E. N. (2016). Content and use of colonial archives: an under-researched issue. *Archival Science*, 16(2), 111-123. <https://doi.org/10.1007/s10502-014-9234-0>
- Nannyonga-Tamusuza, S., & Weintraub, A. N. (2012). The Audible Future: Reimagining the Role of Sound Archives and Sound Repatriation in Uganda. *Ethnomusicology*, 56(2), 206-233.
- Ngoepe, M., & Netshakhuma, S. (2018). Archives in the trenches: repatriation of African National Congress liberation archives in diaspora to South Africa. *Archival Science*, 18(1), 51-71. <https://doi.org/10.1007/s10502-018-9284-9>
- Niessen, J. P. (2018). Heritage and Repatriation in the History of Habsburg and Hungarian Archives. *Hungarian Cultural Studies*, 11, 136-143. <https://doi.org/10.5195/ahca.2018.327>
- Niles, D. (2012). The National Repatriation of Papua New Guinea Recordings: Experiences Straddling World War II. *Ethnomusicology Forum*, 21(2, SI), 141-159. <https://doi.org/10.1080/17411912.2012.689469>
- Nordlinger, R., Green, I., & Hurst, P. (2019). Working at the interface: The Daly Language Project. In R. Nordlinger, Rachel, Ian Green, e Peter Hurst. 2019. «Working at the Interface: The Daly Languages Project». Em *Archival Returns: Central Australia and Beyond*, editado por Linda Barwick, Jennifer Green, e Petronell. Em L. Barwick, J. Green, & P. Vaarzon-Morel (Eds.), *Archival returns: Central Australia and beyond* (pp. 193-216). University of Hawai’i Press & Sydney University Press. <http://hdl.handle.net/10125/24884>
- O’Neill, J. E. (1979). Replevin: A Public Archivist’s Perspective. *College & Research Libraries*, 40(1), 26-30. https://doi.org/10.5860/crl_40_01_26

- Patkus, R. D. (2005). Musical migrations: A Case Study of the Teresa Carreño Papers. *RBM: A Journal of Rare Books, Manuscripts, & Cultural Heritage*, 6(1), 26-33. [http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502953150\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=502953150(=pt-pt&site=ehost-live&scope=site)
- Pennell, C. R. (2019). Digitised, digital and static archives and the struggles in the Middle East and North Africa. *Archives & Manuscripts*, 47(2), 242-259. <https://doi.org/10.1080/01576895.2019.1585266>
- Peterson, T. (2000). Macro archives, micro states. *Archivaria*, 50, 41-51.
- Phillips, D. (2013). The «migrated archives»: the underbelly of colonial rule in Borneo. *Borneo Research Bulletin*, 44, 40-92. [http://search.ebscohost.com/login.aspx?direct=true&db=asn&AN=110084211\(=pt-pt&site=ehost-live&scope=site](http://search.ebscohost.com/login.aspx?direct=true&db=asn&AN=110084211(=pt-pt&site=ehost-live&scope=site)
- Phillips, D. (2016). The ‘Migrated Archives’ and a Forgotten Corner of Empire: The British Borneo Territories. *The Journal of Imperial and Commonwealth History*, 44(6), 1001-1019. <https://doi.org/10.1080/03086534.2016.1251557>
- Porter, M. (1980). An algorithm for suffix stripping. *Program*, 14(3), 130-137. <https://doi.org/10.1108/eb046814>
- Porter, M., & Boulton, R. (2002). *Snowball*. <https://snowballstem.org/>
- Punzalan, R. (2006). Archives of the new possession: Spanish colonial records and the American creation of a «national» archives for the Philippines. *Archival Science*, 6(3-4), 381-392. <https://doi.org/10.1007/s10502-007-9040-z>
- Punzalan, R. (2014a). Understanding virtual reunification. *The Library Quarterly*, 84(3), 294-323. <https://doi.org/10.1086/676489>
- Punzalan, R. (2014b). Archival diasporas: A framework for understanding the complexities and challenges of dispersed photographic collections. *American Archivist*, 77(2), 326-349. <https://doi.org/10.17723/aarc.77.2.729766v886w16007>
- Rawlings, G. (2015). Lost Files, Forgotten Papers and Colonial Disclosures : The «Migrated Archives» and the Pacific, 1963 – 2013. *The Journal of Pacific History*, 50(2), 189-212. <https://doi.org/10.1080/00223344.2015.1048585>
- Rhoads, J. (1966). Alienation and Thievery: Archival Problems. *The American Archivist*, 29(2), 197-208. <https://doi.org/10.17723/aarc.29.2.510581828404555q>
- Ribeiro, F. (1998). *O acesso à informação nos arquivos* [Universidade do Porto]. <https://repositorio-aberto.up.pt/bitstream/10216/7058/3/fribeirovol01000061435.pdf>
- Saini, M., & Shlonsky, A. (2012). Systematic Synthesis of Qualitative Research. Em *Systematic Synthesis of Qualitative Research* (Vol. 9780195387). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195387216.001.0001>
- Sandelowski, M., & Barroso, J. (2010). *Handbook for synthesizing qualitative research*. Springer.
- Sato, S. (2017). «Operation Legacy»: Britain’s Destruction and Concealment of Colonial Records Worldwide. *The Journal of Imperial and Commonwealth History*, 45(4), 697-719. <https://doi.org/10.1080/03086534.2017.1294256>
- Schenkolewski-Kroll, S. (2016). Jewish archives and archival documents: Israel and the Diaspora. *Archival Science*, 16(3), 309-326. <https://doi.org/10.1007/s10502-015-9249-1>
- Sela, R. (2018a). «Imprisoned Photographs»: The Looted Archive of Photo Rissas (Rassas) – Ibrahim and Chalil (Khalil) Rissas. *Intermedialites*, 32, 1-28. <https://doi.org/10.7202/1058469ar>
- Sela, R. (2018b). The Genealogy of Colonial Plunder and Erasure – Israel’s Control over Palestinian Archives. *Social Semiotics*, 28(2), 201-229. <https://doi.org/10.1080/10350330.2017.1291140>

- Shepard, T. (2015). «Of Sovereignty»: Disputed Archives, «wholly Modern» Archives, and the Post-Decolonization French and Algerian Republics, 1962-2012. *American Historical Review*, 120(3), 869-883. <https://doi.org/10.1093/ahr/120.3.869>
- Shepard, T. (2017). Making Sovereignty and Affirming Modernity in the Archives of Decolonisation: The Algeria–France ‘Dispute’ between the Post-Decolonisation French and Algerian Republics, 1962-2015. Em J. Lowry (Ed.), *Displaced Archives* (pp. 21-40). Routledge. <https://doi.org/10.4324/9781315577609-3>
- Silva, J. H., & Albuquerque, M. E. (2018). Abordagens acerca da representação da informação na arquivologia: análise a partir dos Anais da ISKO – Brasil. *Biblionline*, 14(1), 67-79.
- Sinkoff, N. (2016). From the archives: Lucy S. Dawidowicz and the restitution of Jewish cultural property. Em *American Jewish History* (Vol. 100, Número 1, pp. 117-147). Johns Hopkins University Press. <https://doi.org/10.1353/ajh.2016.0009>
- Stein, S. A. (2015). Black Holes, Dark Matter, and Buried Troves: Decolonization and the Multi-Sited Archives of Algerian Jewish History. *American Historical Review*, 120(3), 900-919. <https://doi.org/10.1093/ahr/120.3.900>
- Stoler, A. L. (2002). Colonial Archives and the Arts of Governance. *Archival Science*, 2, 87-109. <https://doi.org/10.1007/BF02435632>
- Thorne, S., Jensen, L., Kearney, M. H., Noblit, G., & Sandelowski, M. (2004). Qualitative metasynthesis: Reflections on methodological orientation and ideological agenda. Em *Qualitative Health Research* (Vol. 14, Número 10, pp. 1342-1365). <https://doi.org/10.1177/1049732304269888>
- Thorner, S., Rive, L., Dallwitz, J., & Inyika, J. (2019). Never giving up: Negotiating, culture-making, and the infinity of the archive. Em L. Barwick, J. Green, & P. Vaarzon-Morel (Eds.), *Archival returns: Central Australia and beyond* (pp. 263-284). University of Hawai'i Press & Sydney University Press. <http://hdl.handle.net/10125/24887>
- Tognoli, N. B. (2012). A representação na arquivística contemporânea. *RICI Revista Ibero-Americana de Ciência da Informação*, 5(2), 1983-5213. <http://periodicos.unb.br/ojs248/index.php/RICI/article/view/7974>
- Tognoli, N. B., & Guimarães, J. A. C. (2012). Challenges of Knowledge Representation in Contemporary Archival Science. *Knowledge Organization*, 13, 267-271. <https://repositorio.unesp.br/handle/11449/117825>
- Treloyn, S., & Emberly, A. (2013). Sustaining Traditions: Ethnomusicological Collections, Access and Sustainability in Australia. *Musicology Australia*, 35(2), 159-177. <https://doi.org/10.1080/08145857.2013.844473>
- Treloyn, S., Martin, M. D., & Charles, R. G. (2016). Cultural precedents for the repatriation of legacy song records to communities of origin. *Australian Aboriginal Studies Journal*, 2, 94–103. <https://search.informit.org/doi/10.3316/informit.521032219518643>
- Vaarzon-Morel, P., & Kelly, L. (2019). Enlivening people and country: The Lander Warlpiri cultural mapping project. Em L. Barwick, J. Green, & P. Vaarzon-Morel (Eds.), *Archival returns: Central Australia and beyond* (pp. 111-138). University of Hawai'i Press & Sydney University Press. <http://hdl.handle.net/10125/24880>
- Vallier, J. (2010). Sound Archiving Close to Home: Why Community Partnerships Matter. *Notes*, 67(1), 39-49.
- Vital, L. P., Medeiros, G. M. de, & Brascher, M. (2017). Classificação e descrição arquivística como atividades de organização e representação da informação e do conhecimento. *Brazilian Journal of Information Science*, 11(4), 40-46.
- Walsh, D., & Downe, S. (2005). Meta-synthesis method for qualitative research: a literature review. *Journal of Advanced Nursing*, 50(2), 204-211. <https://doi.org/10.1111/j.1365-2648.2005.03380.x>

- Winn, S. R. (2015). Ethics of Access in Displaced Archives. *Provenance, Journal of the Society of Georgia Archivists*, 33(1), 6-13. <https://digitalcommons.kennesaw.edu/provenance/vol33/iss1/5>
- Wood, S., Carbone, K., Cifor, M., Gilliland, A. J., & Punzalan, R. (2014). Mobilizing records: re-framing archival description to support human rights. *Archival Science*, 14(3-4), 397-419. <https://doi.org/10.1007/s10502-014-9233-1>
- Wurl, J. (2005). Documenting Displacement: The Migration of Archival Sources From Post-WW II East European Émigré Groups. *Archival Science*, 5(1), 79-92. <https://doi.org/10.1007/s10502-005-5763-x>
- Yakel, E. (2003). Archival representation. *Archival Science*, 3(1), 1-25. <https://doi.org/10.1007/BF02438926>
- Yeo, G. (2009). Custodial History, Provenance, and the Description of Personal Records. *Libraries & the Cultural Record*, 44(1), 50-64.

EPISTEMOLOGIA DA ORGANIZAÇÃO DO CONHECIMENTO: UM NOVO VELHO PARADIGMA

Carlos Cândido de Almeida¹

¹Universidad Estatal Paulista – UNESP, Instituto de Estudios de Género
– Universidad Carlos III de Madrid,
carlos.c.almeida@unesp.br, ORCID iD 0000-0002-8552-1029

Resumo

Pragmatismo e pragmática encontram pontos em comum e algumas divergências. A inclusão dessas correntes na organização do conhecimento pode tratar-se de um paradigma. O objetivo é examinar as linhas pragmatistas e pragmáticas presentes em estudos da organização do conhecimento para entender esse candidato a paradigma. A pesquisa segue um recorte teórico-bibliográfico com a finalidade de reconstituir as premissas dos pioneiros do pragmatismo e da pragmática, bem como estas aplicações no âmbito de uma epistemologia geral para a organização do conhecimento. Como resultados encontrados destacam-se as linhas referentes aos paradigmas pragmáticos à pesquisa na ciência da informação, ao pragmatismo subsidiando a teoria do conceito, à abordagem pragmática na organização do conhecimento, à pragmática linguística, ao pragmatismo de Peirce na organização do conhecimento e aos aspectos culturais em organização do conhecimento. Conclui-se que o paradigma pragmatista-pragmático, nas suas diversas nuances e confrontações, seria um adequado rótulo para incorporar a dinâmica cultural, os movimentos culturais, a igualdade de gênero, as lutas afirmativas, entre outros temas na organização do conhecimento. Tal paradigma, plural e abrangente, pode reconhecer as tendências investigativas da organização do conhecimento mais produtivistas e universalistas (pragmatista) e ainda contemplar os resultados de suas ações às inúmeras comunidades discursivas (pragmático). Entende-se que quando se segue um pressuposto cientificista para descrever os conceitos para compor sistemas universais em áreas científicas, se está sob a égide pragmatista (relativamente peirceano) e quando se dedica à construção de sistemas particulares voltados a comunidades discursivas, se reconhecem pressupostos pragmáticos (relativamente James, Wittgenstein, Morris, Austin etc.). Como essas duas linhas estão presentes e bem ativas em organização do conhecimento, se está diante do mesmo espectro paradigmático pragmatista-pragmático, de certa forma já antecipados por Jacob (2000) e Dousa (2010a, 2010b).

Palavras-chave: Pragmatismo, Pragmática, Epistemologia, Organização do Conhecimento, Paradigma.

Introdução

O campo da organização do conhecimento¹, considerado uma área de ensino, pesquisa e de práticas profissionais, investiga as diversas formas de representação da informação e do conhecimento, bem como seus processos. As perspectivas epistemológicas da organização do conhecimento incluem-na relativamente dentro (Hjørland, 2016) ou relativamente fora da ciência da informação (Dahlberg, 2006, 2014). Dahlberg (2014, p. 89) sustentou a ideia que seria necessário retirá-la da biblioteconomia e da documentação. Contudo, independente da configuração epistemológica, o que parece fundamental é perceber que seus objetos, metodologias e história permanecem mantendo uma proximidade com a ciência da informação.

Nas últimas décadas, a ciência da informação tem adotado abordagens² teóricas, metodológicas e principalmente respostas tecnológicas para solucionar demandas da sociedade. Essas demandas são representadas por problemas a serem atacados: a organização de grandes volumes de informação, o processamento de informações especializadas – compreendendo os contextos militar, científico, industrial e empresarial –, a conservação e a restauração de documentos, a representação e a recuperação da informação em grandes bancos de dados, a disseminação e acesso à informação pela sociedade e a organização do conhecimento que potencializa a recuperação da informação em empresas, instituições públicas, centros de documentação, bibliotecas, arquivos e museus³.

É neste último grupo de problemas que se tem apresentado um conjunto ímpar de respostas teóricas no contexto da ciência da informação. A organização do conhecimento tem-se apropriado de abordagens filosóficas, sociológicas, computacionais, estatísticas, lógicas, linguísticas e semióticas. Entre estas, as contribuições relacionadas à filosofia, à semiótica e à linguística têm especial importância para a configuração paradigmática da área. É na confluência entre filosofia, linguística e semiótica que se encontram os estudos genericamente designados como pragmáticos.

¹ Adotar-se-á organização do conhecimento como a expressão mais ampla que congrega os estudos referentes à organização do conhecimento e à organização da informação, salientando que estas linhas têm memórias acadêmicas distintas e marcadas no tempo, as quais não serão aprofundadas neste projeto. A organização do conhecimento enquanto área acadêmica e de práticas profissionais tem sido promovida pela *International Society for Knowledge Organization* – ISKO, instituição fundada em 1989. O periódico *Knowledge Organization*, principal meio de divulgação das pesquisas do campo teve início em 1993 em substituição da revista *International Classification*, que circulava desde 1974. Ver: <https://www.isko.org/ko.html>

² Entende-se por abordagem uma maneira de atacar um problema de pesquisa, servindo-se de ou sendo equivalente a um ponto de vista, uma perspectiva, uma visão, um enfoque ou um tipo particular de interpretação. Uma abordagem, nesse sentido, pode agregar uma ou mais teorias ou metodologias sobre um objeto de estudo. A ideia de abordagem é mais flexível e abrangente que uma teoria. É possível que a ideia de corrente ou linha de pensamento possa ser tomada como equivalente à abordagem em alguma medida.

³ Em um período como o de pandemia, viu-se a importância de tratar e organizar informações relativas aos estudos de fármacos e vacinas contra vírus em geral, e sobre o Coronavírus, em particular. Há também o trabalho da ciência da informação em propor diretrizes para se combater informações falsas (*fake news*) e promover o conhecimento científico, principalmente para evitar pseudo-tratamentos aplicados contra a COVID-19.

Vale destacar que há uma imbricação temática e teórica entre o que se conhece como pragmática, disciplina que estuda os usos e os usuários da linguagem, e o pragmatismo, corrente filosófica que busca reposicionar e reconciliar os conceitos da tradição filosófica a luz dos efeitos práticos concebidos. A linha vinculante destas correntes de pensamento deve ser investigada pela organização do conhecimento, pois tais abordagens tem sido utilizadas para rotular e adjetivar estudos ou tendência na área.

Nesse sentido, seria adequado conhecer a incursão dessas correntes na organização do conhecimento e ver em que medida trata-se de um “novo” paradigma. O objetivo deste trabalho é examinar as linhas pragmatistas e pragmáticas⁴ presentes em estudos da organização do conhecimento para entender esse candidato ao posto de paradigma. A pesquisa segue um recorte teórico-bibliográfico com a finalidade de reconstituir, ainda que com poucas fontes, as premissas dos pioneiros do pragmatismo e da pragmática, bem como estas aplicações no âmbito de uma epistemologia para a organização do conhecimento.

Pragmatismo e pragmática

O pragmatismo clássico teve como figuras principais Charles Peirce (1839-1914), William James (1842-1910) e John Dewey (1859-1952)⁵. Por sua vez, a pragmática, entendida como um dos ramos da linguística e da filosofia da linguagem, teve sua origem com o trabalho de Charles Morris (1901-1979), o qual ressaltou que pragmática é o estudo de uma das dimensões da semiose, conceito de Charles Peirce associado à ação dos signos. Colapietro (1993, p. 162) foi mais taxativo e afirmou que pragmática é o termo usado por Morris para designar um ramo da semiótica devotado à investigação das relações dos signos e seus vários usuários, produtores e intérpretes. Yule (2008, p. 132) tratou da pragmática como estudo do que os falantes desejam comunicar.

Para Armengaud (2006), a pragmática tem como fundadores principais Charles Peirce, Gottlob Frege (1848-1925), Ludwig Wittgenstein (1889-1951) e Charles Morris. Neves (2012, p. 172) reafirma que a pragmática começa a articular-se com a linguística a partir de Morris e, por sua vez, vincula-se ancestralmente com o pragmatismo de Peirce e James. Também deve-se incluir nesse conjunto, mas não na condição de pioneiros, o trabalho de John Langshaw Austin (1911-1960) sobre atos de fala. Pinto (2012, p. 59) declara que há três grupos de estudos associados à pragmática: os estudos relacionados ao pragmatismo americano, os estudos voltados aos atos da fala e os estudos pragmáticos interdisciplinares.

⁴ A paranomásia faz-se necessária nesse caso, pois entende-se que os pragmatistas seriam os correligionários dos princípios do pragmatismo (corrente filosófica e científica) e os pragmáticos aqueles autores e pesquisadores adeptos à pragmática (enquanto disciplina vinculada à linguística e à filosofia da linguagem).

⁵ Em razão do enfoque deste trabalho que procura relacionar as linhas principais do pragmatismo e sua relação estreita com a pragmática, o pragmatismo de Dewey não será objeto deste projeto, mas se reconhece a sua fundamental importância para o pragmatismo clássico, em especial, o papel que representa nos trabalhos de Richard Rorty (Nascimento, 2020).

Como pode ser notado, há uma implicação epistemológica entre pragmatismo e pragmática, contudo esta consideração nem sempre é destacada quando pragmática e pragmatismo são incorporados a outras áreas. Para qualificar a discussão, é importante retomar às linhas gerais do pragmatismo e da pragmática.

O pragmatismo de Peirce resultou de uma longa discussão dos métodos de fixação das crenças (conhecimento, conceitos, termos aceitos). Seus artigos que marcaram o nascimento do pragmatismo foram, sugestivamente: “A fixação das crenças” (1877) e “Como tornar claras nossas ideias” (1878). Em ambos Peirce procurou recuperar o caráter metodológico das formas de definir os termos especializados. A natureza dos problemas terminológicos presentes nos textos aproxima seu pragmatismo do campo da organização do conhecimento.

Comentadores da obra de Peirce, entre eles Santaella (2004, p. 26), sustentam que o pragmatismo peirceano pode ser dividido em dois grandes períodos: o primeiro, que se refere à máxima, e o segundo, que procura fundamentar um método de determinação de conceitos intelectuais, valendo-se da abdução (criação de hipóteses) e da seleção de hipóteses. O pragmatismo peirceano sugere o afastamento de “filosofias do faz-de-conta” para concentrar os esforços intelectuais em problemas reais.

O pragmatismo, expresso na máxima de 1878, busca verificar os significados dos pensamentos, somando os efeitos práticos concebíveis e resultantes da ação: “Considerar que efeitos – imaginavelmente possíveis de alcance prático – concebemos que possa ter o objeto de nossa concepção. A concepção desses efeitos corresponderá ao todo da concepção que tenhamos do objeto” (Peirce, 1972, p. 59, CP 5.402). Essa regra guia o pensamento no sentido de descobrir como deixar as ideias claras e distingui-las de ideias obscuras.

Nesse sentido, o pragmatismo é um método que consiste na aplicação de regras de ação para garantir a clareza das ideias. O teste proposto pelo método possibilita o estabelecimento do significado dos conceitos a partir dos efeitos sensíveis em um uso concreto e podem ser utilizados para verificar a eficácia da ideia. Nessa época, Peirce dialogou com os clássicos, em especial, com Descartes. O fato é que, para Descartes algumas ideias pareciam nitidamente claras e não havia necessidade de distingui-las. Segundo ele, o processo de análise constante permitiria chegar a estados mínimos das ideias. Contudo, argumenta Peirce, este não seria um método seguro, pois não confronta a ideia com a realidade. Qualquer conceito somente poderá ser distinto de um outro quando seus efeitos práticos ou experimentáveis possam ser imaginados. Com essa regra, pode-se propor um caminho para chegar-se a ideias claras, devendo-se adotar as noções claras e separá-las das obscuras. “Uma idéia clara é definida como aquela apresentada de forma tal que se torna possível reconhecê-la em qualquer situação e não confundi-la com qualquer outra. Se não dotada desta clareza, a idéia é dita como obscura.” (Peirce, 1972, p. 49, CP 5.389).

Os fins últimos de uma ideia não estão na ação; pelo contrário, a ação é um meio cujo fim é a conceptualização do objeto. Peirce opõe-se a William James, pois este defende que a ação é o fim último do sujeito e todo conceito somente tem algum valor caso possa ser revertido em ação ou prática útil. Em outras palavras, um conceito útil é aquele que na ação foi comprovada a sua eficácia para satisfazer a um maior número de pessoas. Posteriormente, Peirce cunhou um termo mais adequado

para ser o instrumento para a filosofia, o “pragmaticismo”, palavra estranha, mas que assegurava o ideal de seu pragmatismo.

Em síntese, o pragmatismo de Peirce é um instrumento ou, mais precisamente, um método para auxiliar a filosofia, dentre outras ciências, a esclarecer os fundamentos de seus argumentos e tornar as ideias distintas e passíveis de experimentação. É um método de reflexão, no sentido de servir para a análise das concepções cotadas para aceitação intelectual. Para Ibri (1992, p. 102), “De modo inequívoco, o Pragmatismo não é um sistema filosófico, mas, tão-somente, *um método de análise filosófica de sistemas teóricos.*” O pragmatismo consiste em estabelecer um método para que se alcance o significado dos conceitos, e seu principal campo de aplicação é a filosofia, segundo Peirce. Contudo, poder-se-á adequar a quaisquer formas de conduzir o pensamento.

Na mesma direção, Silveira (2007, p. 182) ressalta que “[...] o Pragmatismo é um método de elaboração teórica de pensamento. Sua formulação, em mais de um momento, assumiu a forma de uma máxima orientadora da conduta científica.” De longe, o pragmatismo peirceano afasta-se do entendimento vulgar que relaciona a palavra “pragmatismo” às ações que repercutem algum resultado prático. Em um estudo relativo às diferentes abordagens do pragmatismo, De Waal (2007, p. 40) não hesita em afirmar que: “Em suma, para Peirce, o pragmatismo é um método para determinar o significado dos conceitos, idéias, crenças, alegações, proposições etc., de qualquer coisa que pode agir como um signo. Essa visão, Peirce manteria por toda a sua vida.”

O pragmatismo de Peirce não procura explicar ou, em si mesmo, ser a resposta a um problema prático. Doutrinas religiosas, filosóficas e até pseudo-científicas cumprem perfeitamente este papel. Ele é somente uma estratégia para retirar de nossas mentes candidatos a conceitos, crenças ou teorias pouco informativos em relação aos objetos de representação.

William James (1967, p. 44), por outro lado, tomará de empréstimo a mesma ideia de pragmatismo como método de Peirce e aludiu a sua primeira manifestação em 1878, no artigo de Peirce, “Como tornar claras as nossas ideias”, publicado no *Popular Science Monthly*. Segundo aponta James “O método pragmático é, primeiramente, um método de assentar disputas metafísicas que, de outro modo, se estenderiam interminavelmente.[...] O método pragmático nesses casos é tentar interpretar cada noção traçando as suas conseqüências práticas respectivas.” (James, 1967, p. 44). A ideia era pensar no pragmatismo como forma de apaziguar posições pessoais, tendências ou correntes de pensamento filosófico claramente conflitantes: racionalismo e empirismo, intelectualismo e sensacionalismo, religiosidade e irreligiosidade, dogmatismo e ceticismo, idealismo e materialismo, otimismo e pessimismo.

Para cumprir essa função, o método pragmático, para James, poderia admitir como razoáveis conceitos não científicos, desde que produzissem um efeito positivo para o sujeito da experiência, fato questionado por Peirce. Conforme explica James (1967, p. 57) “Se as idéias teológicas provam que têm valor para a vida concreta, são verdadeiras, pois o pragmatismo as aceita, no sentido de serem boas para tanto. O quanto serão verdadeiras, dependerá inteiramente de suas relações com as demais verdades, que têm, também, de ser reconhecidas.” Nesse sentido, abre-se para um entendimento de verdade no plural, isto é, haveria um conjunto de verdades com o

mesmo peso dada pela eficácia prática que cada uma produz na vida diária, dependendo do grupo em questão.

Em outras palavras, o pragmatismo seria mais aberto a outras doutrinas, desde que essas pudessem passar pelo crivo ou o rigor dos efeitos práticos – úteis –, ou seja, dos fatos. Embora os fatos pudessem ser isoladamente determinantes, como é o caso para os empiristas. James voltava-se às múltiplas facetas da verdade, e considerava-as como válidas em alguma medida antes do teste ou da verificação de sua adequação a casos reais. As várias verdades concorrentes, filosóficas, científicas e religiosas teriam o mesmo valor e reconhecimento antes das consequências; e parece que mesmo que não se inclinem às verdades científica, deveriam ser reconhecidas pela sua eficácia no tratamento das questões humanas, morais e religiosas. Tanto em James quanto em Peirce, o teste e a experimentação são elementos essenciais para certificar-se da adequação das ideias frente a realidade.

O pragmatismo de James difere-se em vários aspectos do pragmatismo de Peirce, com especial atenção para o fato do pragmatismo peirceano propor-se como mais científico ou voltado às querelas de cunho filosófico-científicas que a proposta de James, cujo contexto de aplicação inclui a aceitação do discurso religioso e político.

Também associadas às origens da pragmática, mas não vinculado ao pragmatismo clássico, devem-se citar as premissas importantes para a pragmática no trabalho de Wittgenstein. O filósofo austríaco Ludwig Wittgenstein figura entre os mais importantes teóricos da linguagem, em especial, da corrente pragmática. Contudo, a sua primeira contribuição foi no sentido de uma teoria do significado aplicada à linguagem da ciência. A forma lógica da linguagem é a estrutura mesma da realidade, segundo o autor: “[...] A realidade total é o mundo. [...] Figuramos os fatos.[...] A figuração representa a situação no espaço lógico, a existência e inexistência de estado de coisas. [...] A figuração é um modelo da realidade.” (Wittgenstein, 2010, p. 143).

De certa forma, o *Tractatus* se alinha a outras teorias lógicas que explicam o significado da linguagem da ciência, como única linguagem capaz de produzir a figuração, ou seja, o isomorfismo entre linguagem e realidade. “O *Tractatus* leva em consideração a função representativa da linguagem e só admite os enunciados suscetíveis de ser verdadeiros ou falsos, ou seja, suscetíveis de entrar no quadro da lógica proposicional. O que importa é a linguagem da ciência.” (Armengaud, 2006, p. 35).

A virada no pensamento do autor ocorre, de acordo com Armengaud (2006, p. 36-37), na obra *Investigações Filosóficas* (Wittgenstein, 1989). A partir dessa obra seriam produzidos os seguintes avanços para Armengaud (2006): a crítica da teoria mentalista do significado; a indissolubilidade entre pensamento e linguagem, uma vez que estes são gerados simultaneamente; a não existência de uma linguagem privada, na medida em que toda a linguagem é pública e segue regras coletivas; a importância de considerar o uso da linguagem em situações e ações concretas; as unidades transfrásicas devem ser consideradas dentro das situações de uso; a língua é uma atividade e as situações de uso fazem referência a jogos de linguagem, os quais são atividades partilhadas, formas de vida, com regras próprias em inúmeros tipos de jogos utilizados pelos sujeitos (exemplos: cantar cantigas de roda, rezar, amaldiçoar etc.).

Para Marcondes (2009), os jogos de linguagem são múltiplos e multifacetados e a ideia de uma estrutura básica da linguagem ou uma forma lógica não faz sentido nesta abordagem. Nesse sentido, segundo Marcondes (2009), o significado é indeterminado e deve ser considerado a partir do jogo de linguagem. Encontra-se aqui noções fundamentais à pragmática, entre as quais a ideia da linguagem como uma forma de vida e a dificuldade de separar, mesmo para fins de pesquisa, a linguagem da vida social. Aqui se aplica a máxima de que o significado está em seu uso.

Não obstante os trabalhos realizados no sentido de uma guinada pragmática sobre a linguagem e o fazer da filosofia, foi Charles Morris quem projetou uma organização disciplinar para a pragmática. Charles Morris elaborou uma teoria semiótica por volta das décadas de 1930 e 1940. Os fundamentos da teoria de Morris estão enraizados na semiótica de Peirce, no behaviorismo social, na teoria interacional simbólica de G. Mead, no pragmatismo americano, no empirismo e no positivismo lógico (Noth, 2005, p. 181).

Quando se trata da noção de pragmática enquanto estudo dos usos e usuários dos signos, Morris é citado como pioneiro por diversos autores (Noth, 2005; Armengaud, 2006). Uma diferença nem sempre observada é a de que Morris recorreu às divisões da lógica de Peirce para compor a concepção de pragmática. Como bem registrou Noth (2005, p. 190), só recentemente a pragmática, cuja origem muitos supõem estar vinculada diretamente à obra de Morris, tornou-se parte da linguística. No entanto, a concepção de pragmática varia muito no âmbito da linguística, correspondendo a um conjunto de pesquisas que trata de temas nem sempre conexos.

Morris (1976) divide a semiótica em dimensões da semiose: sintaxe, semântica e pragmática. A sintaxe ocupa-se do estudo das relações sintáticas entre os signos. A semântica investiga as relações entre os signos e a realidade, os *designatum* e *denotatum*. E a pragmática, cunhada com base no pragmatismo peirceano, é a ciência das relações dos signos com seus intérpretes, isto é, os usos e os usuários dos signos.

Segundo Armengaud (2006, p. 49), a pragmática para Morris é considerada “Ciência que trata da relação dos signos com seus intérpretes: esta é a definição primitiva de pragmática. Morris esclarece: uma vez que a maioria dos signos tem por intérpretes organismos vivos, pode-se caracterizar perfeitamente a pragmática dizendo que ela trata dos aspectos bióticos da semiose.”

As três dimensões da semiose têm sua origem no *trivium* medieval das artes da linguagem: gramática, dialética (lógica) e retórica. Charles Peirce reinterpretou este *trivium* e dividiu a ciência dos signos também em três ramos: gramática pura, lógica crítica e retórica pura. Um paralelo entre as dimensões da semiótica de Morris é possível quando se concebem a gramática pura como precursora da sintática, a lógica própria, da semântica e a retórica pura como precursora da pragmática (Noth, 2005, p. 186-187). Na opinião de Santaella (2004, p. 178), “Charles Morris popularizou as complexas concepções originais que Peirce tinha dos três ramos da lógica ou Semiótica na conteudística e psicologizante divisão da Semiótica nos níveis sintático, semântico e pragmático. Nada poderia estar mais distante de Peirce do que essa simplificação.”

Evidentemente, as contribuições de Morris à teoria dos signos não se limitam aos elementos indicados anteriormente, mas é possível ter clareza deste ponto de vista

sobre o conceito de signo. Decorre da inclinação aos usuários da linguagem e a seu contexto a teoria dos atos de fala.

A teoria montada pelo filósofo John Austin sustenta que se expressar oralmente é uma forma de ação efetiva e não apenas uma representação. Quando se usa a língua para ordenar, descrever, persuadir, agradecer, desculpar-se, insultar etc., produz-se uma ação. Essa performance envolvida configura os atos de fala.

Essa teoria envolve atos ilocucionários e atos perlocucionários. Ilocucionários referem-se ao que se faz quando se fala e perlocucionários tratam dos efeitos produzidos por nossas sentenças (Armengaud, 2006, p. 100). Por meio da análise do uso da linguagem poder-se-ia identificar como as sentenças funcionam para realizar ações do próprio sujeito ou para demonstrar como o sujeito força outrem a responder a sentença com determinadas ações.

Outro teórico que contribui a teoria dos atos da fala é John Searle (1932-) quem, além de criticar os tipos de atos da fala propostos por Austin, estabeleceu critérios para considerar adequadamente os atos ilocucionários.⁶

Como se pôde notar, “A pragmática está, pois, relacionada com os utilizadores da linguagem, já que a língua em situação de uso está para além do conhecimento gramatical dessa mesma língua” (Neves, 2012, p. 172). De acordo com Pinto, a pragmática é o estudo do uso linguístico e “As pessoas que estudam esperam explicar antes a linguagem do que a língua.” (Pinto, 2012, p. 55-56).

Entre as correntes da pragmática, as quais vinculam-se direta ou indiretamente ao pragmatismo clássico, encontram-se os estudos interdisciplinares que de maneira mais detida voltam-se às questões políticas e sociais. O grupo de estudos pragmáticos interdisciplinares possui a “[...] preocupação firmada nas relações sociais, de classe, de gênero, raciais e entre culturas, presentes na atividade linguística. Este último grupo é especialmente diverso e, sem dúvida, é o que mais tem se desenvolvido nos últimos anos.” (Pinto, 2012, p. 59). Segundo a observação de Pinto (2012), há um entendimento que os estudos pragmáticos interdisciplinares no campo da linguística voltam-se às relações socioculturais e como estas influenciam a comunicação. É possível que as teses desses estudos estejam vinculadas a várias premissas tanto da pragmática quanto do pragmatismo, mas aparecem agora sob uma forma compósita.

Do pragmatismo à pragmática encontram-se pontos em comum e muitas divergências. A recorrência aos efeitos práticos como forma de se definir um conceito, uma ideia ou uma teoria (Peirce), ou para apaziguar tensões entre correntes extremistas dentro e fora do contexto científico (James), bem como a ênfase no contexto, no uso e nas ações dos usuários da linguagem (Wittgenstein, Morris e Austin), revelam mais proximidades mas não uma completa identidade entre pragmatismo e pragmática. É importante saber como a organização do conhecimento tem incorporado tais noções do pragmático como forma de se autoexplicar.

⁶ Além do trabalho de Dewey, também deve-se salientar a importância dos trabalhos de Richard Rorty (1931-2017) sobre o pragmatismo, o papel da filosofia e a importância da linguagem e da hermenêutica. Contudo, em razão da limitação da discussão construída não se abordará as ideias deste autor.

Um “novo” paradigma pragmatista-pragmático na organização do conhecimento?

O argumento defendido é que a área saiu aos poucos de uma corrente mais realista, objetivista, universalista e cientificista na compreensão do conhecimento e das formas de classificá-lo, vinculada precisamente com a abordagem do pragmatismo clássico, especialmente de Peirce, e passa agora por uma fase pragmática, com alguns fundamentos completamente distintos do ciclo anterior, embora esteja vinculada ao mesmo arco temático. A linha vinculante destas correntes de pensamento devem ser investigadas pela organização do conhecimento, na medida em que as referidas abordagens têm sido utilizadas para rotular estudos, tendências e o próprio paradigma atual da área. Considera-se que diversos estudos pertencem a um mesmo arco paradigmático – se é que se pode falar de paradigma na acepção exata de Kuhn –, em que as pontas não compartilham das mesmas premissas, mas há gradações intermediárias que as unem.

De um ponto de vista mais abrangente, a palavra “pragmática” figura na ciência da informação como um paradigma metodológico. Em outras palavras, visa guiar o desenho de pesquisas aplicadas recorrendo a diversas técnicas de investigação, quantitativas e qualitativas. Revez e Borges (2018, p. 517) assinalam que o paradigma pragmático é pouco conhecido na área. Os autores o compreendem como um paradigma para as pesquisas e não uma teoria da linguagem. Nesse sentido, o pragmatismo é a base filosófica para o paradigma pragmático que mobiliza diversos métodos e técnicas mistas de pesquisa apresentando-se como uma solução à dicotomia e às tensões entre opções qualitativas e quantitativas (Revez; Borges, 2018). Esse paradigma não é exclusividade da área e tem sido utilizado por diversas ciências sociais.⁷ A análise empreendida por Revez e Borges (2018, p. 590) revela que a pesquisa em ciência da informação não está consciente do paradigma pragmático como um fundamento metodológico nem mesmo é reconhecido como a base para as pesquisas que combinam métodos mistos.

As oposições entre os paradigmas pragmáticos e positivistas têm sido discutidas na ciência da informação, especificamente na organização do conhecimento⁸. A esse respeito, deve-se fazer referência ao trabalho de San Segundo (2008) que observa:

No contexto da Ciência da Informação, a concepção de paradigma apresenta-se como a confluência de correntes epistemológicas de todas as demais ciências, na atualidade predomina dois paradigmas fundamentais: Paradigma positivista,

⁷ Para Revez e Borges (2018, p. 595) “The pragmatic paradigm is used in several fields of social science research. These applications can be found in ethnography, anthropology, sociology, information science and even in the development of public policies, etc. In practical terms, pragmatism is reflected in the appropriateness of the method to the research question without a priori limitations established.”

⁸ San Segundo (2004) defendeu a necessidade de se basear no pragmatismo americano para construir um novo conceito de representação do conhecimento. A autora reconheceu antes que outros autores que nesta escola de pensamento a ação do sujeito é determinante para as noções de conhecimento, verdade e sua representação (San Segundo, 2004, p. 108).

bibliográfico, de equiparação total, enquadrado nos séculos XIX e XX, que gerou as linguagens documentárias e as classificações universais, e Paradigma pragmatista, subjetivo, de interação, que gerou as interfaces entre computador e usuário inserido nas tecnologias da informação e no contexto epistemológico do início do século XXI. (SAN SEGUNDO, 2008, p. 681, tradução nossa).

A separação entre positivista e pragmático não é tão clara assim quando se entende que o pragmatismo clássico produz uma resposta coerente para o método científico. Nesse sentido, o método científico proposto por Peirce abduutivo-dedutivo-indutivo resumido em hipotético-dedutivo, seria também pragmatista. Por outro lado, a impressão que o pragmatista enfatiza o subjetivo também não encontra respaldo no pragmatismo clássico, mas pode abrigar-se nas linhas pragmáticas interdisciplinares presentes na linguística.

Os paradigmas conduziram as práticas da ciência da informação e não apenas da organização do conhecimento. Pode-se entender esta vertente defendida por Revez e Borges (2018) e San Segundo (2008) como de âmbito geral que incide sobre as metodologias do campo. Há também estudos que têm aproximado pragmatismo à epistemologia da ciência da informação no geral (Saldanha, 2011), mas intenta-se aqui concentrar esforços no pragmatismo e na pragmática como subsídio à organização do conhecimento.

Particularmente na teoria da organização do conhecimento, encontra-se a compreensão do pragmatismo como fundamento e do pragmatismo como subsídio à teoria do conceito. Por um lado, Jacob (2000) levantou a questão do pragmatismo como fundamento da organização do conhecimento justamente pela proximidade de teses pragmatistas com a pós-modernidade, o pós-estruturalismo e a desconstrução. Jacob conclui sua análise ressaltando a importância de se adotar a estrutura pragmática como base para a teoria da classificação e seus relacionamento com as epistemologias pós-modernas. Por outro lado, para Hjørland (2003), a teoria do conceito não tem sido desenvolvida pelos estudos no âmbito da organização do conhecimento, mas é pesquisada mais em outras áreas. Segundo Hjørland (2003, p. 100), dentro da teoria do conceito é importante separar duas tradições: o positivismo lógico e as visões alternativas baseadas nas teorias pragmáticas. Para explicar a influência de diversos paradigmas filosóficos na composição da teoria do conceito, o autor agrupa as definições de conceito nos seguintes paradigmas: 1) empirismo, 2) racionalismo, 3) criticismo de Kant (empirista-racionalista), 4) historicismo-hermenêutico e 5) pragmaticismo e realismo crítico. No paradigma pragmático, os conceitos e o conhecimento são formados nos contextos históricos dos objetos e relacionam-se às atividades práticas das pessoas (Hjørland, 2003, p. 101).

Para Dousa (2010a, 2010b), Hjørland está mais perto do pragmatismo de Dewey que de Peirce, e aponta que tanto o pragmatismo clássico quanto o neopragmatismo tem sido utilizados para justificar as posições perspectivista, construtivista e anti-essencialista na organização do conhecimento. Contudo, Dousa (2010b) conclui que o pragmatismo clássico pode ser mais integrativo e útil à organização do conhecimento.

Weiss e Brascher (2014), por sua vez, fizeram um levantamento em diversas bases de dados (BRAPCI, LISA, ARIST, Repositórios, Web of Science e Plataforma

Lattes) sobre os termos pragmática e pragmatismo visando caracterizar a abordagem pragmática na organização do conhecimento. Organizaram a abordagem pragmática em um conjunto de categorias representativas, a saber: a informação é um fenômeno construído; a informação e conceitos ocorrem em contextos múltiplos e heterogêneos; a qualidade da informação não é constante; a prática é importante para a construção do significado; a interação é importante para a validação; a bibliografia é uma fonte para identificar pontos de vista distintos e construir linguagens de comunicação e guiar o usuário para uma escolha informada.

Esse conjunto de ideias é aplicado não mais à teoria do conceito, mas a todo e qualquer sistema de organização do conhecimento. Segundo as autoras, “Ao adotar princípios da abordagem pragmática, os SOC podem contribuir para uma comunicação efetiva, ao possibilitar a compreensão dos diferentes pontos de vista presentes numa comunidade discursiva e não optar por um como sendo o melhor para representar uma realidade que é, por natureza, heterogênea.” (Weiss; Brascher, 2014, p. 698).

Obviamente, pode-se relacionar as categorias representativas ao pensamento de Morris, Wittgenstein, Morris, Austin, bem como, e talvez principalmente, a linha interdisciplinar da pragmática presente na linguística. Essa linha difusa e heterogênea tem incluído elementos das relações sociais, teoria de gênero, estudos sobre racismo, relações étnicas e culturais (Pinto, 2012, p. 59). Contudo, tomar tal conjunto de categorias como representativo de todo o pragmatismo pode parecer uma simplificação.

A ideia de uma abordagem pragmática geral associada à organização do conhecimento parece ter se inserido definitivamente no debate da área. Não obstante, as premissas precisam ser melhor explicadas e vinculadas às abordagens epistemológicas correspondentes. Nesse sentido, concorda-se com a observação de Kobashi e Fernandes (2009) de que “Trazer a pragmática para a área da Ciência da Informação pode ser uma tarefa complexa [...]”.

As autoras Kobashi e Fernandes (2009) articularam de maneira muito apropriada a pragmática com a ciência da informação, em especial, limitaram-se à pragmática linguística, a qual traz noções e conceitos para se entender melhor o contexto (circunstancial, factual, existencial, referencial situacional ou paradigmático interacional pressuposicional), tal como sistematizado por Armengaud (2006). Neste caso, haveria uma aplicação direta da pragmática linguística à linguagem documentária no que tange a organização da informação documentária.

Sob um linha específica do pragmatismo clássico, no que se refere à incorporação do pragmatismo de Peirce à organização do conhecimento, destacam-se os trabalhos de Thellefsen (2002, 2004, 2009, Thellefsen; Thellefsen, 2004). O autor busca propor uma teoria semiótica da organização do conhecimento calcada no pragmatismo de Peirce. Friedman e Thellefsen (2011) sistematizaram as principais relações teóricas entre teoria do conceito e semiótica, vinculada ao pragmatismo de Peirce. Na mesma vertente encontram-se os trabalhos de Almeida (2012) e Almeida, Fujita e Reis (2013).

Por outro lado, deve-se apontar para os estudos críticos que versam sobre os aspectos culturais em organização do conhecimento como nuances da linha pragmática. Afirmarções como as de Weiss e Brascher (2014), de que os sistemas de organização do conhecimento devem oportunizar diferentes pontos de vista conforme a

comunidade-objeto da organização, contribuem para se pensar nos embates entre os sistemas universais de organização do conhecimento e os sistemas voltados à redescritção do papel e da valorização de determinadas comunidades e grupos na produção social do conhecimento.

Na mesma linha de conta, Weiss, Brascher e Vianna (2016) sustentam que construtivismo e realismo crítico, tal como aludido por Hjørland, são consistentes com a posição pragmática na organização do conhecimento. O construtivismo seria compatível com a abordagem epistemológica do pragmatismo na área. Para os autores, o construtivismo supõe que a inteligência humana seja desenvolvida considerando ações mútuas entre indivíduo e ambiente. A influência e a interação do ambiente no processo de construção e organização do conhecimento podem ser resumidas na relação sujeito-objeto-contexto.

Nesta perspectiva, teses culturalistas e feministas têm sido utilizadas para amparar as práticas de organização do conhecimento. Segundo Lee (2015, p. 302) “Os esquemas de classificação baseados em culturas diferentes incluem e excluem diferentes conceitos.” Nessa linha, podem-se citar os trabalhos de López-Huertas (classificações diferentes de instrumentos musicais conforme a cultura), Bergthol (a necessidade de promover garantia cultural) e Olson (sobre o mito da neutralidade nos esquemas de classificação e descrição de assunto) e Mai (promove uma prática baseada em uma classificação ética que valoriza a cultura e o contexto). Lee (2015, p. 306) introduz o contexto na pesquisa, isto é, as influências entre cultura, percepções e esquema de classificação.

Para San Segundo e Martínez-Ávila (2012), a epistemologia tradicional sofreu muitas objeções e seria importante quebrar tal legitimidade. Nesse contexto, têm surgido outras perspectivas epistemológicas: pragmatismo, positivismo lógico, fenomenologia, hermenêutica e outros. Para os autores, “A epistemologia faz um discurso que analisa a ciência, mas na realidade observa o mundo e termina elaborando a sua construção, pois sempre estamos imersos na episteme vigente.” (San Segundo; Martínez-Ávila, 2012, p. 415). Com base em Foucault, os autores destacam as epistemes que condicionam as formas de pensar as coisas e o conhecimento. A ordem dos saberes também é produto da episteme e como tal determina a produção do conhecimento. Com isso, a organização do conhecimento pode ser pensada como um reflexo ou subproduto das epistemes que se encarregam de guiar a interpretação do mundo. Os sistemas de classificação decorrem da episteme vigente.

O que estaria na base desta crítica à organização do conhecimento seria a contestação das epistemologias modernas. Estas epistemologias (dogmatismo, empirismo, positivismo, positivismo lógico) associaram-se à busca de uma verdade absoluta. De acordo com San Segundo (2008), o sociohistoricismo, a fenomenologia e hermenêutica, ao contrário, procuram validar o conhecimento científico nos simbolismos socioculturais. A autora defende que no contexto pós-moderno, a epistemologia deve ser mais participativa, um exemplo seriam as epistemologias ocidentais que tomaram a mulher como objeto de estudo. Nesse sentido, a produção do conhecimento deve evitar o sexismo e o androcentrismo, bem como questionar os postulados epistemológicos segregadores. Além disso, os desafios da epistemologia na organização do conhecimento são inúmeros e profundos, especialmente na era digital (San Segundo; Codina-Canet, 2018).

Perspectivando sobre as tendências teóricas da comunidade internacional de organização do conhecimento, Guimarães (2017, p. 91) comenta:

Já na dimensão cultural têm lugar as questões sociais, políticas, éticas educativas e contextuais da organização do conhecimento, com especial ênfase ao papel mediador da organização do conhecimento entre distintos contextos culturais. Nessa abordagem tem-se os estudos de Maria José López Huertas, Michèle Hudon, Widad Mustafa El Hadi, Grant Campbell, Hope Olson, Clare Beghtol, Antonio García Gutiérrez e José Augusto Guimarães, entre outros, ao que se aliam os estudos de Birger Hjørland e Joseph Tennis, cuja trajetória científica vem procurando contemplar, ao mesmo tempo, os aspectos epistemológicos e os aspectos culturais dessa área.

Nesta mesma linha, Araújo (2018, p. 56) confirma que a tendência atual em organização do conhecimento é a construção de sistemas de classificação desde um ponto de vista ou perspectiva sociocultural que considera a comunidade de intérpretes. Nota-se que essa abordagem, que se pode designar arriscadamente como sociocultural, recorre à suposta contestação das grandes narrativas epistêmicas da modernidade.

As chamadas epistemologias do sul mais voltadas ao reconhecimento da diversidade de saberes e de soberania intelectual questiona a epistemologia moderna, a qual se converteu em um projeto normativo que procura distinguir a ciência da não ciência, o científico do não científico. Nunes (2009), propondo um resgate da epistemologia no projeto das epistemologias do sul, encabeçado entre outros pensadores como Boaventura Souza Santos, reconhece que há elementos para compreendê-la dentro da base teórica do pragmatismo clássico, especificamente em James e Dewey.

Segundo Nunes (2009, p. 216), “A proposta de Santos configura, explicitamente, um *pragmatismo epistemológico* que apresenta continuidades, mas também importantes elementos de inovação em relação ao pragmatismo clássico e à sua descendência.” Considerando que as abordagens socioculturais da organização do conhecimento voltam-se a uma concepção pluralista de epistemologia, devem-se reconhecer, como fez Nunes, os traços genealógicos das epistemologias historicistas, construcionistas, críticas, feministas e demais conhecimentos epistemológicos subalternos valorizados pelo pensamento do sul, no pragmatismo clássico.

Assim, o paradigma conjuntivo pragmatista-pragmático, nas suas diversas nuances, mais ou menos científicas, seria uma alternativa para reconhecer a dinâmica cultural, os movimentos culturais e de igualdade de gênero, as lutas afirmativas, entre outros. Tal paradigma, plural e abrangente, pode agregar-se às tendências investigativas da organização do conhecimento e melhorar os resultados de suas ações às inúmeras comunidades discursivas.

Nesse sentido, encontram-se, preliminarmente, níveis da aplicação do pragmatismo e da pragmática à organização do conhecimento, os quais devem ser mais bem explorados. Primeiro, o pragmatismo enquanto base filosófica para o paradigma pragmático na ciência da informação de modo a conduzir as pesquisa que mesclam técnicas quantitativas e qualitativas. Segundo, o pragmatismo como paradigma filosófico que influencia a concepção de conceito no interior da organização do

conhecimento. Em terceiro lugar, encontra-se a tentativa de ampliar o pragmatismo e a pragmática como abordagem geral para toda a organização do conhecimento, de modo a balizar a construção dos sistemas de organização mais focalizados. Quarto, a aplicação da pragmática linguística, esta sim mais específica, à construção de linguagens documentárias. Esta, por sua vez, é mais ajustada em duas frentes, tanto em termos disciplinar quanto no campo de aplicação. Em quinto lugar, cita-se a aproximação de uma corrente do pragmatismo clássico – a de Peirce – ao conjunto da organização do conhecimento, de modo a sugerir métodos para organização de conceitos. Por fim, em sexto lugar, pode-se indicar o paradigma pragmatista-pragmático como forma de reconhecer a particularidade da organização do conhecimento das diversas comunidades discursivas ao passo que se respeita a necessidade de produtividade e generalização dos sistemas, quando assim lhe é exigido.

Contudo, a literatura tende a agrupar grosseiramente pragmatismo e pragmática que, em muitos casos, possuem teses aplicadas a setores distintos. Um exemplo é a associação do pragmatismo de Peirce para dar base à ideia genérica de que um conceito está apenas associado ao retorno prático que oferece; ou que a definição do conceito está apenas em seu uso e nos contratos que estabelece linguisticamente, sem referência a objetos do signo.

Concretamente, é importante avançar nas propostas pragmáticas que se têm adotado como base teórica da organização do conhecimento. As premissas pragmáticas não estão claras e as simplificações grosseiras que resumem todo o trabalho da organização do conhecimento como ações voltadas ao contexto, tende a diminuir o potencial das teorias que orbitam em seu entorno e que procuram erigir epistemologicamente este campo. Considera-se que a simplificação dos estudos pragmáticos em organização do conhecimento tem efeitos danosos à área e implica a desconstrução e descontinuidade histórica de teorias lógicas e linguística de fundamental importância para a autonomia do campo nas últimas décadas.

As pesquisas voltadas a repensar o papel dominante e reprodutor de desigualdades dos sistemas de organização do conhecimento, muito utilizados no século XX e ainda vigentes, têm recebido o rótulo de estudos mais “pragmáticos”, ou que focam comunidades discursivas determinadas: mulheres e a agenda feminista⁹, grupos LGBTQIA+ e sua luta pelo reconhecimento público, negros e a necessária discussão do racismo e dos preconceitos na sociedade contemporânea¹⁰, povos originários com os problemas de falta de reconhecimento social, etc. Essa agenda tem influenciado os estudos da organização do conhecimento, mas nem sempre com uma clara distin-

⁹ Alguns estudos têm se voltado ao tema feminista desde uma perspectiva pragmática, especialmente com base na obra de John Dewey. Segundo Vaamonde Gamo (2015, p. 16), “Dewey sustenta, como no feminismo crítico, que a igualdade é indispensável para o desenvolvimento da liberdade. A sua tarefa, como no feminismo crítico, consiste em reconstruir esses ideais levando-os às últimas consequências.” A volta ao pragmatismo clássico para dar conta do pensamento feminista pode configurar-se como uma base teórica aos estudos em organização do conhecimento.

¹⁰ A respeito do assunto na organização do conhecimento, ver Martínez-Ávila, Semidao e Ferreira (2016) e Martínez-Ávila, Ferreira e Magro (2015). No que tange o racismo na sociedade brasileira, vale a pena consultar o estudo de Schucman (2014).

ção dos fundamentos epistemológicos que legitimam tais posturas ou pontos de vista científicos frente a visão anterior e supostamente distinta.

É possível que a rotulação desses estudos como pragmáticos de maneira genérica tende, no lugar de valorizar o potencial da crítica social, a apagar a importância dessas pesquisas, qualificando-as erroneamente como de viés político. Entende-se que a organização do conhecimento sob o arco pragmatista-pragmático, deve produzir explicações realistas e coerentes sobre os esquemas de mundo, ao mesmo tempo que deve responder às demandas sociais, principalmente, as exigências por representação na esfera pública. Critérios realistas do pragmatismo clássico de Peirce e relativistas sustentados por versões da pragmática impõem escolhas paradigmáticas ao campo da organização do conhecimento, mas que não se resolverão com uma decisão por um ou por outro projeto de ciência quando na verdade os dois são necessários. Assim, esses dois polos não derivam de posições completamente distintas, mas de bases comuns associadas ao pragmatismo clássico do qual se derivaram até chegar na versão mais relativista que se encontra atualmente nas abordagens socio-culturais.

Nesse sentido, estar-se-ia diante de um abrangente e compósita proposta de paradigma que poderia fomentar uma teoria geral a ser concebida pela organização do conhecimento. Tal modelo explicativo do campo questiona os pressupostos racionalistas, idealistas e demasiadamente empiristas, mas sem deixar de responder às demandas realistas (pragmatismo clássico), as quais, ao fim e ao cabo, atuam para promover a reedição de sistemas universais de organização do conhecimento. Por outro lado, o reconhecimento das demandas socioculturais e a representação simbólica dos grupos oprimidos estariam na ponta mais pragmática (não confundir com pragmatista) do suposto paradigma. Assim se estaria paradoxalmente alinhando as visões de mundo pragmatista e pragmática, as quais respondem a duas grandes necessidades da organização do conhecimento: garantir a generalização-padroneização de sistemas e promover a visibilidade dos sujeitos e seus contextos.

Considerações finais

Conclui-se que o paradigma pragmatista-pragmático, nas suas diversas nuances e confrontações, seria uma explicação que incorpora a dinâmica cultural, os movimentos culturais, a igualdade de gênero, as lutas afirmativas, entre outros temas na organização do conhecimento. Nesse caso, estar-se-ia diante de um mesmo arco paradigmático, um “novo” e “velho” paradigma, com o desenvolvimento intelectual decorrente de posições tão distantes quanto as de Peirce, James, Wittgenstein, Morris e Austin. Tal arco, plural e abrangente, pode reconhecer as tendências investigativas da organização do conhecimento, das mais produtivistas e universalizantes (pragmatistas ou vinculadas ao pragmatismo mais científico) até os resultados de suas ações legitimadoras das inúmeras comunidades discursivas (pragmáticos).

Em outras palavras, quando se enfatizam os sistemas universais e se baseam na necessária correspondência entre conceito e realidade, “somos pragmatistas”; e no instante em que se sustenta que os conceitos são construções que dependem unicamente da comunidade discursiva e que o fundamental é o uso e o contexto indepen-

dente dos objetos de referência ao conceito, “somos pragmáticos”. Entende-se que quando se seguia um pressuposto cientificista para descrever os conceitos para compor sistemas universais em áreas científicas estava-se sob a égide pragmatista (relativamente peirceano) e quando se dedica à construção de sistemas particulares voltados a comunidades discursivas, reconhecem-se os pressupostos pragmáticos (relativamente James, Wittgenstein, Morris, Austin etc.). Como essas duas linhas estão presentes e bem ativas em organização do conhecimento, se está diante do mesmo espectro paradigmático pragmatista-pragmático.

Nesse sentido, não haveria incoerência em sustentar que se está diante de um mesmo grande paradigma, pois há pontos em comum como: o rechaço a soluções unicamente racionais e idealistas que não respondem a prática; a valorização de uma cientificidade acordada coletivamente, mas respaldada pela realidade (fatos científicos, tais como preconceito, fome, desigualdade, violência são uma realidade e não versões da realidade); a linguagem é fundamental para representar, seja o mundo físico-químico ou as relações sociais; a produção do conhecimento é uma ação e um ato; a experiência é fundamental para corrigir as representações, fato esse que não se encerra; a continuidade do pensamento (semiose) suspende a imposição de leis imutáveis, mas cobra a área por padrões explicativos que ajudam os sujeitos a agir com o mundo e, por fim, quaisquer propostas de representação do conhecimento serão tomadas como hipóteses a serem testadas junto a prática, segundo fins e contextos, não havendo a definição apriorística do objeto representado, embora este sempre venha a corrigir a representação presente nos sistemas.

Agradecimento

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 (Processo: 88887.571329/2020-00).

Referências

- Almeida, C. C. (2012). The Methodological Influence of Peirce’s Pragmatism on Knowledge Organization. *Knowledge Organization*, 39, 204-215.
- Almeida, C. C., Fujita, M. S. L., & Reis, D. M. (2013). Peircean Semiotics and Subject Indexing: contributions of Speculative Grammar and Pure Logic. *Knowledge Organization*, 40, 225-241.
- Araújo, C. A. A. (2018). *O que é ciência da informação*. KMA.
- Armengaud, F. (2006). *A pragmática*. Parábola Editorial.
- Colapietro, V. M. (1993). *Glossary of semiotics*. Paragon House.
- Dahlberg, I. (2006). Knowledge Organization: A New Science? *Knowledge Organization*, 33 (1), 11-19.
- Dahlberg, I. (2014). Brief Communication: What is Knowledge Organization? *Knowledge Organization*, 41 (1), 85-91.
- De Waal, C. (2007). *Sobre pragmatismo*. Loyola.
- Dousa, T. M. (2010a). Classical Pragmatism and its Varieties: On a Pluriform Metatheoretical Perspective for Knowledge Organization. *Knowledge Organization*, 37(1), 65-71.

- Dousa, T. M. (2010b). Whither pragmatism in knowledge organization? Classical pragmatism vs. neopragmatism as KO metatheories. *Advances in Knowledge Organization*, 12, 78-84.
- Friedman, A., & Thellefsen, M. (2011). Concept theory and semiotics in knowledge organization. *Journal of Documentation*, 67(4), 644-674.
- Guimarães, J.A.C. (2017). Organização do conhecimento: passado, presente e futuro sob a perspectiva da ISKO. *Informação e Informação*, 22(2), 84-98.
- Hjørland, B. (2003). Fundamentals of knowledge organization. *Knowledge Organization*, 30(2), 87-111.
- Hjørland, B. (2016). Knowledge organization. *Knowledge Organization*, 43(6), 475-484.
- Ibri I. A. (1992). *Kósmos noetós: a arquitetura metafísica de Charles S. Peirce*. Perspectiva; Hólon.
- Jacob, E. K. (2000). The legacy of pragmatism: implications for knowledge organization in a pluralistic universe. *Proceedings of the International ISKO Conference*, Würzburg, German, 6th. p. 16-22.
- James, W. (1967). *Pragmatismo e outros ensaios*. Lidorador.
- Kobashi, N. Y., & Fernandes, J. C. (2009). Pragmática linguística e organização da informação. *Anais do Encontro Nacional de Pesquisa em Ciência da Informação*, João Pessoa, PB, Brasil, 10.
- Lee, W.-C. (2015). Culture and classification: an introduction to thinking about ethical issues of adopting global classification standards to local environments. *Knowledge Organization*, 42(5), 302-207.
- Marcondes, D. (2009). *Textos básicos de linguagem: de Platão a Foucault*. Zahar.
- Martínez-Ávila, D., Ferreira, M., & Magro, J. L. (2015). Aplicación de la teoría crítica de raza en la organización y representación del conocimiento. *Scire*, 21, 27-33.
- Martínez-Ávila, D., Semidao, R., & Ferreira, M. (2016). Methodological aspects of critical theories in knowledge organization. *Knowledge Organization*, 43, 118-125.
- Morris, C. W. (1976). *Fundamentos da teoria dos signos*. Livraria Eldorado Tijuca; Edusp.
- Nascimento, E. M. (2020). Richard Rorty e a filosofia sem fundamentos. *Cognitio-Estudos*, 17(1), 81-91.
- Neves, M. A. C. (2012). *Semiótica linguística e hermenêutica do texto jurídico* (2.^a ed.). Edições Piaget.
- Nöth, W. (2005). *A semiótica no século XX* (3.^a ed.). Annablume.
- Nunes, J. A. (2009). O resgate da epistemologia. In B. S. Santos & M. P. Meneses (org.), *Epistemologias do sul* (pp. 215-242). Edições Almedina.
- Peirce, C. S. (1972). *Semiótica e filosofia*. Cultrix.
- Pinto, J. P. (2012). Pragmática. In F. Mussalim & A.C. Bentes (Org.), *Introdução à linguística: domínios e fronteiras* (8.^a ed., pp. 55-80). Cortez.
- Revez, J., & Borges, L. C. (2018). Pragmatic paradigm in information science research: a literature review. *Qualitative and Quantitative Methods in Libraries*, 7, 583-593.
- Saldanha, G. S. (2011). Entre a retórica e a filologia: do pragmatismo ao humanismo na epistemologia da ciência da informação. *IncID*, 2(1), 47-67.
- San Segundo Manuel, R., & Martínez Ávila, D. (2012). El orden de los saberes y la organización digital. In C. Pérez Pais & M.A. González Bonome (Org.), *Capítulo Español de ISKO: actas del XX Congreso ISKO-España* (pp. 413-421). Universidade de Coruña.
- San Segundo, R., & Codina-Canet, M.A. (2018). Epistemological challenges in Knowledge Organization in the digital age. In F. Ribeiro, & M. E. Cerveira (Ed.), *Challenges and Opportunities for Knowledge Organization in the Digital Age* (pp. 198-203). ISKO.
- San Segundo, R. (2004). A new conception of representation of knowledge. *Knowledge Organization*, 31(2), 106-111.

- San Segundo, R. (2008). Epistemologías posmodernas en información y documentación. In J. A. Frías Montoya, & C. Travieso Rodríguez (Org.), *Formación, investigación y mercado laboral en Información y Documentación en España y Portugal: III Encuentro Ibérico de Docentes e Investigadores en Información y Documentación* (pp. 675-684). Ediciones Universidad Salamanca.
- Santaella, L. (2004). *O método anticartesiano de C. S. Peirce*. Unesp.
- Schucman, L. V. (2014). Sim, nós somos racistas: estudo psicossocial da branquitude paulistana. *Psicologia & Sociedade*, 26(1), 83-94.
- Silveira, L. F. B. (2007). *Curso de semiótica geral*. Quartier Latin.
- Thellefsen, T. L. (2002). Semiotic knowledge organization: theory and method development. *Semiotica*, 142(1/4), 71-90.
- Thellefsen, T. L. (2004). Knowledge profiling: the basis for knowledge organization. *Library Trends*, 52(3), 507-514.
- Thellefsen, T. L. (2009). *Fundamental signs and significance-effects: a semeiotic outline of fundamental signs, significance-effects, knowledge profiling and their use in knowledge organization and branding*. (Doctoral Dissertation, Department of internationale Culture and Communication Studies, Copenhagen Business School, Denmark).
- Thellefsen, T. L., & Thellefsen, M. M. (2004). Pragmatic semiotics and knowledge organization. *Knowledge Organization*, 31(3), 177-187.
- Vaamonde Gamo, M. (2015). *Debate feminista contemporáneo: aportaciones de John Dewey*. Biblioteca Nueva.
- Weiss, L. C., & Bräscher, M. (2014). Pragmática na organização do conhecimento. *Anais do Encontro Nacional de Pesquisa em Ciência da Informação*, Belo Horizonte, MG, Brasil, 15.
- Weiss, L. C., Bräscher, M., & Vianna, W. B. (2016). Pragmatism, Constructivism and Knowledge Organization. *Advances in Knowledge Organization*, 15, 211-218.
- Wittgenstein, L. (1989). *Investigações filosóficas*. Nova Cultural.
- Wittgenstein, L. (2010). *Tractatus logico-philosophicus* (3.^a ed). EDUSP.
- Yule, G. (2008). *El lenguaje* (3.^a ed). Ediciones Akal.

DISCIPLINAS E TEORIAS DA CIÊNCIA DA INFORMAÇÃO: A ORGANIZAÇÃO DO CONHECIMENTO EM FOCO

Richele Grengue Vignoli¹, Carlos Cândido de Almeida²

¹Universidade Estadual Paulista, r.vignoi@unesp.br, ORCID iD 0000-0003-1550-5258

²Universidade Estadual Paulista, carlos.c.almeida@unesp.br,
ORCID iD 0000-0002-8552-1029

Resumo

Objetivos: O estudo disserta a respeito de disciplinas que fundamentaram o surgimento da Ciência da Informação conforme demonstra a literatura científica, como a Biblioteconomia, a Bibliografia, a Documentação, e em especial, a Organização do Conhecimento. No mesmo sentido, a pesquisa discute acerca de teorias como a Teoria Matemática da Comunicação, a Cibernética, a Recuperação da Informação (*information retrieval*) e a Teoria de Sistemas que contribuíram para o arcabouço teórico da Ciência da Informação. **Métodos:** Trata-se de um estudo bibliográfico com abordagem qualitativa. A busca por literatura foi realizada na base de dados *Web of Science*, Base de dados Referenciais de Artigos de Periódicos em Ciência da Informação e no portal e edições da *International Society for Knowledge Organization*. Não foram determinados períodos específicos para a coleta de literatura científica, que ocorreu em inglês, espanhol e português. **Resultados/Conclusões:** É importante compreender que a Ciência da informação é um campo do conhecimento que apesar de manifestar-se mais amplamente na década de 1960, não nasce sem precedentes históricos, e que a a Biblioteconomia, a Bibliografia, a Documentação e a Organização do Conhecimento fornecem à Ciência da Informação, importante aporte teórico à sua prática e teorização.

Palavras-chave: Fundamentos da Ciência da Informação, Organização do Conhecimento, Epistemologia da Ciência da Informação.

1 Introdução

A Ciência da Informação teve seu ápice no século XX, mais precisamente no contexto pós-guerra durante a década de 1960. Entretanto, sua fundamentação teórica é decorrente de outras disciplinas e teorias que culminaram em seu surgimento. O estudo volta-se à questão das disciplinas e teorias que fundamentaram o surgimento da Ciência da Informação conforme demonstra a literatura científica, como a Biblioteconomia, a Bibliografia, a Documentação, a Organização do Conhecimento, a Teoria Matemática da Comunicação (TMC), a Cibernética, a Recuperação da Informação (*information retrieval*) e a Teoria de Sistemas.

A primeira disciplina abordada ressalta o papel da biblioteca e da Biblioteconomia na constituição da Ciência da Informação e nas influências de seu arcabouço teórico-prático voltado, por exemplo, nas teorias da classificação, estudos de usuários e demais enfoques (Quadro 1). Na atualidade, a Biblioteconomia e Ciência da Informação compartilham teoria, escopo de pesquisa e, especialmente, *locus* profissional.

A Bibliografia procurou criar listas de livros como uma prática possível e necessária para organizar o conhecimento existente. No séc. XVIII, a Bibliografia foi compreendida no escopo dos estudos de organização de manuscritos e/ou livros antigos (Otlet, 1934), o que possibilitou a criação da Documentação e teve suas práticas repercutidas na Ciência da Informação.

A Documentação é referida como grande feito de Paul Otlet (1868-1944) e Henri La Fontaine (1854-1943). Na tessitura de criação da Documentação, os advogados belgas oficializaram a criação do *Office International Bibliographie* em 1892 em Bruxelas, diante da crise causada na Biblioteconomia, que esteve com foco mais voltado as bibliotecas públicas em detrimento do surgimento dos periódicos científicos (Lopez Yepes, 1995). A proliferação de periódicos e do conhecimento científico criou demanda para uma ciência responsável em sistematizar e organizar esse tipo de conhecimento. Nesse cenário, Otlet (1934) expande práticas da Biblioteconomia e Bibliografia na conjectura de uma ciência que visou também, tratar documentos para além de livros, a Documentação.

Após surgimento da Ciência da Informação no pós-guerra, principalmente com a oficialização do primeiro conceito da nova ciência na *Georgia Tech* em 1961 e 1962 e com as publicações dos artigos de Bush (1945) e Borko (1968), a Organização do Conhecimento surge como disciplina que objetiva fundamentar e discutir o conhecimento humano como objeto necessário de aprofundamento a recuperação e disseminação da informação.

Conforme pontuado por Dahlberg (2006, p. 12, trad.), “Conhecimento é a certeza subjetiva e objetivamente bem fundada de alguém sobre a existência de um fato ou assunto. Este conhecimento não é transferível, só pode ser elaborado pela reflexão pessoal de alguém.” Para a autora, somente por meio da linguagem é possível transferir o conhecimento de um indivíduo no tempo e no espaço. O estudo das linguagens nas formas pelas quais podem representar o conhecimento são objetos de investigação da Organização do Conhecimento. Diferente da Biblioteconomia, Bibliografia e Documentação, a Organização do Conhecimento passou a ter mais impacto como área ou subárea de modo mais incisivo a partir do final década de 1980. Por isso, as contribuições interdisciplinares entre a Organização do Conhecimento e Ciência da Informação são destacadas como decorrentes do pós-guerra.

Nas teorias que serviram de base à formação da Ciência da Informação, também como fruto do pós-guerra, destaca-se a influência da TMC de Shannon e Weaver de 1948, que com suporte matemático possibilitou compreensão de um processo comunicacional, em que o objetivo é que a mensagem encontre o seu destino/destinatário (Shannon & Weaver, 1975) de forma linear.

A Cibernética de Wiener (1970) teve sua primeira aparição por meio do livro *Cibernética, ou controle e comunicação no animal e na máquina* de 1950. Para o autor, a comunicação com uma pessoa é possível por meio da transmissão de men-

sagens e quando o contrário acontece, uma mensagem conexa é replicada que lhe é originalmente acessível, e não a outra pessoa, como resquícios da TMC. Diante das possibilidades do controle de máquinas, em 1960, Calvin Mooers cunha o termo Recuperação da Informação (*information retrieval*), que passa a ser uma importante área de estudos e de *práxis* para a Ciência da Informação e Organização do Conhecimento. Mooers (1960) desenvolve um princípio ou lei de comportamento que acreditava governar os sistemas de recuperação da informação, em que sistemas tecnicamente bons não significavam, entretanto, garantia de utilização.

Com base no ideal de sistemas e nos preceitos de Von Bertalanffy, Niklas Luhmann desenvolve a Teoria de Sistemas. Luhmann (2016) tinha intenção de desenvolver uma teoria unificadora para a sociologia e que fosse sistêmica – baseada em sistemas e subsistemas. Na Ciência da Informação, a teoria foi importante para subsidiar estudos de sistemas de informação e do conhecimento.

Apesar da exposição das disciplinas e teorias não demonstrarem a totalidade das inferências sofridas pela Ciência da Informação ao longo de sua trajetória, acredita-se que o recorte elucidado possa configurar-se no entendimento de sua formação em disciplinas e teorias que foram basilares para sua formação e, sobretudo, do papel da Organização do Conhecimento nesse contexto.

2 Metodologia

Para atingir os objetivos da pesquisa, um estudo bibliográfico com abordagem qualitativa foi realizado. A pesquisa bibliográfica é definida por Prodanov e Freitas (2013) como àquela que é desenvolvida a partir de material já publicado, como por exemplo: livros, revistas, periódicos científicos, jornais, teses e dissertações, internet e outras fontes que inserem o pesquisador em contato direto com todo material já escrito e publicado a respeito da temática da pesquisa. A abordagem do estudo é qualitativa, pois “Os objetos não são reduzidos a simples variáveis, mas sim representados em sua totalidade, dentro de seus contextos cotidianos”. (Flick, 2009, p. 24).

A busca por literatura científica foi realizada na base de dados *Web of Science*, Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI) e no Portal e edições da *International Society for Knowledge Organization* (ISKO) como forma de contemplar a literatura especializada da Organização do Conhecimento.

Os principais termos de busca empregados foram perpetuados no âmbito da epistemologia da Ciência da Informação, quais sejam: Ciência da Informação AND Biblioteconomia; Ciência da Informação AND Bibliografia; Ciência da Informação AND Documentação; Ciência da Informação AND Organização do Conhecimento e no mesmo sentido e critérios, com as teorias estudadas: Ciência da Informação AND Teoria Matemática da Comunicação; Ciência da Informação AND Cibernética; Ciência da Informação AND Recuperação da Informação e Ciência da Informação AND Teoria de Sistemas. Não foram determinados períodos específicos para a coleta de literatura científica. As buscas ocorreram em inglês, espanhol e português.

3 As influências disciplinares e teóricas

Esta seção procura apontar algumas contribuições da Biblioteconomia, Bibliografia, Documentação e Organização do Conhecimento a Ciência da Informação durante seu percurso científico. Os resultados da pesquisa indicam que apesar da Ciência da Informação ter surgido oficialmente na década de 1960, outros acontecimentos, disciplinas e teorias foram cruciais para a sua construção e propagação. Sob esses aspectos, é possível que a Ciência da Informação tenha utilizado aportes teóricos e práticos de suas antecessoras como a Biblioteconomia, a Bibliografia e a Documentação e da Organização do Conhecimento, como sua antecessora para sua sustentação enquanto campo de conhecimento. Para explicar como as áreas fazem relação com a Ciência da Informação e a Organização do Conhecimento, uma síntese dessas inferências foi elaborada no Quadro 1.

Quadro 1. Síntese das contribuições disciplinares a Ciência da Informação

Biblioteconomia	<ul style="list-style-type: none"> • Primeiros sistemas de classificação; • Acesso público ao conhecimento a partir de Gabriel Naudé; • Visão social do conhecimento com J. Shera; • Aprimoramento de técnicas e práticas de organização do conhecimento e da informação; • Junção de ensino e pesquisa; • Estudo de usuários; • Criação da LIS; • Estrutura profissional.
Bibliografia	<ul style="list-style-type: none"> • Estudo e organização de manuscritos; • Sistemas de organização do conhecimento; • Surgimento dos catálogos e listas de livros; • Instrumentos para recuperação da informação; • Bases estruturais para a comunicação científica.
Documentação	<ul style="list-style-type: none"> • Intenção de inventariar o conhecimento humano; • Aprimoramento de técnicas e criação de outras para organizar o conhecimento; • Valorização do periódico e conhecimento técnico e científico; • Valorização do conteúdo do objeto (informação) em detrimento de seu suporte; • Início da representação da informação e do conhecimento; • Expansão da concepção de objeto informacional.
Organização do Conhecimento	<ul style="list-style-type: none"> • Preocupação com estudos aprofundados acerca do conhecimento; • Conhecimento compreendido no contexto da Ciência da Informação; • Inserção da Teoria de Conceitos, terminologia, representação da informação e conhecimento, teorias de classificação, indexação; • Inserção da análise de domínio e de conteúdo; linguística; • Estudo e aprimoramento de <i>Knowledge Organization Systems</i> (KOS) • Estudos de organização e representação da informação na <i>web</i>: ontologias, redes semânticas e outras.

Referência: Elaborado pelos autores.

Conforme demonstrado no Quadro 1, a partir do trajeto das bibliotecas e da Biblioteconomia as primeiras noções de organização do conhecimento foram semeadas por meio da classificação. As bibliotecas foram importantes instrumentos

de acesso ao conhecimento e por isso, destaca-se o papel de Gabriel Naudé na luta por tornar esse conhecimento público e aberto. Entre a visão de uma Biblioteconomia mais humanizada em Shera, os instrumentos de organização do conhecimento foram aprimorados, assim como surgiu a preocupação com o usuário da informação. Entre ensino e prática, pesquisa e campo profissional, a Biblioteconomia e a Ciência da Informação continuam a compartilhar de problemas de estudo e de práticas informacionais em comum ou muito próximas. Para Saracevic (1996), a Biblioteconomia e a Ciência da Informação são disciplinas distintas e ao mesmo tempo aliadas, pois se manifestam em campo comum e com forte relação interdisciplinar. Há relação entre a Biblioteconomia e a Ciência da Informação na prática de ensino e pesquisa, pois as áreas se misturam nas temáticas de organização, gestão, mediação, políticas, comunicação científica, tecnologias, memória, estudo de usuários e outros (SILVA, 2017).

A classificação responde por uma das principais teorias de sua história e de contributo teórico e prático tanto para a Ciência da Informação quanto para a Documentação e Organização do Conhecimento. As primeiras classificações foram filológicas, antecedidas por Aristóteles, Platão, Porfírio com sua *Árvore* e classificação binária (BARBOSA, 1969), Cassiodoro e outros como Gessner. Konrad Gessner (1516-1565), baseado no sistema de Trivium e Quatrivium de Cassiodoro criou e organizou um catálogo denominado de *Bibliotheca Universalis*, utilizado para registrar livros em latim, grego e hebraico. Barbosa (1969) explica que as ações de Gessner podem ter sido a primeira tentativa metódica de organização de livros e/ou o primeiro esquema de classificação bibliográfica existente da humanidade com impactos diretos na Biblioteconomia e Bibliografia. Já na esteira da classificação bibliográfica, destacam-se os papéis de Bliss, Cutter, Dewey e Ranganatan que juntos respondem aos principais instrumentos e teorias de classificação da Biblioteconomia e da Organização do Conhecimento.

A esse respeito, ressalta-se a importância da criação da Biblioteconomia e Ciência da Informação (*Library and Information Science* – LIS) nos EUA, que objetivou a junção das áreas. Hjørland (2018a, 2018b) em estudo dedicado a desvendar os princípios da Organização do Conhecimento aplicada a *Library and Information Science* frisa que a LIS é como o nome indica, a combinação da Biblioteconomia com a Ciência da Informação.

Entretanto, apesar da confluência e às vezes, separação dos conteúdos nos programas de pós-graduação em Ciência da Informação, é possível visualizar com alguma clareza pesquisas específicas de cada área. Como disciplinas distintas, a Ciência da Informação se fundamenta na Biblioteconomia principalmente em sua teoria da classificação e em seu lócus profissional. Divide ou compartilha temáticas de pesquisa, de ensino e de prática de seus profissionais, apesar das tentativas em separá-las ou em demonstrar uma suposta ideia de superioridade da Ciência da Informação diante da Biblioteconomia.

Na Bibliografia, sua influência na Ciência da Informação pode ser destacada no princípio da organização do conhecimento humano em listas do que há de publicação e na preservação de sua memória (Pinheiro, 2005). Na Bibliografia, os estudos pautados no registro gráfico e baseado em livros e manuscritos, fortaleceram o desenvolvimento de verdadeiros catálogos dinâmicos do conhecimento disponíveis

em uma ampla gama de canais, especialmente os científicos. A comunicação científica tornou a Bibliografia viva e ainda mais necessária como prática na sociedade da informação e do conhecimento e para a Ciência da Informação. A Ciência da Informação empresta da Bibliografia seu corpo de conhecimentos relacionados à construção da organização dos registros gráficos do conhecimento que objetivam sua disseminação.

A trajetória da Bibliografia em sua práxis na difusão do conhecimento se solidifica na Ciência da Informação especialmente nas ações voltadas a comunicação científica. Para Spinak (1996), a Bibliografia atual resulta de demandas de investigação advindas do desenvolvimento do número crescente de revistas na formação de sociedades científicas e literárias. Não por acaso, a Bibliografia que foi iniciada na descrição de registros bibliográficos, agora tem como propósito auxiliar na disseminação do conhecimento (Spinak, 1996). Para Bálamo (1998), a Bibliografia é um completo sistema de comunicação social, assim como pensa Saldanha (2020), pois a área continua a exercer seu papel por meio de outros produtos derivados de suas práticas, como a comunicação científica (Hjørland, 2018a, 2018b), também objeto de estudo da Ciência da Informação. Segundo Pinheiro (2005), a Ciência da Informação tem como uma de suas raízes a Bibliografia e a Documentação, pensamento compartilhado com Saldanha (2020) que defende que a relação da Bibliografia com a Ciência da Informação é irrefutável, já que a primeira lança na segunda, sua função estrutural.

A Documentação expandiu conceitos, técnicas e objetivou tornar sua práxis uma disciplina científica. Acabou por projetar na Ciência da Informação a compreensão da multiplicidade de formas e formatos da informação, além da importância da informação e seu acesso em detrimento de sua guarda. A partir da noção de Representação da Informação proposta por Otlet, a Documentação fortaleceu a compreensão de relação do objeto de estudo da Ciência da Informação com a sua forma de tratamento, organização e disseminação.

Na Documentação, o foco no acesso a diversos tipos documentais e ao conhecimento científico (Lopez Yopez, 1995), além do início dos estudos de Representação da Informação são ações institucionalizadas na disciplinas que tanto a Ciência da Informação, quanto a Organização do Conhecimento têm incorporado nas últimas décadas. Na Documentação, a informação é valorada e não o suporte do documento ou da obra. Mais importante que possuir uma obra é disponibilizar o seu acesso, aspecto primordial que a Documentação apresenta à Ciência da Informação, e também na condição de tratar o conhecimento científico em disciplina científica.

Dotada de uma frente de trabalho que objetivava possuir uma extensa rede de registros dos acervos custodiados nas instituições, a Documentação não previa armazenar ou constituir uma coleção física (como a Biblioteconomia), mas inventariar a produção do conhecimento (Araújo, 2018). A Documentação traz semelhanças consideráveis com a Bibliografia, se não fosse na pluralidade de suportes documentais tratados e na ressignificação do livro com um dos diversos tipos documentais. Esse parece ser um dos grandes feitos da Documentação que visava ações para além do registro e formato da informação, ou ainda, em tudo o poderia transmitir o conhecimento: livros, revistas, artigos, periódicos, filmes, desenhos, selos, medalhas, quadros, amostras e outros (Vicentini, 1970). A Documentação de Otlet conduziu igualmente, a prática bibliográfica nos registros científicos do conhecimento.

Desse modo, a disciplina se pautou, sobretudo, na organização do conhecimento científico e técnico, em sua projeção como campo científico e nas formas de representar a informação em qualquer suporte e formato. Para Freire (2006), o foco no conteúdo e na informação contida nos documentos discutida na Documentação representou uma mudança paradigmática, uma inovação e novas formas de visualizar a informação, em especial, a registrada. Era objetivo de Otlet criar novas informações baseadas na informação original, já estocada, isto é, conduzir à representação do conhecimento, da informação. Surge nesse âmbito os conceitos de descrição e síntese da informação – a sua representação condensada que tornava o acesso mais dinâmico e otimizado (Tálamo; Smit, 2007). Não se tratava, contudo, de explorar o documento e seus formatos, mas a informação e suas formas de representá-la e acessá-la. Portanto, a defesa é que quando a Ciência da Informação surge na década de 1960, utiliza de toda a trajetória da Biblioteconomia, Bibliografia e Documentação, aliando-se principalmente às questões tecnológicas e científicas, para se valer como campo científico. A Ciência da Informação não surge sem o recurso às disciplinas precedentes.

Durante seu surgimento, a Ciência da Informação busca por outras teorias que pudessem lhe trazer cientificidade. Nesse momento, a Organização do Conhecimento, que também possui culminância mais acentuada no pós-guerra, suplanta na Ciência da Informação, valorosas teorias e aprofundamentos acerca do conhecimento humano. Muito discutido no paradigma cognitivo, o conhecimento esteve como objeto de estudo, fortemente associado as ciências cognitivas. Na Organização do Conhecimento, foi possível a sua extratificação enquanto objeto que conduz a sua tradução e ordenação por linguagens específicas e padronizadas, a sua representação e utilização por usuários com suas necessidades em informação.

A Organização do Conhecimento teve seu auge de propagação mais acentuado partir de 1989 quando sua denominação foi firmada na *International Society for Knowledge Organization* (ISKO) baseada no termo utilizado por Henry Evelyn Bliss em 1929 e 1933, e também a partir da publicação do periódico *International Classification* em 1974, que teve seu título alterado para *Knowledge Organization* em 1993. Entretanto, a simplificação da Organização do Conhecimento apenas na classificação se mostrou inadequada para demonstrar a pluralidade de suas abordagens. Por isso, Dahlberg (2006, p. 12, trad.) explica que na mudança de nome, o periódico enfatizou que a Organização do Conhecimento se definia por: [...] ‘objetos e atividades da teoria do conceito, classificação e indexação e representação do conhecimento’, onde por ‘representação do conhecimento’ não compreendeu apenas a estrutura lógica da representação conceitual, mas também todas as questões de nomear conceitos pelos termos mais adequados, em que questões de terminologia são consideradas. Assim, a classificação incorpora-se a Organização do Conhecimento como uma de suas atividades e métodos, mas não representa sua totalidade. O mesmo caso acontece com a indexação. Em entrevista realizada em 2007, Dahlberg (2007), ao falar a respeito da Organização do Conhecimento e sua nomenclatura, explicou que se trata do campo de atuação da classificação conceitual, da teoria do conceito, da indexação e da representação de conceitos e não somente da classificação. Nesse entendimento, a Organização do Conhecimento fornece a Ciência da Informação aspectos pertinentes aos estudos voltados ao conhecimento humano e

enquanto formas padronizadas de tratá-lo, organizá-lo e compartilhá-lo como informação.

A Organização do Conhecimento para Hider (2018), como uma disciplina filosófica e de outras áreas como da Biblioteconomia e Ciência da Informação, se mescla a outros campos do conhecimento e é associada à compreensão da natureza do conhecimento, da cognição, da linguagem e da organização social do conhecimento (Hjørland, 2008). A Organização do Conhecimento pauta-se no “[...] conhecimento documentado produzido por seres humanos em algum tipo de documento de uso potencial para outros seres humanos” (Hjørland, 2003, p. 93), e por isso, é fundamento pertinente na construção e consolidação teórica da Ciência da Informação.

A menção aos feitos de Otlet e La Fontaine são igualmente referenciadas por Dahlberg (2006), essencialmente no desenvolvimento da Classificação Decimal Universal que visou desde o início, à relação entre conceitos. Hjørland (2003, 2008) demarca a necessidade de estudo da Organização do Conhecimento por conceitos, mas enaltece a disciplina com seu caráter social, ou na organização social do conhecimento. Para Smiraglia (2013b), a organização do conhecimento dedica-se à ordem conceitual do conhecimento e as heurísticas do conhecimento em ordenação são estudadas. Para o autor, a organização do conhecimento é uma comunidade de pesquisa dedicada à classificação, ontologia, tesouros e vocabulário controlado, epistemologia e garantia, e no desenvolvimento de sistemas aplicados para a Organização do Conhecimento. Nesse sentido, a teoria do conceito de Dahlberg, a classificação, a indexação e a terminologia, além de outros estudos como a respeito dos KOS e relações semânticas na *web* são aportes que contribuem com o *corpus* teórico e prático da Ciência da Informação, ao condicionarem a informação a tratamentos e metodologias padronizadas antes de sua recuperação e disseminação.

Apesar da síntese das contribuições e inferências das disciplinas não representarem ordem linear ou cronológica dos fatos, objetivou-se por meio dos pontos destacados, salientar algumas contribuições das áreas para a Ciência da Informação. Contudo, é possível perceber que a Organização do Conhecimento perpassa todo o contexto teórico e prático tanto da Biblioteconomia, quanto da Bibliografia, Documentação e Ciência da Informação. Ainda que a Organização do Conhecimento enquanto área, campo ou disciplina ainda não estivesse institucionalizada, as práticas dos campos discutidos sempre envolveram a preocupação com a organização do conhecimento humano.

Por outro lado, deve-se mencionar as contribuições teóricas, especificamente das Teoria da Matemática da Comunicação, da Cibernética, da Recuperação da Informação e da Teoria de Sistemas à Ciência da Informação. Para tanto, o Quadro 2 foi desenvolvido para demonstrar algumas influências à Ciência da Informação.

O quadro 1 demonstra contribuições basilares da TMC, da Cibernética, da Recuperação da Informação e da Teoria de Sistemas à Ciência da Informação. É possível observar que ocorre uma espécie de evolução entre as teorias, nas quais a Ciência da Informação utilizou de aportes e técnicas em seu escopo de formação. Nota-se da mesma forma, que as teorias contribuíram em diversos vieses na Ciência da Informação e que essas influências continuam a moldar seu núcleo de estudos e práticas. A informação, como objeto de estudo da Ciência da Informação, deve à TMC a sua primeira conceituação enquanto teoria e objeto científico. A Teoria foi ponto de

Quadro 2. Síntese das contribuições de teorias à Ciência da Informação

Teoria Matemática da Comunicação (TMC) (1948)	<ul style="list-style-type: none"> • Processo linear de comunicação; • Transferência de mensagens em qualquer suporte; • Quantificação da informação; • Primeira teoria da informação. 	<i>Transferência de mensagens e/ou informação.</i>
Teoria Cibernética (1950)	<ul style="list-style-type: none"> • Transferência de mensagens com controle do homem sobre máquinas; • Relação de <i>feedback</i> das máquinas em controle humano; • Início da compreensão de retorno do que é enviado/transferido. 	<i>Transferência com uso de máquinas controladas.</i>
Teoria da Recuperação da Informação (1960)	<ul style="list-style-type: none"> • Sistemas de recuperação da informação com aporte tecnológico (uso de máquinas); • Estudos com foco na recuperação da informação com precisão e revocação; • Princípio de <i>feedback</i> aplicado no usuário na recuperação da informação. 	<i>Transferência e possibilidade de autonomia na recuperação da informação.</i>
Teoria de Sistemas (1984)	<ul style="list-style-type: none"> • Concepção de sistemas; • Ampliação dos atores, inclusive de não vivos (máquina) que operam no processo comunicacional; • O princípio de <i>feedback</i> envolve sistemas complexos que se autorreproduzem na autopoiese. 	<i>Transferência, mas com autonomia na recuperação e processo comunicacional ampliado para todo tipo de ator que age e se reproduz em sistemas.</i>

Referência: Elaborado pelos autores.

partida para a compreensão do fluxo e processo comunicacional que envolve a informação e seus sujeitos, assim como de sua quantificação, uma necessidade daquele período.

A Cibernética perpetuou na Ciência da Informação e em outras áreas, o ideal de controle das máquinas pelo humano, o que fez emergir a noção de sistemas automáticos de recuperação da informação e o fortalecimento da tecnologia associada a informação. O controle das máquinas pelo homem tornou possível a organização e recuperação da informação por meio de computadores e o estudo do comportamento humano nessas ações. Para Pinheiro (2002), devido à importância da TMC e da Cibernética para as ciências de forma universal, na Ciência da Informação, as teorias foram determinantes para que a informação passasse a compor debates mundiais e intensos em múltiplas abordagens. Santos Júnior (2012) exprime a importância da Cibernética para a Ciência da Informação por meio da *Informatika* soviética e russa com o aprofundamento dos estudos métricos com a cientometria em seu núcleo.

Na Recuperação da Informação, Mooers descobriu que a interação entre a necessidade de informação e o que se recupera, nem sempre depende de um bom sistema de informação. De extrema importância para a Ciência da Informação, Organização do Conhecimento e outras áreas correlatas, a Recuperação da Informação foi um divisor de águas para os campos do conhecimento que possuem a informação como

objeto de estudo voltado a necessidades informacionais. A Recuperação da Informação é uma subárea que associa a TMC (no fluxo da informação), a Cibernética (na recuperação da informação mecânica, eletrônica, digital e em seu *feedback*) e a Teoria de Sistemas (no *feedback* e autopoiese) a outras possibilidades e práticas na Ciência da Informação. A Teoria de Sistemas de Luhmann de 1984 fez emergir entre outras ações na Ciência da Informação, a noção de sistemas de informação e de conhecimento e a ideia de que as partes representam o todo. Para Capurro e Hjørland (2007) a informação segundo Luhmann não é idêntica para receptor e emissor, e dependerá de sua constituição em processos de comunicação. As teorias foram a centelha de cientificidade que a Ciência da Informação buscava naquele momento para se valer como confiável, útil e de importância à sociedade.

4 Conclusões

As disciplinas e as teorias destacadas são frações de acontecimentos que culminaram no surgimento da Ciência da Informação. Como qualquer outro campo científico, a Ciência da Informação se fortaleceu em decorrência de aportes disciplinares e teóricos de outras disciplinas e/ou áreas do conhecimento. É importante compreender que a Ciência da Informação é um campo do conhecimento que apesar de manifestar-se mais amplamente na década de 1960, não nasce sem precedentes históricos, como decorrentes das disciplinas discutidas. As disciplinas funcionam também como teorias, já que fornecem seus aportes teóricos a Ciência da Informação. Quanto as teorias discutidas no estudo, ressalta-se que foram cruciais para o momento em que a Ciência da Informação buscava por sua cientificidade em seu surgimento, ainda que advinda das ciências duras.

A partir de aportes teóricos da Organização do Conhecimento como provenientes das teorias do conceito, classificação, indexação, terminologia, análise de domínio e de outros, estudos aprofundados acerca do conhecimento humano na Ciência da Informação tornaram-se possíveis e imprescindíveis para organizar a disseminar a informação na sociedade. Também é salutar ressaltar os estudos da Organização do Conhecimento acerca dos KOS, das abordagens semânticas de tratamento do conhecimento na *web* nas ontologias e *web* semântica, além de outras discussões. A teoria de conceitos envolve todas as relações epistemológicas e práticas da Organização do Conhecimento enquanto campo e na Ciência da Informação. Observa-se de igual modo, que todas as disciplinas discutidas compartilham de teorias, técnicas e de literatura em comum na construção e desenvolvimento de seus objetos de estudo.

Referências

- Araújo, C. A. A. 2018. *O que é ciência da informação*. Belo Horizonte, MG: KMA.
- Balsamo, L. 1998. *La bibliografía: historia de una tradición*. Gijón: Trea, 1998.
- Barbosa, A. P. 1969. *Teoria e prática dos sistemas de classificação bibliográfica*. Rio de Janeiro: IBBD.
- Borko, H. 1968. Information science: what is it? *American Documentation*, 19(1), pp. 03-05.
- Bush, V. 1945. As we may think. *Atlantic Monthly*, 76(1), pp. 101-108.

- Capurro, R., Hjørland, B. 2007. O conceito de informação. *Perspectivas em Ciência da Informação*, 12(1), pp. 148-207, jan./jun.
- Dahlberg, I. 2006. Knowledge organization: a new science? *Knowledge Organization*, 33(1), pp. 11-19.
- Dahlberg, I. 2007. Feature: interview with Ingetraut Dahlberg: december 2007. *Knowledge Organization, Germany*, 35(2/3), pp. 82-87.
- Flick, U. 2009. Introdução a pesquisa qualitativa. Porto Alegre: Artmed.
- Freire, G. H. de A. 2006. Ciência da informação: temática, histórias e fundamentos. *Perspectivas em Ciência da Informação*, 11(1), pp. 06-19.
- Hider, P. 2018. The terminological and disciplinary origins of information and knowledge organization. *Education for Information*, 34(1), pp. 135-161.
- Hjørland, B. 2003. Fundamentals of knowledge organization. *Knowledge Organization*, 30(2).
- Hjørland, B. 2018a. "Library and Information Science (LIS). Part 1." *Knowledge Organization*, 45(3), pp. 232-254.
- Hjørland, B. 2018b. "Library and Information Science (LIS). Part 2." *Knowledge Organization*, 45(4), pp. 319-338.
- Hjørland, B. 2005. What is knowledge organization (KO). *Knowledge Organization*, 35(2/3), pp. 86-101.
- Lopez Yepes, J. 1995. *La documentacion como disciplina: teoria e historia*. Navarra: Ediciones Universidad del Navarra.
- Luhmann, N. 2016. *Sistemas sociais: esboços de uma teoria geral*. Petrópolis: Vozes.
- Mooers, C. 1960. Zatocoding and developments ind information retrieval. *Aslib Proceedings*, 1.
- Otlet, P. 2018 [1934]. *Tratado de documentação: o livro sobre o livro teoria e prática*. Brasília: Briquet de Lemos.
- Pinheiro, L. V. R. 2005. Processo evolutivo e tendências contemporâneas da ciência da informação. *Informação & Sociedade: Estudos*, 15(1), 13-48.
- Prodanov, C. C., Freitas, E. C. de. 2013. *Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico*. 2. ed. Novo Hamburgo: Frevale.
- Saldanha, G. S. 2020. Linha cumeada: uma arqueologia dos enunciados epistemológicos da bibliografia na fundamentação da ciência da informação. *Encontros Bibli*, 25, pp. 01-16.
- Santos júnior, R. L. dos. 2012. Os estudos cientométricos na antiga União Soviética e Rússia: origens, desenvolvimento e tendências. In: Pinheiro, L. V. R.; Oliveira, E. da C. P. de. (org.). *Múltiplas facetas da comunicação e divulgação científicas: transformações em cinco séculos*. Brasília: IBICT.
- Saracevic, T. 1995. Interdisciplinary nature of information science. *Ciência da Informação*, 24 (1), pp. 36-41.
- Shannon, C. E.; Weaver, W. 1975. *A teoria matemática da comunicação*. São Paulo: DIFEL.
- Silva, J. L. C. 2017. As interfaces entre biblioteconomia e ciência da informação. *Pesq. Bras. Ci. Inf. Biblio. DISCIPLINAS E TEORIAS DA CIÊNCIA DA INFORMAÇÃO: A ORGANIZAÇÃO DO CONHECIMENTO EM FOCO*, 12(2), pp. 435-444.
- Smiraglia, R. P. 2013. The epistemological dimension of knowledge organization. *IRIS*, 2(1), pp. 2-11, jan./jun.
- Spinak, E. 1996. *Diccionario enciclopédico de Bibliometría, Cienciometría e Infometría*. Caracas: UNESCO.
- Tálamo, M. de F. G. M., Smit, J. W. 2007. *Ciência da informação: a transgressão metodológica*. Fortaleza: UFC.
- Wiener, N. 1970. *Cibernética ou controle e comunicação no animal e na máquina*. São Paulo: Polígono; USP.
- Vicentini, A. L. C. 1970. Da biblioteconomia à informática: evolução do conceito de documentação. *Revista do Serviço Público*, 105(3), pp. 251-296.

(Página deixada propositadamente em branco)

REFLEXÕES DOS CONCEITOS DE ORGANIZAÇÃO DO CONHECIMENTO E ORGANIZAÇÃO DA INFORMAÇÃO PELO PRISMA DE BRÄSCHER E CAFÉ: A EMERGÊNCIA DE UM PARADIGMA NO GT2 DO ENANCIB?

Marcos Oliveira da Costa¹, Thiago Henrique Bragato Barros²

¹Universidade Federal do Pará, marcos.costaoliver@gmail.com, 0000-0001-6877-6612.

²Universidade Federal do Rio Grande do Sul, bragato.barros@ufrgs.br, 0000-0001-7439-5779.

Resumo

Discute sobre a noção de conceito de Organização do Conhecimento e Organização da Informação no âmbito do Grupo de trabalho 2 do Encontro Nacional de Pesquisa em Ciência da Informação pelo prisma de Bräscher e Café (2008), para isso, utiliza-se da concepção de conceito apontada por Dahlberg (1978). **Objetivos:** Tem o intuito de compreender como os pesquisadores que publicam no GT2 do Enancib, enxergam e utilizam em seus textos os termos Organização da Informação e Organização do Conhecimento e se essa utilização condiz com os conceitos apontados pelas autoras citadas. **Métodos:** No que diz respeito ao método da pesquisa, a recuperação de dados fora feita por meio de levantamento bibliográfico nas bases Brapci e na base do Enancib, utilizou-se da noção de análise de domínio apontada por Hjørland (2002) e do entendimento de paradigma no âmbito da CI discutido por Capurro (2003) para compreender e analisar os dados obtidos. **Resultados:** identificou-se, por meio dos pressupostos epistemológicos apontados pela investigação, duas comunidades que trabalham a organização no contexto da Ciência da Informação: a Comunidade Discursiva de Organização da Informação (CDOI) e a Comunidade Discursiva de Organização do Conhecimento (CDOC), sobre a análise dos resultados, notou-se que o número de pesquisadores que publicam no evento e que identificam seus trabalhos dentro de Organização da Informação vem diminuindo, em contrapartida, o quantitativo de pesquisadores que designam suas pesquisas no contexto da Organização do Conhecimento tem uma crescente constante. Identificou-se também que a comunidade de pesquisadores que publicam em Organização da Informação compartilha do mesmo pressuposto epistemológico de Bräscher e Café (2008), e estas autoras são tidas como autoridades da área, pois seus conceitos de OI tem um quantitativo relativamente auto de citação feitas pela CDOI. Já na comunidade discursiva de organização do conhecimento (CDOC), nota-se que apesar da visão epistemológica apresentados nos trabalhos desenvolvidos por essa comunidade estarem alinhados com os conceitos das autoras, no sentido de serem idênticos, ou similares, as mesmas não são as autoridades com mais destaque nos trabalhos publicados em OC, ocupando os primeiros lugares como autoridades que embasaram os conceitos de orga-

nização do conhecimento dos pesquisadores desta área, Hjørland, seguido de Dahlberg. **Conclusão:** Por fim, percebeu-se que a Organização do Conhecimento pode estar se tornando um novo paradigma epistemológico no domínio do GT2 do Enancib.

Palavras-Chave: Organização do Conhecimento, Organização da Informação, Análise de Domínio, Epistemologia da Ciência da Informação.

Introdução

A Organização do Conhecimento (OC) e da Informação (OI) estão presentes desde a gênese da Ciência da Informação (CI), esta afirmação pode ser observada por meio do próprio contexto no qual esta ciência emergiu, sendo impulsionada pela necessidade de se organizar o conhecimento produzido pós segunda guerra mundial para fins de disseminação científica. Tal necessidade de organização manifestou-se por meio do emblemático artigo *As We May Think* de Vannevar Bush (1945) que tinha por objetivo automatizar os processos de organização do grande fluxo informacional gerado naquela época, fazendo assim, com que estas áreas não só fossem ainda mais reconhecidas como importantes para a CI, mas constituídas como intrínsecas a esta, inclusive no âmbito do desenvolvimento de pesquisas.

À vista disso, no que concerne ao cenário atual destas áreas para a CI e sobre sua relação com os domínios de pesquisa existentes no âmbito desta ciência, Francelin e Kobashi (2012) apontam para um desenvolvimento significativo de pesquisas com o enfoque em organização da informação e do conhecimento no contexto das investigações em Ciência da Informação. Todavia, demonstram também que existem vários discursos baseados na influência teórico-metodológica dos autores que publicam nestas áreas, discursos estes que são influenciados por tendências de correntes de pensamentos assumidas pelos pesquisadores.

Esta dicotomia sobre o consenso dos conceitos de OC e OI apontada pelos autores citados anteriormente, não é algo recente, pois Brascher e Café (2008) em artigo publicado no Encontro Nacional de Ciência da Informação (ENANCIB), já propunham um modelo conceitual de definição para Organização do Conhecimento e Organização da Informação com o intuito de amenizar as discrepâncias conceituais existentes nestas áreas.

Dessarte, o artigo visa discutir no âmbito do domínio do grupo de trabalho 2 (GT2) que é relativo a organização do conhecimento no ENANCIB, qual a concepção que os autores que publicam neste espaço têm dos conceitos de OC e OI, visando analisar essas comunidades discursivas da CI com o intuito de identificar se os pesquisadores da área estão chegando a um consenso e se existe um modelo (paradigma) sobre os conceitos mencionados.

Com isso, para dar aporte a pesquisa e fundamentar a discussão levantada ao decorrer do trabalho, se faz necessário ressaltar que para este artigo os conceitos de OC e OI se basearam em Brascher e Café (2008), ou seja, os dados levantados foram analisados levando em consideração a proposta de modelo conceitual que estas autoras apontaram para a área no intuito de identificar se o conceito destas pesquisadoras está se tornando um modelo aceito para as comunidades discursivas de OI e OC. Ainda sobre isso, para embasar a escolha conceitual e compreender a utilização

destes termos no contexto do domínio escolhido, foi utilizado Dahlberg (1978) e suas contribuições, mais especificamente as relacionadas a teoria do conceito. Destarte, outra soma importante para o processo de construção do arcabouço teórico foi Capurro (2003) e sua ideia de conceito de paradigma para a CI.

Por fim, o principal método utilizado no estudo é o de análise de domínio (AD) de Hjørland (2002) onde o autor aponta 11 abordagens para se fazer esse tipo de análise, especificamente para esta pesquisa foram utilizadas a abordagem bibliométrica, de indexação e recuperação de especialidades, estudos epistemológicos e críticos.

Dessarte, após os dados levantados, identificou-se duas comunidades discursivas a Comunidade discursiva de Organização da Informação (CDOI) e a Comunidade Discursiva de Organização do Conhecimento (CDOC), percebeu-se que a CDOI tem uma produção relativamente baixa se comparado aos trabalhos da CDOC o que aponta para a emergência de um paradigma no âmbito do Grupo de Trabalho 2 do Encontro Nacional de Pesquisa em Ciência da Informação.

Metodologia

O escopo metodológico fora dividido da seguinte forma: a) pesquisa bibliográfica, b) levantamento de dados, c) análise dos dados. Segue alguns quadros explicativos sobre os métodos, critérios e suas aplicações na pesquisa:

Quadro 1. Métodos e suas aplicações

MÉTODOS E CRITÉRIOS	APLICAÇÃO NA PESQUISA
1. PESQUISA BIBLIOGRÁFICA E MARCO TEÓRICO	Selecionou-se as bibliografias concernentes a temática do estudo, utilizando obras bem constituídas na área como livros e artigos científicos. Das diversas obras encontradas optou-se para fazer parte do escopo principal da pesquisa Brascher e Café (2008) por demonstrarem a relação dicotômica entre OC e OI bem como uma proposta de conceitualização destas áreas e por servir como parâmetro de paradigma desta investigação sobre a produção nacional publicada no GT2 do Enancib, utilizou-se também Dahlberg (1978) para embasar a concepção de conceito do estudo e entender melhor esta temática, Capurro (2003) fora escolhido, pois o estudo se apoiou na sua noção de paradigma e por fim, Hjørland (2002) e seu método de análise de domínio que foi primordial para análise e compreensão dos dados desta investigação.
2. LEVANTAMENTO DE DADOS	O levantamento dos dados foi realizado na base do Enancib e na Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (Brapci) considerando os anos de 2014 a 2019 especificamente o grupo de trabalho 2 deste encontro, utilizou-se os seguintes termos nestas bases: Organização da Informação, Organização do Conhecimento, Representação da Informação, Representação do Conhecimento. No que concerne ao critério de seleção dos dados, este levou em consideração a abordagem de indexação e recuperação de especialidades Hjørland (2002), pois optou-se por selecionar os estudos por meio das palavras chaves dos trabalhos levantados, onde teriam que apresentar os termos “Organização do Conhecimento” e ou “Organização da Informação” para serem considerados válidos durante a fase do levantamento de dados.

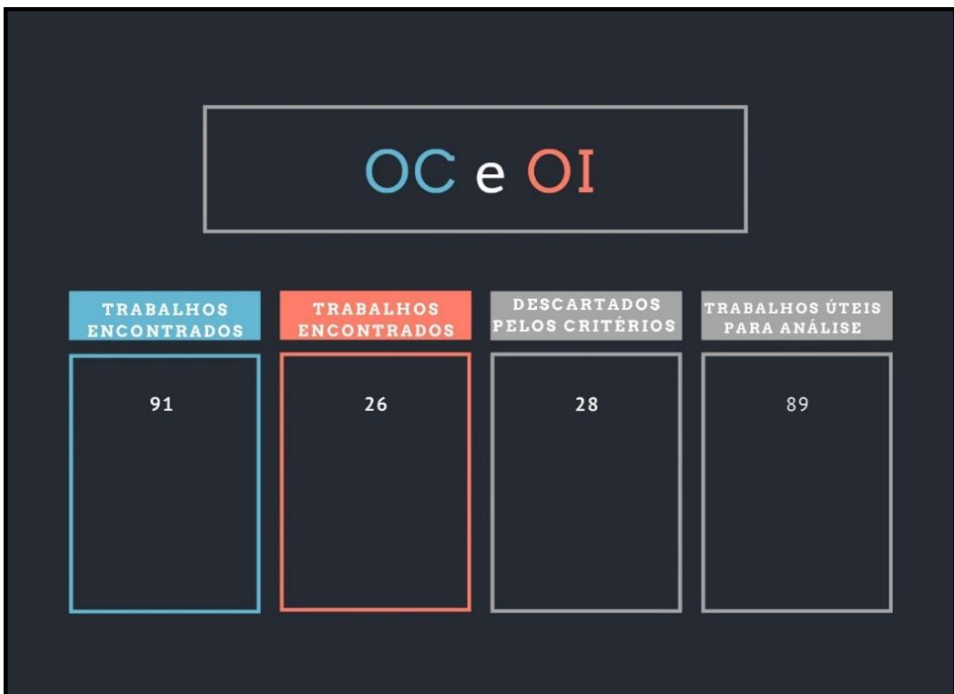
2.1 CRITÉRIOS DE AMOSTRAGEM 1	Optou-se também por definir critério de amostragem, baseado na tipologia dos trabalhos publicados no Enancib, sendo assim, foram escolhidos apenas artigos que estavam dentro da modalidade de apresentação oral, não sendo considerados os artigos em formato de pôster.
2.2 CRITÉRIOS DE AMOSTRAGEM 2	O critério de amostragem escolhido para a elaboração dos dados de autoridades, foi selecionar os autores que tiveram o quantitativo de citação maior do que 1 trabalho, ou seja, aqueles que foram utilizados como autoridades em apenas 1 trabalho da Comunidade discursiva de organização da informação e da Comunidade discursiva de organização da informação não foram considerados para o levantamento dos dados.

Fonte: Elaborado pelos autores (2021)

Como fora visto, o quadro 1 apresenta o método, os critérios, bem como o âmbito da pesquisa. No tópico 1, explana-se sobre o principal método de coleta de dados utilizado na pesquisa, o levantamento bibliográfico, bem como as bibliografias que embasaram a pesquisa como Marco Teórico. No tópico 2, aborda-se como fora realizado o processo de levantamento de dados do estudo, abordando os critérios de amostragem que foram selecionados.

A vista disso, sobre os dados levantados, segue-se uma figura:

Figura 1. Dados levantados



Fonte: Elaborado pelos autores (2021)

Análise dos dados:

Segue um quadro explicativo sobre os conceitos e suas definições aplicados na pesquisa para fomentar a análise dos dados e orientar o processo de construção teórico-metodológico deste trabalho.

Quadro 1. Conceitos utilizados na pesquisa

CONCEITO	DEFINIÇÃO	FONTE
OC	Compreende-se organização do conhecimento como o processo cognitivo de representação da realidade	BRASCHER e CAFÉ, (2008)
OI	Organização da informação, pode ser entendida como o esforço de representação física das conclusões e conceitos gerados pela organização do conhecimento, com o propósito de se fazer entendido por alguma comunidade, ou domínio	BRASCHER e CAFÉ, (2008)
Conceito	A definição de conceitos é uma espécie de limitação que visa fixar os limites destes conceitos, e os conceitos individuais podem ser compreendidos como válidos para os indivíduos que compreendem a sua aplicação mais geral no domínio	DAHLBERG (1978)
Comunidade discursiva	As comunidades discursivas são grupos científicos formados por autores, ou instituições que compartilham do mesmo pressuposto epistemológico	HJØRLAND (2002)
Abordagem Indexação e Recuperação de especialidades	A indexação pode ser vista como a forma pela qual uma comunidade discursiva representa seu domínio, no âmbito	HJØRLAND (2002)
Estudos Epistemológicos	Os estudos epistemológicos são estudos que examinam as suposições explícitas ou implícitas por trás das tradições de pesquisa”, para o autor, estes estudos estão diretamente ligados a investigação dos paradigmas que influenciam a construção científica em determinado campo de investigação	HJØRLAND (2002)
Estudo Bibliométrico	Método empírico baseado em análises detalhadas de conexões entre indivíduos e documentos	HJØRLAND (2002)
Paradigma	Tendência histórica de pesquisa de uma determinada área de domínio	CAPURRO (2003)

Fonte: Elaborado pelos autores (2021)

Como visto, a pesquisa baseou-se nos conceitos de OC e OI de Brascher e Café (2008) como parâmetro, visando compreender se estes estariam se tornando paradigmas (modelo) para os pesquisadores que publicam no GT2 do Enancib, a definição de conceito apoiou-se em Dahlberg (1978) e o principal método da pesquisa é o de Análise de Domínio de Hjørland (2002) especificamente as abordagens de inde-

ção, visto que os dados levantados foram escolhidos baseados nas palavras chaves que os próprios autores dos trabalhos selecionados, escolheram para designar o domínio de suas pesquisas a abordagem epistemológica, visto que este estudo investiga como anda o desenvolvimento científico de áreas específicas da CI a OC e OI e quais os paradigmas que influenciam os pesquisadores da área e por fim, a abordagem bibliométrica, pois demonstra os resultados das análises e dos dados levantados através de dados empíricos e de forma principalmente quantitativa.

Resultados

Sobre os resultados, é importante ressaltar algumas abreviações criadas para fins de organização dos mesmos, sendo elas:

Comunidade Discursiva de Organização do Conhecimento (CDOC): representada pelos autores que indexaram seus trabalhos em OC;

Comunidade Discursiva de Organização da Informação (CDOI): representada pelos autores que indexaram seus trabalhos em OI;

Sem autoridade definida (SAD): São os trabalhos onde os autores explanam sobre o conceito de OI ou OC, mas não apontam nenhuma autoridade como marco teórico de suas pesquisas.

Com isso, no que concerne aos dados levantados sobre a noção epistemológica do conceito da CDOI:

Conceito similar ou idêntico: 24

Conceito de OC em OI: 1

Os dados demonstram que dos 25 trabalhos publicados em OI no âmbito do Encontro Nacional de Pesquisa em Ciência da Informação no contexto do Grupo de trabalho 2, 24 compartilham do mesmo pressuposto epistemológico de Brascher e Café (2008), enquanto 1 trabalho publicado, vê o conceito de OI como sendo ligado ao conceito de organização do conhecimento destas pesquisadoras.

Sobre as autoridades mais citadas pela CDOI foram:

- Brascher e Café: 7
- Dahlberg: 7
- Sad: 5
- Fujita Mariângela: 2
- Hjørland: 3
- Lancaster: 3
- Otlet P: 2

Como se vê Brascher e Café juntamente com Dahlberg são as autoridades mais citadas pela CDOI seguidas de SAD, mas ainda que não tenham citado nenhuma autoridade para embasar seus trabalhos, os mesmos compreendem a OI de forma similar ou idêntica aos conceitos de Bräscher e Café.

Sobre isso, no que concerne à comunidade discursiva de organização do conhecimento segue os dados sobre sua concepção epistemológica de conceito:

Conceito similar ou idêntico: 53

Conceito de OI em OC: 16

Mostra-se que dos 69 artigos analisados da CDOC 53 destes se identificam com o conceito de Organização do Conhecimento proposto pelas autoras já citadas e 16 entendem a OC como tendo as características do conceito de Organização da Informação proposta por elas.

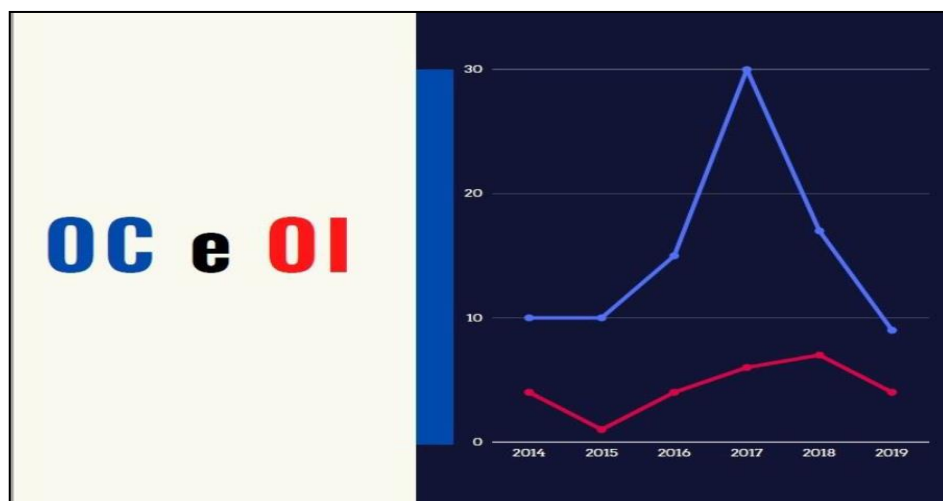
Sobre as principais autoridades citadas segue:

- Hjørland: 21
- Dahlberg: 19
- SAD: 19
- Brascher e Café: 17
- Smiraglia: 2

Percebe-se uma mudança de padrão se comparado com o gráfico da CDOI, pois a autoridade que teve mais constância nas citações foi Horland seguido de Dalgberg e SAD. Sobre o SAD desta comunidade 11 apresentaram conceitos similares ao de Brascher e Café e 8 conceitos diferentes aos destas autoras.

Destarte, para concluir o encadeamento de ideias propostas pela pesquisa nessa seção, vê-se como importante explanar sobre o gráfico geral de publicação por ano da CDOC e da CDOI:

Figura 1. Publicação por ano das CDOC e CDOI



Fonte: Elaborado pelos autores (2021)

Na primeira linha que representa as publicações em OC, percebe-se que existe um quantitativo de trabalhos bem maior da comunidade discursiva de organização do conhecimento se comparado as publicações da comunidade discursiva de organização da informação, em CDOC nota-se que existe uma crescente de publicação de 2014 a 2015, tendo o seu pico em 2017 com um total de 30 trabalhos indexados

dentro de OC. Em contrapartida, identifica-se uma decrescente de trabalhos publicados entre os anos de 2014 a 2015 que se identificam dentro da comunidade discursiva da organização da informação, tendo uma crescente em 2016 até 2018 onde foi o seu pico, ainda assim, ficou na linha de menos que 10 trabalhos publicados por ano.

Conclusões

A pesquisa visou compreender por meio da análise de domínio de Hjørland e dos embasamentos de autores como Dahlberg e principalmente da concepção de Bräscher e Café sobre os conceitos de OC e OI se os pesquisadores das Comunidades Discursivas de Organização da Informação e Organização do Conhecimento compartilham de algum pressuposto epistemológico comum, a hipótese foi de que esses pressupostos estariam alinhados com os das autoras citadas acima e que seus conceitos estariam se tornando paradigmas no sentido restrito citado nesta pesquisa, que está ligado a um modelo de tendência de pesquisa e não a uma concordância unânime dos pesquisadores da área.

Pode-se perceber, baseando-se na análise dos dados que para a Comunidade Discursiva de Organização da Informação (CDOI) o conceito das autoras é bastante utilizado e há um consenso sobre eles nos trabalhos desenvolvidos em OI, isso é percebido até mesmo pelo fato de que as autoras são as mais citadas nos trabalhos desta comunidade no ano de 2014 a 2019, juntamente com Dahlberg.

Já na Comunidade Discursiva de Organização do Conhecimento (CDOC), nota-se que apesar dos conceitos apresentados nos trabalhos desenvolvidos por essa comunidade estarem alinhados com os conceitos das autoras, no sentido de serem idênticos, ou similares, as mesmas não são as autoridades com mais destaque nos trabalhos publicados em OC, ocupando os primeiros lugares como autoridades que embasaram os conceitos de Organização do Conhecimento dos pesquisadores desta área, Hjørland, seguido de Dahlberg.

Notou-se também, que existe uma discrepância relativamente grande com relação a produção de trabalhos e por sua vez, ao número de pesquisadores que identificam suas publicações no âmbito da Organização do Conhecimento em comparação com a área de Organização da Informação, tendo a primeira um quantitativo de pesquisas consideravelmente maior do que a segunda.

Dessarte, no que concerne ao contexto e aos escopos selecionados por esta investigação, que para a Comunidade Discursiva de Organização da Informação, os conceitos de OI apontados por Bräscher e Café, podem ser considerados uma espécie de paradigma no sentido apontado por Capurro (2003), que se refere a uma tendência de pesquisa de uma determinada área dentro da Ciência da Informação.

Já no que se refere a Comunidade Discursiva de Organização do Conhecimento, percebe-se que os principais marcos teóricos apontados por esta comunidade para respaldar seus conceitos de OC é Hjørland e Dahlberg e que apesar de haver algum consenso com relação ao conceito de Organização do Conhecimento de Bräscher e Café nos trabalhos publicados pela CDOC, não se pode dizer que os conceitos destas autoras se aproximaram de se configurar como um paradigma para a área, sendo o

mais próximo do conceito de paradigma pra OC os dois autores citados no início desta explicação.

Por fim, analisando os anos de 2014 a 2019 identificou-se uma crescente de pesquisadores que indexam seus trabalhos como sendo de Organização do Conhecimento e uma constante decrescente de investigadores que se identificam-se como Organização da Informação no âmbito da Ciência da Informação.

Conclui-se então, baseado nos dados e análises levantadas por esta pesquisa, que de forma mais restrita, pode-se dizer que no domínio do Enancib no contexto dos anos levantados, o conceito de Bräscher e Café se configuram como um paradigma para a área da Organização da Informação, mas para a Organização do Conhecimento os conceitos e teorias de Hjørland e Dahlberg parecem assumir esse papel e que de forma mais geral a CDOC parece assumir um papel de maior importância para os estudos em organização no âmbito da Ciência da Informação, o que aponta para a Organização do Conhecimento emergindo como um paradigma quase que definitivo da área.

Portanto, percebe-se por meio do contexto do Enancib levando em consideração a relevância deste encontro para a Ciência da Informação no Brasil que a comunidade de pesquisadores está alinhando o seu discurso com a comunidade internacional em CI, substituindo a OI pela OC, essa afirmação pode ser percebida por meio do próprio nome do GT2 deste encontro que se denomina “Organização do Conhecimento” e em âmbito internacional a *International Society for Knowledge Organization* (ISKO) já ressalta a importância da Organização do Conhecimento para as pesquisas em Ciência da informação demonstrando assim a similaridade e alinhamento dos discursos da comunidade da CI brasileira com a comunidade da CI na esfera internacional.

Referências

- Bush, V. (1945). As we may think. *The atlantic Monthly*, 176 (1), 101-108.
- Brascher, M., & Café, L. (2008). Organização da Informação ou Organização do Conhecimento? Anais do 9.º Encontro nacional de pesquisa em ciência da Informação. USP.
- Capurro, R. (2003). Epistemologia e Ciência da Informação. Anais do 5.º encontro nacional de pesquisa em ciência da informação. Belo Horizonte.
- Dahlberg, I. (1978). Teoria do conceito. *Ciência da informação*, 7(2).
- Francelin, M. M., & Kobashi, N. Y. (2011). Concepções sobre o conceito na organização da informação e do conhecimento. *Ciência da Informação*, 40(2).
- Hjørland, B. (2002). Análise de domínio em ciência da informação: Onze abordagens – tanto tradicionais quanto inovadoras”, *Journal of Documentation*, 58(4), 422-462.

(Página deixada propositadamente em branco)

A ORGANIZAÇÃO DO CONHECIMENTO EM REPOSITÓRIOS INSTITUCIONAIS: UMA ANÁLISE DA LITERATURA RECENTE PUBLICADA EM PERIÓDICOS DE BIBLIOTECONOMIA E CIÊNCIA DA INFORMAÇÃO

Mariângela Spotti Lopes Fujita¹, Carmen Agustín-Lacruz²,
Jéssica Beatriz Tolare³, Ana Lúcia Terra⁴, Gema Bueno-de-la-Fuente⁵

¹ Universidade Estadual Paulista “Júlio de Mesquita Filho”, Unesp,
mariangela.fujita@unesp.br, <https://orcid.org/0000-0002-8239-7114>

² Universidad de Zaragoza, cagustin@unizar.es, <https://orcid.org/0000-0003-2577-1998>

³ Universidade Estadual Paulista “Júlio de Mesquita Filho”, Unesp, jessica.tolare@unesp.br,
<https://orcid.org/0000-0002-8637-7989>

⁴ Universidade de Coimbra, anaterra@fl.uc.pt, <https://orcid.org/0000-0003-1292-2849>

⁵ Universidad de Zaragoza, gbueno@unizar.es, <https://orcid.org/0000-0001-5400-1488>

Resumo

É apresentada uma pesquisa de revisão da literatura de artigos dedicados à organização do conhecimento em repositórios institucionais, publicados em revistas de Biblioteconomia e Ciência da Informação, no período de 2015 a 2020. O objetivo é realizar um estudo exploratório e descritivo sobre o uso de processos e sistemas de organização do conhecimento em pesquisas atuais aplicadas no contexto de repositórios. Este estudo foi estruturado em quatro fases: Busca, recuperação e seleção de artigos sobre o tema em bases de dados; Elaboração de tabela para extração e codificação dos dados; Determinação do tema dos artigos de periódicos na área de Organização do Conhecimento; e Análise dos dados extraídos dos artigos.

A etapa inicial foi a busca e seleção de artigos nas bases de dados Scopus, Web of Science, e na plataforma Capes (Brasil), tendo sido selecionado um corpus final de análise de 44 artigos.

Foi elaborada uma grelha de análise para recolha de dados nos artigos, os quais foram avaliados sob a lente de três categorias: disciplinaridade e interdisciplinaridade, temática e metodologia. Os artigos da amostra foram analisados usando os critérios incluídos em cada uma destas categorias, apresentados com detalhe no corpo do texto. Nesta etapa, foram retirados 11 artigos à amostra de 44 artigos por não abordarem especificamente a organização do conhecimento em repositórios. Assim, a amostra final sujeita a análise foi de 33 artigos.

É apresentada detalhadamente a metodologia concebida e aplicada na análise do corpus de textos, incluindo o uso da Classification System for Knowledge Organization Literature (CSKOL). Os dados ilustrativos da disciplinaridade e interdisciplinaridade presentes nos artigos objeto deste estudo são apresentados e discutidos, no que toca ao

predomínio de certas classes da CSKOL e quanto à diversidade dos temas representativos dos conteúdos.

Palavras-chave: organização do conhecimento, repositórios institucionais, revisão de literatura

Introdução

Os repositórios acadêmicos apresentam-se como sistemas de informação digital que têm como missão principal divulgar e preservar a produção científica dos membros da instituição de ensino ou investigação a que pertencem. Surgiram por iniciativa de pesquisadores das áreas da matemática, física e informática, na última década do século XX, para facilitar o intercâmbio dos avanços científicos, acelerando o ciclo de publicação e comunicação científica, aproveitando o ímpeto das políticas de acesso aberto. Nestes sistemas, é necessário organizar o conhecimento através de atividades de representação que são essenciais para o acesso e recuperação da informação e também contribuem para a organização intelectual e social do conhecimento (Hjørland, 2001 & 2021).

Os repositórios, tal como as bibliotecas digitais, apresentam abordagens específicas à gestão da informação. Enquanto as bibliotecas físicas se foram tornando digitais, as suas práticas, ferramentas e padrões de organização e representação da informação, construídos ao longo de séculos, foram-se adequando, mas sempre mantendo processos e sistemas tradicionalmente utilizados em Biblioteconomia e Ciência da Informação. O que aproxima as bibliotecas digitais dos repositórios é a gestão de recursos de informação, no contexto específico da Web. A Web, segundo Finemann (2020), é um meio de distribuição, uma plataforma para interconexões e interações entre todos os tipos de agências, bem como um meio com seus próprios tipos distintos de conteúdo. Os conteúdos disponíveis na Web não foram reunidos sob uma política global específica e não existem procedimentos específicos para a recolha, tratamento e distribuição tal como em bibliotecas digitais e repositórios. Por outro lado, a capacidade de armazenamento da Web em relação a repositórios e bibliotecas digitais é consideravelmente maior, assim como é maior de repositórios em relação às bibliotecas digitais. Com o advento da web, tornou-se possível a existência de repositórios institucionais, com boa capacidade de armazenamento e consequente preservação digital, que vieram dar mais visibilidade às produções intelectuais e científicas de instituições bem como favorecer a comunicação científica entre a comunidade acadêmica (Saini, 2018).

Com a possibilidade de armazenamento de conteúdos em maior quantidade, as bibliotecas passaram a gerir repositórios em formato digital, porém mantendo bibliotecas digitais e catálogos em diferentes formatos. O problema da fragmentação de conjuntos de informação em diferentes formatos, baseados em diferentes sistemas de gestão, logo veio à tona por conta da falta de interconexão e interação entre esses diferentes sistemas. Essa é uma realidade que resulta em dificuldades de interoperabilidade tanto tecnológica quanto de controlo de vocabulário e semântico, até porque não sabemos se e como estão sendo aplicados os conceitos e a prática da organização do conhecimento, tradicionais da Biblioteconomia e da Ciência da Informação, no que se refere aos processos e sistemas de organização do conhecimento.

Atualmente, as bibliotecas acadêmicas coexistem com repositórios institucionais e frequentemente partilham funções e estruturas administrativas. Embora respondam originalmente a necessidades diferentes, esta convivência próxima nas estruturas administrativas produz sinergias e melhorias mútuas. Assim, enquanto as bibliotecas são altamente orientadas para o processo técnico de gestão do acervo (seleção bibliográfica para apoiar serviços de referência e catalogação, para organização e disponibilização do acervo bibliográfico), os repositórios têm como foco a preservação digital e a disponibilidade e acessibilidade. Contudo, ambos os sistemas de informação desenvolvem estratégias comuns de acessibilidade e valorizam a interoperabilidade e a compatibilidade de padrões. Por isso, compartilham processos básicos de organização do conhecimento, como a atribuição de metadados para favorecer o acesso às coleções.

Aprofundando e alargando uma linha de trabalho sobre a organização do conhecimento em repositórios acadêmicos (Terra, Agustín-Lacruz, Bernardes, Fujita & Bueno-de-la-Fuente, 2021), na presente pesquisa, pretende-se realizar uma revisão da produção científica recente sobre a organização do conhecimento em repositórios acadêmicos publicados em periódicos indexados em bases de dados, desde 2015. O objetivo é realizar um estudo exploratório sobre o uso de processos e sistemas de organização do conhecimento em pesquisas atuais aplicadas no contexto de repositórios, cujos autores não necessariamente publicam em eventos e periódicos específicos da área da Organização do Conhecimento mas antes da área mais abrangente da Biblioteconomia e Ciência da Informação. Para isso, foram identificados em cada artigo os objetivos propostos e estudadas as metodologias e as técnicas utilizadas. Após leitura de cada artigo e a identificação dos citados elementos, foram atribuídos os temas e subtemas com base na *Classification System for Knowledge Organization Literature* (CSKOL) disponível na página web da sociedade científica *International Society for Knowledge Organization* (ISKO).

Essa delimitação de tempo a partir de 2015 deixou de fora trabalhos interessantes feitos anteriormente como, por exemplo, a série de pesquisas publicadas por Zavalina (2011; 2014). As pesquisas e trabalhos exibidos nos diferentes congressos da ISKO em seus respectivos capítulos nacionais ou em encontros internacionais também foram excluídos da seleção, por não constarem nas bases de dados indexadas.

Metodologia

O estudo exploratório e descritivo de análise dos artigos de periódicos, publicados de 2015 a 2020, foi realizado a partir da sistematização de quatro fases: Busca, recuperação e seleção de artigos sobre o tema em bases de dados; Elaboração de tabela para extração e codificação dos dados; Determinação do tema dos artigos de periódicos na área de Organização do Conhecimento; e Análise dos dados extraídos dos artigos.

Na fase de busca, recuperação e seleção de artigos, a etapa inicial foi o levantamento bibliográfico nas bases de dados Scopus, Web of Science, e na plataforma Capes (Brasil) as bases de dados *Library and Information Science Abstracts* (LISA), *Library, Information Science and Technology Abstracts* (LISTA), *Computers and*

Applied Sciences e *Scimedirect*, *SAGE Journals Online*, *Information Science and Technology Abstracts*, *OAIster*, *Directory of Open Access Journals (DOAJ)*. A equação de busca utilizada incluiu os termos “repositories” e “Knowledge Organization” com aspas nos campos de título, resumo e palavras-chave para as bases de dados Scopus e Web of Science e no Portal Capes a mesma equação nos campos de assunto e título. A somatória final de artigos obtidos recuperados foi 150, os quais foram analisados a partir dos resumos disponíveis nas bases de dados, para verificar se os termos de pesquisa usados eram semanticamente relevantes, com os resumos a abordar especificamente trabalhos sobre a organização do conhecimento em repositórios. Desta operação de seleção, resultou um corpus final de análise de 44 artigos.

De seguida, passou-se à leitura de cada artigo com o objetivo de extrair conteúdos representativos das seguintes categorias de análise: Âmbito disciplinar, Objetivo da análise, Metodologia utilizada, Métodos de coleta/Métodos de análise, Instrumentos e ferramentas de análise utilizados. Com essas categorias de análise foi elaborada uma planilha de cálculo em Excel para extrair e codificar os dados, facilitando a análise e síntese dos textos, tendo sido possível incluir trechos de cada artigo representativos das respectivas categorias.

Após leitura de cada artigo e a identificação dos citados elementos, foram atribuídos os temas e subtemas com base na *Classification System for Knowledge Organization Literature (CSKOL)*, disponível na página web da ISKO (<https://www.isko.org/scheme.php>), para analisar a que classes de assuntos se referia cada artigo e, possivelmente, averiguar tendências de temas e subtemas.

Esta etapa de leitura mais detalhada dos textos, mostrou a necessidade de retirar 11 artigos à amostra de 44 artigos por não abordarem especificamente a organização do conhecimento em repositórios. Assim, a amostra final sujeita a análise foi de 33 artigos.¹

Com as descrições das categorias de análise obtidas dos artigos, foram realizados três tipos de análise: disciplinaridade e interdisciplinaridade, temática e metodologia. A disciplinaridade e interdisciplinaridade foi averiguada pela área de atuação dos autores em consulta à afiliação institucional informada dentro do artigo e contida na categoria de análise “Âmbito disciplinar”, depois comparada com a “*UNESCO nomenclature for fields of Science and Technology*”. Na análise temática, foram verificados os temas e subtemas da CKSOL atribuídos e sua pertinência com os objetivos dos artigos para verificar e discutir as orientações e motivações das áreas disciplinares e interdisciplinares. A análise metodológica verificou quais processos e métodos da área de Organização do Conhecimento, de Biblioteconomia e Ciência da Informação ou de outras áreas interdisciplinares foram utilizados por meio das categorias de análise “Metodologia utilizada”, “Métodos de coleta/Métodos de análise” e quais sistemas de organização do conhecimento foram usados por meio da categoria de análise “Instrumentos e ferramentas de análise”.

¹ A lista com a identificação do corpus de artigos analisados está disponível em https://drive.google.com/file/d/1MOJoiSWDprTSCI_wc2RpbGMqRX9xJKYm/view?usp=sharing

Resultados

Como referido, para realizar a análise dos dados, foram atribuídos temas representativos do conteúdo dos artigos com base no título, resumo e introdução.

Foram analisados os temas dos artigos selecionados e atribuídas classes e subclasses representativas da CSKOL, no qual foi analisada a pertinência dos objetivos dos artigos, observando as áreas disciplinares e interdisciplinares.

Para realizar a atribuição de classes aos artigos analisados, foi utilizada a classificação estabelecida pela Knowledge Organization Literature. As classes são estabelecidas as seguintes: 1– Theoretical foundations and general problems, 2 – Classification systems and thesauri, structure and construction, 3 – Methodology of classification and indexing, 4 – On Universal Classification Systems and Thesauri, 5 – On special objects classifications – taxonomies, 6 – On special subjects classifications and Thesauri, 7 – Knowledge Representation by language and terminology, 8 – Applied Classing and indexing e 9 – Knowledge Organization Environment. Cada classe possui uma grande diversidade de subclasses e trabalhos atribuídos a elas. Foi realizada a atribuição de subclasses aos artigos utilizando como base as subclasses já atribuídas aos trabalhos disponíveis no site da ISKO.²

Os temas representativos do conteúdo dos artigos representados por classes e subclasses da CSKOL podem ser observados no Quadro 1 e no Gráfico 1.

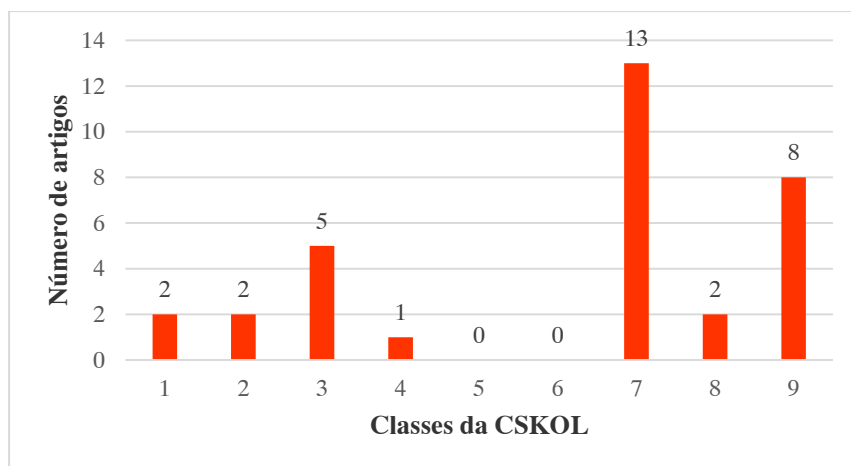
Quadro 1. Artigos representados por classes e subclasses da CSKOL

Classes CSKOL	Artigos 2015-2020	Percentuais
7 – Knowledge Representation by language and terminology Online systems – indexing in geography (1) Language processing – online retrieval problems (1) Evaluation of online information retrieval systems and techniques (5) Semiotics (1) Online retrieval systems and technologies (1) Integration – KO metadata (2) Linguistics and knowledge organization (1) Classification and thesaurus based access. Vertical and horizontal portals (1)	13	39,39%
9 – Knowledge Organization Environment Archival description – KO metadata (1) Bibliographic record display – online systems evaluation (1) Bibliographic descriptors – online systems access and retrieval (1) Organization of work (1) Cataloguing and indexing in general (2) Transfer of data in KO (1) Education and training knowledge organization (1)	8	24,24%

² <https://www.isko.org/kolit.php?cl=1>

3 – Methodology of Classification and indexing Authority control – archive indexing – archival description (1) Comparative analysis (1) Semi-automatic indexing methods (1) Manual and automatic order techniques (1) KO metadata (1)	5	15,15%
1 – Theoretical foundations and general problems Information science. Knowledge Organization in general. Classification and indexing theory in general (2)	2	6,06%
2 – Classification systems and thesauri, structure and construction Conceptual models – KOS conceptual structures (1) KOS construction – cluster analysis thesauri (1)	2	6,06%
8 – Applied classing and indexing Classification and indexing in subject fields (1) Classification and indexing of library catalogues (1)	2	6,06%
4 – On Universal Classification Systems and Thesauri	1	3,03%
Total:	33	100%

Gráfico 1. Número de artigos por classes da CSKOL



De acordo com o Quadro 1 e o Gráfico 1, as classes que mais representaram o conteúdo dos artigos da CSKOL foram: a classe “7 – Knowledge Representation by language and terminology” (13 trabalhos; 39,39%); “9 – Knowledge Organization Environment” (8 artigos; 24,24%) e “3 – Methodology of Classification and indexing” (5 estudos; 15,15%). As classes “1 – Theoretical foundations and general problems”, “2 – Classification systems and thesauri, structure and construction” e “8 – Applied classing and indexing” tiveram 2 artigos atribuídos em cada classe, representando 6,06% da amostras cada. A classe com um único trabalho atribuído foi a “4 – On Universal Classification Systems and Thesauri”, o qual representou 3,03% da amostragem.

As classes “5 – On special objects classifications – taxonomies” e “6 – On special subjects classifications and Thesauri” não foram utilizadas para representar nenhum artigo. Interessantemente, no estudo desenvolvido por Tolare, Fujita e Takahira

(2021) sobre a representação das classes e subclasses da CSKOL nos trabalhos do GT2 do ENANCIB (2016-2018), mostrou que essas mesmas classes também não tiveram trabalhos atribuídos. Dessa forma, abre para questionamentos acerca das razões para não haver trabalhos desenvolvidos nas temáticas das classes 5 e 6 da CSKOL. Fica nítido que existe uma tendência maior em desenvolver estudos que envolvam as temáticas das classes 7, 9 e 3.

A análise dos artigos permitiu também identificar as temáticas dos artigos e observar os assuntos abordados (Quadro 2).

Quadro 2: Temática dos artigos

Temática dos artigos	Quantidade	Percentuais
Visualização da informação	2	6,06%
Organização do conhecimento para a curadoria digital	2	6,06%
Consistência da informação	2	6,06%
Análise da recuperação da informação em repositório com vocabulário controlado	1	3,03%
Interface de busca visual em repositório educacional	1	3,03%
Análise da qualidade de metadados de assuntos de recursos digitais da Biblioteca Digital Europeana	1	3,03%
Indexação de assuntos em humanidades	1	3,03%
Sistema de Organização automatizado em informação nuclear para indexação de assuntos	1	3,03%
Organização do conhecimento na Geoweb	1	3,03%
Registros de autoridade para indexação de metadados em bancos de dados remotos	1	3,03%
Organização do conhecimento em humanidades digitais	1	3,03%
Coleções de fontes de pesquisa temática em humanidades	1	3,03%
Modelos de organização do conhecimento nas coleções de dados vinculadas ao patrimônio cultural	1	3,03%
Perfil de dados vinculados em Sistemas de Organização do conhecimento no domínio do patrimônio cultural	1	3,03%
Descritores para descrição e depósito de conjuntos de dados em repositórios de dados	1	3,03%
Avaliação da qualidade de Linked Open Data	1	3,03%

Acesso à forma e gênero em coleções digitais	1	3,03%
Integração do sistema de classificação DDC repositórios institucionais com DSpace	1	3,03%
Indexação em repositórios digitais	1	3,03%
Desenvolvimento profissional em aplicativos da web semântica e dados vinculados	1	3,03%
Recursos terminológicos na organização e recuperação da informação	1	3,03%
Classificação bibliográfica em ambientes digitais	1	3,03%
Centros de serviços de conhecimento, orientados para o conhecimento	1	3,03%
Métodos de organização do conhecimento dos arquivos nacionais e dos repositórios de conjuntos de dados	1	3,03%
Metodologia para detectar inconsistências Geoespaciais em coleções de metadados de assuntos de materiais cartográficos	1	3,03%
Criação de metadados para objetos de patrimônio cultural digital em 3D	1	3,03%
Sistema automático para gerenciar múltiplos servidores de ontologias e torna-los interoperáveis	1	3,03%
Construção de vocabulário e ontologia para repositório de agronomia	1	3,03%
Repositório de dados de pesquisa	1	3,03%
Repositório de ontologias de objetos de aprendizagem	1	3,03%
Total:	33	100%

No Quadro 2 são apresentadas as temáticas dos artigos analisados. Para realizar a identificação das temáticas dos artigos foram observados seus resumos com o intuito de encontrar os temas que eram abordados no texto. Analisando os resumos, foi possível identificar que muitos continham informações escassas, que são consideradas essenciais para compor um resumo, sendo necessário consultar, além das palavras-chave, o próprio texto.

A partir desta leitura, verificou-se que há uma diversidade de temas nos artigos analisados que possuem diferenças e similaridades entre si. Muitos estudos possuem como produto final o desenvolvimento de uma ferramenta pertencente a Organização do Conhecimento (como por exemplo: ontologia), mas o trabalho em si faz parte de outra área do conhecimento, como Gestão do Conhecimento, Geografia, Repositórios, Agronomia, Arquivos e Informação Nuclear.

Temas interdisciplinares e temas da área da Organização do Conhecimento obtidos da fundamentação teórica analisada

A partir da análise dos artigos foram identificados temas interdisciplinares (Quadro 3) e temas da área de Organização do Conhecimento na fundamentação teórica de cada artigo (Quadro 4).

Para realizar a identificação dos temas interdisciplinares e dos temas da área da organização do conhecimento, foram analisadas as temáticas discutidas nas fundamentações teóricas de todos os artigos e observados a compatibilidade com as referências bibliográficas. Houve alguns artigos que tinham como proposta um viés mais prático em que a fundamentação teórica estava inserida na sua introdução, não havendo, dessa forma, um espaço reservado especificamente para discorrer sobre ela.

Quadro 3: Temas da fundamentação teórica interdisciplinar

Temas da fundamentação teórica interdisciplinar	Quantidade de vezes que apareceu nos artigos	Percentuais
Web semântica	8	10,96%
Linked open data	6	8,22%
Repositórios	4	5,48%
Gestão de dados	3	4,11%
Metadados	3	4,11%
Interoperabilidade	2	2,74%
Sistema de informação específica (Gramatical e Nuclear)	2	2,74%
Sistema de gerenciamento	2	2,74%
Sistema de suporte de decisão	2	2,74%
Visualização da informação	2	2,74%
Abordagem sociotécnica	1	1,37%
Acessibilidade tecnológica	1	1,37%
Adoção e uso do EAC-CPF	1	1,37%
Análise de dados das humanidades	1	1,37%
Arquivos e bibliotecas digitais	1	1,37%
Bibliotecas digitais	1	1,37%

Coleções de patrimônio cultural em repositórios	1	1,37%
Coleções digitais	1	1,37%
Competência em biblioteconomia	1	1,37%
Conjunto de dados	1	1,37%
Dados de produção	1	1,37%
Dados e metadados de repositórios	1	1,37%
Domínio agrônômico	1	1,37%
DSpace	1	1,37%
Educação continuada	1	1,37%
Elemento semântico Europeana	1	1,37%
Epistemes e ordenação do conhecimento museológico	1	1,37%
Esquemas de modelos de metadados digitais em 3D	1	1,37%
Esquemas de metadados descritivos	1	1,37%
Estado da arte	1	1,37%
Estrutura de avaliação	1	1,37%
Forma e gênero	1	1,37%
Gestão do conhecimento do usuário	1	1,37%
Gestão do conhecimento pelas bibliotecas	1	1,37%
Gestão terminológica	1	1,37%
Integração de dados	1	1,37%
Inteligência de negócios	1	1,37%
Interação humana-computador	1	1,37%
Interface de busca visual em repositórios e bibliotecas digitais	1	1,37%
Modelos de metadados Europeana	1	1,37%
Organização do conhecimento indígena, feminista e pós-colonial	1	1,37%
Pesquisa facetada	1	1,37%
Recuperação de recursos na internet	1	1,37%

Semiótica	1	1,37%
Serviço de informação	1	1,37%
Sistemas OPACs	1	1,37%
Tratamento da informação em repositório educacional digital	1	1,37%
Validação de metadados da informação digital	1	1,37%
Total:	73	100%

O Quadro 3 apresenta temas interdisciplinares encontrados na fundamentação teórica dos artigos, a quantidade de vezes em que apareceu neles e sua percentagem. É possível observar uma grande quantidade de temas variados como “Web semântica” (8 artigos abordam essa temática, 10,96% da amostragem), “Sistema de informação específica (Gramatical e Nuclear) com 2 estudos (2,74%), “Estado da Arte” (1 artigo, 1,37% das amostras) e entre outras temáticas diversificadas.

A maioria dos temas interdisciplinares apontados possuem relação com a área de tecnologia, abordando temas como “Interoperabilidade” (2 estudos, 2,74%), “Repositórios” (4 trabalhos, 5,48%) e “Validação de metadados da informação digital” (1 artigo, 1,37%). Outras temáticas encontradas tiveram relação com a área da Gestão do Conhecimento: “Gestão do conhecimento usuário” (1 trabalho, 1,37%), Gestão terminológica (1 estudo, 1,37%) e “Gestão do conhecimento pelas bibliotecas” (1 artigo, 1,37%). Há uma temática que aborda movimentos sociais como “Organização do conhecimento indígena, feminista e pós-colonial” (1 estudo, 1,37%).

A análise dos assuntos dos artigos destaca quantidade e diversidade nos temas representativos dos conteúdos dos artigos sem coincidências. Por outro lado, destacam-se os temas interdisciplinares coincidentes, “Web semântica” (11%) e “Linked open data” (6,22%) e os temas da organização do conhecimento com mais incidência: “Sistemas de organização do conhecimento” (19,4%), “Catalogação” (9%) e indexação (9%).

Já os temas da organização do conhecimento abordados na fundamentação teórica dos artigos analisados foram apresentados no Quadro 4.

Quadro 4: Temas da fundamentação teórica da KO

Temas da fundamentação teórica da KO	Quantidade de vezes que apareceu nos artigos	Percentuais
Sistema de Organização do Conhecimento	13	19,40%
Catalogação	6	8,96%
Indexação	6	8,96%

Recuperação da informação	5	7,46%
Classificação	4	5,97%
Tesauros	3	4,48%
Vocabulários controlados	3	4,48%
RDF	2	2,99%
Esquemas de classificação	2	2,99%
Sistemas de classificação	2	2,99%
Organização da informação	2	2,99%
Organização do conhecimento	2	2,99%
Ontologias	2	2,99%
Representação da informação	2	2,99%
Sistemas bibliográficos	1	1,49%
FRBR	1	1,49%
Taxonomias	1	1,49%
Hierarquia do conhecimento	1	1,49%
Tecnologias da informação e comunicação	1	1,49%
Catálogo cartográfica crítica	1	1,49%
Metadados – Dublin Core	1	1,49%
Metadados descritivos	1	1,49%
Organização de textos arquivísticos	1	1,49%
CDD	1	1,49%
Análise e tratamento de arquivo	1	1,49%
Terminologia	1	1,49%
Controle de autoridade	1	1,49%
Total:	67	100%

De acordo com o Quadro 3, a temática da organização do conhecimento que mais apareceu foi “Sistemas de Organização do Conhecimento” (13 artigos, 19,40%), seguido por “Catalogação” e “Indexação” (ambas com 6 trabalhos cada, represen-

tando 8,96% das amostras). Recuperação da informação possuiu 5 estudos (7,46%) e Classificação 4 artigos (5,97%).

Os temas menos abordados, todos com um único trabalho (1,49% cada), foram: “Controle de autoridade”, “Terminologia”, “Análise e tratamento de arquivo”, “CDD”, “Organização de textos arquivísticos”, “Metadados descritivos”, “Metadados – Dublin Core”, “Catalogação cartográfica crítica”, “Tecnologia da informação e comunicação”, “Hierarquia do conhecimento”, “Taxonomias”, “FRBR” e “Sistemas bibliográficos”. Observando as temáticas apontadas, foi possível perceber que muitos desses temas poderiam ser incluídos como parte do tema “Sistemas de Organização do Conhecimento”, mas essa inclusão não ocorreu devido às fundamentações teóricas terem discorrido acerca das temáticas de forma muito específica, não abordando de forma mais geral.

Outro ponto pertinente e relevante para citar como exemplo é o facto de haver trabalhos cujas propostas foram de criar um modelo de ontologia dentro de um determinado contexto, mas a fundamentação teórica da organização do conhecimento discorria sobre a recuperação da informação, representação da informação, tecnologias da informação e comunicação, da área de gestão do conhecimento, entre outros, mas não abordava a fundamentação teórica sobre ontologia especificamente.

Conclusões

Com os dados da pesquisa obtidos e analisados até ao momento, é possível concluir que o uso da CSKOL para a análise da literatura recente sobre a organização do conhecimento em repositórios institucionais se afigura como uma lente adequada capaz de sustentar um melhor entendimento da literatura publicada neste âmbito. O enquadramento dos artigos nas classes e subclasses da CSKOL permite identificar abordagens predominantes, criando uma representação estruturada da área. Recorde-se que a classificação dos temas em classes e subclasses revelou que dos 33 artigos analisados e classificados na CSKOL, 3 classes contém a maioria dos artigos, quais sejam a classe 7 (Representação do Conhecimento pela linguagem e terminologia) contém 13 dos temas dos artigos, seguindo-se a classe 9 (Ambiente de organização do conhecimento) com 8 temas e a classe 3 (Metodologia de classificação e indexação) com 5 temas, demonstrando uma tendência de temas de investigação no período de tempo analisado.

Considerando estes dados, verifica-se que três quartos dos trabalhos do corpus estão concentrados em três categorias, as quais remetem para abordagens concretas, como seria de esperar, tratando de aspectos metodológicos da classificação e indexação, da representação do conhecimento através da linguagem e de terminologia, atendendo à organização do conhecimento em contextos/ambientes específicos, já que o foco desta pesquisa são os trabalhos relativos a repositórios institucionais.

No que toca à identificação dos temas interdisciplinares e temas da área de Organização do Conhecimento, na fundamentação teórica de cada artigo, destaca-se a quantidade e diversidade nos temas representativos dos conteúdos dos artigos sem coincidências, por outro lado, sobressaem os temas interdisciplinares coincidentes “Web semântica” (11%) e “Linked open data” (6,22%) e os temas de organização do

conhecimento com mais incidência, “sistemas de organização do conhecimento” (19,4%), “catalogação” (9%) e “indexação” (9%).

Existe, portanto, uma consistência entre os temas de organização do conhecimento da fundamentação teórica com as classes e subclasses de assunto como, por exemplo, o tema “sistema de organização do conhecimento” com a classe 7 (Representação do conhecimento pela linguagem e terminologia); demonstra ainda que nesses 33 artigos os temas de organização do conhecimento estão presentes em todos os artigos combinados com temas interdisciplinares de outras áreas do conhecimento, o que, ao nosso ver, oferece enriquecimento e aporte teórico para o desenvolvimento das investigações.

Sendo uma pesquisa em andamento, os dados resultantes do levantamento das metodologias usadas nos trabalhos dedicados à organização do conhecimento em repositórios institucionais representarão um complemento para traçar um quadro mais completo da investigação recente nesta área. Poderá também ser explorada a filiação institucional dos autores dos artigos, identificando a sua proveniência em termos de área científica do ponto de vista da autoria.

Referências

- FINNEMANN, N. O. (2020). Web archive. In: HJORLAND, B.; GNOLLI, C. ISKO *Encyclopedia of knowledge organization*. ISKO. Disponível em: Encyclopedia of KO (isko.org). Acesso em: 16 de março de 2021.
- HJØRLAND, B. (2002). Epistemology and the socio-cognitive perspective in information science. *Journal of the American Society for Information Science and Technology*, 53 (4), 257-270.
- HJØRLAND, B. (2021). Information Retrieval and Knowledge Organization: A Perspective from the Philosophy of Science. *Information*, 12 (135). Disponível em: <https://www.mdpi.com/2078-2489/12/3/135/htm>. Acesso em: 15 de março de 2021.
- SAINI, O. P. (2018). Understanding the role of institutional repository in digital preservation in academic libraries: a review of literature. *Library Philosophy and Practice*. Disponível em: <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=5249&context=libphilprac>. Acesso em: 12 de março de 2021.
- TERRA, A. L., AGUSTÍN-LACRUZ, C. BERNARDES, O., FUJITA, M. S. L., BUENO-DE-LA-FUENTE, G. (2021). Subject-access metadata on ETD supplied by authors: A case study about keywords, titles and abstracts in a Brazilian academic repository. *Journal of Academic Librarianship*, 47 (1). Disponível em: <https://www.sciencedirect.com/science/article/pii/S0099133320301592>. Acesso em: 16 de março de 2021.
- TOLARE, J. B., FUJITA, M.S.L., KATAHIRA, I. (2021). O desenvolvimento da organização e representação do conhecimento no ENANCIB. *XXI Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB)* (no prelo).
- UNESCO *nomenclature for fields of Science and Technology*. Disponível em: <https://skos.um.es/unesco6/00/html>. Acesso em: 16 de março de 2021.
- ZAVALINA, O. L. (2011). Free-text collection-level subject metadata in large-scale digital libraries: A comparative content analysis. *DCMI. International conference on Dublin Core and metadata applications. Papers, Project Reports and Posters for DC-2011 at The Hague*, 147–157. Disponível em: <https://dcpapers.dublincore.org/pubs/article/view/3630/1856>
- ZAVALINA, O. L. (2014). Complementarity in subject metadata in large-scale digital libraries: A comparative analysis. *Cataloging & Classification Quarterly*, 52(1), 77-89.

O CONCEITO DE INFORMAÇÃO PEIRCEANA E SUA ATUALIDADE NOS ESTUDOS DA INFORMAÇÃO

Valdirene Aparecida Pascoal¹, Maria Eunice Quilici Gonzalez²,
Carlos Cândido de Almeida³

¹UNESP (FFC) Marília-SP, *vadirenepascoal@gmail.com*,
ORCID iD 0000-0002-3695-6560 ²UNESP (FFC) Marília – SP,
eunicegonzalez@unesp.br, ORCID iD 0000-0002-3837-4644 ³UNESP (FFC)
Marília – SP, *carlos.c.almeida@unesp.br*, ORCID-iD 0000-0002-8552-1029

Resumo

A filosofia peirceana tem como uma de suas características o pensamento diagramático e expansivo, o que possibilita compreender o conceito de informação em duas fases distintas: proposicional e semiótico. O objetivo desta pesquisa é apresentar as duas fases distintas da proposta elaborada por Charles S. Peirce sobre o conceito de informação, ressaltando sua atualidade diante de problemáticas contemporâneas no contexto das tecnologias de informação e comunicação. Apresentamos, inicialmente, o conceito de informação no sentido lógico-proposicional caracterizado como o produto que emerge da relação entre extensão e profundidade de proposições. Em um segundo momento, discutimos a etapa investigativa em que Peirce expande o domínio da informação para além do lógico-proposicional através da semiótica, argumentando que o processo de comunicação de signos e formas possibilita a significação da percepção por meio da aprendizagem. Argumentamos que a identificação de informação em um processo comunicacional pode ser o diferencial entre uma ação deliberada e uma ação direcionada. Esses aspectos são de interesse da Organização do Conhecimento.

Palavras-chave: Informação, Semiótica Peirceana, Ação, Organização do Conhecimento.

Introdução

Charles Sanders Peirce (1839-1914) elaborou o conceito de informação, no interior da semiótica, em dois momentos distintos da sua vida: entre os anos de 1865 e 1900 (*cf.* Nöth, Gurick, 2011). Seus escritos ficaram conhecidos como a Teoria da Informação, embora o filósofo não a tenha nomeado propriamente dessa forma. O objetivo deste trabalho é apresentar as duas fases do conceito de informação na filosofia peirceana e indicar sua importância para a Organização do Conhecimento. Por se tratar de um estudo interdisciplinar, entre Filosofia e Ciência da Informação, o método utilizado é o levantamento bibliográfico das principais obras que envol-

vem a temática da informação elaborada por Charles Sanders Peirce e seus comentadores. O arcabouço teórico e conceitual tem inspiração nos estudos dos seguintes autores: Peirce (CP e EP); Gonzalez (2004 a 2019); Santaella (2004 a 2017), Nöth (2011 a 2017), Silveira (2007 a 2018) e Almeida (2011 a 2016).

Em contraposição às abordagens matemáticas e técnicas da comunicação, Peirce argumenta que a informação é um processo de *semiose*, em que o crescimento de signos gera outros signos, representando objetos e propiciando interpretantes de ação. Assim, a investigação desenvolvida por Peirce sobre informação abarca aspectos pragmáticos, lógicos, semânticos e cognitivos. Com base em sua teoria, analisamos a seguinte questão que motiva este trabalho: Qual é a relevância do conceito de informação semiótica peirceana para o estudo da Organização do Conhecimento?

Peirce abrange três problemáticas fundamentais no estudo da comunicação e da informação: transmissão, significado e impacto na conduta. Posteriormente ao desenvolvimento da teoria da informação peirceana, Shannon e Weaver (1949, p. 49) desenvolvem A Teoria Matemática da Comunicação (*TMC*) e também elucidam que há três tipos de problemas ao se tratar da informação: Problema Técnico (quão rigorosamente os símbolos envolvidos no processo de comunicação podem ser transmitidos?), Semântico (quão precisamente os símbolos transmitidos carregam o significado desejado?) e da Eficácia (em que medida o significado recebido no processo de comunicação afeta a conduta de modo desejado?). No entanto, os autores focalizam o problema técnico, pois no que concerne à *TMC*, informação se restringe à eficácia entre transmissão e recepção de mensagens. Nesse contexto, revisar a teoria peirceana de informação, capaz de enaltecer as três problemáticas da *TMC*, pode desvelar alternativas diante de temáticas da área.

A primeira fase da Teoria da Informação peirceana foi desenvolvida por volta de 1865 e 1867 (*cf.* Nöth e Gurick) e se restringe à abordagem da lógica-proposicional: informação é caracterizada a partir do produto de duas dimensões lógico-semânticas: extensão e profundidade de termos ou conceitos. A segunda teoria expande o universo teórico da primeira e contempla outros domínios, a informação passa a ser concebida como um fluxo através do qual signos podem ser transmitidos, de modo verbal (ou não). Esse fluxo sígnico não está atrelado apenas ao significado, mas às diversas formas e meios possíveis de comunicação.

Peirce, há muito compreende informação como elemento essencial no que tange à conduta e à geração de interpretantes. Nesse sentido, informação, veiculada no processo de semiose, é fundamental no processo de desenvolvimento de hábitos de conduta, isto é, hábitos responsáveis pela fundamentação de ações deliberadas com alto grau de autonomia e também na fundamentação e Organização do Conhecimento. Nesse cenário, entendemos que a informação semiótica pode contribuir para o questionamento e análise de problemas que surgem no âmbito da Sociedade da Informação.

O presente trabalho resulta de uma pesquisa interdisciplinar, que envolve temas e problemas da Filosofia e da Ciência da Informação, de escopo teórico-conceitual; a metodologia utilizada inclui levantamento bibliográfico das principais obras indicadas nas referências, utilizando bases de dados nacionais e internacionais, tais como:

Base de Dados em Ciência da Informação (BRAPCI), Periódico CAPES e Biblioteca Científica Eletrônica Online (SciELO) e Scopus, Web of Science, Scopus, Web of Science (WoS), Library and Information Science Abstract (PROQUEST). No que concerne à metodologia específica, o acesso bibliográfico aos textos disponíveis em The Collected Papers of Charles S. Peirce será privilegiado, bem como os artigos de seus comentadores e de textos desse filósofo disponíveis na base de dados Digital Companion to C. S. Peirce (Commens).

O problema que direciona o presente estudo é assim formulado: Qual é a relevância do conceito de informação semiótica peirceana para o estudo da Organização do Conhecimento? Tal problemática se justifica diante do cenário atual, em sociedades digitais, em que o compartilhamento massivo de informação e o desenvolvimento contínuo e acelerado de tecnologias de informação impactam ações individuais e coletivas, modificando as formas tradicionais de Organização do Conhecimento. A hipótese é que a relevância da teoria da informação peirceana se deve a que ela abre caminhos para ações deliberadas, por meio de obtenção de informação genuína, propiciando fundamentos para a Organização do Conhecimento.

Teoria da informação de Peirce e contribuições à organização do conhecimento

Como indicamos, a primeira teoria da informação peirceana se restringe à abordagem da lógica proposicional tradicional, focalizando mensagens no domínio das proposições. Nesse contexto, informação é um produto de duas dimensões lógico-semânticas: *extensão* e *profundidade*, ou seja, informação só existe na medida que esses dois elementos se encontram estruturados a partir de uma sintaxe, permitindo a organização de sentenças definidoras de conceitos. Peirce, em The Collected Papers (CP 2.407), define *extensão* e *profundidade*, ressaltando que a *extensão* de termos ou conceitos delimita classes de objetos e/ou sujeitos atribuídos a um conceito ou termo. Assim, por exemplo, a extensão de termos que caracterizam animais terrestres, deve incluir, entre vários outros, o termo animal humano. Nos escritos de Peirce, *extensão* também aparece como denotação, referência e largura. Em contrapartida, *profundidade* refere-se à significação, sentido, intenção e compreensão. No caso da caracterização do animal humano, o domínio da extensão inclui suas propriedades distintivas de animal bípede, racional, entre outras.

A equação que determina a conceituação lógica de informação peirceana pode ser assim representada:

Informação = Extensão x Profundidade de um termo ou conceito que pode ser expresso *através da sintaxe*.

Em síntese, a equação acima caracteriza informação como um conjunto de proposições sintéticas; esse conjunto, segundo Nöth e Gurick (2005), pode ser obtido por meio do encontro entre extensão e profundidade. No mesmo contexto, Vicentini *et al* (2019) consideram que para Peirce só são informativas as sentenças ampliativas; as sentenças tautológicas não fazem parte desse conjunto e, como tal, não transmitem informação. Ou seja, a novidade contida em uma sentença declarativa é

condição necessária para que ela seja informativa. Nesse sentido, Peirce (EP 463: 13) afirma textualmente: “[...] se você me informa a respeito de qualquer verdade que eu já saiba, então não há informação [...]. Ao aproximar informação de conhecimento, Peirce expande a Teoria da Informação.

A partir de 1900, o filósofo amplia sua abordagem inicial de informação, que concentrava inicialmente aspectos lógico-proposicionais e semânticos, desenvolvendo uma teoria mais geral que abarca aspectos cognitivos e pragmáticos, com enfoque em contextos que vão além da linguagem. A caracterização de informação passa, então, a ser focalizada no estudo da natureza semiótica e informacional do signo. Signo é, segundo Peirce, um “meio para a comunicação de uma forma” (EP2, p. 477) que se efetua em um universo triádico; o signo comunica, medeia e representa a forma do objeto para um interpretante.

No que concerne à caracterização Semiótica da informação, Peirce (CP 2. 309) focaliza sua análise no Signo Dicente, definido como o tipo de signo que veicula informação. O Dicente é um signo duplo composto pela junção de um ícone e de um índice. No processo de intersecção entre Ícone e Índice que compõem um signo Dicente, o Ícone incorpora características e propriedades que delimitam o objeto que será informado, e o Índice referencia e indica a localização espaço-temporal do objeto. A combinação desses dois signos confere ao Dicente a possibilidade de transmitir informação (SILVEIRA, 2008).

Santaella (2017) explica o conceito de signo Dicente, no domínio da linguagem, ressaltando a noção de proposição: Na lógica clássica, a unidade mínima, capaz de externar ideias verdadeiras ou falsas, é uma proposição, contanto que combine ao menos um sujeito e predicado. O arquétipo de uma proposição tem a forma “A é B”. Somente esse signo pode afirmar algo ao conectar predicado e sujeito. Dessa forma, o Signo Dicente pode transmitir informação (SANTAELLA, 2017, p. 60). No entanto, o processo só se completará com a Sintaxe que emerge da conjunção entre Ícone e Índice. A sintaxe é o elemento que confere a existência factual da informação para um receptor possível.

Em suma, para que um signo veicule informação, na perspectiva semiótica, é necessário, em primeiro lugar que ele apresente uma qualidade do objeto; em segundo lugar, que aponte para a existência real deste mesmo objeto e, por último, que se apresente uma estrutura sintática factual. Uma vez que o Signo Dicente cumpra tais exigências há a veiculação de informação (SILVEIRA, GONZALEZ, 2014).

Os estudos acerca da informação têm destaque na Organização do Conhecimento na aplicação da Teoria Geral dos Signos, principalmente a concepção de semiótica presente na Gramática Especulativa. Nesse sentido, Almeida (2009, p. 268) compreende que algumas contribuições de Peirce para a Organização do Conhecimento estão relacionadas com modelos de indexação a partir da semiótica, a organização semiótica do conhecimento, semiótica documental e a própria teoria da informação elaborada pelo filósofo. Como apontamos, a concepção de informação está estritamente ligada ao conceito de signo. Além do mais, Almeida (2011, p. 106) defende que a Filosofia e a Semiótica peirceana contribuem para a Organização do Conhecimento pois sugerem explicações consistentes para problemáticas da área, são elas: reflexões epistemológicas sobre fundamentos teóricos, análise da informação, processos modelos e instrumentos que operacionalizam a Organização do Conhecimento.

A Organização do Conhecimento adota como pressuposto noções de conhecimento que se adequam a sua realidade. O produto do conhecimento social materializado em documentos e institucionalizado por organizações que o coletam, selecionam, organizam e disseminam, é tomado como pivô do trabalho da área. Não obstante, a informação que sustenta a noção de conhecimento coletivizado não pode ser considerada apenas como subproduto da conversão e da representação linguística como tem sido considerado por tradições não semióticas do signo, a exemplo da Linguística. No momento em que o fenômeno das fake news cobra uma teoria mais consistente que separe informação de não informação, representação adequada dos fenômenos das representações limitadas e menos propícias à geração de conhecimento seguro, a Organização do Conhecimento deve retomar a discussão sobre qual base conceitual tem construído a sua discussão. Nesse sentido, se informação, especialmente a informação científica, está relacionada à verdade das pesquisas, a área deve reconsiderar a necessária conexão entre informação e verdade, tal como já antecipadamente concebida por Peirce, antes mesmo da teoria matemática da informação.

Os conhecimentos operados pelos processos da Organização do Conhecimento, tais como tradução, representação, classificação; especialmente no campo científico, não são fruto de mera convenção entre falantes de uma mesma língua. Os conhecimentos fazem referência a informações que conjugam objeto e signo denotando aspectos da realidade ordinária do pesquisador.

Café, Barros e Santos (2014, p. 204) compreendem que a Organização do Conhecimento¹ abarca análises de conceito, relações e delimitações da linguagem, significados e terminologias, com o intuito de representar e aproximar de forma apropriada seus domínios. Os conceitos, para os autores, são a representação de uma ideia que estrutura o próprio conhecimento, sendo elemento fundamental na Teoria da Cognição. Os autores (2014, p. 205) consideram que essas representações, na Ciência da Informação, implementam Sistemas de Organização do Conhecimento, sendo: “tesauros, taxonomias, ontologias e sistemas de classificação”. Tais sistemas em áreas técnicas referem-se a conceitos que informam elementos da realidade e somente por esta capacidade podem ser utilizados como representantes. Caso não houvesse relação fenomênica entre conceito e informação, a possibilidade de representar ocorrências de secundidade, declaradas ou não via *sinsignos* dicentes, estaria comprometida e com isso a própria utilidade do conceito dentro de um sistema de Organização do Conhecimento.

Entendemos que a relevância do conceito de informação semiótica peirceana para o estudo da Organização do Conhecimento na atualidade reside, justamente, na

¹ Muitas vezes os termos Organização do Conhecimento e Organização da Informação são empregados sem delimitações de fronteiras, como se ambos tivessem o mesmo significado. Café, Barros e Santos (2014) reforçam a importância de distinguir seus fundamentos e aplicações, pois há dois conceitos importantes que as áreas englobam: informação e conhecimento. Neste trabalho, priorizamos a Organização do Conhecimento, por entender sua capacidade de englobar fundamentos que dão sustentação ao objetivo aqui defendido: apresentar as duas fases distintas da proposta elaborada por Charles S. Peirce sobre o conceito de informação, ressaltando sua atualidade diante de problemáticas contemporâneas no contexto das tecnologias de informação e comunicação.

possibilidade de aplicação de informação confiável a certos campos do conhecimento. A identificação do signo Dicente, no domínio da linguagem, em contextos específicos pode auxiliar não apenas na tarefa de estruturação de fontes informacionais, através da classificação de índices pertinentes à certos domínios do conhecimento, mas também no próprio desenvolvimento de modelos explicativos de conceitos em áreas específicas do conhecimento. Conforme ressaltam Café, Barros e Santos (2014): um Sistema de Organização do Conhecimento pode não ser capaz de atender as demandas de uma cultura ou contexto. Dessa forma, o processo de organização do conhecimento, nas palavras dos autores: "1) é um processo de modelagem do conhecimento e 2) gera representações que se concretizam como Sistemas de Organização do Conhecimento". (Café, Barros, Santos, 2014, p. 205).

Em síntese, a Teoria da Informação peirceana representa um retorno aos fundamentos da relação entre informação e realidade e contribui com elementos para a compreensão da tríade informação-conceito-conhecimento indispensável à Organização do Conhecimento à medida que contribui com a otimização de modelagens, além de possibilitar discussões relevantes para a área, principalmente na explicitação dos conceitos de informação e conhecimento. Peirce considera que o sustentáculo do conhecimento é o signo Dicente, estruturado na secundidade, no real e no existente. As relações signícas constituem elementos que propiciam conhecimento, fundamentando crenças, bem como processos de comunicação, que alicerçam a capacidade de direcionar a conduta.

A Organização do Conhecimento quando explorada na Ciência da Informação possibilita o estudo dos procedimentos de análise, forma e conteúdo da informação. Nesse sentido, autores como Jens-Erik Mai (2001) e Torkild Leo Thellefsen (2002) elaboram aproximações entre hipóteses da Semiótica Peirceana e aquelas do domínio da Organização do Conhecimento. Um exemplo ilustrativo dessa aproximação é dado por Mai (2001) que desenvolve um modelo de indexação utilizando a semiótica e as classes de signos, com o objetivo de analisar problemas centrais da representação de documentos e seus objetos, e apontar caminhos de investigação para a compreensão de como os indivíduos interpretam e representam documentos. O autor parte da premissa de que qualquer teoria de indexação abarca uma teoria de significado no domínio da linguagem. Dessa forma, ao analisar semioticamente a indexação, caracterizada enquanto um processo, Mai (2001) desenvolve a análise de documentos por meio de interpretações que dependem de contextos sociais e culturais do indivíduo responsável pela indexação e do próprio processo de indexação.

Entendemos que no contexto de mídias digitais, a análise semiótica da indexação, fundada na Teoria da Informação proposta por Peirce, abre também caminhos, ainda pouco explorados para a investigação de efeitos desencadeados por documentos digitais na conduta. Embora não seja objeto da Organização do Conhecimento, a investigação sobre a conduta mental e sua relação com a informação nos ajuda a compreender como os conceitos são formados e quais variáveis atuam para a admissão ou não de um conceito a partir de informações disponíveis. Além do mais, é possível fazer uma aproximação entre o processo de indexação e o processo sógnico informacional ao avaliar o papel da semiose presente nos múltiplos processos interpretativos desses elementos. Mai (2001) afirma que qualquer estudo de busca, recuperação e avaliação de sistemas de informação devem ser considerados a partir

de sua natureza interpretativa presente inevitavelmente no processo de indexação. Tal natureza interpretativa pode contribuir para discussões recentes acerca da veracidade dos conteúdos dispostos em mídias, inclusive documentos.

Como aponta Almeida (2011), Thellefsen propõe uma teoria da organização semiótica do conhecimento enquanto uma nova metodologia para a própria Organização do Conhecimento, com base no conceito de signo, nos fundamentos da Semiótica Peirceana e em sua estrutura enraizada na evolução contínua do conhecimento. O autor questiona como o conhecimento é organizado na Ciência da Informação, sugerindo indicações relevantes para a arquitetura do conhecimento, dentre elas: (i) esboçar o perfil do conhecimento e suas implicações, (ii) apontar as bases teóricas do objeto de pesquisa, (iii) refinar a forma geral dos conceitos, sufixando e prefixando-os, (iv) encontrar possibilidades de tornar o conceito mais preciso, explorar se é o momento adequado de precisar um conceito ou indicar suas implicações. (Thellefsen, 2004, p. 512-513 apud Almeida, 2011, p. 109).

Por fim, a informação-processo, veiculada pelo signo dicente, permite compreender acontecimentos reais sobre objetos reais. O efeito produzido pelo signo dicente perpassa elementos qualitativos e de existência, e a sintaxe que se estrutura nessa relação indica padrões que podem direcionar processos, pensamentos e ações. A partir da caracterização de informação peirceana aqui esboçada, cabe refletir sobre as suas possibilidades de auxiliar no entendimento de problemáticas contemporâneas da Sociedade da Informação. A Organização do Conhecimento, enquanto campo, precisa participar do debate público sobre a circulação de informações recomendando aos sujeitos a necessidade de basear-se em conceitos bem estruturados, pois tanto os sistemas de organização do conhecimento quanto a recuperação da informação dependem de conceitos bem construídos semioticamente.

Conclusões

Ao retomar a questão central inicial sobre a relevância do conceito de informação semiótica peirceana para o estudo da Organização do Conhecimento, sugerimos que somente o adequado acesso e identificação de informação pode contribuir para a formação do conceito e conseqüente construção de sistemas que representam conceitos. Para tanto, entender melhor essa questão é uma tarefa não apenas do âmbito da Organização do Conhecimento, mas para o campo da Ciência da Informação. Sugerimos também que o entendimento do conceito de informação semiótica pode iniciar estudos sobre a relação entre informação, conceito e conhecimento. Porém, mais do que compreender se um conteúdo é não ou informativo, é importante que o usuário consiga identificar símbolos de narrativas manipuladas que incidem na formação de conceitos adequados..

Ainda, concordamos com a hipótese de Almeida (2016) que a Semiótica peirceana pode ser utilizada para revisar conceitos estabelecidos nessa área e auxiliar no desenvolvimento de novas teorias para pensar questões urgentes da contemporaneidade. Também julgamos que cabe à Ciência da Informação buscar abordagens científicas consistentes que compreendam fenômenos acerca da noção de informação, cujo assunto “[...] está ligado ao estabelecimento de um conjunto de princípios

fundamentais que direcionam o comportamento em todo processo de comunicação”. (Moura, 2006, p. 4).

Concluimos que a Teoria da Informação peirceana, associada às preocupações da Organização do Conhecimento com o processo de formação de conceitos, pode auxiliar no entendimento dos fluxos de conteúdos disseminados em ambientes digitais, conteúdos esses, muitas vezes não-informativos. E mais do que isso: cumpre as áreas indicar alternativas para o acesso e a identificação da informação de qualidade.

Agradecimento

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Referências

- Almeida, C. C. A. (2016) Semiótica Na Ciência Da Informação Brasileira: Ideias E Tendências. *Tendências Da Pesquisa Brasileira Em Ciência Da Informação*, V. 9, N. 2. Disponível Em: <Http://Hdl.Handle.Net/20.500.11959/Brapci/119663>. Acesso Em: 29 Jun. 2021.
- Almeida, C. C. (2016). A Biblioteconomia E A Ciência Da Informação Na Taxionomia Das Ciências De Charles Sanders Peirce. *Rdbci: Revista Digital De Biblioteconomia E Ciência Da Informação*, V. 3, N. 1, P. 1-19, 30 Set. 2005.
- Almeida, C. C. A. (2011). Sobre o pensamento de Peirce e a organização da informação e do conhecimento. *Liinc em Revista*, v.7, n.1, p. 104-112, mar. 2011. Disponível em: <http://www.ibict.br/liinc>. Acesso em 27 ago. 2021.
- Café, L.; Barros, C. M.; Santos, V. C. D. (2014) O conceito de organização do conhecimento nas revistas brasileiras de ciência da informação. *Revista Interamericana de Bibliotecología* (Colombia), v. 37, n. 3, p. 201-214, 2014. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/84836>. Acesso em: 01 set. 2021.
- Gonzalez, M.E.Q., Nascimento, T.C.A., Haselager, W.F.G. (2004). Informação E Conhecimento: Notas Para Uma Taxonomia Da Informação. In A. Ferreira, M.E.Q. Gonzalez & J.G. Coelho (Eds.). *Encontros Com As Ciências Cognitivas*, Volume 4. São Paulo: Coleção Estudos Cognitivos (Pp. 195-220).
- Gonzalez, M. E. Q.; Souza, R. S.; Silveira, L. F. B. (2018). Admirabilidade E Ação Autônoma: Uma Reflexão Filosófica Na Era Dos Big Data. In: Eluíza Bortolotto Ghizzi; Lúcia Ferraz Nogueira De Souza
- Dantas; Marcelo S. Madeira; Maria Eunice Quilici Gonzalez; Mônica Aiub. (Org.). *Sementes De Pragmatismo Na Contemporaneidade*. 1ed.São Paulo: Filoczar, V., P. 257-277.
- Gonzalez, M. E. Q.; Silveira, L. F. B. (2014). Instinct And Abduction In The Peircean Informational Perspective: Contributions To Biosemiotics. In: Romanini, Vinicius; Fernández, Eliseo (Eds.). (Org.). *Peirce And Biosemiotics: A Guess At The Riddle Of Life*. 1ed, P. 151-169.
- Nöth, W.; Gurick, A. (2011). A Teoria Da Informação De Charles S. Peirce. In: *Revista Digital De Tecnologias Cognitivas*. Edição 5, Pp. 4-29.
- Nöth, W. (2016). Habits, Habit Change, And The Habit Of Habit Change According To Peirce. In: WEST, D.E.; Anderson, M (Eds.). *Consensus On Peirce's Concept Of Habit: Before And Beyond Consciousness*. [S.L.]: Springer, 2016. P. 35-63.

- Mai, J-E. (2001). Semiotics And In Dexin G: An An Alysis Of The Subject Indexing Process. *Journal of Documentation*, vol. 57, no. 5, p. 591-622, Sep. 2001.
- Moura, M. A. (2006). Ciência Da Informação e Semiótica: Conexão De Saberes. *Encontros Bibli: Revista Eletrônica De Biblioteconomia E Ciência Da Informação*, Florianópolis, P. 1-17. ISSN 1518-2924. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2006v11nesp3p1/430>>. Acesso em: 24 Ago. 2020.
- Peirce, C. S. (1988). *The Essential Peirce*, V. 2, Peirce Edition Project (Ed.).Bloomington, In: Indiana University Press. [Ep]
- Peirce, C. S. (2008). A Fixação Da Crença. In: *Ilustrações Da Lógica Da Ciência*. Trad: Renato Rodrigues Kinouchi. 2 Ed. Aparecida, Sp: Ideias E Letras.
- Peirce, C. S. (1931-1958). *Collected Papers Of Charles S. Peirce*. 8 Vols. Ed. By Hartshorne, Charles, Paul Weiss And Arthur Burks. Cambridge, MA. Harvard University Press
- Santaella, L. (2004). O Papel Da Mudança De Hábito No Pragmatismo Evolucionista De Peirce. *Cognitio*, São Paulo, V. 5, N. 1, P. 75-83, Jan./Jun.
- Santaella, L.; Nöth, W. (2017) *Introdução À Semiótica: Passo A Passo Para Compreender Os Signos E A Significação*. São Paulo: Paulus.
- Vicentini, M. Rogério. Pascoal, V. A.; Gonzalez, M. E. Q (2019). Impactos Das Tecnologias Informacionais De Comunicação Na Conduta: Contribuições Da Teoria Peirceana De Informação. *COGNITIO (PUCSP)*, V. 20, P. 429-445.
- Vitti-Rodrigues, M. Gonzalez, M. E. Q.; Matulovic, M. (2017). Informação-Processo E Abdução. In: Alves, M. A.; Cabrini, M. C. G.Martínez-Ávila, D. (Org). *Informação, Conhecimento E Modelos (Cle)*, V. 78, Pp. 131-152.

(Página deixada propositadamente em branco)

SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO E HUMANIDADES DIGITAIS: POSSÍVEIS INTERLOCUÇÕES A PARTIR DA ABORDAGEM DA ANÁLISE DO DOMÍNIO

Ana Cristina de Albuquerque¹, Ania Rosa Hernandez Quintana²

¹Universidade Estadual de Londrina (UEL), albaunati@uel.br,
ORCID iD <https://orcid.org/0000-0003-3506-0479>

²Universidad de la Habana, aniahernandez.quintana@gmail.com,
ORCID iD 0000-0002-1484-8892

Resumo

O objetivo do presente trabalho é discutir quais são os elementos dos Sistemas de Organização do Conhecimento que demonstram uma interlocução com as Humanidades Digitais. Para tanto, serão apresentados a partir do enfoque específico das ontologias. Com a discussão apresentada, percebe-se um caminho para traçar algumas interlocuções a partir da Análise de Domínio e contribuir com o debate sobre recuperação e disseminação do conhecimento, assim como discutir o papel das Humanidades Digitais no campo da Organização do Conhecimento e consequentemente o diálogo com os Sistemas de Organização do Conhecimento.

Palavras-chave: Organização do Conhecimento, Sistemas de Organização do Conhecimento, Humanidades Digitais, Análise de Domínio.

Introdução

Um domínio específico, ou um conjunto de temas que tem aspectos semelhantes trabalhados por grupos de uma mesma comunidade, é compreendido como a unidade de análise que servirá de alicerce para a elaboração de Sistemas de Organização do Conhecimento (SOC). Lidando com o conceito e com os “lugares” que este conceito pode ocupar no âmbito de um domínio, a OC se caracteriza como um processo de modelagem para elaboração da representação do conhecimento, que, por sua vez, são modelos que vão da abstração ao mundo materializado através dos SOC.

Os SOC, mostram sua amplitude ao congregar esquemas que organizam o conhecimento como: taxonomias, sistemas de classificação, tesouros, ontologias.

Por sua vez, as Humanidades Digitais (HDs), se caracterizam como diferentes práticas que convergem interdisciplinarmente no âmbito das aplicações das tecnologias computacionais. Nas Ciências Humanas e Sociais, Artes e Letras, as HDs movimentam instrumentos e perspectivas referentes ao mundo digital, formando comunidades de práticas que têm a finalidade de contribuir com o avanço, disseminação e

compartilhamento do conhecimento, com foco na qualidade das pesquisas que podem ser realizadas e procurando enriquecer os saberes coletivos (MANIFESTO DE HUMANIDADES DIGITAIS, 2011). De um modo geral, as HDs podem ser um conjunto de práticas ou uma comunidade de práticas de pesquisa onde a ligação é o uso de tecnologias digitais.

A relação entre HDs e Ciência da Informação vem sendo trabalhada por autores brasileiros como Pimenta (2016), Castro e Pimenta (2018), Paletta (2018), Pimenta (2020), que propõe discussões e ressaltam as especificidades de cada área, mas que sugerem uma aproximação ou possibilidades de interesses que sejam comuns, tendo como mediador o uso das tecnologias. No entanto, no âmbito da OC e conseqüentemente dos SOC, essa relação ainda carece de estudos que problematizem, por exemplo, como o aumento dos espaços informacionais e de acervos digitais podem otimizar a relação do usuário com o acesso à informações que correspondam ao seu universo ou como as questões referentes aos domínios do conhecimento e das comunidades discursivas serão representadas a partir da interrelação dos produtos resultantes das HDs com os processos de tratamento informacional que garantam as dimensões social, histórica e cultural da informação (Hjørland e Albretchen, 1995).

Diante do exposto, o objetivo do presente trabalho é discutir quais são os elementos dos SOC que demonstram uma interlocução com as HDs, considerando-se que estas se dedicam ao compartilhamento e melhores formas de atingir o usuário por meio da recuperação de informações que dialoguem com suas escolhas e contexto.

O estudo se apresenta como exploratório e bibliográfico. Foi realizada a sistematização da literatura referente aos SOC, com ênfase em ontologias e HDs que serviu de base essencial para discutir e mapear os elementos de cada um, assim como para ter um panorama das discussões realizadas pelos pesquisadores da área.

Como percurso metodológico, foi utilizada a Análise de Domínio, considerando-se as onze abordagens de Hjørland e Albretchen (1995), Hjørland (2002) e Hjørland (2004). As discussões a respeito dos pontos comuns ou não dos SOC e das HDs, servirão como mais um tópico de reflexão que auxiliará nos aspectos teóricos e aplicados com o fim de oferecer informação e conhecimento contextualizado e de qualidade para os usuários.

Os elementos de interlocução são apresentados a partir do enfoque específico das ontologias consideradas um tipo de SOC que podem representar e organizar o conhecimento em diferentes domínios, no âmbito da Ciência da Informação, portanto, sua definição foi baseada em Brascher e Carlan (2010), Moura (2011) e Barité (2011). Os elementos das HDs foram baseadas em Galina Russel (2011), Koh (2015), Pimenta (2016), Medeiros et al (2017), Paletta (2018) e Pimenta (2020), que conceituam e determinam as HDs em diálogo com a Ciência da Informação.

Foram elencados cinco elementos constitutivos das ontologias e cinco das Humanidades Digitais e, a partir das onze abordagens definidas por Hjørland (2002), são discutidas suas interlocuções.

Considera-se este estudo relevante pelas perspectivas que vem se consolidando sobre a discussão da informação sob o viés sociológico, abrindo-se a possibilidades que vão além das orientações e abordagens teóricas, que a delimitam como objeto físico ou matéria processada pelos mecanismos cognitivos da mente humana.

Apresenta-se um mapeamento inicial das características que emergem das ontologias e das HDs com vistas a servir de instrumentalização para futuras discussões.

Dessa forma, espera-se contribuir com o debate sobre recuperação e disseminação do conhecimento, a importância social do uso do conhecimento e das fontes documentais, assim como discutir o papel das Humanidades Digitais no campo da Organização do Conhecimento e conseqüentemente o diálogo com os Sistemas de Organização do Conhecimento.

Ontologias e humanidades digitais: elementos característicos

Os Sistemas de Organização do Conhecimento são ferramentas semânticas com vocabulários estruturados e formalizados, usados para o tratamento e a recuperação do conhecimento disponível em recursos informacionais. Atuam como delimitadores quanto a utilização e definição de termos e conceitos em áreas específicas do conhecimento, o que os faz ter função importante por permitirem uma padronização terminológica direcionada à organização e representação de informações.

Para Barité (2011), Sistemas de Organização do Conhecimento é um termo geral que abarca tanto as linguagens elaboradas para indexar e classificar, como as linguagens que representam referenciais conceituais em estruturas semânticas relativas à documentação de qualquer área do conhecimento. O autor explica que o principal objetivo destas linguagens é a representação temática de conteúdos de quaisquer tipos de documentos, por isso a opção ou delineamento de uma definição genérica nomeada como Sistemas de Organização do Conhecimento (Barité, 2011), pois, a diversidade apresentada norteia para uma nomenclatura mais ampla, que possa englobar as semelhanças e definir as particularidades. As características principais dos SOC são definidas pelo referido autor como:

Su referencia al conocimiento especializado, la que se establece de diversas maneras (apelación a clasificaciones científicas, garantía literaria, opinión de expertos, etc.).

Su estructura lógica, que se construye de acuerdo a un método y a una teoría de organización del conocimiento.

El control de vocabulario, que contribuye a la selección, depuración, formalización y normalización de la terminología que incluye, así como al establecimiento relaciones recíprocas entre los términos, considerando criterios semánticos, lingüísticos y disciplinarios. (Barité, 2011, p. 127).

Os Sistemas de Organização do Conhecimento, assim, têm a função de propiciar o uso social do conhecimento quando permitem a circulação da informação através do relacionamento conceitual. Essa função pode ser percebida também a partir do direcionamento a serviços de referência, da utilização específica de conceitos em um domínio determinado e como mediadores entre documento e usuário, conhecimento científico e documentação e classificadores e indexadores no sentido de dar coerência e subsídios plausíveis para a construção de linguagens em ambientes controlados. (Barité, 2011).

Desta forma, considerada um SOC, as ontologias determinam os conceitos, estes definidos por Hjórland (2009), como significados que são construídos de forma

dinâmica, capazes de classificar o mundo por meio da intencionalidade com que foram idealizados, portanto, não podem ser compreendidos de forma isolada, posto que há diferentes concepções e diferentes conceitos em cada domínio do conhecimento, necessitando assim de um contexto específico. De acordo com Brascher e Carlan (2010, p. 160), a ontologia: “define os termos usados para descrever e representar uma área do conhecimento. Podem ser usados por pessoas, bancos de dados e aplicações que necessitam compartilhar informações em um domínio.

Tratando especificamente das ontologias, Barité (2011), explica que estas não são apenas estruturas de conhecimento mas sim desenhos de estruturas funcionais, com elementos e entidades que se relacionam e que convergem em objetivos e funções determinadas (Barité, 2011). Toda a estrutura das ontologias está baseada no relacionamento entre as entidades e assuntos, fazendo com que se tornem claros os trajetos que configuram um determinado domínio. O autor argumenta que: “Más que una estructura de conocimiento, una ontología es ante todo un sistema relacional de acciones que persigue tanto una gestión corporativa de calidad como la satisfacción plena del usuario”. (Barité, 2011, p. 132). O autor cita três níveis de conhecimentos que podem figurar em uma ontologia, esses níveis determinados por autores como Lacroix, Marty y Roche (1998):

- a) el dominio, que contiene el conocimiento que constituye el núcleo del dominio, los conceptos y las relaciones permanentes entre ellos; b) el conocimiento que describe situaciones del mundo real. Las relaciones entre objetos son aplicadas a objetos o casos que son reales en un determinado momento, y no en otro, por eso son eventuales; y, c) el conjunto de deducciones e inferencias de se extraen del dominio y de las situaciones del mundo real. (Barité, 2011, p. 132)

O referido autor explica que uma ontologia bem delimitada deve ter como um dos principais objetivos, sobre o que os usuários se interessam e seus comportamentos, isso vai auxiliar na determinação de objetivos concretos e possibilitar o trabalho com melhor direcionamento dentro de um determinado domínio (Barité, 2011).

Nesse sentido Brascher e Carlan (2010), explicam que para domínios diferentes a abordagem e metodologia de construção de ontologias são específicas e também dependem dos objetivos que terão. No entanto, seja qual for o domínio do conhecimento que as ontologias servirão como ferramenta, sua construção apresenta um alto grau de complexidade, visto que é direcionada e pensada para refletir a rede conceitual do domínio.

De acordo com Moura (2011), as utilidades das ontologias podem ser percebidas através da possibilidade de reuso e armazenamento de forma semântica dos conhecimentos, assim como o uso na interoperabilidade semântica. A autora, baseada em literatura específica da área, cita algumas características das ontologias:

- Elas se dividem em: **ontologias estáticas** que descrevem entidades existentes em um contexto de atributos e relações; **ontologias dinâmicas** que incorporam os aspectos dinâmicos dos contextos sociais e possuem foco nos estados e processos; **ontologias intencionais**, que expressam tarefas, objetivos

e crenças de agentes situados em um determinado ambiente e as **ontologias sociais** que cobrem os aspectos sociais. (Moura, 2011, p. 170)

Pode-se inferir que as ontologias são instrumentos que organizam conceitualmente o conhecimento de determinado domínio, com fins para representação das informações e têm, em um de seus objetivos principais, não só a recuperação das informações pelos usuários mas também ser fiel e coerente com o contexto específico que será construída e utilizada.

Desta forma, de acordo com autores como Hjórland (2009), Brascher e Carlan (2010), Moura (2011) e Barité (2011), os principais elementos característicos das ontologias podem ser apontados como: relacionamento conceitual em relação a um determinado domínio; estruturada em classes; vocabulário formal e complexo em relação ao controle e á representação coerente do domínio; garantia literária; uso social do conhecimento.

Por sua vez, as HDs podem ser compreendidas como uma comunidade de práticas ou práticas mediadas pelas tecnologias de informação. De acordo com Hockey (2004), as HDs se caracterizam como uma nova área acadêmica fortemente interdisciplinar que é capaz de oferecer metodologias específicas oriundas das tecnologias digitais que podem ser trabalhadas nos mais variados aspectos de investigação nas Humanidades. Moura (2019), explica que as HDs, enquanto campo, se tornou viável por conta da excessiva carga de digitalização de dados que favoreceu uma expansão da produção científica, no sentido de maior relacionamento colaborativo e a distância, fato que altera significativamente as relações e as mídias antes utilizadas de forma analógica.

Em 2010 foi lançado o Manifesto das Humanidades Digitais onde, algumas definições são:

- a opção da sociedade pelo digital altera e questiona as condições de produção e divulgação do conhecimento.
- as humanidades digitais referem-se ao conjunto das Ciências Humanas e Sociais, às Artes e às Letras. As humanidades digitais não negam o passado; apoiam-se, pelo contrário, no conjunto dos paradigmas, saber fazer e conhecimentos próprios dessas disciplinas, mobilizando simultaneamente os instrumentos e as perspectivas singulares do mundo digital.
- as humanidades digitais designam uma transdisciplina, portadora dos métodos, dos dispositivos e das perspectivas heurísticas ligadas ao digital no domínio das Ciências Humanas e Sociais.

O termo, segundo Rodríguez-Yunta (2014) e Galina Russel (2011), apresenta um crescimento de uso em todo o mundo, especialmente porque delimita o aspecto interdisciplinar que demonstra os processos relacionados ao uso de tecnologias digitais no âmbito das humanidades e, conseqüentemente, incide sob a demanda acadêmica, cultural e social de acessibilidade de fontes ligadas às humanidades que merecem reflexão, pois, permitem a discussão em torno da humanização das tecnologias por meio dos trabalhos críticos realizados a partir da disponibilização dessas fontes.

Também Almeida e Damian (2015, p. 08), definem o termo Humanidades Digitais considerando que por vezes este tem a tendência de:

[...] ser encarado como mais um rótulo generalista, surge como um campo interdisciplinar disposto a dar guarida às reflexões e às práticas suscitadas pelas mudanças decorrentes da introdução das tecnologias digitais no universo da cultura e das Unidades de Informação e Cultura brevemente descritas acima.

Considerando as Humanidades Digitais não como algo apartado, mas sim com uma continuidade dos processos de desenvolvimento de unidades informacionais, Almeida e Damian (2015), expõe que provavelmente pesquisadores, trabalham com as Humanidades Digitais mas que não utilizam a denominação de tal. Os projetos, bancos de dados, bibliotecas digitais, repositórios, curadoria digital incidem em um terreno que as Humanidades digitais se colocam e o esforço da percepção de que a transferência de suportes e ambiências vão além na direção de métodos articulados.

Neste sentido, Paletta (2018, p. 02), define Humanidades Digitais como:

[...] uma área de atividade acadêmica na interseção da computação ou tecnologias digitais e as disciplinas das humanidades. Inclui o uso sistemático de recursos digitais nas humanidades, bem como a reflexão sobre sua aplicação. Humanidades Digitais pode ser definida como novas formas computacionais de estudos acadêmicos transdisciplinar, que envolvem pesquisa, ensino e publicação colaborativa. Apresenta ferramentas e métodos digitais para o estudo das humanidades com o reconhecimento de que a palavra impressa não é mais o principal meio de produção e distribuição de conhecimento. Ao produzir e usar novas aplicações e técnicas, a Humanidades Digitais possibilita novos tipos de ensino e pesquisa, ao mesmo tempo que estuda e crítica como elas afetam o patrimônio cultural e a cultura digital. Assim, uma característica distintiva da Humanidades Digitais é o cultivo de uma relação bidirecional entre as humanidades e o digital: o campo emprega tecnologia na pesquisa e no questionamento de temas humanísticos.

Dessa forma, é possível pontuar que as Humanidades Digitais são percebidas como uma nova área, que em sua delimitação tem como objetivo congrega as áreas de Humanidades e Computacionais não só para melhor oferecer acesso às informações e documentos mas principalmente para possibilitar novas formas de disseminação e até de produção do conhecimento e que há um espaço para discussões acerca de aspectos sociais, culturais e de tratamento das informações contidas nos documentos disponibilizados.

Os estudos de Pimenta (2016), demonstram uma preocupação em definir aspectos teóricos e metodológicos das Humanidades Digitais. O autor explica que o campo se configura como espaço de estudo e pesquisa de ensino e de forma predominante como acesso à informação e inovação, se caracterizando assim como um campo híbrido (Pimenta, 2016).

Essas definições implicam principalmente na busca e na mudança do modo de trabalho, pois a partir da alteração das condições de divulgação do conhecimento, mudam-se também os objetivos de como se chega a esse conhecimento. Isso fica

delineado em Koh (2015), quando explica a necessidade de elaborar pesquisas que se referem especificamente as Humanidades Digitais para usuários que não dominam, por exemplo, as linguagens do campo da Ciência da Computação pois, de acordo com a autora, essas pesquisas devem se preocupar em como se fará uso e quais os efeitos das HDs quanto ao armazenamento documental.

Assim, as Humanidades Digitais podem contribuir em discussões sobre as implicações das tecnologias no dia a dia das pessoas.

Os objetivos das HDs, de acordo com Galina Russel (2011), versam sobre a criação de bases de dados digitais que contenham sistemas de recuperação, preservação, armazenamento da documentação e disseminação das informações dispostas; desenvolvimento de estruturas metodológicas que permitam criar novos documentos e dados sobre a documentação armazenadas e constituir incentivo a pesquisas que elucidem o entendimento do novo campo.

Desta forma, de acordo com autores como Galina Russel (2011), Pimenta (2016), Koh (2015), Paletta (2018) e Medeiros et al (2017), os principais elementos característicos das Humanidades Digitais podem ser apontados como: ferramentas de aprendizagem, por permitir interação entre usuário e ambiente; possibilidade de acesso a grande quantidade de dados, sendo essencialmente analítica; têm o caráter colaborativo, por apresentarem documentos digitalizados e bancos de dados interligados e passíveis de discussão quanto aos conteúdos; acesso às fontes documentais das humanidades; uso social do conhecimento.

Com os elementos das ontologias e das Humanidades Digitais, parte-se para a discussão das interlocuções a partir da abordagem da Análise de Domínio.

Interlocuções a partir das abordagens da análise de domínio

O estudo de um domínio, é proposto por Hjørland (2002), a partir de onze abordagens que podem ser compreendidas como um método a ser utilizado por profissionais da informação pois, ao aplicar, em suas análises, mais de uma abordagem, será possível compreender melhor a delimitação de determinado domínio. As onze abordagens, de acordo com Hjørland (2002), são: 1 guias de literatura e portais de assuntos especializados; 2 elaboração de tesouros e classificações especializadas; 3 indexação e recuperação especializada; 4 estudos empíricos de usuários; 5 estudos bibliométricos; 6 estudos históricos; 7 estudos de gênero e documentais; 8 estudos críticos e epistemológicos; 9 estudos terminológicos, de linguagem e de discurso; 10 estudos das estruturas e das comunidades científicas; 11 análise de domínio em cognição profissional e inteligência artificial.

O referido autor, explica que a Análise de Domínio é um estudo do desenvolvimento de discursos representados através da literatura de uma comunidade de pesquisa, o que considera como um produto formal. Neste sentido, a Análise de Domínio busca indicar primeiro os princípios e as regras de cada domínio, para distinguir a maneira de organização.

Seguindo esta premissa, os elementos das ontologias e das Humanidades Digitais, no presente trabalho, podem ser caracterizados conforme cada abordagem e relacionados tanto quanto a abordagem como entre si. A tabela a seguir demonstra as relações definidas:

Tabela 1: Abordagens da Análise de Domínio e elementos das ontologias e Humanidades Digitais.

ABORGADEM	ONTOLOGIAS	HUMANIDADES DIGITAIS
1 guias de literatura e portais de assuntos especializados	– uso social do conhecimento.	– acesso às fontes documentais das humanidades; – uso social do conhecimento.
2 elaboração de tesouros e classificações especializadas	– vocabulário formal e complexo; – garantia literária; – uso social do conhecimento.	– acesso às fontes documentais das humanidades; – uso social do conhecimento.
3 indexação e recuperação especializada	– relacionamento conceitual em relação a um determinado domínio; – estruturada em classes; – vocabulário formal e complexo; – garantia literária; – uso social do conhecimento.	– acesso a grande quantidade de dados; – colaborativas; – acesso às fontes documentais das humanidades; – uso social do conhecimento.
4 estudos empíricos de usuários	– garantia literária; – uso social do conhecimento.	– uso social do conhecimento.
5 estudos bibliométricos	– relacionamento conceitual em relação a um determinado domínio.	– acesso a grande quantidade de dados; – uso social do conhecimento.
6 estudos históricos	– relacionamento conceitual em relação a um determinado domínio; – estruturada em classes; – vocabulário formal e complexo; – garantia literária; – uso social do conhecimento.	– ferramentas de aprendizagem; – acesso a grande quantidade de dados; – colaborativas; – acesso às fontes documentais das humanidades; – uso social do conhecimento.
7 estudos de gênero e documentais	– relacionamento conceitual em relação a um determinado domínio; – estruturada em classes; – uso social do conhecimento.	– acesso a grande quantidade de dados; – acesso às fontes documentais das humanidades; – uso social do conhecimento.
8 estudos críticos e epistemológicos	– relacionamento conceitual em relação a um determinado domínio; – estruturada em classes; – vocabulário formal e complexo; – garantia literária; – uso social do conhecimento.	– ferramentas de aprendizagem – acesso a grande quantidade de dados; – colaborativas; – acesso às fontes documentais das humanidades; – uso social do conhecimento.
9 estudos terminológicos, de linguagem e de discurso	– relacionamento conceitual em relação a um determinado domínio; – estruturada em classes; – vocabulário formal e complexo; – garantia literária; – uso social do conhecimento.	– acesso a grande quantidade de dados; – colaborativas; – acesso às fontes documentais das humanidades; – uso social do conhecimento.

10 estudos das estruturas e das comunidades científicas	- garantia literária; - uso social do conhecimento.	- acesso às fontes documentais das humanidades; - uso social do conhecimento.
11 análise de domínio em cognição profissional e inteligência artificial	- relacionamento conceitual em relação a um determinado domínio; - estruturada em classes; - vocabulário formal e complexo; - garantia literária; - uso social do conhecimento.	- ferramentas de aprendizagem acesso a grande quantidade de dados; - colaborativas; - acesso às fontes documentais das humanidades; - uso social do conhecimento.

Fonte: elaborado pelas autoras

De acordo com a Tabela 1, as interlocuções das ontologias com as Humanidades Digitais são perceptíveis nas onze abordagens apontadas por Hjørland (2002). Os destaques podem ser verificados quanto as ontologias apresentarem o relacionamento conceitual em relação a um determinado domínio enquanto as HDs prezam por terem apelo ao acesso a grande quantidade de dados, serem colaborativas e oferecerem possibilidade de acesso às fontes documentais das humanidades. Todos esses elementos se relacionam nas abordagens no sentido de envolver o rigor metodológico de construção de uma ontologia, onde são preservados os sentidos dos domínios e de suas comunidades discursivas com o fim no usuário.

Um elemento comum é o uso social do conhecimento. Considerando que todas as abordagens delimitam um relacionamento com os usuários e uma garantia de que os conteúdos e documentos tratados serão coerentes com o contexto apresentado pelas comunidades discursivas, as ontologias como sistemas que servirão para a organização e representação das informações e as HDs que objetivam a aprendizagem, assim como o conhecimento e competências dos usuários em relação a novas metodologias e reflexões quanto a área de humanidades, podem ser percebidas em seu princípio, como ferramentas específicas direcionadas a socialização do conhecimento, visto que são colaborativas. As ontologias fornecem as bases classificatórias, taxonômicas e de relacionamento conceitual, enquanto as HDs apresentam os documentos digitalizados, as fontes documentais e o conteúdo propriamente dito que necessita de tratamento para poder ser disponibilizado.

Nesse sentido, as interlocuções a partir das onze abordagens apresentadas, enfatizam a importância da observação e estudo acerca de elementos que influenciarão diretamente na organização e na recuperação da informação, considerando as especificidades e as diferenças nas comunidades discursivas que são o espaço onde a informação é efetivamente construída e compartilhada, nesse caso, na intância das humanidades.

A combinação dos interesses das Humanidades Digitais quanto as fontes necessárias para se pensar um domínio específico e dar acesso ao conhecimento, com os objetivos das ontologias, inserindo o relacionamento conceitual, pode-se promover um caminho para direcionar de forma coerente e crítica as necessidades de diferentes usuários e assim, conduzir a maior acessibilidade e maneiras de gerir, disseminar e criar novos conhecimentos.

Conclusões

A articulação de conhecimentos no universo digital e das Ciências Sociais e Humanas é elemento de conciliação quando observado a centralidade na digitalização e disponibilização de fontes, mas também a procura da sistematização e análise dessas fontes para a recuperação e compreensão de forma adequada aos usuários e seus contextos

Nesse sentido, as ontologias permitem o controle da linguagem utilizada nos espaços de informação, como os ambientes digitais onde os objetos passam a um estado múltiplo de fixação, ou seja, as coisas podem estar em mais de um lugar ao mesmo tempo fazendo com que as informações não fiquem armazenadas de forma contígua. Assim, a implementação de uma ontologia combinada com as características e funções das HDs têm o potencial de elaboração de um controle nas relações conceituais, além de materializar alguns objetivos das HDs como o de disponibilizar e disseminar o conhecimento por meio digital, fornecer bases integradas às Ciências Humanas e Sociais no sentido de metodologias e teorias e potencializar as atividades digitais diante das perspectivas da humanização visando os usuários, seus relacionamentos com os conceitos e contextualização dos documentos.

Estas interlocuções aparecem nas onze abordagens de Hjørland (2002): guias de literatura e portais especializados; classificações especiais e *thesaurus*; Especialidades da indexação e recuperação; Estudo empírico de usuários; Estudos bibliométricos; Estudos históricos; Documentos e estudos de gêneros; Estudos críticos e epistemológicos; Estudos terminológicos, de linguagem e de discurso; Estudos de instituições e estruturas das comunidades científicas; Cognição profissional e inteligência artificial. A partir das onze abordagens foi possível verificar o objetivo proposto pelo autor, como um método de treinar profissionais especialistas em informação (Hjørland, 2002). Estas abordagens fornecem o elo principal para o início das discussões sobre as interlocuções entre os SOC e as HDs

Conforme o encaminhamento do estudo buscou-se um paralelo dos objetivos das ontologias e das HDs, e a pontuação de alguns elementos característicos que os agregam, a partir da Análise de Domínio, enquanto campos preocupados com a fundamentação digital para acesso ao conhecimento.

As ontologias encontram uma interlocução com as HDs em diversos momentos e abarcam projetos e perspectivas interdisciplinares que vão ao encontro da necessidade de reflexão mais profunda em torno dos aspectos como a produção de registros digitais para a recuperação e disseminação do conhecimento, indexação, classificação e representação de forma coerente desses recursos.

Referências

- Barité, M. (2011). Sistemas de Organização do Conhecimento: uma tipologia atualizada. *Informação & Informação*, 16(2), 122-139. doi:<http://dx.doi.org/10.5433/1981-8920.2011v16n2p122>
- Bräscher, M.; Carlan, E. (2010). Sistemas de organização do conhecimento: antigas e novas linguagens. In: Jaime Robredo; Marisa Bräscher (Orgs.). *Passeios no Bosque da Infor-*

- mação: Estudos sobre Representação e Organização da Informação e do Conhecimento. Brasília DF: IBICT, 147-176 Edição eletrônica. Disponível em: <http://www.ibict.br/publicacoes/eroic.pdf>. (Edição comemorativa dos 10 anos do Grupo de Pesquisa EROIC).
- Castro, R. M., & Pimenta, R. M. (2018). Novas práticas informacionais frente às humanidades digitais: a construção de acervos digitais como suporte para as digital humanities. *Informação & Informação*, 23(3), 523-543. <http://dx.doi.org/10.5433/1981-8920.2018v23n3p523>
- Galina Russell, I. (2011). ?Qué són las humanidades digitales?. *Revista Digital Universitaria*, 12(7). <http://eprints.rclis.org/19368/1/037-043-Rz-Yunta-Humanidades-digitales.pdf>
- Hjørland, B., & Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*, 46(6), 400-425. [https://doi.org/10.1002/\(SICI\)1097-4571\(199507\)46:6<400::AID-ASIS2>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASIS2>3.0.CO;2-Y)
- Hjørland, B. (2002). Domain analysis in information science: eleven approaches: traditional as well as innovative. *Journal of Documentation*, 58(4), 422- 462.
- Hjørland, B. (2004). Domain analysis: a socio-cognitive orientation for Information Science research. *Bulletin of the American Society for Information Science and Technology*, 30(3).
- Hjørland, B. (2009). Concept theory. *Journal of the American Society for Information Science and Technology*, 60 (8), 1519-1536. <https://doi.org/10.1002/asi.21082>
- Hockey, S. (2004). The history of humanities computing. In: S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A companion to digital humanities* (pp. 3-19). Malden, MA: Blackwell Publishing.
- KOH, A. (2015). *A letter to the humanities: dh will not save you*. Disponível em: <<http://www.digitalpedagogylab.com/hybridped/a-letter-to-the-humanities-dh-will-not-saveyou/>>.
- Manifesto de Humanidades Digitais, 2011. <https://humanidadesdigitais.org>.
- Medeiros, A. L. S.; Trézze, L. J. L.; Leite, T. P.; Melo, E. S.; Oliveira, A. C.; Teixeira, M. F. (2017). Humanidades digitais na fundação casa de Rui Barbosa: um estudo aplicado de seu conceito. *Informação & Tecnologia*, 4(2), p. 243-259. <http://hdl.handle.net/20.500.11959/brapci/101598>.
- Moura, M. (2011). Interoperabilidade Semântica e Ontologia Semiótica: a construção e o compartilhamento de conceitos científicos em ambientes colaborativos online. *Informação & Informação*, 16(2), 165-179. doi:<http://dx.doi.org/10.5433/1981-8920.2011v16n2p165>
- Moura, M. (2019). Ciência da Informação e humanidades digitais: mediações, agência e compartilhamento de saberes. *Perspectivas em Ciência da Informação*, 24(esp.), 57-69. <https://doi.org/10.1590/1981-5344/3893>
- Paletta, F. C. (2018). Ciência da Informação e humanidades digitais: uma reflexão. In *Anais do XIX Encontro Nacional de Pesquisa em Ciência da Informação* (pp. 147-162). Londrina, PR.
- Pimenta, R. M. (2016). As rugosidades do Ciberespaço: um contributo teórico aos estudos dos web espaços informacionais. *Informação & Sociedade: Estudos*, 26(2), 77-90. Recuperado de <https://periodicos.ufpb.br/index.php/ies/article/view/28116>
- Pimenta, R. M. (2020). Por que Humanidades Digitais na Ciência da Informação? Perspectivas progressas e futuras de uma prática transdisciplinar comum. *Informação & Sociedade: Estudos*, 30(2). <https://doi.org/10.22478/ufpb.1809-4783.2020v30n2.52122>
- Tang, MC., Cheng, YJ & Chen, KH. (2017). Um estudo longitudinal de coesão intelectual em humanidades digitais usando análises bibliométricas. *Scientometrics* 113, 985-1008. <https://doi.org/10.1007/s11192-017-2496-6>

(Página deixada propositadamente em branco)

APROXIMAÇÕES ENTRE ORGANIZAÇÃO DO CONHECIMENTO E HUMANIDADES DIGITAIS

Laura Mariane de Andrade¹, Paula Regina Dal'Evedove²

¹Universidade Federal de São Carlos (UFSCar), marapan@gmail.com, ORCID 0000-0001-7675-1546

²Universidade Federal de São Carlos (UFSCar), dalevedove@ufscar.br, ORCID iD 0000-0003-1521-2201

Resumo

Nas últimas décadas um novo campo de investigação surgiu e tem se consolidado, nomeadamente, Humanidades Digitais, que na integração das humanidades tradicionais e da tecnologia digital, estabelecem a sua própria identidade aplicando a inovação digital na prática humanística. Em razão de sua característica transdisciplinar, cada vez mais se apresenta como um tema emergente para a Ciência da Informação, idealmente na Organização do Conhecimento. Este estudo exploratório objetiva analisar as iniciativas que aproximam Organização do Conhecimento e Humanidades Digitais de modo a compreender como as teorias e tendências de se pensar as tecnologias digitais nos fazeres humanos são dialogadas na área. Para tanto, foram selecionadas as bases de dados LISA, LISTA e WoS para localizar documentos provenientes da Ciência da Informação sobre Humanidades Digitais publicados entre 2010 e 2020. A partir da análise de conteúdo, foram aplicados os critérios de inclusão e exclusão e obtido um *corpus* de pesquisa composto por 24 artigos científicos, em que a leitura dos textos completos viabilizou o tratamento, inferência e interpretação dos resultados, os quais revelam como a Organização do Conhecimento está incorporando discussões sobre tecnologias digitais acrescidas das humanidades pela perspectiva das Humanidades Digitais. De modo geral, constata-se que a inserção terminológica das Humanidades Digitais em discussões inerentes à Organização do Conhecimento não é escassa, mas que os estudos se direcionam para as questões de natureza aplicada e não discutem ou teorizam questões próprias dessa aproximação entre Organização do Conhecimento e Humanidades Digitais. Conclui que este movimento de aproximação se revela importante para subsidiar perspectivas científicas comuns.

Palavras-chave: Interdisciplinaridade, Organização do Conhecimento, Humanidades Digitais, Revisão bibliográfica.

Introdução

Nas últimas décadas um novo campo de investigação surgiu e tem se consolidado, nomeadamente Humanidades Digitais. Configura-se como a intersecção das

ciências humanas, ciências sociais e tecnologia digital (Moura, 2019), ao passo que aspira a criação de ferramentas para novas formas de representações (Sousa, 2015), além de traçar caminhos da tecnologia digital sob perspectiva transdisciplinar.

Na integração das humanidades tradicionais e da tecnologia digital, as Humanidades Digitais estabelecem sua própria identidade aplicando a inovação digital na prática humanística. Porém, vão além quando o impacto da inovação tecnológica nos campos acadêmico e social é compreendido por meio de metodologias, teorias e procedimentos práticos (Santos & Venâncio, 2017). O Manifesto das Humanidades Digitais (Dacos, 2010), declaração que lidera a discussão global neste campo de investigação, aponta que as Humanidades Digitais integram métodos, equipamentos e perspectivas de pesquisa das Ciências Humanas e Sociais, bem como recursos semelhantes criados pela tecnologia digital. O manifesto surge em 2010 como um marco para o fortalecimento desta nova abordagem investigativa.

Humanidades Digitais é um conceito em construção que, além de seu desenvolvimento natural com a aceitação da tecnologia digital pela sociedade, também está se desenvolvendo em um curto espaço de tempo. Embora ainda não esteja plenamente consolidado como campo científico, apresenta discussões cada vez mais sólidas e em número crescente. Os antecedentes e os temas trazidos para a presente investigação referem-se às abordagens e uso das Humanidades Digitais na área de Organização do Conhecimento. As discussões teóricas e filosóficas sobre as potencialidades e limites das Humanidades Digitais estão em franco desenvolvimento, sendo notório o aumento de publicações na Ciência da Informação brasileira dedicadas ao tema (Almeida & Damian, 2015; Santos & Venâncio, 2017; Paletta, 2018; Moura, 2019; Siqueira & Flores, 2019; Andrade & Dal'Evedove, 2020; Pimenta, 2020, dentre outros).

Em atenção ao exposto, questiona-se como as Humanidades Digitais são dialogadas e integradas em pesquisas que discutem os problemas, desafios e possibilidades da Organização do Conhecimento frente às tecnologias digitais? Neste sentido, objetiva-se analisar as iniciativas que aproximam Organização do Conhecimento e Humanidades Digitais mediante o exame da literatura científica disponível. A discussão aqui estabelecida busca compreender como as teorias e tendências de se pensar as tecnologias digitais nos fazeres humanos são dialogadas na Organização do Conhecimento. Contribui para tal direcionamento investigativo a incipiente discussão envolvendo as Humanidades Digitais em Organização do Conhecimento, sobretudo, que vinculem os desafios de se pensar o tecnológico e o humano como elementos articulados.

Trata-se de um estudo exploratório que recorre à literatura internacional para identificar pesquisas científicas de Ciência da Informação que trabalham com questões próprias da Organização do Conhecimento e, ao mesmo tempo, contemplam Humanidades Digitais, sendo pautada em uma metodologia de abordagem qualitativa que emprega o método da análise de conteúdo para a coleta, leitura, inferência e interpretação dos dados. Os principais resultados apontam que as Humanidades Digitais estão presentes, implícita ou explicitamente, na Organização do Conhecimento, sendo a contextualização desse movimento oportuno para o entendimento de seus limites e potencialidades, visando uma aproximação mais consistente e proveitosa.

Metodologia

Para atingir o objetivo proposto, em janeiro de 2021 foram realizadas buscas de artigos científicos sobre Humanidades Digitais datados de 2010 a 2020 nas bases de dados *Library and Information Science Abstracts* (LISA), *Library, Information Science & Technology Abstracts with Full Text* (LISTA) e *Web of Science* (WoS). A Tabela 1 detalha a estratégia de busca empregada com os respectivos resultados.

Tabela 1. Busca e Resultados obtidos

Bases de dados	Estratégia de busca	Publicados 2010-2020
LISA	nof(“ <i>Digital Humanities</i> ” OR “ <i>Humanidades Digitales</i> ” OR “ <i>Humanidades Digitais</i> ”) AND stype.exact(“ <i>Scholarly Journals</i> ”) AND at.exact(“ <i>Article</i> ”), além dos limites adicionais de Tipo de fonte (periódicos acadêmicos); e Tipo de documento (artigo)	441 registros
LISTA	TI (“ <i>Digital Humanities</i> ” OR “ <i>Humanidades Digitales</i> ” OR “ <i>Humanidades Digitais</i> ”) OR SU (“ <i>Digital Humanities</i> ” OR “ <i>Humanidades Digitales</i> ” OR “ <i>Humanidades Digitais</i> ”) OR AB (“ <i>Digital Humanities</i> ” OR “ <i>Humanidades Digitales</i> ” OR “ <i>Humanidades Digitais</i> ”) e com os filtros de Tipo de documento (<i>article</i>) e Tipo de publicação (<i>periodical</i>)	373 registros
WoS	Tema: “ <i>Digital Humanities</i> ” OR “ <i>Humanidades Digitales</i> ” OR “ <i>Humanidades Digitais</i> ”, refinada por Tipos de documento (artigo) e Categoria (<i>Information Science Library Science</i>)	252 registros
Total		1.066 registros

Para a definição do *corpus* de pesquisa, foram removidos os registros que não eram apresentados na língua inglesa, registros que não possuíam resumos, as duplicações e inconsistências, ocasionando a baixa de 540 registros. Procedeu-se, então, com a aplicação do método de análise de conteúdo (Bardin, 2016) mediante as fases de pré-análise e exploração do material nos 526 registros selecionados. Nesta fase, a leitura do resumo, título e palavras-chave foi direcionada para a identificação de pesquisas na Ciência da Informação que vinculam as Humanidades Digitais ao universo científico da Organização do Conhecimento. Para tanto, tomou-se como base interpretativa e de confirmação o estudo de Dahlberg (2006), os artigos de Humanidades Digitais selecionados pela literatura de Organização do Conhecimento¹, bem como dos princípios estabelecidos pelo *Classification System for Knowledge Organization Literature* (CSKOL). Neste momento foi possível observar a ocorrência comum de termos provenientes do universo científico da Organização do Conhecimento, como por exemplo: *categorise, classification, classification system, controlled vocabularies, descriptors, domain, index, information retrieval, knowledge management, knowledge organization, ko, markup, ontology, problems of natural language, semantic, standardization, subject, tagging, taxonomies, thesauri e vocabularies*. Esta manobra resultou em um total de 24 documentos elegíveis e que

¹ Disponível em: <https://www.isko.org/lit.html>

compuseram o *corpus* final da pesquisa. Para a fase de tratamento dos resultados, inferência e interpretação, realizou-se a leitura completa e atenta dos documentos, ação que oportunizou às pesquisadoras a compreensão de aspectos mais profundos dos textos para desvendar conteúdos latentes. Com isso, a metodologia aqui empregada mostrou-se adequada ao permitir identificar os escopos investigativos trabalhados nos estudos que integram Organização do Conhecimento e Humanidades Digitais; e conhecer as concepções de Humanidades Digitais empregadas pelos autores.

Resultados

O Quadro 1 apresenta o escopo investigativo de cada um dos 24 artigos científicos analisados de forma pormenorizada, conforme segue:

Quadro 1. Escopo investigativo dos artigos

N.	Autores	Escopo investigativo
1	(Pertsas & Constantopoulos, 2016).	Apresenta uma ontologia para modelagem de práticas acadêmicas chamada <i>Scholarly Ontology</i> (SO). Apontam características das estruturas ontológicas, principalmente os benefícios frente às novas tecnologias das humanidades, além de descrever a modelagem, uso e reuso de dados pela ferramenta.
2	(Vlachidis & Tudhope, 2015)	Apresenta a técnica de detecção de negação e desambiguação de sentido de palavras automática, um desafio apresentado pela variedade em que a negação pode ser expressa em linguagem natural no contexto de indexação semântica de literatura cinza arqueológica para recuperação de informação. Aponta que os resultados são encorajadores, apresentando 89% de precisão.
3	(Chen, 2019)	Usa o Banco de Dados de Nomes e Biografias na transformação de dados originais em dados vinculados (<i>linked data</i>) para explorar o enriquecimento semântico através do aumento do conteúdo das entidades com recursos externos, aumentando as relações hierárquicas entre as propriedades e mostrando seu domínio para enriquecimento de pesquisas que se utilizam de uma plataforma de humanidades digitais chinesa.
4	(Deng, 2020)	Apresenta um modelo conceitual para o desenvolvimento da memória oral chinesa com base nos métodos das humanidades digitais, ao se assemelhar a uma rede relacionada de conhecimento, útil tanto para acadêmicos neste campo quanto para a pesquisa sobre memória oral das humanidades digitais.
5	(Liang, Wang, & Li, 2020)	Baseia-se na teoria arquivística e analisa automação de texto por OCR e ferramentas de crowdsourcing (como <i>captcha</i>) para realizar o gerenciamento de atribuição de tarefas a voluntários na digitalização de arquivos da biblioteca de Xangai, como transcrição, pontuação, revisão, descrição de metadados, segmentação e anotação de atributos.
6	(Li, Zhu, Shen, Du, Guan, & Deng, 2020)	De acordo com as etapas de construção de um modelo de ontologia, o artigo estudou e analisou os elementos de eventos, baseado em muitas ontologias e vocabulários de eventos existentes, e modela a ontologia de eventos históricos da República da China, que representa eventos e personalidades históricas com relacionamentos ricos e de entidades.
7	(Arano-Poggi & Bel, 2014)	Apresenta o projeto europeu do Clarin IULA-UPF Competence Center que compila e inter-relaciona as informações necessárias aos humanistas digitais em um catálogo de dados vinculados (<i>linked data</i>) e a seleção dos

		materiais e instrumentos descritivos utilizados na elaboração do catálogo, também chamado de repositórios de informação especializada.
8	(Heuwing, Mandl, & Womser-Hacker, 2016)	Observa e discute seu o comportamento informacional de historiadores em suas pesquisas em mídias educacionais históricas para apresentar um processo de design contextual ao definir requisitos informacionais que estes buscam para ferramentas que possibilitem a análise de texto automática e multidimensional.
9	(Marsh, 2017)	Descreve como o sistema de gerenciamento de conteúdo Omeka apóia a representação e apresentação das exposições digitais da National Agricultural Library (NAL), apresentando como a ferramenta trabalha a representação, organização e acesso de objetos digitais com metadados de assunto, autoria, contexto ou outros recursos.
10	(Wettlaufer & Westphal, 2014)	Apresenta alguns dos métodos e práticas textuais das humanidades digitais em relação a bibliotecas e especialmente arquivos, principalmente em linguística computacional e PNL. Aborda as Humanidades Digitais como transdisciplinar e apresenta o histórico do termo e práticas como “humanities computing”.
11	(Tomić & Willer, 2013)	Aponta infra estruturas colaborativas como soluções de organização para reuso de dados e interoperabilidade de ferramentas digitais no contexto das humanidades digitais e tecnologia da informação para digitalização de incunábulo e manuscritos medievais, neste caso, especificamente os croatas, através de tecnologias semânticas e dados abertos vinculados (<i>linked open data</i>).
12	(Melgar Estrada, 2011)	Revisa a literatura que aborda os Mapas de tópicos (<i>Topic Maps</i>), buscando a resposta conceitual sobre a intersecção com a Organização do Conhecimento, mas também sua aplicabilidade em Bibliotecas, Arquivos, Museus e Bibliotecas Digitais em Humanidades. Aponta que os mapas de tópicos fornecem os mecanismos necessários para levar os sistemas de organização do conhecimento existentes para a <i>Web</i> através de ontologias de forma mais natural que as tecnologias existentes.
13	(Ross, Terras, Warwick, & Welsh, 2011)	Investiga o uso do Twitter pela comunidade acadêmica de humanidades digitais em conferências e aponta que um backchannel (recurso de comunicação não verbal, em tempo real, que não interrompe um apresentador ou evento) possibilita a expansão da comunicação e participação em eventos entre os membros da comunidade.
14	(Portier, Chatti, Calabretto, Egyed-Zsigmond, & Pinon, 2012)	Trata da questão central da codificação de documentos multi estruturados e propõe um modelo para construção de edições digitais de grandes arquivos de manuscritos, denominado <i>Multi-Structured Document Model</i> (MSDM), que tem como diferencial sua base em estrutura documentária, estrutura básica e correspondências de relações.
15	(Gill, Hinrichs-Krapels, Blanke, Grant, Hedges, & Tanner, 2017)	Combina abordagens computacionais de processamento de linguagem natural e mineração de texto com análise humana para demonstrar como eles podem explorar um conjunto de dados textuais de forma otimizada, aplicando mineração de texto e utilizando desde palavras-chave em contexto e informações de frequência a modelagem de tópicos.
16	(Simon, Barker, Isaksen, & Soto, 2017)	Apresenta o Recogito 2, uma ferramenta de código aberto para anotação semântica, além de discorrer sobre sua história e atividade atual, e discutir o papel da anotação de dados vinculados. Examina alguns trabalhos relacionados no campo de aplicativos de dados vinculados, ferramentas de anotação semântica e integração com tesouros e listas de autoridade.
17	(Yaco Ramaprasad, 2019)	Sugere uma estrutura de uma linguagem comum para melhorar a conexão entre os domínios dos artefatos de patrimônio cultural e proporciona mapear a literatura científica atual para descobrir áreas com muitas e poucas pesquisas em determinados domínios de instrução. O mapeamento pode ser usado para gerar gráficos ontológicos que iluminarão claramente

		os elementos ou caminhos de informática para o ensino de patrimônio cultural.
18	(Robledano-Arillo, Navarro-Bonilla, & Cerdá-Díaz, 2019)	Apresenta um modelo conceitual e ontologia de dados abertos vinculados (<i>linked open data</i>) para a codificação e disseminação de metadados de arquivos fotográficos históricos, buscando abranger exaustivamente a representação de conteúdos e contextos que este suporte não costuma ter devido às suas peculiaridades através de métodos de representação textual que permitam apoiar sistemas de consulta de conteúdos gráficos potentes.
19	(Potha & Stamatatos, 2019)	Apresenta um estudo sistemático que examina a utilidade da modelagem de tópicos quando combinada com métodos de verificação de autor, cobrindo paradigmas baseados em instância e baseados em perfil, bem como paradigmas intrínsecos e extrínsecos, focando em Indexação Semântica Latente e Alocação de Dirichlet Latente. Conclui que os métodos baseados em perfis, extrínsecos e Alocação de Dirichlet Latente são mais eficazes.
20	(Martin-Rodilla & Gonzalez-Perez, 2019)	Trata do panorama dos metadados nas Humanidades Digitais a partir da modelagem conceitual. Descreve 30 abordagens existentes e propõe uma abordagem para diminuir a proliferação de padrões de metadados. Uma abordagem simples e expressiva da metainformação pode ser usada ao tratar a metainformação em Humanidades Digitais voltada para indexação e busca, representação, integração e geração de conhecimento.
21	(Kalita & Deka, 2020)	Desenvolve e apresenta uma ontologia para preservação da base de conhecimento das práticas de dança tradicional, testada usando as danças das tribos Rabha do Nordeste da Índia. Considera a importância de sistemas de informação para a preservação coletiva de bases de conhecimento cultural, seguindo sistemas de organização de conhecimento.
22	(Hocker, Schindler, & Rittberger, 2020)	Apresenta uma nova abordagem para a criação de ontologias no campo da ciência aberta através do design participativo em entrevistas, considera que os esquemas de codificação são um resultado importante da pesquisa qualitativa e a reutilização de dados pode gerar um grande potencial para a ciência aberta.
23	(Daquino, 2020)	Apresenta uma análise comparativa de dados abertos vinculados sobre história da arte com o objetivo de avaliar a credibilidade de fontes secundárias que registram atribuições de obras de arte. Desenvolve um modelo de classificação baseado em ontologia para recomendar atribuições de autorias para obras de arte e apoiar o processo de tomada de decisão de historiadores e catalogadores.
24	(Niu, 2020)	Apresenta a ferramenta Mapa de Conhecimento em um projeto de memória urbana e propõe um modelo conceitual ao usar os portões da cidade de Pequim como um caso focal, pois é útil para a organização semântica e visualização de dados heterogêneos. A partir da mineração de pontos de memória de dados de memória urbana, busca organizar o conhecimento e construir relações semânticas entre os pontos de memória.

Fonte: Dados da pesquisa.

Na análise dos resultados obtidos, observa-se como primeira convergência dos estudos em Humanidades Digitais o enfoque em esquemas, ontologias, modelos conceituais e ferramentas no geral que se encontram no universo das tecnologias digitais de humanidades abordadas na Organização do Conhecimento, geralmente na dimensão aplicada. Os estudos nesta direção observam objetos culturais ou históricos em contextos semânticos, com destaque para a aplicação em ou para bibliotecas, arquivos ou museus nacionais da China, Espanha, Croácia, dentre outros.

A análise também viabilizou o agrupamento temático dos estudos no que concerne a origem dos dados ou informações tratadas, que divergem entre conteúdo coletado da *Web* (através de mídias sociais ou não), catálogos de bibliotecas ou arquivos, repositórios digitais para análise; e do movimento inverso, de digitalização, que busca capturar e digitalizar dados de acervos que tem menor projeção para disponibilizar na Internet e, por isso, objetivam criar ou aprimorar ferramentas que se preocupam com o reuso e a recuperação da informação através de camadas semânticas estruturadas. Constata-se, também, que os estudos que abordam ontologias, tecnologias semânticas e *linked open data* se classificam como pesquisas de Humanidades Digitais, seja por se especializarem em informações e dados das ciências humanas para aplicar as tecnologias abordadas, seja por dar enfoque à temática do desenvolvimento da indexação de assunto, contextualização de dados e da própria organização do conhecimento para o enriquecimento das construções semânticas.

Apesar de todos os estudos analisados abordarem as Humanidades Digitais, curiosamente, nem todos a conceituam, o que pode refletir três possíveis entendimentos:

- Humanidades Digitais se apresenta como um conceito estabelecido e o cenário não demanda uma conceitualização pelos pesquisadores para a sua vinculação com a Organização do Conhecimento;
- Os estudos consideram ambientes específicos de humanistas digitais, deslegitimando a necessidade desta explanação; ou
- Os autores a consideram um conceito dedutivo pela simplicidade da própria terminologia, com vinculação natural às questões que integram o ambiente digital aplicado às humanidades.

Referente às concepções de Humanidades Digitais, constata-se que dos 24 artigos científicos analisados, apenas 7 apresentam uma conceitualização sobre Humanidades Digitais. Esses conceitos estão listados no Quadro 2, de modo a evidenciar a abordagem empregada em cada um dos estudos e, idealmente, revelar como as discussões sobre Humanidades Digitais estão sendo conduzidas no âmbito da Organização do Conhecimento.

Quadro 2. Concepções de Humanidades Digitais nos estudos analisados

Autores	Concepções
(Deng, Wang, Song, & Zhang, 2020)	As humanidades digitais integram os quocientes da computação e da tecnologia com a pesquisa das humanidades tradicionais e, portanto, permite que os pesquisadores abordem “atomização acadêmica e questões não correlacionadas” (BURDICK et al., 2018). Por meio de tecnologias como computação, análise de dados e visualização, o conhecimento humanístico pode ser filtrado, extraído, explicado, conectado e reduzido, o que remodela o conhecimento humanístico, “redesenha o contorno” do setor acadêmico e constrói um novo campo acadêmico interativo que expande a esfera de uma disciplina acadêmica.
(Liang, Wang, & Li, 2020)	Na atualidade, as humanidades estão estabelecendo novas metodologias para a era digital que utilizam as tecnologias da informação, dando origem ao emergente campo interdisciplinar das humanidades digitais. Uma área de pesquisa nas humanidades digitais busca aplicar ferramentas de análise digital para processar os materiais de pesquisa em massa, e a observação geral por máquina é uma das filosofias centrais no projeto do sistema (CHEN et al., 2011).
(Li, Zhu, Shen, Du, Guan, & Deng, 2000)	Área de pesquisa interdisciplinar que surgiu nos últimos anos. A pesquisa nesta área não apresenta apenas aspectos teóricos, mas também tem aplicações práticas específicas em diversas disciplinas de humanidades, como história, literatura e filosofia. Algumas dessas aplicações práticas envolvem arquivos digitais, coleções de textos, edições acadêmicas digitais e ferramentas para pesquisar dados em textos digitalizados.
(Martin-Rodilla & Gonzalez-Perez, 2019)	As Humanidades Digitais, como um campo relativamente novo e interdisciplinar (PARK, 2009), apresenta significativa fragmentação institucional, tanto em programas de financiamento quanto em custódia de informações e políticas de uso. Humanidades Digitais também envolvem uma ampla gama de trabalhos auto reflexivos sobre como a disciplina está trabalhando para produzir novos conhecimentos a partir de dados.
(Wettlaufer & Westphal, 2014)	Humanidades digitais estão em fase de definição desde o início de sua implantação como disciplina especializada, o que dificulta a delimitação da área temática e dos métodos. Ainda não existe uma definição geral aplicável, mas há uma série de tentativas de definição que mais ou menos giram em torno das humanidades digitais como “prática”. Isso descreve uma situação em que uma disciplina especializada se relaciona com uma área de assunto (humanidades) que não pode ser definida em mais detalhes e na qual os processos auxiliados por computador devem ser usados. O elemento de conexão encontra-se atualmente mais no uso de métodos de ciência da computação do que no campo extraordinariamente amplo de pesquisa em humanidades, que é dividido em pelo menos 40 disciplinas individuais com diferentes métodos de pesquisa e áreas de pesquisa.
(Tomić & Willer, 2013)	As humanidades digitais estavam no início da disciplina chamada computação em humanidades. Conforme definido por seu “pai” Roberto Busa, as humanidades digitais são “a automação de toda análise possível da expressão humana (portanto, é uma atividade primorosamente “humanística”), no sentido mais amplo da palavra, da música ao teatro, do design e pintura à fonética, mas cujo núcleo continua sendo o discurso dos textos escritos (BUSA, 2004).
(Ross, Terras, Warwick, & Welsh, 2011)	Campo interdisciplinar de pesquisa e ensino preocupado com a intersecção das disciplinas de computação e humanidades.

Fonte: Dados da pesquisa.

Em Wettlaufer & Westphal (2014), os autores enfatizam que as Humanidades Digitais estão em processo de estabelecimento como disciplina independente que prefere se definir como uma prática ou conjunto de métodos. Essa situação reflete a ausência de definições sobre Humanidades Digitais no *corpus* de pesquisa analisado. Como observado no Quadro 2, seus métodos e aplicabilidade não encontram muitos obstáculos, e o grau de persistência em pesquisas é maior. A única exceção apontada é o próprio artigo em destaque, cujas definições têm aspectos mais teóricos do que práticos. Seus autores explicam que isso ocorre porque ainda não existe uma definição geral aplicável e comumente aceita sobre Humanidades Digitais, mesmo com tentativas crescentes nesta direção.

Os estudos analisados não fazem menção ao Manifesto das Humanidades Digitais enquanto declaração coletiva oficial de acadêmicos e profissionais atuantes em Humanidades Digitais. Porém, em um dos estudos, cita-se a obra “*A Companion to Digital Humanities*” publicada em 2004 por Susan Schreibman, Ray Siemens e John Unsworth, comumente referenciada em pesquisas sobre Humanidades Digitais. Alguns estudos recorrem aos conceitos históricos das Humanidades Digitais propostos por Busa (2004), chamado de pai da primeira aplicação das Humanidades Digitais e da sua predecessora, as Humanidades Computacionais.

Araújo (2017) destaca que o principal diálogo da Ciência da Informação com as Humanidades Digitais compreende discussões sobre preservação e patrimônio cultural na sociedade atual; promoção do acesso universal de forma democrática; e formulação de políticas públicas de desenvolvimento tecnológico. Como os objetos de estudo das Humanidades Digitais e da Ciência da Informação se aproximam de forma orgânica e caracterizam alguns dos novos problemas enfrentados pelas humanidades, é fácil observar a correlação entre os campos. Idealmente, essa aproximação acaba se materializando no universo investigativo da Organização do Conhecimento e a inserção de problemas e debates provenientes do ambiente digital. No entanto, embora o *corpus* de pesquisa congregue temas e questões próprias da Organização do Conhecimento, essa disciplina raramente é mencionada, sendo observado que a sua aproximação com as Humanidades Digitais se concentra nos aspectos culturais e tecnológicos, sem que os autores expressem preocupações com questões de natureza teórica e epistemológica desse movimento de aproximação.

Humanidades Digitais não se limitam ao desenvolvimento tecnológico das humanidades, e ao longo de seu processo de desenvolvimento, entre muitas características, a natureza transdisciplinar do campo é extremamente notável. O surgimento das Humanidades Digitais não só permite e aprimora a compreensão da tecnologia, como possibilita que as pessoas entendam a tecnologia como um meio social, sendo a compreensão do seu impacto, importância e as diferentes possibilidades de uso questões emergentes para os pesquisadores que atuam no campo da informação (Araújo, 2017). Os métodos empregados nos estudos analisados ilustram a extensão desse entendimento, permitindo-nos observar aspectos como a transformação da pesquisa tradicional em humanidades, sua dinâmica, o ambiente transdisciplinar e interativo e a forma como utilizam a tecnologia digital em benefício social.

Conclusões

A era digital oferece novas práticas informacionais e, por decorrência, a construção de debates acerca de suas implicações e potencialidades tornam-se cada vez mais necessários. No presente estudo, buscou-se identificar os estudos em Ciência da Informação que trabalham com as Humanidades Digitais, a fim de apresentar as iniciativas direcionadas à Organização do Conhecimento. Os resultados revelam como os estudos no universo da Organização do Conhecimento estão incorporando as tecnologias digitais acrescidas das humanidades. Em outras palavras, permitem compreender a presença do viés humanístico em estudos pautados nas tecnologias digitais a partir da vinculação das Humanidades Digitais nas pesquisas analisadas.

Humanidades Digitais mostra-se um novo paradigma que afeta temas dialogados na Organização do Conhecimento.

Conclui-se que a maior parte dos estudos que vinculam Organização do Conhecimento e Humanidades Digitais contempla a dimensão aplicada e servem como vitrine para projetos futuros de Humanidades Digitais. As especificidades que acompanham essa aproximação são importantes e devem ser pontualmente observadas, a fim de conhecer as iniciativas e oportunidades provenientes da vinculação do digital e humanístico em Organização do Conhecimento, em que discussões na área envolvendo Humanidades Digitais podem subsidiar o desenvolvimento de ambientes digitais mais acessíveis. Com efeito, novos entendimentos para questões trabalhadas no universo científico da Organização do Conhecimento podem ser amparados pela perspectiva das Humanidades Digitais diante da necessidade de se articular o tecnológico aos fazeres humanos, cuja aproximação revela-se importante para subsidiar perspectivas científicas comuns.

Referências

- Almeida, M. A. de., & Damian, I. P. M. (2015). Humanidades Digitais: um campo praxiológico para mediações e políticas culturais? *Anais, 16 Encontro Nacional de Pesquisa em Ciência da Informação*, João Pessoa, Paraíba, Universidade Federal da Paraíba. Recuperado em 15 janeiro, 2021, de <https://bdpi.usp.br/item/002749636>
- Andrade, L. M. de., & Dal'Evedove, P. R. (2020). Humanidades Digitais na Ciência da Informação brasileira: análise da produção científica. *Revista Ibero-Americana de Ciência da Informação*, 13 (1), 439-451.
- Araújo, C. A. A. (2017). Teorias e tendências contemporâneas da Ciência da Informação. *Informação em Pauta*, 2 (2), 9-34.
- Bardin, L. (2016). *Análise de conteúdo*. São Paulo: Edições 70.
- Dacos, M. (2010). Manifesto das humanidades digitais. Recuperado em 08 dezembro, 2020, de <https://humanidadesdigitais.org/manifesto-das-humanidades-digitais>
- Dahlberg, I. Knowledge Organization: a new science? *Knowledge Organization*, [S.l.], v. 33, n. 1, p. 11-19, 2006. Recuperado em 10 janeiro, 2021, de <https://www.nomos-elibrary.de/10.5771/0943-7444-2006-1-11/knowledge-organization-a-new-science-volume-33-2006-issue-1>
- Deng, J., Wang, R., Song, X., & Zhang, Z. (2000). A conceptual model of chinese oral memory based on digital humanities. *Library Trends*, 69 (1), 223-249.

- Liang, J. H., Wang, H., & LI, X. J. (2020). Task design and assignment of full-text generation on mass chinese historical archives in digital humanities: a crowdsourcing approach. *Aslib Journal of Information Management*, 72 (2), 262-286.
- Li, H., Zhu, L., Shen, W., Du, X., Guan, S., & Deng, J. (2000). Research on Knowledge Organization and visualization of historical events in the Republic of China Era. *Library Trends*, 69 (1), 138-163.
- Martin-Rodilla, P., & Gonzalez-Perez, C. (2019). Metainformation scenarios in Digital Humanities: characterization and conceptual modelling strategies. *Information Systems*, 84, 29-48.
- Moura, M. A. (2019). Ciência da Informação e humanidades digitais: mediações, agência e compartilhamento de saberes. *Perspectivas em Ciência da Informação*, 24 (spe), 57-69.
- Paletta, F. C. (2018). Ciência da Informação e Humanidades Digitais: uma reflexão. *Anais, 19 Encontro Nacional de Pesquisa em Ciência da Informação*. Londrina, Paraná, Universidade Estadual de Londrina. Recuperado em 10 janeiro, 2021, de <http://enancib.marilia.unesp.br/index.php/XIXENANCIB/xixenancib/paper/view/1531>
- Pimenta, R. M. (2020). Por que Humanidades Digitais na Ciência da Informação? Perspectivas progressas e futuras de uma prática transdisciplinar comum. *Informação & Sociedade: Estudos*. 30 (2), 1-20.
- Ross, C., Terras, M., Warwick, C. & Welsh, A. (2011). Enabled backchannel: conference Twitter use by digital humanists. *Journal of Documentation*, 67 (2), 214-237.
- Santos, H. P., & Venâncio, R. P. (2017). Redocumentarização, historiografia e Humanidades Digitais: o caso do arquivo público mineiro. *Anais, 18 Encontro Nacional de Pesquisa em Ciência da Informação*, Marília, São Paulo, Universidade Estadual Paulista. Recuperado em 6 janeiro, 2021, de <http://hdl.handle.net/20.500.11959/brapci/104021>
- Schreibman, S., Siemens, R., & Unsworth, J. (2004). *A companion to digital humanities*. Oxford: Blackwell.
- Siqueira, M. N. de., & Flores, D. (2019). Ciência da informação e Humanidades Digitais: diálogos possíveis de uma relação em desenvolvimento – artigos científicos no Brasil. *Liinc em Revista*. 15 (1), 82-107.
- Sousa, M. C. P. (2015). *As Humanidades Digitais Globais?* anotações. Recuperado em 15 fevereiro, 2021, de <https://humanidadesdigitais.org/hd2015/anotacoes/>
- Tomić M., & Willer M. (2013). Possibilities of extending digital humanities into semantic web environment: the case study of Croatian Medieval manuscripts and incunabula and their fragments. *Knygotyra*, 61, 254-277.
- Wettlaufer, J., & Westphal, S. (2014). Digital Humanities. *Der Archivar*, 67 (3), 270-277.

(Página deixada propositadamente em branco)

EPISTEMOLOGIA SOCIAL E ORGANIZAÇÃO DO CONHECIMENTO: CONTRIBUIÇÕES AS ABORDAGENS CULTURAIS

Wilson Roberto Veronez Júnior¹, Carlos Cândido de Almeida²,
Daniel Martínez-Ávila³, Sonia Maria Troitino Rodriguez⁴

¹Universidade Estadual Paulista, wilson.veronez@unesp.br,
<https://orcid.org/0000-0003-2939-1917>

²Universidade Estadual Paulista, carlos.c.almeida@unesp.br,
<https://orcid.org/0000-0002-8552-1029>

³Universidad Complutense de Madrid, dmartinezavila@gmail.com,
<https://orcid.org/0000-0003-2236-553X>

⁴Universidade Estadual Paulista, sonia.troitino@unesp.br,
<https://orcid.org/0000-0002-7204-3283>

Resumo

Este trabalho apresenta uma reflexão do ponto de vista teórico, conceitual e histórico sobre a vinculação entre Epistemologia Social de Jesse Shera e Margaret Egan e a Organização do Conhecimento. O objetivo da pesquisa foi investigar a relação entre essas duas teorias no campo da Ciência da Informação e demonstrar de que forma elas poderiam contribuir para a abordagens culturais no processo de organização, representação e socialização do conhecimento. Como metodologia, foi utilizada a pesquisa bibliográfica para a coleta, seleção, análise, fichamento, leitura e interpretação dos materiais localizados em base de dados, periódicos e dicionários especializados em Epistemologia, Sociologia, Ciência da Informação e Organização do Conhecimento, eventos científicos e revistas científicas nacionais e internacionais. Os resultados obtidos demonstram que a relação entre Epistemologia Social e Organização do Conhecimento ainda se trata de uma discussão que não foi aprofundada de maneira eficiente por pesquisadores da área, sobretudo a partir da década de 1980. Os elementos que vinculam essas duas teorias são os registros gráficos e a comunicação gráfica, conceitos entendidos como primordiais na socialização do conhecimento, bem como às abordagens culturais na concepção de Shera e Egan.

Palavras-chave: *Epistemologia Social, Organização do Conhecimento, Shera, Jesse; Egan, Margaret; Abordagens Culturais.*

1 Introdução

O campo da Ciência da Informação, segundo Borko (1968), Saracevic (1995) e Le Coadic (1996), é interdisciplinar por natureza, em que esse campo estabelece

diálogos com a Epistemologia, bem como em assuntos relacionados a Sociologia e Organização do Conhecimento. A noção de Epistemologia Social foi apresentada pelos bibliotecários e cientistas da informação estadunidenses, Jesse Hauk Shera (1903-1982) e Margaret Elizabeth Egan (1905-1959) em um famoso artigo denominado *Foundations of a Theory of Bibliography*, elaborado em 1952, na revista científica *The Library Quarterly*. Vieira (2019), destaca que esse projeto disciplinar foi desenvolvido por estudiosos da *Graduate Library School* (GLS), onde na década de 1950 estavam incluídos Shera e Egan, entre outros pesquisadores das áreas de Biblioteconomia, Educação e Sociologia. Por sua vez, Zandonade (2003, 2004), em sua tese de doutorado, defende que esse projeto já havia sido discutido por Shera em trabalhos publicados entre os anos de 1949 e 1950. De acordo com Furner (2004), o termo foi concebido por Egan, porém, o mérito acabou ficando com Shera, visto que foi ele quem deu maior visibilidade e desenvolveu um número maior de estudos sobre esse novo projeto epistemológico, ainda mais depois da morte de sua companheira de trabalho, Margaret Egan. Essa proposta de disciplina científica, desenvolvida na *GLS*, teve por objetivo analisar e compreender a produção e organização da informação e do conhecimento em seu aspecto humanístico e social. Contudo, é preciso lembrar que ambos os autores são provenientes da Escola Sociológica de Chicago, importante instituição que contribuiu no estabelecimento e desenvolvimento de trabalhos com viés filosófico, humanístico, histórico e cultural para a Ciência da Informação (Araújo et al., 2010).

A relação entre Epistemologia Social e Organização do Conhecimento é uma discussão que foi ensaiada por alguns autores, porém, ainda não finalizada no campo da Ciência da Informação. No entanto, com base em Albrechtsen (1997) & Hartel (2018), podemos destacar alguns autores como Rafael Capurro, Birger Hjørland, Bernd Frohmann, Soren Brier, entre outros pensadores que procuraram estabelecer essa relação do ponto de vista histórico, teórico, interdisciplinar e epistemológico. Esses estudos começaram a ter maior visibilidade a partir dos anos 2000.

Com base em autores consagrados como Hjørland (2003, 2008), García Marco (1997), Lima & Alvares (2012), Barité (2012, 2015) e Guimarães (2015), entende-se que a Organização do Conhecimento é um campo científico, epistemológico e interdisciplinar que visa contribuir nos processos de organização, representação e no estabelecimento de mecanismos de acesso à informação e conhecimento produzido em unidades de informação, como Arquivos, Bibliotecas, Centros de Informação e Documentação, Memoriais e Museus. Além disso, procura contribuir na elaboração de Tesouros, Vocabulários Controlados e Sistemas de Organização do Conhecimento, que são elaborados com base em ontologias, taxonomias, folksonomias, conceitos, terminologias e linguagens documentárias/documentais.

2 Epistemologia social

Nesta seção, serão apresentados os principais aspectos conceituais sobre Epistemologia Social e seus principais personagens, ou seja, Jesse Shera e Margaret Egan. A começar por Shera, que foi um importante bibliotecário e cientista da informação norte-americano nascido em Oxford, Ohio, em 8 de dezembro de 1903. Graduou-se

em literatura inglesa em 1925, dois anos mais tarde, ele concluiu o mestrado em inglês na Universidade de Yale. Anos mais tarde, ele concluiu o doutorado em biblioteconomia em 1944 na *Graduate Library School* da Universidade de Chicago. Shera foi um pesquisador que acreditava nos valores morais, éticos e sociais da Biblioteconomia, levando em consideração os principais aspectos do campo: sociológicos, linguísticos, antropológicos, éticos, culturais e científicos, aliado ao impacto das atividades das bibliotecas na sociedade, sobretudo na aquisição e no uso do conhecimento em sua dimensão e aspecto social, isto é, o aspecto técnico da área era algo que deveria ser mais aprofundado do ponto de vista epistemológico, social e humanístico. Nesse sentido, podemos compreender que na concepção de Shera (1972, 1977), a Biblioteconomia não deveria se aprofundar somente nos aspectos técnicos e procedimentais, era necessário que essa área estivesse voltada também para os estudos em que os seus usuários pudessem ter uma maior autonomia e que fossem mais participativos no que se refere à produção, à representação e à organização da informação e do conhecimento. O elo mediador entre a biblioteca e sociedade era uma necessidade maior no processo de produção, organização e representação da informação e do conhecimento.

Para Shera, a maior relevância e contribuição das bibliotecas, independentemente de sua natureza, seja ela pública, particular, universitária, comunitária ou especializada, assim como a preocupação dos bibliotecários e gestores da informação deveria estar atrelada à produção do conhecimento para a sociedade, o principal foco de Shera estava nesse ideal (Araújo et al., 2010).

Além de Shera, outra pioneira que foi mencionada no início deste capítulo, se trata de Margaret Elizabeth Egan. De acordo com Furner (2004), “Margaret Egan (1905-59) chegou a ser professora na instituição *Graduate Library School* da Universidade de Chicago (1946-55) e na *School of Library Science da Western Reserve University* em Cleveland, Ohio (1955-59)”.

Em 1943, Egan juntou-se ao Centro de Relações Industriais da Universidade de Chicago como bibliotecária e ocupou a função de ensino em tempo parcial na GLS. Segundo Furner (2004), Margaret Egan desempenhou um papel muito maior no desenvolvimento do conceito da Epistemologia Social do que normalmente é reconhecido pela comunidade acadêmica e científica. Para Egan, no entendimento de Furner (2004) e Morán (2015), a Epistemologia Social baseava-se no processo de produção, distribuição e utilização de produtos intelectuais em bibliotecas.

A partir da concepção de Morán (2015), compreende-se que a ideia de Epistemologia Social que trabalhou Egan, era voltada para a tradição bibliotecária e as discussões de Bliss e Danton, porém Shera rompeu esse entendimento e desviou essa tradição”. Desse modo, isso só reforça a ideia de que os autores discordavam em determinados aspectos sobre a concepção da Epistemologia Social no campo da Biblioteconomia e Ciência da Informação. No entanto, apesar de todo o seu esforço intelectual para o desenvolvimento da Epistemologia Social, Egan não ficou com o mérito em relação à constituição do termo “Epistemologia Social”, na verdade, Jesse Shera acabou ficando com o bônus. Além disso, ao contrário de Shera, Egan tinha a sua formação no campo da Biblioteconomia e da Ciência da Informação, neste caso ela tinha mais subsídios do ponto de vista teórico e prático adquiridos por meio de sua experiência atuando neste campo científico e aplicado no decorrer de muitos

anos. Aliado a isso, segundo Oddone (2007), outro fator que não deixou com que esse projeto fosse adiante foi a desconfiança da comunidade científica da época em que a proposta foi apresentada.

De acordo com Wilson (2008), Egan e Shera enxergavam a Epistemologia Social como uma teoria baseada no processo de produção, distribuição e utilização dos produtos intelectuais. Shera mais tarde contrastou a Epistemologia Social no que diz respeito ao conhecimento na sociedade, muito mais do que conhecimento no indivíduo. Para Moreno-Jiménez (2008), a Epistemologia Social tem por finalidade analisar, examinar, estudar e investigar o conhecimento a partir de uma perspectiva social. Logo, o aspecto coletivo é fator determinante no entendimento e desenvolvimento dessa teoria do conhecimento com propostas que pudessem estabelecer um diálogo mais profícuo da sociedade com as bibliotecas do ponto de vista social e humano.

Segundo Lamar (2007) e Marteleto (2015), essa nova teoria possui uma natureza teórica e prática, e que ela se vale da interdisciplinaridade com outras áreas do conhecimento, como os estudos linguísticos, comunicacionais e da lógica. Isso é um aspecto importante, no sentido de que ela trouxe uma enorme contribuição no campo da Ciência da Informação, e que na opinião de Martínez-Ávila (2018), o projeto de Epistemologia Social foi um dos mais sérios apresentados enquanto uma teoria social. E também, fez com que a área tivesse a condição e oportunidade de expandir os seus estudos a partir da vinculação teórica e prática com outras perspectivas. Nessa perspectiva interdisciplinar, destacamos que a Epistemologia Social está inserida nos estudos sociais de ciência e tecnologia, e que abrange campos como a Sociologia, Economia Política, Retórica, História e a Filosofia da Ciência, e ainda, os estudos de Ciência-Tecnologia-Sociedade e Análise de Discurso (Lamar, 2007, p. 107). Para corroborar essa interdisciplinaridade, determinadas disciplinas são essenciais para a sua constituição, podemos citar a Filosofia, Antropologia, Epistemologia, Comunicação e a Psicologia (Linares Columbié, 2019). Assim, Fallis (2006), complementa que essa teoria incorpora um grande número de pesquisas de áreas da Ciência da Informação, bem como áreas relacionadas à Bibliometria, à Recuperação da Informação e à Sociologia do Conhecimento. Com relação a essa última área, Shera foi um dos maiores defensores dela, sobretudo a partir das análises feitas por ele em trabalhos do sociólogo húngaro Karl Mannheim. Ainda no aspecto interdisciplinar, Fallis (2006), complementa que na verdade, a Epistemologia Social está interligada a quase todas as áreas de informação científica. Como resultado, uma grande variedade de aplicações da Epistemologia Social tem sido discutida na literatura da Ciência da Informação nos últimos anos, e que teve como fator preponderante o paradigma social amplamente discutido por Rafael Capurro em seus trabalhos.

3 Organização do conhecimento

Sobre a Organização do Conhecimento, se trata de um termo elaborado por bibliotecários, como Henry Bliss, Shiyali Ranganathan, Dagobert Soergel, Julius Kaiser e que foi difundido com mais ênfase a partir da segunda metade do século

XX por pesquisadores, como Birger Hjørland, Ingetraut Dahlberg, Richard Smiraglia, Rosa San-Segundo, Mario Barité, García-Marco, García-Gutierrez, entre outros (Guimarães, 2015 et al, 2017).

No sentido geral, a Organização do Conhecimento, para Hjørland (2003), é tratada em dois aspectos: 1) a organização intelectual do conhecimento também pode ser denominada como a organização cognitiva do conhecimento. Isso é basicamente a organização do conhecimento em sistemas conceituais e teorias. Se usarmos os sistemas periódicos da química ou a taxonomia zoológica da biologia como bases dos sistemas de indexação, usaremos sistemas intelectuais ou cognitivos de organização do conhecimento; 2) a organização social do conhecimento, por outro lado, é basicamente as organizações das profissões, ofícios e disciplinas. Se nos referirmos a disciplinas em nossas representações de conhecimento, estamos usando sistemas sociais de organização do conhecimento.

De acordo com San-Segundo (2012), a evolução histórica da Organização do Conhecimento está ligada à trajetória cultural, política, econômica, tecnológica, de acordo com a evolução das práticas sociais e culturais dos grupos. Na opinião da autora espanhola, a Organização do Conhecimento tem o seu desenvolvimento atrelada a inúmeros fatores. Nessa perspectiva, segundo Barité (2015), podemos analisar a Organização do Conhecimento como uma disciplina que estuda as leis, procedimentos, normas e técnicas específicas que os profissionais da informação, sobretudo os bibliotecários deverão adotar esses procedimentos no processo de organização e representação da informação e do conhecimento.

Entre as tendências de estudo no campo organização do conhecimento, segundo Guimarães (2015 et al), estão os trabalhos sobre sistemas de organização do conhecimento, estudos sobre gênero, além dos de caráter epistemológicos, linguísticos, semióticos, culturais etc. Ao longo dos últimos anos, algumas sociedades científicas foram criadas para contribuir na propagação do conhecimento científico, assim, em 1974, na Alemanha, da *Society for Classification* e, em 1989, da *International Society for Knowledge Organization* (ISKO). A missão dessa última instituição consiste em promover o avanço teórico e aplicado da Organização do Conhecimento em diversos campos e em diferentes maneiras, congregando pesquisadores de distintos países, em áreas como Ciência da Informação, Linguística, Filosofia e Ciência da Computação, entre outros campos interdisciplinares, transdisciplinares, multidisciplinares e pluridisciplinares (Guimarães et al 2015).

No âmbito da Ciência da Informação, a Organização do Conhecimento é um domínio em contínua construção e um espaço relativamente autônomo, como também dialógico e interdisciplinar. Faz interface com outros domínios (áreas do conhecimento) e se preocupa com as questões de natureza teórico-metodológicas para contribuir na sistematização, produção, classificação, organização, disseminação, representação e recuperação da informação nos diferentes contextos científicos (Alves; Oliveira, 2016). Além disso, é preciso ressaltar que Hjørland (2003) destaca que a Organização do Conhecimento é um campo científico amplo, e conforme foi aludido, é onde outros elementos são considerados, como: a divisão social do trabalho (disciplinas); instituições sociais (universidades, escolas, instituições de ensino, organizações); linguagem e os sistemas simbólicos; sistema conceitual e teorias, e os gêneros literários.

Para Barité (2015), podemos analisar a Organização do Conhecimento como uma disciplina, pois estuda as leis, procedimentos, normas e técnicas específicas que os profissionais da informação, sobretudo aqueles procedimentos que arquivistas e bibliotecários deverão adotar no processo de organização e representação da informação e do conhecimento. Ainda, com base no autor supracitado, o objeto de estudo da Organização do Conhecimento é o conhecimento socializado e registrado em um suporte, sendo ele físico, digital ou híbrido. Sendo que este suporte deverá ser conservado e preservado para mediar o acesso imediato e futuro pelo usuário. Além disso, no entendimento de Friedman e Thellefsen (2011, p. 645), a “Organização do Conhecimento é entendida como uma subárea dentro da Biblioteconomia e Ciência da Informação e que se trata de uma comunidade que se concentra em representação bibliográfica”. Assim, Ohly (2015), argumenta que Dahlberg analisa a Organização do Conhecimento como uma subdisciplina da Ciência da Ciência, em um metanível crítico para outras ciências, porém ela destaca a orientação de uma investigação mais descritiva, como a cientometria e menos sua abordagem tradicional. Segundo Ohly (2012), existem muitas disciplinas que são parcialmente aplicadas na Organização do Conhecimento ou que estão operando em estreita conexão com ela, como é o caso da Linguística, Ciências Cognitivas e a Filosofia. Para o autor, a Organização do Conhecimento é mais voltada para rotular, organizar e recuperar informação, ou seja, conhecimento em documentos registrados e materializados em uma instituição social, como é o caso da biblioteca.

Segundo o autor Garcia Marco (1997), a Organização do Conhecimento é um campo interdisciplinar, isto é, podemos associar essa semelhança à natureza da Ciência da Informação, pois ela recebe influência de outras áreas do conhecimento, principalmente as que se ocupam dos modelos computacionais, como a Ciência da Computação, Linguística, Lógica e a Matemática teórica e aplicada. Nesse seguimento, no entendimento de Lima e Alvares (2012, p. 27), com relação ao seu caráter interdisciplinar, a Organização do Conhecimento é estudada em outros campos do conhecimento, como a antropologia, computação, filosofia, linguística, psicologia, sociologia, entre outras que colaboram na sistematização e na elaboração de métodos, técnicas e ferramentas específicas para a organização e representação do conhecimento registrado em unidades de informação, como os arquivos, bibliotecas e museus. Desse modo, segundo Barité (2012), uma das principais características da Organização do Conhecimento como domínio é a sua interdisciplinaridade, sem prejuízo de se aceitar que sua base seja principalmente em Biblioteconomia, Documentação e Ciência da Informação. Melhor dizendo, ela extrapola outros campos de conhecimento e aplicação, e que de alguma forma acaba se confluindo com a epistemologia da Ciência da Informação. Nessa perspectiva, conforme aponta Barité (2015, p. 120), a Organização do Conhecimento é um campo científico e epistemológico que recebe aportes teóricos e práticos de áreas como a Informática, Linguística, Terminologia e da Ciência da Informação.

Em relação às dimensões, para Guimarães (2017), a Organização do Conhecimento se baseia em três, sendo a aplicada, a cultural e a epistemológica: a primeira estuda os modelos, formatos, instrumentos e estruturas na organização do conhecimento. A segunda tem lugar nas questões sociais, políticas, étnicas, linguísticas, educativas e contextuais da organização do conhecimento. A terceira tem as bases conceituais, históricas e metodológicas da organização do conhecimento.

Em síntese, a Organização do Conhecimento ao longo do tempo vem contribuindo para a elaboração de linguagens documentais, sistemas de informação do Conhecimento e colabora com a Ciência da Informação na recuperação da informação.

4 Metodologia

Esta investigação de caráter teórico foi desenvolvida por meio de revisão de literatura dos conceitos que norteiam os trabalhos mais relevantes de Shera e Egan, como Epistemologia Social, Registro Gráficos, Comunicação Gráfica, Ciência da Informação, Classificação Bibliográfica, Documentos, Documentação e Organização e Representação do Conhecimento (Egan & Shera, 1952).

Para o desenvolvimento desta proposta foram consultados trabalhos que relacionavam os principais aspectos sobre Epistemologia Social e Organização do Conhecimento, no primeiro caso, trabalhos desenvolvidos por Jesse Shera, Margaret Egan, Steve Fuller, John Budd, Alvin Goldman, Linares Columbié, entre outros comentaristas que trabalham o tema no escopo da Biblioteconomia, Ciência da Informação e Documentação. No segundo caso, foi realizado um levantamento bibliográfico sobre as principais produções científicas de Birger Hjørland, Ingetraut Dahlberg, Mario Barité, Rosa San-Segundo, José Guimarães, Richard Smiraglia, entre outros teóricos da Organização do Conhecimento em âmbito nacional e internacional que trabalham como o aspecto epistemológico e interdisciplinar desse campo científico. Foram consultadas bases de dados, como a *Scopus*, *Web of Science*, *Lisa*, *Brapi* e *SciELO*, revistas, periódicos e dicionários especializados em Ciência da Informação, Organização do Conhecimento e Epistemologia da Ciência da Informação, além de eventos científicos, como o Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação (ENANCIB), a *International Society for Knowledge Organization* (ISKO), *International Society for Knowledge Organization Brasil* (ISKO-Brasil), bem como outros capítulos, como a ISKO-Ibérico (Portugal e Espanha) e NASKO (Canadá e Estados Unidos. Não houve a delimitação de um período em específico para a coleta, análise, seleção e recuperação dos materiais. Os termos pesquisados foram: Epistemologia Social, Organização do Conhecimento, Epistemologia da Ciência da Informação, Registros Gráficos, Comunicação Gráfica, Jesse Shera, Margaret Egan e Escola Sociológica de Chicago.

Para a escrita do trabalho, deu-se preferência para as citações indiretas dos trabalhos e argumentos recuperados nas bases de dados, o que contribuiu na formulação dos discursos ao longo do texto.

5 Análise dos resultados

Com base nos materiais recuperados e analisados, sustenta-se do ponto de vista teórico que a Organização do Conhecimento, na concepção de Albrechtsen (1997), podemos perceber a relação com a disciplina Epistemologia Social, sobretudo graças à conexão com disciplinas centrais das ciências humanas e sociais, como antropologia, sociologia, linguística, comunicação, entre outras.

Outra disciplina interessante que de certa forma busca realizar uma aproximação entre Epistemologia Social e Organização do Conhecimento, mas que foi pouco aprofundada por Shera, é a Sociologia do Conhecimento, que para Outhwaite & Bottomore (1996, p. 745), é uma disciplina que “investiga as interligações entre a categoria de pensamento, reivindicações do conhecimento e realidade social”. Para Japiassu (1986, p. 35), no entendimento de Karl Marx, Émile Durkheim, Max Weber e Karl Mannheim, o conhecimento não é considerado como construção autônoma e individual, mas como atividades sociais, inseridas num contexto sociocultural. Isto é, o conhecimento só tem efeito a partir da relação com a sociedade, visto que ele é produzido por ela.

Nessa perspectiva, no que entendem Bortoli & Galon (2015), o foco da Sociologia do Conhecimento também está voltado para a compreensão dos modos de pensamento, sendo eles compreendidos a partir das suas origens sociais. Sobre esse pensamento, podemos compreender que a essência da vinculação entre a Epistemologia Social e a Sociologia do Conhecimento provêm das bases e das origens sociais de uma determinada comunidade. Por mais que a teoria proposta por Shera e Egan não apresentava uma tendência mais profunda quanto a Sociologia do Conhecimento, podemos apontar que ambas estão relacionadas por um aspecto primordial, ou seja, a análise do conhecimento em sua dimensão social, cultural e epistemológica, e que ambas poderão fornecer uma base teórica e metodológica para os estudos sociais e culturais na Ciência da Informação. Somado a isso, na década de 1950, os estudos sobre Ciência da Informação estavam associados aos elementos físicos, ou seja, a elaboração de mecanismos e sistemas de recuperação da informação, logo, o aspecto social era pouco valorizado naquele contexto. Já na década de 1980, são elaborados os primeiros trabalhos sobre a necessidade de se analisar a cognição dos sujeitos informacionais, esse movimento ficou conhecido como paradigma cognitivo. Nos anos 2000, tem-se as produções voltadas para o que pode ser denominado de paradigma social baseado em uma abordagem sociocultural.

Retomando a discussão sobre a vinculação entre Epistemologia Social e Organização do Conhecimento, ela tem conexão com o que pode ser denominado de Responsabilidades Sociais, que na concepção de Garcia, Targino & Dantas (2012), supõe os conceitos de Biblioteca, Bibliotecário e o campo da Ciência da Informação. Esses três conceitos foram discutidos de maneira mais aprofundada por Shera (1968), em que ele entende que a Biblioteca como instituição política e social desempenha um papel fundamental na produção, organização e disseminação do conhecimento para a sociedade.

Sobre o papel do bibliotecário, Shera (1972) destaca que esse profissional é o supremo “ligador do tempo”, visto que a sua responsabilidade é mediar os registros gráficos (conhecimento registrado em documentos) junto a sociedade. O autor entende que o bibliotecário não deve se ater apenas às questões técnicas realizadas nas bibliotecas, mas deve estar atento aos aspectos sociais, culturais e políticos que ocorrem na sociedade.

No que tange a Ciência da Informação, Shera (1968) entende que se trata de uma ciência social e que ela havia sido definida assim pelos russos, e que também tem um caráter interdisciplinar, não só no aspecto teórico, mas prático, oriunda de processos e práticas informacionais realizadas em bibliotecas, como a recuperação

da informação e documentação. O autor defende que a Ciência da Informação não está estritamente relacionada ao aspecto físico dos processos que estabelecem interface com os sujeitos informacionais, mas que esse campo tem muito a contribuir com o avanço da sociedade, principalmente no que concerne à socialização do conhecimento.

Partindo de um pressuposto de que o conhecimento é um fenômeno construído por relações sociais e científicas, ele deve ser organizado e sistematizado para com que a sociedade possa acessá-lo por meio de sistemas de organização do conhecimento, como tesouros, vocabulários controlados, entre outros. Nesse sentido, inúmeros são os aspectos, que deverão ser levados em consideração pela instituição produtora (arquivos, bibliotecas e museus, assim como os seus agentes culturais (arquivistas, bibliotecários e museólogos), para a organização, representação e elaboração desses sistemas, tais aspectos podem ser considerados: relação social de uma determinada comunidade, a cultura, a linguagem e a forma de representação e os mecanismos de acesso ao conhecimento que essa comunidade dispõe, seja manual ou automatizado. Em virtude da miscigenação cultural e étnica entre os povos, é uma tarefa difícil analisar de maneira geral como esses povos se relacionam culturalmente. No campo da Ciência da Informação, é comum analisarmos as diferentes culturas, bem como os seus agentes sociais a partir dos estudos de usuários. Essa é uma das principais maneiras de se compreender as relações sociais e culturais entre as comunidades.

Conclusões

Os resultados obtidos até então, apontam que a relação entre Epistemologia Social e Organização do Conhecimento ainda se trata de uma discussão que não foi aprofundada tanto no aspecto teórico, epistemológico e prático no decorrer do século XX, sobretudo no período de 1980 (época do falecimento de Shera) e início do século XXI (época em que houve a retomada dos estudos de caráter social e cultural no campo da Ciência da Informação por meio do paradigma social).

Os elementos que vinculam essas duas teorias são os registros gráficos e a comunicação gráfica, conceitos entendidos como primordiais nas abordagens culturais no contexto da Ciência da Informação. Esses conceitos foram apresentados por Shera e Egan no artigo que deu fundamento e início a discussão sobre epistemologia social. Conforme defendia Shera, outro elemento que contribui com os estudos em Epistemologia Social é a Sociologia do Conhecimento, especificamente de Karl Mannheim.

No decorrer das décadas de 1950 a 1980 não foram discutidos aspectos relacionados a epistemologia social e organização do conhecimento de forma mais específica. No entanto, de forma clara e objetiva, por registro gráfico e comunicação gráfica, é possível compreender que estamos analisando um processo que procura esclarecer como se dá a produção, organização e socialização do conhecimento na sociedade como um todo. Outro elemento que procura inserir os estudos em Epistemologia Social e Ciência da Informação, é a Sociologia do Conhecimento, mas não houve avanços nesse sentido, e que de certa forma poderia contribuir para uma vinculação mais substancial entre Epistemologia Social e Organização do Conhecimento.

Além da contribuição no entendimento do processo de produção, organização e representação da informação e do conhecimento em unidades de informação, a relação entre Epistemologia Social e Organização do Conhecimento pode estabelecer um diálogo profícuo no que tange aos aspectos culturais no campo da Ciência da Informação, principalmente a partir da ideia de organização social do conhecimento.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento: 88887.609892/2021-00

Grupos de Pesquisa:

- a) Fundamentos Teóricos da Informação (GPFTI)
- b) Acervos, Dimensões do Documentos, Memória e Patrimônio (GPADDMP)

Referências

- Albrechtsen, H (1997). Knowledge organization in the humanities. *Knowledge Organization* 24:2, 61-63.
- Alves, B. H., & de Oliveira, E. F. T. (2016). O desenvolvimento do domínio da "Organização do Conhecimento" no contexto da Ciência da Informação a partir da ISKO-Brasil. *Brazilian Journal of Information Science*, 10(2), 103-108.
- Araújo, C. A. Á., de Souza Lage, D. F., Souza, R. M. F., & Assis, R. A. (2010). A contribuição de JH Shera para a Ciência da Informação no Brasil JH Shera's contribution to the information science in Brazil. *Revista ACB*, 15(2), 71-89.
- Barité, M. G. R (2015). **Diccionario de Organización del Conocimiento: Clasificación, Indización y Terminología** / Mario Barité et al. 6.a ed. corregida y aumentada. Montevideo: csic.
- Borko, H. (1968). Information Science: what is it?. *American documentation*, 19(1), 3-5.
- Bortoli, L. V; Gallon, S (2015). A repercussão da Sociologia do Conhecimento de Karl Mannheim no Brasil: uma análise da presença do autor no país e nos estudos de administração. IBEPES Curitiba – PR, Brasil. **RECADM**, 2015, 14:3 (set/dez) 166-181.
- Egan, M. E.; Shera, J. H (1952). Foundations of a theory of bibliography. **Library Quarterly**, 22:1 (apr. 1952) 125-137.
- Fallis, D (2002). Introduction: Social Epistemology and Information Science. **Social Epistemology**, 16:1, 1-4, 2002.
- Furner, J (2004). "A Brilliant Mind": Margaret Egan and Social Epistemology. **LIBRARY TRENDS**, 52:4, Spring 2004, 792–809.
- Garcia Marco, F. J. G. (1997). Avances en organización del conocimiento en España: los II encuentros sobre organización del conocimiento en los sistemas de información y documentación. In *Organización del conocimiento en sistemas de información y documentación: actas del II Encuentro de ISKO-España*. Getafe, 16 y 17 de noviembre de 1995 (pp. 7-12). Universidad de Zaragoza.
- Garcia, J. C. R; Targino, M. das G; Dantas, E. R. F (2012). Conceito de responsabilidade social da Ciência da Informação. **Informação & Informação** 17:1 1-25.

- Guimarães, J. A. C.; Matos, D. F. de O.; Santos, A. Y.; Sales, R. de (2015). A dimensão conceitual da organização do conhecimento no universo científico da ISKO: uma análise de domínio a partir dos congressos... // *Scire*. 21:2 (jul.-dic. 2015) 13-26.
- Guimarães, J. A. C (2017). Organização do conhecimento: passado, presente e futuro sob a perspectiva da ISKO (2017). *Informação & Informação* 22:2, 84-98.
- Hartel, J (2018). Social Epistemology as Theoretical Foundation for Information Science: Supporting a Cultural Turn, Copenhagen. *KNOWLEDGE ORGANIZATION*, 45:1 (August-2018) 79-84
- Hjørland, B (2003). Fundamentals of knowledge organization. *Knowledge organization*, 30:2, 87-111.
- Japiassu, H. (1977). **Introdução ao pensamento epistemológico**. Livraria F. Alves Editora.
- Lama, A. R. (2007). Epistemologia social: possível origem e alguns momentos de seu percurso. *Pro-posições*, 18(1), 103-113.
- Le Coadic, Y. F. (1996). **A ciência da informação**. Briquet de Lemos.
- Marteletto, R. M. (2015). Epistemologia social e cultura digital: reflexões em torno das formas de escritas na web. *Em Questão*, 21(3), 9-25.
- Martínez-Ávila, D. (2018). Hacia una base teórica social de la ciencia de la información. *Anuario Think EPI*, 12, 83-89.
- Moreno, P. M. (2008). *Epistemología social y estudios de la información* (No. 4). El Colegio de México AC.
- Oddone, N. (2007). Revisitando a” epistemologia social”: esboço de uma ecologia sociotécnica do trabalho intelectual. *Ciência da Informação*, 36, 108-123.
- Outhwaite, W.; Bottomore, T (1996). (Org). Dicionário do pensamento social do século XX. Rio de Janeiro: Jorge Zahar.
- San Segundo, R (2012). **Visibilidad científica y diálogos internacionales en Organización del conocimiento**. Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade. [recurso eletrônico] José Augusto Chaves Guimarães, Vera Dodebei (organizadores). – Marília: ISKO-Brasil: FUNDEPE, 285.
- Saracevic, T. (1995). Interdisciplinary nature of information science. *Ciência da informação*, 24(1), 36-41.
- Shera, J. H (1968). Sobre bibliotecología, documentación y ciencia de la información. *Boletín de la Unesco para las bibliotecas*, 1968, 22:2 62-70.
- Vieira, K. R. (2019). *As contribuições da Escola de Chicago para a Ciência da Informação. 2019a. 157 f* (Doctoral dissertation, Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de Santa Catarina, Florianópolis).
- Wilson, T. D. (2008). A dimensão epistemológica da ciência da informação e seu impacto sobre o ensino em arquivologia e biblioteconomia. *Brazilian Journal of Information Science*, 2(1), 1-15.
- Zandonade, T (2003). **As implicações da Epistemologia Social para uma teoria da recuperação da informação**. 184 fls. Tese (Doutorado em Ciência da Informação). Faculdade de Economia, Administração, Contabilidade e Ciência da Informação e Documentação Departamento de Ciência da Informação e Documentação Programa de Pós-Graduação em Ciência da Informação. Brasília: Universidade de Brasília.
- Zandonade, T (2004). Social Epistemology from Jesse Shera to Steve Fuller. *Library Trends*, 2004, 52:4 810-832.

(Página deixada propositadamente em branco)

A TEORIA DO CONCEITO EM HEGEL APLICADA À ORGANIZAÇÃO DO CONHECIMENTO: UM ESTUDO DIALÓGICO E DIALÉTICO

Marco André Feldman Schneider¹, Marco Antonio de Carvalho Bonetti²,
Gustavo Silva Saldanha³, Fernanda Valle⁴, Diogo Xavier⁵

¹IBICT- UFF, art68schneider@gmail.com, ORCID 0000-0001-5053-9491

²UFJF, marco_bonetti@id.uff.br, ORCID 0000-0002-9130-0054

³IBICT – UNIRIO, *gustavosalदानha@ibict.br*, ORCID 0000-0002-7679-8552

⁴PPGCI IBICT UFRJ, fvallegalvao@gmail.com, ORCID 0000-0002-4156-027X

⁵PPGCI IBICT UFRJ, diogo.biblio@gmail.com, ORCID 0000-0002-2745-7196

Resumo

Investigou-se como as diferentes abordagens do conceito do próprio termo “conceito”, conhecimento, universalidade e totalidade expostas na obra de Georg Hegel contribuem para aspectos teórico-epistemológicos da Organização do Conhecimento, dada a relevância da noção para o desenvolvimento epistemológico do domínio. De base teórica e do tipo exploratório, o trabalho inseriu a filosofia hegeliana no debate sobre as teorias da classificação, e os desafios do contemporâneo sobre os usos e limites dos artefatos infocomunicacionais. A partir dos métodos dialético e biobibliográfico, dialogou-se com Hope Olson, Rodrigo de Sales e Shachar Freddy Kislev acerca da possível influência do pensamento de Hegel nos modelos de classificação bibliográfica de William Harris e Melvil Dewey. Foram propostos desdobramentos dialéticos entre os estudos hegelianos e a Teoria da não-conceitualidade, presente em Hans Blumenberg.

Palavras-chave: Conceito, Verdade, Organização do Conhecimento, Sistemas de Organização do Conhecimento, Georg W. F. Hegel.

Introdução: Seria Hegel uma das fundações teóricas da OC?

Immanuel Kant e Georg Hegel são como dois continentes epistemológicos capazes de influenciar toda tradição filosófica posterior. Por isso, é relevante averiguar as visões de mundo que ajudaram a formar e consolidar os modelos, teorias e métodos legítimos em Biblioteconomia e Ciência da Informação (BCI). Centralizamos o debate em Georg Hegel (1770-1831), autor ainda pouco explorado na história do campo em contraste a outros, conforme análise cartográfico-epistemológica de Joseph Nitecki na década de 1990. Pelo mesmo motivo, enfatizamos a reflexão sobre Sistemas de Organização do Conhecimento (SOCs), em especial linguagens documentárias.

A pesquisa se justifica ao observar que, desde os primeiros movimentos bibliográficos e enciclopédicos aos atuais SOCs identifica-se o conjunto de problemas do singular, particular, universal, presente na filosofia de Hegel e na historiografia da BCI. Se observadas as construções de sistemas bibliográficos até o século XIX, e a era das grandes sistematizações totalizadoras universais, como em Dewey e Bruxelas, podemos deslindar um debate existente do hegelianismo na esfera da Organização do Conhecimento (OC): por um lado, Olson (2011) e Ferreira e Sales (2019) examinam as afirmações de Eugene Graziano que apontam para a influência de Hegel na construção da Classificação Decimal de Dewey (CDD). De outro, Kislev (2021) contesta os autores, ancorado na própria lógica hegeliana.

Sem pretensão de esgotar a questão, nossa proposta é ampliar a discussão dada à complexidade do pensamento de Hegel e os limites de operacionalização dos conceitos. Para os fins da pesquisa, consideramos ainda a Agenda 2030, da Organização das Nações Unidas (ONU), uma herança iluminista que suscita a ideia de progresso e harmonia social mediante o bom uso da razão pública, como o dilema histórico vivido por Hegel. Narrativa progressista igualmente presente em Paul Otlet, cada qual com suas especificidades teóricas e sociais, e sua interpretação sobre os livros serem instrumentos determinantes do processo civilizatório e, conseqüentemente, da paz mundial. Nesse horizonte, a visão continuísta da História sobre o protagonismo dos artefatos infocomunicacionais na transformação social pretendida e a intencionalidade de Dewey quanto à ordenação bibliográfica torna frutífera a inserção da dialética hegeliana na OC de modo mais profundo.

Vale ressaltar que Hegel estava dedicado à construção ontológica da chamada consciência e sua busca atravessa os conceitos de razão, conceito, verdade e conhecimento. Já Dewey e os demais autores da BCI operam no âmbito da representação dos saberes. Contudo, nosso posicionamento é orientado ao entendimento de que toda e qualquer representação parte de pressupostos filosóficos, ainda que nem sempre enunciados, e instituem existência social aos seres, com impacto direto na cotidianidade dos sujeitos e saberes representados e documentados. Dito isto, a BCI se dedicará, também, a estabelecer métodos e sistemas que abarquem a totalidade dos objetos infocomunicacionais e tornem universais os assuntos. Aplicado à linguagem – matéria-prima de qualquer representação –, à expressão do real pelo universo dos símbolos –, o modelo conceitual hegeliano concebe uma estrutura ontológica para os SOCs – diferentemente da posição de Dahlberg perante o silogismo aristotélico.

Para os fins do artigo, elegemos a Fenomenologia do Espírito, das obras mais relevantes de Hegel (2014), onde o autor explicita o percurso da experiência rumo ao saber universal. Além de Sales, Olson e Kislev, anteriormente citados, como percurso teórico dialógico-complementar, consultamos Charles Taylor, relevante comentador de Hegel, e Solange Mostafa (1985,1986), autoridade brasileira reconhecida nos estudos epistemológicos do campo, em seu diálogo com o filósofo alemão.

Conhecimento e verdade como fruto da luta entre consciências

Em 1997, Nitecki analisou a ancestralidade filosófica da BCI para compreender quais escolas filosóficas mais influenciaram direta e indiretamente a epistemologia do campo e indicou Hegel como um dos filósofos influentes nas classificações bibliográficas de Harris e Dewey. Similarmente, Solange Mostafa (1985,1986) discutiu a história da teoria do conhecimento de Platão a Marx, passando por Kant e Hegel. Em sua perspectiva, o pensamento kantiano, também inspirado no sistema categorial aristotélico, fez surgir determinadas correntes adotadas pelo campo da BCI que parecem replicar a ideia de um sujeito contemplativo sem ação no mundo concreto – indo de encontro ao pensamento hegeliano.

Para verificar as possíveis influências hegelianas nos referidos sistemas bibliográficos, Rodrigo de Sales (2017) avalia especificamente a categoria temática sobre Artes. Em linhas gerais, William Harris foi estudioso da obra hegeliana e publicou o *Book Classification*, de onde Dewey teria se inspirado para elaborar seu próprio esquema decimal. Conforme Sales (2017, pp. 203), o autor teria herdado de Hegel “o foco centrado na Ideia, na razão modelada pela Ciência, o que tornou sua classificação logicamente idealista”, concluindo, portanto, que o universalismo intentado pelo classificacionista pretendia representar o mundo de forma racional em direção “à totalidade da inteligência humana materializada nos livros”. Sales cita Eugene Graziano, também presente na revisão de Nitecki. É em Graziano que indicam ser o início do discurso acerca da fundamentação hegeliana em ambos os sistemas de classificação, e não somente Francis Bacon.

Ferreira e Sales (2019) explicam que, no sistema baconiano, Artes consta como subclasse de História; Arquitetura, da Matemática; Pintura, da Filosofia, revelando relação direta às partes do entendimento, categorias filosóficas: na concepção de Bacon, história–memória, poesia–imaginação, filosofia–razão. Dentre essas, a poesia – oriunda de *poiesis*, criação – estaria relacionada à liberdade e à retórica, ocupando uma posição de mediação entre a história e a filosofia (Ferreira & Sales, 2019). Apesar de Hegel não enunciar um sistema classificatório formal sobre Belas Artes, os autores consultaram seus discursos sobre estética, organizados por alunos do filósofo alemão, para compará-los às etapas do entendimento de seu sistema lógico, tal qual fizeram com Bacon.

Na tabela apresentada, elencam a seguinte ordem: arquitetura, escultura, pintura, música e poesia. A natureza da arquitetura seria a própria natureza, de onde advém as formas e os materiais para representação das formas. A escultura, em segundo lugar, possuiria uma individualidade espiritual, um “aspecto corporal imanente do espírito” (Ferreira & Sales, 2019). Pintura e música se ocupariam da interioridade subjetiva, ou seja, do absoluto voltado ao exterior, permitindo contato e sensações. Posteriormente, a poesia – possível por meio da palavra – seria a arte absoluta, o espírito manifestando-se como espírito (Ferreira; Sales, 2019, pp. 116). Diferente de Bacon, a poesia em Hegel não é intermediária, mas categoria final e, mais importante, nas palavras dos autores, elevada – o que dimensiona um tipo de transcendentalidade da arte dentro da perspectiva da totalidade em Hegel.

No *Book Classification*, Harris inverte a ordem baconiana, alocando a Ciência como primeiro estágio e substitui a palavra Poesia por arte (Ciência – Arte – Filosofia), mas mantém a ordem observada em Hegel, começando na Arquitetura e terminando na Poesia. Quando os autores, apoiados em Graziano, apresentam a tabela comparativa entre Bacon, Hegel, Harris e Dewey, percebe-se que também Dewey seguiu a estrutura hegeliana (Arquitetura na classe 700 e Literatura, 800). Os autores concluem que “o pensamento idealista de Hegel focado no ‘como’ e no ‘processo’ talvez tenha encontrado na interpretação de Harris o diálogo para a concepção Baconiana de ‘forma’ e ‘forma de conhecimento’” (Ferreira & Sales, 2019, pp;123); ou seja, a definição das classes seria de ordem baconiana e a ordenação, hegeliana, equiparando a classe Arte em Harris à Essência, de Hegel (Ferreira & Sales, 2019, pp. 125)

Olson (2011) não focaliza uma classe temática, mas parte dos conceitos de hierarquia e ordem para refletir sobre a lógica hegeliana e a influência não enunciada por Harris e Dewey. Segundo Olson, Dewey interpretava que a classificação era “necessária para suplantar o caos” e que, por meio de um cérebro inteligente, as palavras seriam ordenadas e adquiririam sentido. Na metáfora de Dewey, uma multidão desordenada é diferente de um exército disciplinado (Olson, 2011). Nesse contexto, Olson argumenta que o discurso sobre hierarquia relacionava os conceitos de lógica e razão – princípio este que será a base da digitalização do real representado tanto nas classificações universais do século XIX como no mecanicismo eletrônico computacional e computacional em rede do século XX. Nesse fio, surge a lógica hegeliana (sob um ponto de vista aplicado, uma operacionalização da estrutura conceitual da lógica em sistemas empíricos).

De acordo com quadro adaptado por Olson (2011), a inversão baconiana em Harris é mantida (Ciência – Arte – História). Já em Hegel, a ordem é retratada como Ideia-Ideia imperfeita – Essência (*Wesen*) – Ser (*Sein*) e, por fim, na CDD: Filosofia e Psicologia – Religião – Ciências Sociais – Linguagem – Ciências Naturais e Matemática – Tecnologia – Artes – Literatura e Retórica – Geografia e História. Assim, na lógica de Bacon, o conhecimento seria adquirido por meio da “História, como uma emanção da Memória, da Poesia como uma emanção da Imaginação, e da Filosofia como uma emanção da Razão. Bacon, portanto, via o conhecimento como o fim definitivo deste desenvolvimento. Basicamente, um alicerce epistemológico para a classificação” (Olson, 2011, pp. 7).

Na sequência, destacamos o que Olson retrata da interpretação de Harris sobre Hegel e a trilogia Essência, Ser e Ideia: a) Ser como condição ilusória, “a pessoa percebe a realidade somente como se fosse o que aparenta ser”, percepção imediata, prescindindo de mediação; b) a Essência como dependência de diferentes forças sob a ilusão do ser independente e opera por mediação; c) Ideia como categoria superior de entendimento. Nesse sentido, Olson (2011) assevera:

[...] as belas artes e as artes mecânicas e a Essência como uma compreensão de ilusões; e as ciências e a religião e a filosofia e a Ideia como um estado que está além e é independente, tanto do ser material quanto da ilusão. De forma interessante, Harris e outros classificacionistas colocaram o estado mais elevado no início de sua classificação, em vez de no final. Embora descreva

que a progressão ontológica de Hegel vá de Ser para Essência para Ideia, Harris reverte esta sequência em sua classificação, colocando o ápice da compreensão humana em uma posição de primazia no início (assim como as classificações soviéticas traziam o Marxismo-Leninismo em primeiro lugar) (Olson, 2011, pp. 8)

Complementa com o fato de que Dewey utilizou os discursos dos especialistas, baseado em garantia literária, para definir o conteúdo de cada classe temática. A hierarquia é criada por ordem de citação, e a ordem de citação encontra seu caminho de volta à lógica” (OLSON, 2011, pp. 8). Em relação a Harris, aponta que a lógica hegeliana pressupõe uma hierarquia, de finalidade quantitativa – individual, particular e universal, “uma ordem crescente de generalidades”, onde cada parte interage entre si, em relação hierárquica de todo / parte. À etapa seguinte, relata que Harris denominou classificação o que Hegel chamou de divisão, a relação exaustiva das “determinações do universal” (Olson, 2011, pp. 10). A hierarquia determina as generalidades do que existe e às partes, para compor as classes gerais, é exigido que tenham características generalizáveis representativas. “A ideia de que era uma escolha binária é um dos problemas que continuamos a enfrentar. Um exército representa um tipo particular de organização hierárquica, dividida, restrita” (Olson, 2011, pp. 12-13).

Embora Olson concorde com a presença hegeliana nos sistemas classificatórios, suas críticas sobre limitação, exclusão e preconceito servem à justificação contrária: a de que a CDD não é necessariamente hegeliana, mas a análise de Graziano refletiu um olhar sobre o sistema lógico de Hegel. Existem outros. A confirmação está em Kislev (2021). Para o autor, a suposta influência de Hegel é frágil e sua leitura representa “uma imagem reacionária do conhecimento”, incoerente com o “sistema sofisticado e método detalhado de organização do conhecimento” hegeliano. Ao inverter o raciocínio: “se a filosofia de Hegel resultasse em um esquema de classificação, a CDD seria o modelo?”, para Kislev (2021), a resposta seria não. Em outros termos, a operacionalização potencial – do “sistema total filosófico hegeliano” para o “sistema universal bibliográfico deweyano” – representaria uma redução da teoria crítica presente em Hegel, aberta a um caminho de inversão do propósito da totalidade dialética.

Hegel não propôs um sistema de conhecimento disciplinar, apesar de seu sistema filosófico ter um caráter especulativo de universalidade, tal qual Aristóteles ou Bacon, mas sim, um sistema de percurso do conhecimento em si. Sua intenção não era criar um sistema novo, mas dialeticamente propor um sistema que organizasse todos os sistemas filosóficos. Kislev equipara Hegel a um bibliotecário em busca do melhor método de organizar conhecimentos variados.

Na Fenomenologia do Espírito, Hegel questiona a maneira da epistemologia kantiana, quando esta defende que para conhecer algo estamos condicionados a um instrumento presente de modo *a priori* na mente, suas faculdades e seu poder de síntese com objetos. Esta visão, segundo Hegel, coloca o Absoluto fora do âmbito do que é dado ao homem conhecer, por meio da famosa divisão entre o objeto, passível de conhecimento pela razão, e a coisa-em-si, sempre exterior, impossível de ser conhecida sem exorbitar nos limites do que é dado à razão conhecer com verda-

de. Hegel critica a fissura intransponível entre objeto e coisa-em-si, que torna o conhecimento uma coisa sempre aproximativa, nunca necessária e universal. A solução hegeliana contra o ceticismo kantiano é colocar, desde o princípio, a Coisa mesma em análise. Todo saber é saber de uma coisa para uma consciência, numa unidade. Se a consciência percebe que seu saber está em desacordo com a Coisa mesma, a própria consciência percebe seu erro e atualiza seu saber. Reelaborando o saber, a própria coisa já é outra coisa. Essa evolução do saber é necessária. A própria relação entre Coisa, saber e consciência exige seu desenvolvimento.

Conceito, em uma das acepções empregadas por Hegel, é esse saber que se desenvolve em busca de formas mais desenvolvidas, até que a consciência seja capaz de constatar a coincidência entre saber e coisa. Desse modo, um conceito pode estar num estágio desenvolvido ou pouco desenvolvido. Conceito implica uma noção de transformação que é posta pela própria relação Coisa, saber, consciência. Há estágios distintos. Por vezes ele é imediato, outras, mediação. Conforme Taylor (2011), há pelo menos três acepções de conceito em Hegel. O Conceito [maiusculo] 1) é expressão da Ideia ou “a necessidade subjacente que trouxe consigo o que ela [a Ideia] aplica ao ser” (pp. 138). Nesse sentido, conceito é a cópula necessária entre a razão de ser da coisa e sua efetivação; o conceito [minúsculo] pode ser 2) o estado de potência que se opõe ao de efetividade, que é a realização última do ser, ou seja, o que a coisa pode ser, é propensa a ser, mas ainda não é; ou 3) o estado de potência que se opõe à realidade tal qual entendida pelo senso comum, com a realidade aparente, em dado momento de desenvolvimento / atualização / efetivação seja do que for.

Quando uma consciência sabe que está frente a alguma coisa, surge um tipo de objeto, a coisa que a consciência percebe num aqui e agora, e surge um tipo de consciência, a certeza sensível. Essa é sua experiência mais básica. A verdade desse nível de consciência é que ela sabe que é algo e que ela sabe que está frente a algo no espaço e tempo, conforme suas sensações indicam com certeza. Este nível de verdade é muito rasteiro, porque o conceito da coisa na consciência está em seu nível mais baixo, um nível do conceito a partir do qual ainda não conseguimos saber quase nada. Mas a consciência quer mais, o próprio conceito exige seu desenvolvimento a partir do que a própria Coisa mesma faz questão de oferecer como conhecimento a respeito de si. É nessa necessidade de desenvolvimento que a consciência sofre um *Aufhebung*, uma suprassunção, que significa dar um salto para outra figura (Gestalt) da consciência num nível mais elevado. Percebe-se que a força (Hegel chama de violência) que impulsiona a consciência para um patamar mais elevado de consciência é inerente ao processo, é interno, nasce de uma dúvida ou desespero que se apodera da consciência que quer saber mais. Nesse movimento, como em todos os momentos posteriores de *Aufhebung*, a nova figura emergente da consciência ao mesmo tempo nega o momento anterior, mas conserva algo e supera o nível anterior. Como o marceneiro em relação a uma peça bruta de madeira, fazendo dela uma mesa ou cadeira. Como a crítica teórica que depura um argumento, mediante a negação de alguns de seus momentos, a conservação de outros e a proposição de um novo argumento, melhor.

Na Fenomenologia, a consciência passa então da certeza sensível à percepção, uma segunda experiência. Na sequência da obra, temos um processo dialético de

exposição das figuras subsequentes da consciência que é a própria obra, a Fenomenologia do Espírito. O método dialético, portanto, é imanente à própria relação Coisa, saber, consciência que compõem o conceito de ciência. Não é preciso sair dali para a evolução ir se dando patamar por patamar, formando a famosa figura de apresentação da Fenomenologia presente no prefácio, de que a obra é uma espécie de escada para a consciência. Os passos que a consciência vai saltando degrau por degrau é o que Hegel chama de experiência. A epistemologia hegeliana propõe superar esta ordem de subjetivismo e relativismo da verdade e da razão. A ciência para Hegel é o conjunto de saber gerado a partir da experiência (tomada no sentido estreito de Hegel enunciada acima).

Em vez de discutir condições transcendentais para garantir limitações da razão humana para descobrir a verdade a respeito de objetos, a filosofia em Hegel deve ser o lugar que estimula a ciência a seguir seus processos de desenvolvimento dialético a partir das relações Coisa, saber, consciência. Esse ‘nós’ genérico que são todas as auto-consciências escreve a história da ciência na substância de um povo e de sua história da ciência. Superadas as tensões dialéticas que Hegel resume nas figuras do senhor e do escravo, do céptico, do crente (“consciência infeliz”), e do salto no vazio do cepticismo, não mais como vazio, mas como negativo determinante, Hegel chega à figura da razão. Um novo patamar das experiências consciência.

A verdade do conceito é histórica, a totalidade do conhecimento sobre cada Ser, enquanto este Ser se desenvolve Em-si, depois Para-si, até chegar a um desenvolvimento Em-si e Para-si como unidade no conceito; uma espécie de projeto enciclopédico do saber em um saber universal. Só existe uma Verdade, porém ela é complexa, histórica, de modo que só a ciência é capaz de levá-la a suas formas mais desenvolvidas. Com efeito, a razão está no mundo, na consciência como autoconsciência racional. Mas está na caneta, no livro, na história, na política, nas leis. Sua forma é o conceito, que é a expressão da Razão. A Razão não é só o que existe na forma do pensamento na cabeça do homem, mas também é a força suprassensível que faz as coisas serem o que são, a finalidade que está dentro de cada coisa para que seja o que é, a ideia platônica, porém não imobilizada, mas sim com desenvolvimento numa historicidade.

Para exemplificar a instabilidade que deveria existir numa classificação em moldes hegelianos, citamos uma passagem da Lógica. Dentro de uma classificação devem existir forças internas tensionadas, assim como há uma força atômica, molecular, no elemento do hidrogênio. Essa força cria uma espécie de instabilidade interna do hidrogênio, que pode ser estabilizada provisoriamente, assim como uma palavra numa classificação. No dizer de Hegel, “Enquanto cada objeto está, pelo seu conceito em contradição com unilateralidade própria de sua existência, esforçando-se, portanto, para supassumi-la” (HEGEL, 2018, pp. 206). Força que leva à reação química quando a estabilidade se perde: “está imediatamente posto o esforço para supassumir a unilateralidade do outro e, através desse igualmente recíproco e dessa combinação, para tornar a realidade adequada ao conceito que contém ambos os momentos”(idem: *ibidem*).

O que ocorre quando moléculas de hidrogênio encontram oxigênio é uma reação que forma água, como se um exercesse sobre o outro uma violência exterior, mas que dispara uma potência preexistente em cada um deles. Na reação, forma-se um

silogismo dialético, um elemento num termo, outro elemento no outro, um termo médio que será a mediação da reação, no *devoir* inorgânico, posto que dialético, do mundo material. Derrida (1969) mostra que essa reatividade é a marca do signo linguístico de Hegel, pois se trata de um elemento de comunicação. Mais radical, na verdade, porque, como produção intelectual humana, cada signo tem uma potência de criação infinita dentro de si, que se desdobra na historicidade da palavra, da cultura, da história... do Espírito.

Visto que a diferença real pertence aos extremos, assim esse meio-termo é apenas a neutralidade abstrata, a possibilidade real dos mesmos – quase que o elemento teórico da existência dos objetos químicos, do seu processo e do seu resultado – no corpóreo, a água tem a função desse meio; no espiritual, enquanto nele se realiza o análogo de uma tal relação, deve ser visto como meio o signo em geral e, mais precisamente, a linguagem.(Derrida, 1969, pp. 39)

É essa plasticidade histórica da palavra que não existe nas classificações. Dewey e Harris invertem a lógica hegeliana ao iniciar do todo para as partes de forma imóvel, apesar das atualizações e da figura metodológica da hospitalidade classificatória, em forma de edições, que o sistema possa sofrer, a sua condição estática, morosa, a faz parecer um instrumento “morto”, fora da vida dialética e movimentada do real. É certo que, conforme vimos com Olson, a garantia literária oferece aos classificacionistas as partes que serão generalizadas para emergência das classes, em um processo indutivo generalizável. Contudo, a partir do momento em que as classes se consolidam, o processo indutivo não mais existe: as partes devem se adequar ao todo, no caso, as dez classes amplas de Dewey. A classificação se transforma em dedutiva (todo para as partes). Da perspectiva hegeliana, um processo incompleto na aquisição e no entendimento da categoria conhecimento, já que em Hegel as etapas da consciência são fluidas e não fixas, e o conhecimento, a categoria ampla que remete ao conceito de verdade, se consolida apenas ao fim do processo, ao fim da luta entre as consciências ao longo da história humana.

Ao situar o cérebro inteligente como sinônimo de razão, o cérebro humano do classificacionista ou de um grupo de especialistas, Harris e Dewey parecem não considerar o fato de que a razão em Hegel não é a cognição humana, e nem está situada nela, embora seja o homem que organize o conhecimento via pensamento e linguagem. Sendo assim, com base em Hegel, não seria o assunto preconcebido que ditaria a norma de inclusão ou exclusão de uma obra bibliográfica em determinada prateleira, nem um ser humano ou apenas um grupo específico de especialistas, mas o resultado da relação mental e material entre os seres e o livro a ser classificado, entre o livro e a cultura que o produziu, entre o livro e os outros livros do acervo, no jogo dos conflitos históricos que reposicionam mundo e representação. Se fosse possível uma classificação final ou fechada, com ares de universalidade, seria também uma estabilidade transitória, orientada por um determinado tempo histórico. Embora o idealismo hegeliano permita a inferência de que o assunto do livro já está contido no livro, bastando se revelar via luta entre as consciências, parece que esse desvelar não seria dedutivo como a CDD pressupõe.

Ao discorrer sobre as categorias da consciência, as autorias aparentam abordagem quantitativa, quando o caminho da experiência das consciências hegeliana é primordialmente qualitativo, do contrário, não haveria progresso – ou conhecimento do Absoluto. O *Aufhebung* é sempre qualitativo em *devoir* constante. Nesse contexto, o saber verdadeiro associado à ideia de progresso não está resolvido via tabela documentária, nem estará pela simples existência do ciberespaço. Ao situar a Europa como categoria anterior, e América do Sul como a última, a CDD não está apenas organizando livros em prateleira, mas reforçando um discurso valorativo. E ao interpretar a efetivação da coisa na Ideia hegeliana na inversão baconiana (Ciência – Arte – História) não considera que é a História o motor da análise de Hegel, a História na construção da Ciência e não a Ciência antes da História – esta última, derradeira e isolada classe no sistema de Dewey.

À guisa de conclusão

A trajetória total de nossa investigação não está, e nem poderia estar, contida neste artigo. No entanto, já apresenta as primeiras impressões, provocativas, que a filosofia hegeliana é capaz de imprimir nas pesquisas da OC. Os indícios apontam que a CDD, em sua razão e essência, para além da ordenação das classes, conforme apresentado, não é um sistema hegeliano, mas tal interpretação deriva de uma determinada interpretação acerca de Hegel aplicada aos SOCs, e carece de estudos mais profundos. A replicação de categorias temáticas, como as Belas Artes, presentes em Hegel, Harris e Dewey não parece ser suficiente para afirmar que trata-se de uma classificação temática hegeliana dada as divergências lógicas entre os dois sistemas do conhecimento, conforme preconiza uma leitura interpretativa de Kislev (2021).

Observa-se a possibilidade de construção de uma teoria do conceito hegeliana, ainda em aberto no campo, capaz de pensar uma epistemologia da organização do conhecimento com base nos conceitos de razão, conceito e verdade conforme o diálogo com Hegel, aprofundado, e pautado na historicidade e na dialética; um pensamento que supere a tradição kantiana contemplativa do mundo, como Mostafa (1985, 1986) verificou, e que inclua o real em movimento, como motor da Organização do Conhecimento. Uma verificação das influências hegelianas na Classificação Decimal Universal e na Colon Classification ainda está por ser desenvolvida. Faz-se relevante observar que, dada a análise desenvolvida até o momento, tanto a influência hegeliana problemática apropriada em Dewey e várias outras mentalidades do século XIX, igualmente problemática se torna a construção da crítica pós-estruturalista em organização do conhecimento não observando a dialética e a historicidade em Hegel como ferramentas filosóficas já portadoras de identificação da abertura para uma teoria crítica do conceito. Em geral, as produções bibliográficas de crítica sociocultural na era das universalizações classificatórias partem já de leituras pós-estruturalistas como caminho único de recontar a história da formação e da repercussão das classificações Oitocentistas.

Como desdobramento do diálogo entre as leituras contemporâneas da crítica sociocultural em organização do conhecimento e o pensamento hegeliano, emerge

como possibilidade de debate um diálogo entre a Teoria da não-conceitualidade de Hans Blumenberg e a fenomenologia hegeliana. Para Blumenberg, não existe identidade entre conceito e razão, pois o conceito se relaciona com a ausência do objeto. Destarte, tanto Blumenberg quanto Hegel são contrários ao espelhamento entre conceito e objeto, ocupando posição discursiva divergente de Dahlberg e tradições canônicas da OC. Se, para Hegel, a verdade é a totalidade do conhecimento sobre cada Ser, enquanto este Ser se desenvolve Em-si, depois Para-si, até chegar a um desenvolvimento Em-si e Para-si como unidade no conceito Para-nós, desenvolvendo-se no trabalho da experimentação, divorcia-se também de Blumenberg, pois a presença, em Hegel, é fulcral, não sendo o conceito uma operação à distância, mas resultado de relações entre opostos que produzem diferenças, moldando as singularidades, as particularidades e o universal. Todavia, a despeito da divergência entre as premissas de cada um, ambos assentam a noção de um conceito que não se dá *a priori*, transcendental. Entendemos que um debate cruzado sobre as autorias pode indicar pistas de uma teoria da classificação mais condizente com a diversidade, transformações e contradições humanas.

Referências

- Blumenberg, Hans. (2013). Teoria da não-conceitualidade. Editora UFMG.
- Dahlberg, Ingetraut. (1978). Teoria do Conceito. *Ciência da Informação*, 7(2), 101-107.
- Derrida, Jacques. (1979) O poço e a pirâmide: Introdução à semiologia de Hegel. In: J. Derrida et al. (Eds.) Hegel e o pensamento Moderno. Rés Editora.
- Hegel, Georg (2014). Fenomenologia do Espírito. Vozes & Editora Universitária São Francisco.
- Hegel, G. (2018). *Ciência da Lógica*. Vozes.
- Harris, William. (1870). Book Classification. *The Journal of Speculative Philosophy*, 4(2), 114-129.
- Kislev, Shachar F. (2021). Hegel and Knowledge Organization, or Why the Dewey Decimal Classification Is Not Hegelian. *Information & Culture*, 56(1), 32-48.
- Mostafa, Solange. (1985). Epistemologia da Biblioteconomia [Doctoral Dissertation]. Pontifícia Universidade de São Paulo. <https://tede2.pucsp.br/handle/handle/10705>.
- Mostafa, Solange. (1986). Ainda sobre metodologia. *Revista da Escola de Biblioteconomia UFMG, Belo Horizonte*, 15 (2), 171-202.
- Nitecki, Joseph. (1997). *Philosophical Ancestry of the American Library Information Science*. Washington, DC; Educational Resources Information Center (ERIC).
- Olson, Hope (2007). *Cultural infrastructure: The story of how classification came to shape our lives*. 18th Annual ASIS SIG/CR Classification Research Workshop.
- Olson, Hope (2011). A potência do não percebido: Hegel, Dewey e seu lugar na corrente principal do pensamento classificatório. *Revista de Ciência da Informação e Documentação* 13 (2), 3-15.
- Sales, Rodrigo de. (2017). A Classificação de Livros de William Torrey Harris: influências de Bacon e Hegel nas classificações de biblioteca. *Encontros Bibli*, 22(50), 188-204.

- Ferreira, Verônica de Sá; Sales, Rodrigo de. (2019). Espelho de uma filosofia: a presença de Hegel nas Classificações de Harris e Dewey. *Revista Informação e Informação*; 24 (3), 104-128.
- Peignot, Gabriel. (1802). *Dictionnaire raisonné de bibliologie*, tomo II. Paris: Chez Villier.
- Taylor, Charles. (2011). *Hegel: sistema, método e estrutura*. Realizações Editora.

(Página deixada propositadamente em branco)

ARQUIVOLOGIA E ORGANIZAÇÃO DO CONHECIMENTO: UMA ANÁLISE NOS ANAIS DA ISKO BRASIL, ISKO INTERNACIONAL E ISKO IBÉRICO

Wilson Roberto Veronez Júnior¹, Daniel Martínez-Ávila²,
Sonia Maria Troitiño Rodríguez³

¹Universidade Estadual Paulista, wilson.veronez@unesp.br,
<https://orcid.org/0000-0003-2939-1917>

²Universidad Complutense de Madrid, danima27@ucm.es,
<https://orcid.org/0000-0003-2236-553X>

³ Universidade Estadual Paulista, sonia.troitino@unesp.br,
<https://orcid.org/0000-0002-7204-3283>

Resumo

Esta proposta de trabalho tem por objetivo apresentar uma análise teórica, qualitativa e quantitativa sobre as produções científicas entre Arquivologia e Organização do Conhecimento nos eventos da ISKO Brasil, ISKO Internacional e ISKO Ibérico, no período de 2010-2020. A metodologia de pesquisa bibliográfica consistiu nas etapas de levantamento, seleção, quantificação e análise dos materiais produzidos por pesquisadores nacionais e internacionais nos referidos eventos, como estratégia de recuperação dos trabalhos, foram utilizadas palavras-chave, como Arquivologia, Arquivística, Organização do Conhecimento, Classificação Arquivística, Princípio da Proveniência, Diplomática, nos idiomas português, espanhol e inglês. Neste trabalho, é destacado o quantitativo de trabalhos recuperados, assim como apresentada a evidência na crescente de trabalhos relacionados à discussão sobre Arquivologia e Organização do Conhecimento no período supracitado, tanto em âmbito nacional e internacional. Os assuntos são bem variados, como Classificação Arquivística, Descrição Arquivística e Diplomática, Princípio da Proveniência Arquivística até Sistemas de Organização do Conhecimento, entre outros temas relacionados às Tecnologias da Informação em Arquivos. No total, foram recuperados 33 trabalhos nos três eventos realizados no período investigado.

Palavras-chave: *Arquivologia, Arquivística, Organização do Conhecimento, Diplomática, Princípio da Proveniência.*

1 Introdução

A arte de organizar o conhecimento, no sentido geral, pode ser entendida como uma prática milenar, seja por meio de seus instrumentos de representação do conhecimento adotados em arquivos, bibliotecas, centros de informação e documentação,

museus e outras instituições que produzem, tratam, classificam, organizam e disseminam o conhecimento para usuários especializados, comunidades específicas e para a sociedade como um todo. Apesar de ser vista como um procedimento prático, é preciso ressaltar que a organização do conhecimento não deve ser considerada somente como uma prática funcional, em outras palavras, uma prática documental utilizada nessas unidades de informação, visto que ela possui uma função histórica, teórica, epistemológica, institucional e de grande importância para os campos da Ciência da Informação, Biblioteconomia, Documentação e Arquivologia.

No contexto arquivístico, é imprescindível a abordagem dos arquivos como o objeto a ser classificado, preservado e disseminado. A Arquivística pode ser um meio, entre outras coisas, para efetivar a organização, disseminação da informação e do conhecimento, visto que o documento arquivístico possui a capacidade de cumprir a função para o qual ele foi criado, por esse ângulo, ele também tem a competência de produzir e organizar o conhecimento a partir de discursos e por meio de registros gráficos materializados ao longo do tempo em um suporte informacional, sendo ele certificado por uma autoridade ou um grupo que produz, registra, materializa e representa esse documento em um suporte físico, digital ou híbrido (Ribeiro, 2011). Além disso, o documento arquivístico é produzido a partir de uma necessidade individual, coletiva, institucional, e tem em sua essência a finalidade jurídica, administrativa, científica, histórica, social e cultural. Segundo Albuquerque e Souto (2013), independente de seu suporte, o documento arquivístico é prova das atividades e das funções exercidas pela instituição/entidade que o produziu e pertence a toda uma cadeia que engloba outros documentos. Assim, um documento está relacionado a outro, logo, ele não pode ser analisado de forma isolada, com base no Princípio da Proveniência e do Princípio de Respeito aos Fundos. Nessa discussão, Tognoli (2013, p. 107), lembra que, segundo Bautier, um documento não pode ser verdadeiramente compreendido fora do contexto do fundo arquivístico ao qual pertence.

No contexto arquivístico, a organização do conhecimento é uma prática que carece de mais reflexões e aplicações, considerando-se que os arquivos são instituições que ainda estão discutindo a possibilidade de mediação, acesso à informação e estudos de usuários de maneira mais aberta, visto que os arquivos estão vinculados a níveis de classificação e de sigilosidade da informação, nesse sentido, a ideia de organização do conhecimento em arquivos ainda pode ser considerada inicial, apesar de alguns estudos terem sido desenvolvidos, mas que ainda não avançaram em pontos específicos que mereçam uma atenção mais crítica. Neste caso, os arquivos permanentes são os que mais contribuem no processo de mediação da informação arquivística (Albuquerque & Freitas, 2017). Os arquivos permanentes ou históricos, são imprescindíveis para usuários específicos, como historiadores, sociólogos, museólogos, entre outros pesquisadores.

O conhecimento arquivístico ainda é um enigma a ser debatido no escopo da Organização do Conhecimento Arquivístico, é preciso destacar que, Tognoli (2015) ao se basear em Duchain (1983), entende que ele pode ser analisado do ponto de vista dos Princípios Arquivísticos, com destaque para o da Proveniência ou do Respeito aos Fundos, que em linhas gerais, é o principal mandamento que estabelece a origem e a forma como os documentos deverão ser organizados, classificados e representados em instituições arquivísticas. Para Tognoli, Rodrigues & Guimarães

(2019), o conhecimento arquivístico é definido a partir de três conceitos: Fundo Documental, Forma Documental e o Contexto de Produção dos Documentos.

Partindo para a relação entre Arquivologia e Organização do Conhecimento, segundo Barros & Moraes (2012), no âmbito da Arquivologia, a temática sobre Classificação Arquivística é um assunto que pode ser considerado pioneiro no escopo da Organização do Conhecimento, já no entendimento de Tognoli & Guimarães (2019), o Princípio da Proveniência se configura como um assunto que merece destaque nessa relação dialógica e interdisciplinar entre os dois campos científicos, visto que se trata do principal princípio norteador da produção do documento e do conhecimento arquivístico. Na concepção de Fonseca & Rodriguez (2017), a identificação documental, diplomática e os arquivos pessoais também podem ser incorporados nessa discussão.

A problemática desta pesquisa consiste nas seguintes questões: Quais os principais temas que foram produzidos sobre a relação entre Arquivologia e Organização do Conhecimento nos eventos da Isko Internacional, Isko Brasil e Isko Ibérico no decorrer de 2010 a 2020? Quem são os principais pesquisadores? Quantos trabalhos de pesquisadores nacionais e internacionais foram produzidos no decorrer desse período?

O objetivo geral deste trabalho é realizar uma análise teórica, qualitativa e quantitativa dos trabalhos, pesquisadores e dos temas produzidos sobre Arquivologia e Organização do Conhecimento nos eventos da Isko Internacional, Isko Brasil e Isko Ibérico.

2 Organização do conhecimento

A presente seção teórica tem por objetivo apresentar o campo científico da Organização do Conhecimento, seus principais fatos históricos, epistemológicos, bem como realizar um diálogo acerca da Organização do Conhecimento Arquivístico.

A Organização do Conhecimento (letra maiúscula) se trata de um campo científico, epistemológico e interdisciplinar. Já a organização do conhecimento (letra minúscula) tem relação com processos e atividades desempenhadas no contexto das unidades de informação por bibliotecários, arquivistas, documentalistas, museólogos, especialistas em informação e cientistas da informação por intermédio de instrumentos sistemas de organização do conhecimento (Hjørland, 2003).

No âmbito da Ciência da Informação, a Organização do Conhecimento é um domínio em contínua construção e um espaço relativamente autônomo, como também dialógico e interdisciplinar. Faz interface com outros domínios (áreas do conhecimento) e se preocupa com as questões de natureza teórico-metodológicas para contribuir na sistematização, produção, classificação, organização, disseminação, representação e recuperação da informação nos diferentes contextos científicos (Alves & Oliveira, 2016). Além disso, é preciso ressaltar que Hjørland (2003) destaca que a Organização do Conhecimento é um campo científico muito amplo, em que outros elementos são considerados, como: a divisão social do trabalho (disciplinas); instituições sociais (universidades, escolas, instituições de ensino, organizações); linguagem e os sistemas simbólicos; sistema conceitual e teorias, dos gêneros literários.

A organização do conhecimento vale-se de ferramentas como sistemas de classificação, tesouros, vocabulários controlados e outras linguagens de indexação. Como abordagens predominantes destacam-se as questões relativas a aspectos cognitivos, tecnológicos, lógico-conceituais, socioculturais e de gestão, o que sinaliza para os diálogos mais presentes na área (Guimarães et al, 2015). No entendimento de Barros & Sousa (2019), a Organização do Conhecimento é um campo científico que contribui de forma essencial para a elaboração de bases teóricas, epistemológicas e metodológicas para o acesso à informação.

San Segundo (2012), menciona que a evolução histórica da Organização do Conhecimento está ligada à trajetória cultural, política, econômica, tecnológica, de acordo com a evolução das práticas sociais e culturais dos grupos. No entendimento da pesquisadora espanhola, inúmeros são os fatores que contribuíram no desenvolvimento da Organização do Conhecimento enquanto um campo científico, epistemológico e interdisciplinar. Devemos lembrar que ao tratarmos sobre o assunto organização do conhecimento, não podemos esquecer de citar um dos maiores pensadores do campo da Biblioteconomia, ou seja, Shiyali Ramamrita Ranganathan (1892-1972), que contribuiu de forma significativa para a área, podemos citar a elaboração da *Colon Classification* e as cinco leis da Biblioteconomia (Almeida Júnior & Santos Neto, 2014, p. 105). Desse modo, além de Hjørland e Dahlberg, Ranganathan pode ser considerado um dos maiores estudiosos da Organização do Conhecimento e da Ciência da Informação sobre diversos assuntos, porém, nosso estudo não será focado neste último autor.

Para Guimarães et al (2015a), embora o termo Organização do Conhecimento remonte a Henry Evelyn Bliss, na década de 1930 do século passado, sua constituição como campo de estudos interdisciplinar é recente, podendo-se atribuir seu nascimento quatro décadas após, quando a partir da década de 70 Dagobert Soergel e Ingetraut Dahlberg buscaram conceituá-lo. Essa tentativa levou à criação, em 1974, na Alemanha, da *Society for Classification* e, em 1989, da *International Society for Knowledge Organization-ISKO*. Na opinião dos autores, a missão dessa instituição consiste em promover o avanço teórico e aplicado da Organização do Conhecimento em diversos campos e em diferentes maneiras, congregando pesquisadores de distintos países, em áreas como Ciência da Informação, Linguística, Filosofia e Ciência da Computação, entre outros campos interdisciplinares, transdisciplinares, multidisciplinares e pluridisciplinares.

Guimarães et al (2015b), do ponto de vista histórico, declaram que o termo Organização do Conhecimento foi introduzido pelo bibliotecário estadunidense Henry Evelyn Bliss (1870-1955), a partir de suas obras *'The organization of Knowledge and systems of sciences'*, de 1929, e *'The organization of knowledge in libraries'*, de 1933.

No sentido geral, Lima & Alvares (2012, p. 27), destacam que a “organização do conhecimento é o modo como ele é disposto em assuntos em toda parte onde se deseja a sua sistematização ordenada para atingir determinado propósito”. Ademais, as autoras lembram que Henry Evelyn Bliss foi o pioneiro em registrar o termo “Organização do Conhecimento” em suas obras bibliográficas. E depois de um longo período, Dagobert Soergel, precisamente no ano de 1971, adota a expressão na sua tese de doutorado, nesse sentido, em 1973, Ingetraut Dahlberg passa a fazer uso do termo (Lima & Alvares, 2012).

2.1 Organização do Conhecimento Arquivístico

Passando para o contexto arquivístico, em meados da década de 1990, Esteban Navarro propõe a intersecção dos estudos arquivísticos à Organização do Conhecimento, considerando esse último um campo científico integrador, capaz de ir além da divisão entre as áreas de Arquivologia, Biblioteconomia, Ciência da Informação e Documentação, pressupondo uma integração entre elas fundamentado nos estudos dos métodos e técnicas de armazenamento, classificação, tratamento, representação e recuperação de documentos (Tognoli, Silva & Silva, 2019).

Para corroborar, na visão de García Marco (1997), a Organização do Conhecimento se trata de uma disciplina científica dedicada à investigação e desenvolvimento de fundamentos e técnicas de planejamento, construção, gestão, uso e avaliação de sistemas de descrição, indexação, catalogação, ordenação, classificação, armazenamento, comunicação e recuperação de documento criados pela humanidade a fim de testemunhar, conservar e transmitir o seu conhecimento e suas práticas, a partir de seu conteúdo, com a finalidade de garantir sua conversão em informação capaz de gerar novo conhecimento e assim sucessivamente. Além de tudo, é um campo epistemológico e interdisciplinar, onde se relaciona com a Arquivologia, Biblioteconomia, Ciência da Informação, Epistemologia, Filosofia, Lógica, Psicologia, Ciências Cognitivas, Ciências da Computação, Semiótica, Linguística, Terminologia, Cibernética, Inteligência Artificial, entre outras, uma vez que permite distinguir os materiais substanciais, do que são somente instrumentos e outros que são analisados do ponto de vista reflexivo.

Com base em Hjørland (2008), Tognoli & Barros (2015, p. 95), acreditam que seja possível compreender o trabalho de representação arquivística como uma forma de organização do conhecimento, ou seja, a Arquivística como área com relação interdisciplinar com a Organização do Conhecimento. Além disso, os autores destacam a classificação e a descrição como ponto de intersecção entre Arquivologia e Organização do Conhecimento. Na visão de Navarro & García Marco (1995), em virtude da emergência na abordagem dos arquivos e da Arquivologia no escopo da Organização do Conhecimento, Tognoli, Rodrigues & Guimarães (2019), complementam que não havia motivos para que os investigadores, as sociedades científicas e os congressos sobre organização do conhecimento ignorassem os assuntos relacionados aos arquivos, e que era necessário que o tema fosse tratado de forma mais aprofundada no escopo dos eventos da *International Society for Knowledge Organization*. A partir dessa discussão, outras edições da ISKO passaram a adotar temáticas voltadas para a relação entre Arquivologia e Organização do Conhecimento, porém, conforme foi indicado neste trabalho, o foco de recuperação das produções científicas será centrado entre os anos de 2010 a 2020. Os principais temas apresentados e publicados, bem como os pesquisadores e as instituições ao quais eles estão vinculados serão destacados na seção 4.

Em síntese, a Organização do Conhecimento é a forma como o conhecimento é disposto e como ele pode ser acessado por meio de sistemas de organização do conhecimento. Por mais que seja entendido como um procedimento prático, a organização do conhecimento também possui um caráter científico. No contexto arqui-

vístico, os estudos referentes à organização do conhecimento tiveram mais êxitos em meados da década de 1990, e que ao longo dos anos, essa relação vem tendo uma frequência de produção científica exitosa, no sentido apresentar os principais estudos sobre esses campos. Vale ressaltar que os estudos sobre organização do conhecimento estão mais sedimentados no contexto da Biblioteconomia, sobretudo com destaque para a elaboração de linguagens documentárias, folksonomias, taxonomias, tesauros, entre outros instrumentos que realizam a interface entre sujeito informacional e o conhecimento.

3 Metodologia

A presente investigação foi realizada com base na pesquisa bibliográfica com foco nos trabalhos sobre Arquivologia e Organização do Conhecimento no âmbito da *International Society for Knowledge Organization* (ISKO). Deu-se preferência para a recuperação dos trabalhos nos idiomas português, espanhol e inglês. A estratégia de busca foi realizada por meio de palavras-chave, como Arquivologia, Arquivística, *Archival Science*, *Archivist*, *Archivística*, Organização do Conhecimento, *Knowledge Organization*, *Organización del Conocimiento*, Organização do Conhecimento Arquivístico, *Knowledge Archivist Organization*, *Organización del Conocimiento Archivístico*, Classificação, *Classification*, *Clasificación*, Princípio da Proveniência, *Principle Provenance*, *Principio de Procedência*, Diplomática, *Diplomatic Science e Diplomática*, Descrição Arquivística, *Archival Description*, *Descripción de Archivo*, Identificação Documental, *Document Identification e Identificación de Documento*.

Para o desenvolvimento desta pesquisa, a parte teórica que deu embasamento epistemológico para a definição dos conceitos propostos foi realizada nas seguintes bases de dados: Brapci, *Web of Science*, *LISA*, *SCOPUS*, e na revista *Knowledge Organization* (KO), e por meio de dicionários especializados em Arquivologia, Organização do Conhecimento e Ciência da Informação, artigos científicos, dissertações, teses, livros e o *Google Scholar*, este último foi utilizado para a localização de artigos, resumos e trabalhos de modo amplo e que pudessem estar relacionados a outras áreas do conhecimento. Já a análise quantitativa foi realizada nos anais dos eventos da ISKO Internacional, ISKO Brasil e ISKO Ibérico, realizados nos períodos de 2010-2020. Na elaboração do texto e dos argumentos, deu-se preferências para as citações indiretas com base em fontes primárias (originais) e fontes secundárias (comentadores).

Em virtude da extensão dos trabalhos, bem como o curto tempo disponível, não foi possível analisar todos os trabalhos em sua íntegra, mas, esta pesquisa restringiu-se a analisar os resumos dos trabalhos recuperados nos anais dos três eventos. Desse modo, foi possível ter um panorama geral acerca das temáticas e dos desdobramentos das produções científicas.

Os elementos representados nas tabelas 1, 2 e 3 (abaixo), são os locais e anos dos eventos, assuntos e os pesquisadores que desenvolveram os trabalhos. Com relação aos assuntos, para que eles coubessem nas tabelas, os títulos foram suprimidos, mantidos apenas as palavras-chaves para a recuperação dos artigos nos anais.

4 Análise dos resultados

Diante dos fundamentos teóricos apresentados nas seções anteriores, na continuação desta pesquisa, serão apresentadas três tabelas que representam a sistematização dos temas discutidos entre esses dois campos científicos nos anais dos encontros de ISKO Brasil, ISKO Internacional e ISKO-ibérico, respectivamente.

Na tabela 1, são apresentados os eventos ocorridos no Brasil, com destaques para as regiões do Distrito Federal (centro-oeste), São Paulo e Rio de Janeiro (sudeste), Pernambuco (nordeste) e Pará (norte).

Tabela 1 – Arquivologia e Organização do Conhecimento no ISKO Brasil

Ano/Local	Assuntos	Autores	Total
2011 – Brasília (DF)	Não encontrados	Não encontrados	0
2012 – Marília (SP)	Não encontrados	Não encontrados	0
2013 – Rio de Janeiro (RJ)	Não encontrados	Não encontrados	0
2017 – Recife (PE)	1) Classificação, organização do conhecimento arquivístico. 2) Classificação Arquivística, Teoria do Conceito, Organização do Conhecimento.	1) Tognoli (UFF); Vitoriano (UNESP); Leme (UNESP); 2) Albuquerque (UEL);	2
2017 – Marília (SP)	1) Identificação documental, organização do conhecimento, arquivos pessoais; 2) Organização, representação do conhecimento em arquivística;	1) Fonseca (UNESP); Rodríguez (UNESP); 2) Suenaga (UEL); Cervantes (UEL);	2
2019 – Belém (PA)	1) Fotografia, organização do conhecimento arquivístico; 2) Sistemas de Organização do Conhecimento, Arquivologia; 3) Arquivologia, Organização do Conhecimento; 4) Semântica, Organização do Conhecimento Arquivístico; 5) Identificação, classificação, organização do conhecimento arquivístico; 6) Valores éticos, Organização do Conhecimento, Arquivologia	1) Machado (UNESP); Semidão (FURG); Madio (UNESP); Martínez-Ávila (UCM); 2) Lehmkuhl (UFSC); Linden (UFSC); Barros (UFSC); Silva (UFSC); Vital (UFSC); 3) Silva (UFF); Tognoli (UFF); 4) Alencar (UEL); Cervantes (UEL); 5) Faben (UFF); Rodrigues (UFF); 6) Silva (UNESP); Evangelista (UNESP); Guimarães (UNESP);	6

Fonte: elaborado pelos autores (2021)

Com base na Tabela 1, os trabalhos publicados nos anais da ISKO Brasil, evento que teve a sua primeira edição no ano de 2011, sediado na cidade de Brasília-DF, seguido do evento sediado em Marília-SP, em 2012, e o de 2013, realizado no Rio

de Janeiro-RJ, demonstram que não foram publicados trabalhos sobre Arquivologia e Organização do Conhecimento nas primeiras edições. Entre os anos de 2017 e 2019, foram desenvolvidos e publicados temáticas bem variadas que vinculam Arquivologia e Organização do Conhecimento, como: classificação arquivística, identificação documental, fotografia, valores éticos em Arquivologia e Sistemas de Organização do Conhecimento. No total, entre 2011 e 2019 foram identificados 10 trabalhos, em sua maioria teóricos, sendo que os principais pesquisadores estão vinculados a universidades públicas estaduais/federais brasileiras, como a Universidade Estadual Paulista (UNESP), Universidade Estadual de Londrina (UEL), Universidade Federal Fluminense (UFF), Universidade Federal de Santa Catarina (UFSC) e da Universidade Federal do Rio Grande do Sul (UFRGS). Geograficamente, essas instituições universitárias estão localizadas nas regiões sul e sudeste do Brasil.

A seguir, na Tabela 2, são apresentados os trabalhos publicados nos anais dos encontros da ISKO Internacional, com destaques para os eventos ocorridos na Itália, Índia, Polônia, Brasil, Portugal, Bélgica e Dinamarca, congregando países da Europa e da Ásia Meridional.

Tabela 2 – Arquivologia e Organização do Conhecimento na ISKO Internacional

Ano/Local	Assuntos	Autores	Total
2010 – Roma (ITA)	1) Postmodern archival science, contemporary diplomatics, archival KO;	1) Tognoli (UFF); Guimarães (UNESP);	1
2012 – Mysore (IND)	1) Archival Classification, Knowledge Organization, Archival Field; 2) Knowledge Organization, Records, Archives Management;	1) Barros (UFRGS); Moraes (UNESP); 2) Henttonen (U.Tampere)	2
2014 – Cracóvia (POL)	1) Classification, Archives, Organization, Representation, Retrieval of Information; 2) Provenance, Archival Knowledge Organization;	1) Ribeiro (U. Porto) 2) Guimarães (UNESP); Tognoli (UFF);	2
2016 – Rio de Janeiro (RJ)	1) Knowledge Organization, Archives	1) Orrico (UNIRIO); Silva (UNIRIO);	1
2018 – Porto (POR)	1) Parameters, Knowledge Organization, archive; 2) Diplomatics, archival description, Knowledge Organization; 3) Classification, Knowledge Organization Systems, archival classification	1) Rodriguez (UNESP); 2) Tognoli (UFF); Rodrigues (UFF); 3) Barros (UFRGS); Gomes (UFRGS);	3
2019 – Bruxelas (BEL)	1) Archival Science, Knowledge Organization; 2) Knowledge Organization, Archival Arrangement and Description; 3) Archival Knowledge, Terminology, KO Domain; 4) Provenance, Knowledge Organization	1) Barros (UFRGS); Sousa (UNB); 2) Bak (U.Manitoba); Allard (U.Manitoba); Ferris (U.Manitoba); 3) Tognoli (UFF); Rodrigues (UFF); Guimarães (UNESP); 4) Tognoli (UFF); Guimarães (UNESP);	4

2020 – Aalborg (DIN)	1) The Classification, Portuguese Archives, Knowledge Organization; 2) Provenance in KO; 3) Provenance, Archival Knowledge Organization, Photographs; 4) Knowledge Organization Systems, Archival Science.	1) Silva (ULisboa); Corujo (ULisboa); Revez (ULisboa); 2) Tognoli (UFF); Milani (UFF); Guimarães (UNESP); Moraes (UNESP); 3) Machado (UNESP), Semidão (FURG); Madio (UNESP); Martínez-Ávila (UCM) 4) Tognoli (UFF); Corrêa (UFF);	4
-----------------------------	---	--	---

Fonte: elaborado pelos autores (2021)

Na Tabela 2, apresentam-se as principais produções sobre os temas relacionados a Arquivologia e Organização do Conhecimento, nomeadamente: Arquivologia na pós-modernidade, gerenciamento de arquivos, descrição arquivística, Princípio da Proveniência, Classificação Arquivística, Diplomática, Sistemas de organização do conhecimento, Representação do Conhecimento e Recuperação da Informação. Entre 2010 e 2016, teve uma média de 1 ou 2 trabalhos apresentados, porém, entre 2018 e 2020, houve uma crescente no que tange a produção, com uma média de 3 a 4 trabalhos apresentados. É importante destacar que muitos dos autores que produziram trabalhos no evento da ISKO Brasil são aqueles que também publicaram no evento da ISKO internacional. Em sua maioria, esses autores estão vinculados a Instituições Públicas Estaduais Federais Brasileiras, como a Universidade Estadual Paulista (UNESP), Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Universidade Federal do Rio Grande do Sul (UFRGS), europeias, como a Universidade do Porto (Portugal), Universidade de Lisboa (Portugal), Universidade Carlos III de Madrid (Espanha), Universidade de Tampere (Finlândia) e da América do Norte, representada pela Universidade de Manitoba (Canadá). No total, foram recuperados 17 trabalhos científicos sobre Arquivologia e Organização do Conhecimento.

Na tabela 3, são apresentados os trabalhos apresentados na Isko Ibérico, evento realizado na Espanha e Portugal.

Tabela 3 – Arquivologia e Organização do Conhecimento na Isko-Ibérico

Ano/Local	Assuntos	Autores	Total
2011 – Ferrol (ESP)	1) Archival field and the impact on knowledge organization	Orrico (UNIRIO), Silva (UNIRIO);	1
2013 – Porto (Por)	1) Descrição arquivística e a organização do conhecimento;	Silva (UNIRIO), Orrico (UNIRIO);	1
2015 – Murcia (ESP)	1) Organização do conhecimento arquivístico: um estudo terminológico comparativo;	Silva (UNESP), Moreira (UNESP), Guimarães (UNESP), Moraes (UNESP);	1

2017 – Coimbra (POR)	1) Representação da informação na Organização do Conhecimento, vocabulários controlados em arquivos;	Souza (UFF), Botão (UNIRIO), Silva (UNIRIO)	1
2019 – Barcelona (ESP)	1) Tipologia documental em arquivos pessoais, organização do conhecimento arquivístico; 2) Organização do conhecimento e identificação tipológica, arquivos pessoais.	1) Mota (UNEB), Cunha (UFBA); 2) Oliveira (FCRB), Sobral (UFF), Panisset (UNIRIO), Silva (UFF)	2

Fonte: elaborada pelos autores (2021)

Na tabela acima, nota-se que os trabalhos apresentados nos eventos são produzidos por autores brasileiros, sendo que todos esses pesquisadores estão vinculados a instituições públicas federais, com destaques para a Universidade Federal Fluminense (UFF), Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Universidade Federal da Bahia (UFBA) e a Universidade Estadual Paulista (UNESP). No total, foram recuperados 6 trabalhos relacionados a Arquivologia e Organização do Conhecimento. Essas instituições universitárias estão localizadas nas regiões sudeste e nordeste do Brasil. Com relação aos assuntos abordados, eles são diversos, como o impacto da Arquivologia na Organização do Conhecimento, Descrição Arquivística e Organização do Conhecimento, estudos terminológicos entre Arquivologia e Organização do Conhecimento, Organização do Conhecimento e vocabulários controlados em arquivos e tipologia documental, arquivos pessoais e Organização do Conhecimento.

Conclusões

A partir da revisão de literatura, bem como a realização das análises nos três eventos realizados pela ISKO no decorrer de 2010 a 2020, os resultados demonstram que a temática sobre a relação entre Arquivologia e Organização do Conhecimento tem sido uma crescente ao longo da última década, mas que requer mais aprofundamentos em estudos sobre o aspecto epistemológico, teórico, científico e interdisciplinar entre os dois campos. Embora Esteban Navarro e García Marco (1995) tenham apresentado uma proposta de discussão sobre os dois campos, é preciso dar continuidade às produções que serão apresentadas nos próximos anos.

No contexto da ISKO Brasil, evento realizado a partir de 2011, foi constatado que o primeiro trabalho sobre o assunto foi abordado no evento de 2017, e depois disso, no evento de 2019, ou seja, existe uma produção incipiente sobre a temática no contexto brasileiro. Autores como Tognoli (UFF), Barros (UFRGS), Moraes (UNESP), Fonseca (UNESP), Rodriguez (UNESP), Guimarães (UNESP) são os que mais produziram trabalhos sobre Arquivologia e Organização do Conhecimento.

No âmbito da ISKO Internacional, houve a publicação de trabalhos em todos os anos analisados, ou seja, considera-se que países da Europa, América do Norte, entre outros continentes estão mais avançados em pesquisas sobre o assunto. Interessante destacar que muitos pesquisadores brasileiros também produziram trabalhos neste evento.

No escopo da ISKO Ibérico, constatou-se que o primeiro trabalho sobre o assunto foi apresentado no ano de 2011 e que foi dada uma continuidade de produções sobre o assunto em eventos posteriores até o mais recente, em 2019. Todos os trabalhos recuperados foram produzidos por autores brasileiros.

A partir da análise nos três eventos, constata-se que os temas são bem variados, onde os pesquisadores desenvolveram trabalhos que abordaram assuntos como a Descrição Arquivística, Diplomática, Tipologia Documental, Proveniência Arquivística, Representação Arquivística e que se estende até tópicos sobre Sistemas de Organização do Conhecimento, entre outros temas relacionados às Tecnologias da Informação e Comunicação em Arquivos. Quantitativamente, em todo período analisado, foram produzidos, analisados e recuperados um total de 33 trabalhos. Não só levando em consideração o aspecto quantitativo, mas de certo modo visando analisar a qualidade dos trabalhos analisados, a Arquivologia é uma ciência que ao longo do tempo vem ocupando o seu espaço em diversos contextos, além da Ciência da Informação, a Arquivologia tem estreitado laços com a Organização da informação e a Organização do Conhecimento, e que esse diálogo tem gerado frutos para o fazer arquivístico, bem como nas possibilidades de discussão sobre meios de representação, organização, recuperação e acesso à informação.

Na continuidade desta pesquisa, o objetivo é dar ênfase aos estudos epistemológicos, científicos e interdisciplinares entre Arquivologia e Organização do Conhecimento, em que a ideia é abordar a Classificação Arquivística como componente que vincula esses dois campos científicos a partir da concepção de Organização e Representação do Conhecimento Arquivístico. O intuito é dar aprofundamento aos princípios arquivísticos para estabelecer interdisciplinaridades entre esses dois campos científicos. Além disso, este trabalho procurou dar embasamento para futuros estudos sobre a Arquivologia e a sua relação com outras disciplinas científicas.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível

Superior – Brasil (CAPES) – Código de Financiamento: 88887.609892/2021-00

Grupos de Pesquisa:

a) Fundamentos Teóricos da Informação (GPFTI)

b) Acervos, Dimensões do Documentos, Memória e Patrimônio (GPADDMP)

Referências

- Albuquerque, A. C. Freitas, L. M. de (2017). Concepções acerca da classificação arquivística à luz da teoria do conceito. *VII Seminário em Ciência da Informação (SECIN) – 21 a 23 de agosto de 2017*. Universidade Estadual de Londrina (UEL), Londrina, Paraná, 2017.
- Barros, T. H. B; Moraes, J. B. E. De (2012). Archival classification and knowledge organization: theoretical possibilities for the archival field. In: **Categories, Contexts And Rela-**

- tions In Knowledge Organization. Proceedings of the Twelfth International ISKO Conference.** 6-9.
- Duchain, M (1983). O respeito dos fundos em Arquivística: princípios teóricos e problemas práticos. *Arquivo & Administração*, 10:14, n. 2, 1986, 01-16.
- Esteban Navarro, M.A. & García Marco, F.J. (1995). Las “primeras jornadas sobre organización del conocimiento: organización del conocimiento e información científica”. *Scire: representación y organización del conocimiento* (ISSNe 2340-7042; ISSN 1135-3716), 149-157.
- Fonseca, G. F; Rodríguez, S. M. T (2017); The contribution of archival identification to knowledge organization in personal archives. Knowledge Organization and Cultural Diversity [recurso eletrônico] / José Augusto Chaves Guimarães, Vera Dodebei, organizadores. – Pernambuco: ISKO-Brasil; UFPE, 2017. 760 f. 30 cm. **INTERNATIONAL SOCIETY KNOWLEDGE ORGANIZATION**, Marília – SP. 322-327.
- Garcia Marco, F. J. G. (1997). Avances en organización del conocimiento en España: los II encuentros sobre organización del conocimiento en los sistemas de información y documentación. In *Organización del conocimiento en sistemas de información y documentación: actas del II Encuentro de ISKO-España*. Getafe, 16 y 17 de noviembre de 1995 (pp. 7-12). Universidad de Zaragoza..
- Hjørland, B. (2003). Fundamentals of knowledge organization. *Knowledge organization*, 30:2, 87-111.
- Hjørland, B. (2008). What is knowledge organization (KO)?. *KO Knowledge Organization*, 35(2-3), 86-101.
- Ribeiro, F. (2011). A arquivística como disciplina aplicada no campo da ciência da informação. *Perspectivas em gestão & conhecimento*, 1(1), pp. 59-73.
- San Segundo, R (2012). **Visibilidad científica y diálogos internacionales en Organización del conocimiento**. Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade. [recurso eletrônico] José Augusto Chaves Guimarães, Vera Dodebei (organizadores). – Marília: ISKO-Brasil: FUNDEPE, 285.
- Tognoli, N. B. (2013). **A construção teórica da Diplomática: em busca de uma sistematização de seus marcos teóricos como subsídio aos estudos arquivísticos**. 161 fls. Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, Marília, 2013.
- Tognoli, N. B., & Barros, T. H. B. (2015). Os processos de representação do conhecimento arquivístico: elementos históricos e conceituais da classificação e descrição. *Dodebei*, V.
- Tognoli, N. B. (2015). A organização do conhecimento arquivístico a partir da Diplomática: elementos históricos e conceituais de um método em revisão. *Scire: representación y organización del conocimiento* (ISSNe 2340-7042; ISSN 1135-3716), 35-40.
- Tognoli, N. B; Guimarães, J. A. C (2019). Provenance as a Knowledge Organization Principle. *Knowledge Organization*, 46:7, 558-568.
- Tognoli, N. B., da Silva, A. M. S., & da Silva, A. P. (2019). Organização do Conhecimento e Arquivologia: uma análise de domínio nos periódicos Knowledge Organization e Scire. *Informação & Informação*, 24(3), 52-77.
- Tognoli, N. B., Rodrigues, A. C., & Guimarães, J. A. C (2019). Definindo o conhecimento arquivístico: estruturas conceituais. *Informação & Informação*, 24(2), 58-75.

PRODUÇÃO DO CONHECIMENTO CIENTÍFICO EM ORGANIZAÇÃO DO CONHECIMENTO DA ARQUIVOLOGIA NO BRASIL

Rosale de Mattos Souza¹, Natália Araujo Lima²

¹Universidade Federal do Estado do Rio de Janeiro- UNIRIO,
rosale.m.souza@unirio.br, ORCID iD0000-0002-5302-5629

²Universidade Federal do Estado do Rio de Janeiro-UNIRIO,
natalialima@edu.unirio.br, ORCID iD 0000-0001-8669-4020

Resumo

Este estudo visa dar visibilidade à produção científica da Organização do Conhecimento (OC) em Arquivologia no Brasil. Assinala-se que a Organização do Conhecimento foi inaugurada e mais desenvolvida na Ciência da Informação e na Biblioteconomia do que na Arquivologia. Mostra-se a importância dessa linha de pesquisa para o desenvolvimento da Arquivologia. A metodologia utilizada será a bibliográfica-documental e descritiva, identificando as pesquisas em OC de Arquivologia no Brasil. Como resultados irá apresentar a participação de autores, linhas e grupos de pesquisa brasileiros na OC. Observa-se a crescente produção científica da Arquivologia na organização do conhecimento, recomenda-se inserir a OC nas agendas de pesquisas da área.

Palavras-chave: Organização do Conhecimento, Arquivologia, Produção Científica, Linhas de pesquisa, Interdisciplinaridade

Introdução

Ingetraut Dahlberg fundou a Organização do Conhecimento (OC) no ano de 1989, preocupada com estudos sobre classificação, termos e conceitos. Segundo Hjørland (2008, p. 86), a OC se relaciona tanto em bibliotecas como em arquivos, e preocupa-se com a descrição, indexação, classificação, sistemas de organização do conhecimento, palavras e conceitos, estruturas coletivas e significados sociais.

[...] A organização do conhecimento se relaciona com atividades como: descrição de documentos, indexação e classificação em Bibliotecas, bases de dados bibliográficas, arquivos e outros tipos de memória institucional especialistas no assunto, bem como por algoritmos de computador. KO como um campo de estudo está preocupado com a natureza e a qualidade de tais processos de organização do conhecimento (KOP), bem como os sistemas de organização do conhecimento (KOS) usados para organizar

documentos, documentar representações e conceitos. Existem diferentes abordagens históricas e teóricas e teorias sobre KO, que estão relacionadas a diferentes visões de conhecimento, cognição, linguagem e organização social. (HJORLAND, 2008, p. 86)

Silva e Albuquerque (2020) apresentam uma interessante abordagem sobre a importância da organização e representação nos arquivos ao remeter à mitologia grega, o “fio de Ariadne”. Na mitologia, Ariadne ajudou Teseu, por quem era apaixonada, a sair do labirinto e derrotar o Minotauro através de um fio de ouro. Assim é visualizada a organização e representação nos arquivos, “como um norte, um fio para guiar usuários de diferentes concepções a encontrar o seu fim ou o começo da exploração: a informação, o conhecimento, os significados, a saída do labirinto” (SILVA; ALBUQUERQUE, 2020, p. 69).

Com o impacto das tecnologias da informação e comunicação, o advento dos sistemas de recuperação da informação, oriundos das novas necessidades das instituições arquivísticas, surgem pesquisadores preocupados com Organização do Conhecimento na Arquivologia. Segundo Vogel; Kobashi, (2019) existe uma iniciativa importante, que é a do Arquivo Nacional da Austrália, que vem construindo tesouros funcionais, com um manual que orienta às ações necessárias para a sua construção. Outras iniciativas importantes são os estudos e produção intelectual das bibliotecárias e cientistas da Informação Smit; Kobashi (2003), Smit (2018) com vocabulários controlados na área de arquivo.

Conforme Sousa (2002), a função arquivística classificação de documentos, considerada matricial para o *fazer arquivístico*, é a explicitação de funções ou estruturas dentro de um contexto de produção documental. É a representação do mapeamento deste contexto produtivo que subsidia as demais etapas da gestão de documentos. Classifica-se o conjunto documental pensando no todo e relações entre documentos, no papel desempenhado pela informação dentro do organismo. Ao contrário do livro para a biblioteca – unidade de informação que independe das outras unidades, o documento ou a informação arquivística, pensada em nível de unidade de informação, só pode ser significada se relacionada com outras unidades informacionais que representam uma atividade, subfunção ou função específica. À medida que vão sendo acumuladas, as informações estabelecem relações entre si. Elas estão unidas por um elo criado no momento em que são produzidas e recebidas, determinado pela razão de sua elaboração e que é necessário à própria existência e à capacidade de cumprir seu objetivo. Elas são um conjunto indivisível de relações intelectuais.

Não se pode deixar de considerar na Organização do Conhecimento em Arquivologia os princípios arquivísticos. No exterior, Princípios arquivísticos como o de Respeito aos Fundos documentais, Proveniência, ordem original, territorialidade, nos séculos XIX e XX, foram primaciais para o norteamento de uma política de organização do Conhecimento na Arquivologia. As normas de descrição arquivísticas divulgadas pelo Arquivo Nacional do Brasil foram decisivas no avanço da OC no Brasil.

Duchein (1986) em seu texto emblemático sobre o respeito aos fundos documentais, trata o assunto alertando-nos para alguns requisitos, tais como: a legítimi-

dade oficial do órgão produtor através de instrumentos legais que os criaram por meio de leis, decretos, portarias, resoluções nas mais diversas esferas de poder, o contexto, a criação de atividades ou supressão delas, atividades com certa autonomia administrativa que levem às decisões sobre determinado tema, tendo uma chefia ou direção do órgão produtor de documentos. Com relação aos princípios hierárquicos e classificatórios pertinentes a um arquivo, deve-se levar em consideração os seguintes princípios: Categorias de Funções (meio e fim); Áreas Funcionais; Função; Subfunção; Atividade; Subatividade; Tarefa.

No estudo de Pazos e Schmidt (2021), sobre a função classificação nos modelos de ensino dos cursos de Arquivologia do sudeste do país, foi possível verificar que o ensino arquivístico em classificação nessas universidades esteve focado na metodologia de arranjo e sua relação com a descrição para organizar arquivos permanentes. As atividades de classificação foram abordadas sob a perspectiva da gestão de documentos a partir da década de 90, provavelmente a partir da aprovação da Lei 8.159/1991, “Lei de Arquivos”, alcançando maior autonomia ao constituir disciplinas obrigatórias nos cursos de Arquivologia nas quatro entre cinco universidades pesquisadas.

Este trabalho visa dar visibilidade à produção científica sobre Organização do Conhecimento em Arquivologia no Brasil, pois ainda são poucos os pesquisadores e teóricos dedicados a esta temática.

Metodologia

A abordagem metodológica utilizada neste estudo foi a revisão bibliográfica-documental, descritiva no mapeamento de pesquisadores e linhas de pesquisa em Organização do Conhecimento através da Plataforma Lattes e do site da ISKO-Brasil.

Resultados

No Brasil existem pesquisadores e grupos de pesquisa que vêm estudando a Teoria do Conceito, de Ingetraut Dhalberg, classificação, indexação, thesaurus funcionais, representação da informação, que são temáticas relacionadas à OC. Podemos mencionar os esforços de pesquisadoras em Londrina quanto à indexação e representação da informação e destacar a experiência de elaboração de um Manual de Indexação do Arquivo Nacional.

No decênio anterior do século XXI, observamos uma mudança de paradigma relacionada com a Organização do Conhecimento e os estudiosos sobre a temática no país. José Augusto Guimarães e Vera Dodebei, bibliotecários, foram pioneiros em relação à inserção do capítulo da Sociedade Brasileira para Organização do Conhecimento em 2007. Hoje, temos arquivistas se destacando, como podemos mencionar Thiago Bragato Barros, presidente eleito da International Society for Knowledge Organization (2020-2024), e Natália Bolfarini Tognoli, presidente do Capítulo Brasileiro da ISKO (2020-2021).

A participação de arquivistas como membros de pesquisas no país está se fortalecendo. Na relação Organização do Conhecimento e Arquivologia no Brasil há

destaque para Renato Tarciso de Sousa, com enfoque em Classificação, sendo vice-líder de grupo de pesquisa sobre Estudos de Representação e Organização da Informação e do Conhecimento. Tem experiência na área de Ciência da Informação, com ênfase em Organização de Arquivos, atuando principalmente nas seguintes áreas: Arquivologia, Organização e Recuperação de Arquivos, Gestão de Documentos e Teoria do Conceito.

Thiago Bragato Barros, Membro do Scientific Advisory Council da ISKO-Internacional, o pesquisador tem experiência na área de Arquivologia e Ciência da Informação, atuando principalmente nos seguintes temas: Abordagens linguísticas em Ciência da informação, Representação e Organização do Conhecimento. Outros destaques são Natália Bolfarini Tognoli e José Augusto Chaves Guimarães, que atuam pesquisando em Diplomática, Tipologia documental e Análise de Domínio, áreas correlatas da Organização do Conhecimento.

Podemos citar alguns grupos de pesquisa liderados por arquivistas que constam no capítulo ISKO-Brasil: “Arquivos, Bibliotecas e Organização do Conhecimento (GPABOC)”, da Universidade Federal Fluminense (UFF), liderado por Natália Bolfarini Tognoli e Suellen Oliveira Milani <http://dgp.cnpq.br/dgp/espelhogrupo/346110>; “Formação e Atuação Profissional em Organização da Informação”, Universidade Estadual Paulista (UNESP), liderado por José Augusto Chaves Guimarães e Natália Bolfarini Tognoli dgp.cnpq.br/dgp/espelhogrupo/0526582900186426; “Gênese Documental Arquivística”, também da Universidade Federal Fluminense (UFF), liderado por Ana Célia Rodrigues e Danilo André Cinacchi Bueno, <http://dgp.cnpq.br/dgp/espelhogrupo/579004>; “Imaginário, Representações Sociais e Informação: estudos culturais e comparativos”, Universidade Federal Fluminense (UFF), liderado por Rosa Inês de Novais Cordeiro, <http://dgp.cnpq.br/dgp/espelhogrupo/578864>; “Organização e Representação do Conhecimento: Abordagens Linguísticas em Arquivos e Bibliotecas (ORCALAB)”, da Universidade Federal do Rio Grande do Sul (UFRGS), liderado por Thiago Henrique Bragato Barros e Rita do Carmo Ferreira Laipelt, <http://dgp.cnpq.br/dgp/espelhogrupo/473980>, <http://www.ufrgs.br/orcalab>.

Observamos o crescente número de pesquisadores, com atualmente cinco grupos de pesquisa no site da ISKO-Brasil, liderados por arquivistas. É notada a frequente publicação de trabalhos em periódicos, assim como a apresentação em eventos nacionais e internacionais das áreas de Organização do Conhecimento e Arquivologia.

Vimos desenvolvendo o projeto de pesquisa “Linguagens documentárias, Semântica e Representação da Informação Arquivística: Impactos da análise, tratamento e disseminação da informação na sociedade do conhecimento.”, cadastrado na Divisão de Pesquisa, da Pró-Reitoria de Pesquisa e Pós-Graduação, associado ao grupo de pesquisa cadastrado no Conselho Nacional de Pesquisa e Tecnologia – CNPQ, o Grupo de Pesquisa Organização do Conhecimento para Representação da Informação da Universidade Federal do Estado do Rio de Janeiro (UNIRIO), liderado pelos professores Marcos Luiz Cavalcanti de Miranda e Brisa Pozzi de Souza, conforme consta no site da ISKO BRASIL.

Destacamos que participamos com trabalhos em eventos da ISKO Internacional desde 2013 até os dias atuais, uma publicação na Knowledge Organization no dossiê

Arquivologia em 2019, a nossa participação como membro da ISKO capítulo brasileiro, e a nossa bolsista de pesquisa participou com pôster no I Simpósio Internacional de Arquivos em 2020.

Conclusões

Finalmente, recomenda-se a Organização do Conhecimento nas agendas de pesquisas da Arquivologia no país, pois apesar da evolução da pesquisa, percebe-se que ainda é mais pesquisada em Biblioteconomia e Ciência da Informação do que na área arquivística. Portanto, faz-se importante o diálogo da Arquivologia com estas áreas para o seu desenvolvimento de forma inter e transdisciplinar, assim como para o progresso do pensar e do fazer arquivísticos na Organização do Conhecimento.

Referências

- Dahlberg, I. 2006. Knowledge organization: A new science?. *Knowledge Organization*. 33. 11-19.
- Hjørland, B. 2008 What is Knowledge Organization (KO)? *Knowledge Organization*, 35 (3/2), 86-101.
- Pazos, J. M.; Schmidt, C. M. S. 2021. A função classificação na formação do arquivista: uma análise histórica dos moldes de ensino dos cursos de Arquivologia do sudeste do Brasil. *Revista do Arquivo*, São Paulo, ano VII, n. 12, p. 44-55.
- Silva, J. H. da; Albuquerque, M. E. B. C. de. 2021 A organização e representação na Arquivologia: o fio de Ariadne dos arquivos. *Archeion Online*, João Pessoa, v. 8, n. 1, p. 68-94, jul./set. 2020. Disponível em <https://periodicos.ufpb.br/index.php/archeion/article/view/55272/31533> Acesso em 17 ago. 2021.
- Sociedade Brasileira para Organização do Conhecimento. 2021 <https://isko.org.br/grupos-de-pesquisa/> Em ISKO BRASIL. Disponível em 24.06.2021.
- Sousa, R.T.B. 2002. *As bases do processo classificatório em Arquivística: um debate metodológico*. São Paulo: Associação dos Arquivistas de São Paulo
- Sousa, R.. T. B. Classificação de documentos arquivísticos: trajetória de um conceito. *Arquivística.net*, Rio de Janeiro, v. 2, n. 2, p. 120-142., ago./dez. 2006. Disponível em <<https://core.ac.uk/download/pdf/33535530.pdf>> Acesso em 30 mar. 2021.
- Smit, J. W.; Kobashi, N. Y.. 2003. *Como elaborar vocabulário controlado para aplicação em arquivos*. São Paulo: Arquivo do Estado e Imprensa Oficial do Estado.
- Smit, J. W. Vocabulário Controlado e Controle de Vocabulário em Arquivos. 2018. *Acervo*. Rio de Janeiro, v. 31, n. 03, p. 46 -56.
- Souza, R.de M. 2019 The Representation of Archival Information in Controlled Vocabularies in Rio de Janeiro. *Knowledge Organization*. Org.46, no.7
- Tognoli, Natália B.; Silva, Amanda M. S.; Silva, Andrieli P. 2019 Organização do conhecimento e Arquivologia: uma análise de domínio nos periódicos Knowledge Organization e Scire. *Inf. Inf.*, Londrina, v. 24, n. 3, p. 52-7, Disponível em <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/38726/pdf>> Acesso em 30 mar. 2021.
- Vogel, M. J. M. |; Kobashi, Nair Yumiko. 2019 Tesouro Funcional para Organização de Arquivos Administrativos *Páginas a&b*. S.3, n.º 12 48-62 | <https://doi.org/10.21747/21836671/pag12a4>

(Página deixada propositadamente em branco)

A POLÍTICA DE INDEXAÇÃO PARA A ORGANIZAÇÃO DO CONHECIMENTO EM MUSEUS: APLICAÇÃO DO PROTOCOLO VERBAL INDIVIDUAL

Raul de Azevedo Carvalho¹, Luciana Di Paula Andrade da Fonseca²,
Franciele Marques Redigolo³, Mariângela Spotti Lopes Fujita⁴

¹Universidade Federal do Pará, raulcarvalho.trab@gmail.com,
ORCID iD 0000-0002-5655-9482

²Universidade Federal do Pará, luhandra16de@gmail.com, ORCID iD 0000-0002-9009-8712

³Universidade Federal do Pará, franciele@ufpa.br, ORCID iD 0000-0001-6277-2960

⁴Universidade Estadual Paulista Júlio de Mesquita Filho, mariangelaslf57@gmail.com,
ORCID iD 0000-0002-8239-7114

Resumo

Aborda a política de indexação como um elemento importante na organização e representação da informação e do conhecimento em diferentes tipos de unidades de informação, incluindo os museus, por contribuir para a otimização dos sistemas de recuperação da informação por meio dos seus requisitos, elementos e variáveis. No entanto, observa-se uma escassa abordagem sobre a temática no contexto museológico, sendo necessário o debate sobre a política de indexação na gestão do acervo no âmbito dos museus. Assim, foi levantada a seguinte questão: como as diretrizes de uma política de indexação podem auxiliar a prática da documentação museológica dentro dos museus? **Objetivo:** Identificar os elementos de uma política de indexação durante a prática da documentação museológica dentro dos museus, pois seus sistemas de documentação têm como função registrar, organizar e difundir as informações de suas coleções. **Métodos:** Consiste em um estudo de caso, de cunho exploratório e análise qualitativa dos dados, que tem como instrumentos de coleta a entrevista semiestruturada e a técnica do Protocolo Verbal Individual (PVI), aplicados com quatro profissionais que realizam a prática de documentação museológica em quatro museus da Universidade Federal do Pará (UFPA). **Resultados:** Identificou-se na prática da documentação museológica requisitos e elementos que constituem uma política de indexação, como: cobertura de assunto, linguagem documentária, processo de indexação na documentação museológica, estratégia de busca e a clientela da instituição; prática esta realizada sem o conhecimento ou aplicação de uma política de indexação. **Conclusão:** Apesar de detectarmos aspectos importantes de uma política que se assemelha a política de indexação existente em bibliotecas, não há diretrizes que guiam esse processo, na forma de um documento dentro das instituições analisadas.

Palavras-chave: Política de indexação, Documentação museológica, Protocolo verbal individual, Museus universitários.

Introdução

Os museus são instituições intimamente ligadas às atividades informacionais, pois os seus objetos são portadores e veículos de informações e, portanto, possuem o relevante papel social ao aproximar o seu usuário de suas coleções, gerando um processo de ensino-aprendizagem e produção do conhecimento (Ferrez, 1994; Desvallées & Mairesse, 2013). Dessa forma, os estudos sobre suas formas de preservação, representação e disseminação da informação têm ocupado um espaço cada vez maior nas discussões no campo da Ciência da Informação.

A própria Museologia, enquanto área do conhecimento, busca organizar e representar a informação contida nos objetos das coleções dos museus, através da chamada documentação museológica, processo este que consiste em um conjunto de pressupostos teórico-metodológicos para a organização e contextualização das informações dos objetos museológicos (Monteiro, 2010), dividido em: aquisição, identificação, registro fotográfico, catalogação, classificação e pesquisa (Ferrez, 1994, Padilha, 2014, Reis et al., 2019). Caracterizando os museus não apenas como um local para guarda de objetos, mas sim uma “estrutura organizacional de referência” (Mensch, 1994, p. 15).

Nesse sentido, observa-se que a política de indexação se revela como uma ferramenta aplicável e importante para a construção dos sistemas de informação dos museus, visto que pode atuar como um instrumento para definir os estágios que serão utilizados para representação e recuperação da informação, pois funciona como uma ferramenta gerencial dos sistemas de informação, elaborada de acordo com os objetivos, recursos financeiros, materiais e a clientela da instituição (Carneiro, 1985; Fujita & Santos, 2016; Rubi, 2008), contendo os elementos de cobertura de assunto, seleção e aquisição de documentos-fontes, o processo de indexação (exaustividade, especificidade, revocação e precisão), escolha da linguagem, estratégias de busca, avaliação do sistema e forma de saída (Fujita & Santos, 2016; Rubi, 2012), que podem ser estabelecidos para qualquer tipo de unidade de informação.

Dito isso, nota-se que a própria análise de assunto e tradução realizadas no processo de indexação (Dias & Naves, 2013; Lancaster, 2004) encontram semelhanças na forma de representação da informação no processo de documentação museológica (Reis et al., 2019). É por meio da análise de assunto que é possível representar e organizar por áreas específicas os documentos, por meio dos Sistemas de Organização do Conhecimento (SOCs), como tesouros, índices alfabéticos e cabeçalhos de assunto (Cintra *et al.*, 2002; Kobashi, 2007; Hjørland, 2018), possibilitando a sua recuperação nos catálogos gerados pela instituição, tornando-os acessíveis aos usuários.

No âmbito dos museus, percebe-se que aspectos deste processo estão presentes dentro do processo de documentação museológica, especificamente durante a etapa da catalogação do seu acervo, sendo esta uma atividade importante para a preservação e disseminação da informação, assim como a indexação que ocorre na catalogação de assunto realizada em bibliotecas universitárias (Lancaster, 2004; Dias & Naves, 2013; Fujita, 2003; 2013; Carvalho, Melo & Linhares, 2015). Observa-se, portanto, que a indexação se insere no processo de documentação museológica a partir do processo de catalogação nos museus.

No entanto, ainda existem poucos estudos sobre a temática tanto na área da Museologia, quanto na Ciência da Informação, ainda é preciso buscar suporte nos pressupostos teórico-metodológicos da Biblioteconomia para se pensar em diretrizes que possam guiar essa atividade no universo dos museus.

Isto posto, este trabalho busca iniciar discussões sobre uma política de indexação voltada para os museus que auxilie o profissional na prática da documentação museológica, a partir da seguinte questão problema: Como as diretrizes de uma política de indexação podem auxiliar a prática da documentação museológica dentro dos museus? Para responder essa pergunta, tem-se como objetivo identificar os elementos de uma política de indexação durante a prática da documentação museológica, especificamente durante o processo de catalogação do acervo, no ambiente organizacional dos museus.

Para alcançar o objetivo proposto, o percurso metodológico consiste em um estudo de caso, de cunho exploratório e análise qualitativa, com uso de entrevistas semi-estruturadas e da técnica do Protocolo Verbal Individual (PVI), aplicados no âmbito dos museus da Universidade Federal do Pará (UFPA), *campus* Belém, pois permite gerar categorias para análise, possibilitando uma investigação a partir dos processos mentais dos profissionais.

Como resultado, observou-se requisitos e elementos da política de indexação na prática da documentação museológica, são eles: cobertura de assunto, linguagem documental, o processo de indexação, estratégia de busca e a clientela. Identifica-se que a temática é importante, pois contribui no aperfeiçoamento dos processos informacionais dentro das instituições museológicas brasileiras em seus principais meios de disseminação (exposições e catálogos), buscando ampliar o escopo de seu campo de atuação, no que diz respeito o desenvolvimento de diretrizes que norteiam o processo de documentação museológica, sendo necessário um maior aprofundamento nessa temática.

Metodologia

A pesquisa foi dividida em dois momentos: a) pesquisa bibliográfica; b) aplicação de entrevistas semiestruturadas e Protocolos Verbais Individuais (PVI) com profissionais que realizam a documentação museológica no seu ambiente de trabalho. Para coleta de dados, primeiramente foi realizada a entrevista semiestruturada, elaborada a partir de um rol de questões abertas, como pode ser observado no Quadro 1, e aplicada com quatro profissionais de quatro instituições museológicas da UFPA, *campus* Belém, a saber: Museu da UFPA, Museu de Geociências, Museu de Zoologia e Museu de Anatomia, sendo duas entrevistas realizadas via Google Meet e duas *in loco*.

Quadro 1 – Rol de perguntas da entrevista semiestruturada baseados em Bottallo (2010), Fujita e Santos (2016), Monteiro (2010) e Padilha (2014)

Categorias	Objetivo da categoria	Fonte
Dados gerais dos museus 1 Quais áreas do conhecimento são cobertas pela coleção?	Conhecer a instituição e suas características institucionais, recursos	Fujita e Santos (2016) e

<p>2 Quantos funcionários trabalham na representação da informação museológica?</p> <p>3 Qual a quantidade de documentos e objetos museológicos do acervo?</p> <p>4 O museu possui algum sistema de informação ou base de dados para registro dos documentos? Se sim, qual seria?</p>	<p>humanos, financeiros e materiais, o sistema de informação usado, o tipo de acervo e a área temática das coleções.</p>	<p>Monteiro (2010)</p>
<p>Documentação museológica</p> <p>5 Quais são as etapas realizadas na documentação museológica?</p> <p>6 Vocês utilizam algum instrumento terminológico para a determinação da linguagem documentária e classificação dos assuntos? Se sim, quantos termos são usados na representação temática da informação? A linguagem documentária é geral ou específica?</p> <p>7 A instituição contém manuais, políticas ou diretrizes que auxiliem na prática da documentação museológica?</p>	<p>Identificar os elementos de uma política de indexação na execução da documentação museológica e uso de linguagem documentária (geral e/ ou específica).</p>	<p>Bottallo (2010), Fujita e Santos (2016), Monteiro (2010) e Padilha (2014),</p>
<p>Perfil do usuário</p> <p>8 Como é feita a consulta do usuário ao acervo da instituição? Explique.</p> <p>9 Qual é o perfil do usuário da sua instituição?</p>	<p>Investigar as características dos usuários da instituição museológica.</p>	<p>Bottallo (2010), Fujita e Santos (2016), Monteiro (2010)</p>

Fonte: Elaborado pelos autores (2021).

Posteriormente, foi aplicado o PVI, uma “técnica de recolha de dados introspectiva [...] baseada em teorias cognitivas e de processamento de informação cujo método é a análise de tarefas” (Alonso-Arroyo et al., 2016, p. 61, tradução nossa), mediante o processo do think aloud, o “pensar alto” (Fujita, 2009; Redigolo, 2014). Para a aplicação do PVI em museus foi realizada uma adaptação da técnica, no qual em vez de um texto escolheu-se um objeto museológico para observar suas verbalizações na atividade da documentação museológica, uma vez que se trata de uma tipologia de acervo diferente daquela encontrada em bibliotecas. Para essa adaptação utilizou-se as pesquisas de Bottallo (2010), Fujita (2009), Monteiro (2010) e Padilha (2014). Assim, para a aplicação do PVI foram utilizados os procedimentos metodológicos abaixo:

Procedimentos anteriores à coleta:

- Definição do universo de pesquisa: quatro museus universitários da UFPA
- Seleção do texto-base: os participantes usaram objetos das suas respectivas coleções museológicas no lugar do texto-base.
- Definição da atividade: os participantes praticaram a documentação museológica.
- Seleção dos participantes: foram selecionados quatro profissionais que realizam a atividade da documentação museológica em museus universitários da UFPA.

- Conversa informal com os participantes: foi feita uma conversa informal com os sujeitos para explicar o objetivo e a importância da pesquisa para os campos da Museologia e Ciência da Informação.

Procedimentos durante a coleta:

- Gravação do “pensar alto”: as falas do sujeito durante a realização da coleta foram gravadas por um gravador digital.
- Entrevista retrospectiva (opcional): foi utilizada para esclarecimento de dúvidas que surgiram durante a aplicação do PVI.

Procedimentos posteriores à coleta:

- Transcrição literal das falas dos participantes: realizou-se a transcrição das falas de cada sujeito, sendo separadas por categorias de análise.
- Leitura detalhada dos dados coletados para a construção das categorias de análise: realizou-se uma leitura detalhada das transcrições dos PVIs em busca de assuntos pertinentes para a construção das categorias de análise mediante a percepção de fenômenos significativos e recorrentes.
- Construção das categorias de análise: após a leitura detalhada construiu-se as categorias de análise para facilitar o entendimento sobre a prática da documentação museológica (ver Quadro 2).
- Retorno aos dados para retirar trechos da discussão que exemplifiquem cada categoria de análise: efetuou-se uma releitura das transcrições para retirada de trechos que exemplificam os fenômenos das categorias e a síntese dos principais aspectos observados.

Quadro 2 – Categorias de análise do protocolo verbal individual.

Categorias	Objetivo da categoria	Fonte
Interpretação de informações intrínsecas e extrínsecas	Identificar como é feita a leitura documentária do objeto museológico a partir da interpretação de informações relativas ao próprio objeto (informações intrínsecas) ou exteriores a ele (informações extrínsecas) pelo profissional.	Bottallo (2010), Ferrez (1994) e Padilha (2014)
Indexação na Documentação Museológica	Identificar se o profissional identifica e seleciona conceitos a partir da análise do objeto museológico para determinar seu assunto (exaustividade, especificidade, precisão e revocação), e se existe alguma preocupação com o usuário no momento da seleção desses conceitos.	Gil Leiva e Fujita (2012), Fujita (2003), Fujita e Santos (2016) e Lancaster (2004)
Uso de Linguagem Documentária	Identificar o uso de alguma linguagem (natural ou controlada) na inserção das informações no catálogo.	Bocato (2009) Cintra <i>et al.</i> (2002),
Catálogo do acervo	Identificar se o processo de descrição do objeto museológico nos catálogos da instituição é padronizado.	Bottallo (2010), Monteiro (2010), Padilha (2014), Reis <i>et al.</i> (2019)

Forma de análise

Para os procedimentos de análise dos dados foram elaboradas categorias, tendo como base a literatura apresentada, o objetivo da pesquisa e os dados gerados pelos profissionais participantes, que foram identificados como M1, M2, M3 e M4. Os dados foram sistematizados em duas seções: análise da entrevista e análise do PVI, onde as subseções foram fracionadas por categorias (ver Quadro 1 e Quadro 2).

Resultados

Este tópico trata-se da discussão dos resultados que foram obtidos por meio das entrevistas semiestruturadas e aplicações da técnica do Protocolo Verbal Individual.

Análise da entrevista

Nesta categoria, que diz respeito aos Dados gerais dos museus, pode-se inferir que os assuntos cobertos pela coleção museológica do sujeito M1 são de obras históricas e de artes visuais. A coleção do M2 é da área de Geociências, como rochas e minerais, além de instrumentos antigos (lupas e microscópios). O M3 abrange assuntos da Zoologia e Ecologia e o M4 da área de Ciências da Saúde e de Ciências Biológicas. Para a organização dos acervos somente o M1 possui um profissional museólogo, os demais museus não contêm museólogos, pois quem faz a organização e preservação dos objetos são técnicos administrativos, professores e seus alunos de pós-graduação e graduação, demonstrando que “grande parte de museus e curadorias de pequeno e médio porte concentra várias especialidades em um número reduzido de profissionais e, não raro, em um único museólogo” (Bottallo, 2010, p. 48) ou nenhum profissional graduado em Museologia. Além disso, os museus M1, M2, M3 e M4 não possuem sistemas de informação, porém, o M1 utilizou o Sistema Donato do Museu de Belas Artes, mas o uso foi interrompido no ano de 2016. A partir da análise realizada, o elemento identificado nesta categoria de acordo com Fujita e Santos (2016) e Rubi (2012), foi o de cobertura de assunto.

Na categoria da Documentação museológica, o sujeito M1 realiza algumas etapas da documentação museológica, como a leitura documentária e registro do objeto em ficha catalográfica, utilizando o sistema Donato do Museu de Belas Artes para consulta de vocabulários controlados e determinação da linguagem documentária, sendo que os termos utilizados são gerais e não específicos. Os sujeitos M3 e M4 não usam vocabulário controlado, tampouco realizam a documentação museológica, apesar de o M4 saber da necessidade da utilização de linguagens documentárias. O M2 utiliza uma ficha de registro dos objetos que é preenchida pelo pesquisador durante a doação do objeto ao museu, não caracterizando a documentação museológica. Por isso, os sujeitos entrevistados relatam que os museus não possuem diretrizes ou manuais que guiam a prática da documentação museológica. Deste modo, foi possível identificar o uso de linguagem documentária (Kobashi, 2007), na documentação museológica realizada no M1 é considerada um dos elementos da política de indexação (Fujita & Santos, 2016).

Na categoria de Perfil do usuário procurou-se identificar as necessidades recorrentes dos usuários dos Museus da UFPA, já que a representação da informação é feita para recuperação, tornando “o usuário um pré-requisito para o planejamento de qualquer sistema de informação” (Carneiro, 1985, p. 223), que na Museologia a formulação de “catálogos e estratégias que possibilitam o usuário realizar uma busca direta” (Reis et al., 2019, p. 99). O M1 narra que o serviço ao usuário deve ser solicitado pelo mesmo, uma vez que a consulta de usuário será avaliada pelo museólogo responsável para verificar a disponibilidade de realização da consulta, já que as consultas de usuários dependem do estado de conservação dos objetos, caracterizando uma busca delegada. E o público visitante são alunos e professores da universidade, prevalecendo usuários da área de Museologia e Artes Visuais por realizarem mais pesquisas nas dependências do museu. O M2 relata que nenhum usuário da comunidade acadêmica ou externa solicitou consulta ou empréstimo de objetos do seu acervo, mas o museu recebe visitas da comunidade acadêmica e do público externo, principalmente, de escolas públicas e privadas, e as visitas devem ser agendadas com antecedência. O M3 recebe demandas para empréstimo das suas coleções para exposição externa, para isso necessita de um agendamento e envio de um ofício por *e-mail* para reserva da coleção. E o sujeito M4 conta que a instituição tem dois públicos, a comunidade acadêmica dos cursos da área da saúde e o público externo que são alunos de escolas públicas e privadas. Notou-se que um dos museus realiza a busca delegada, dado que a estratégia de busca realizada nos espaços museológicos é um elemento da política de indexação, que pode ser delegada ao funcionário da instituição pelo usuário ou o próprio usuário realiza a busca no acervo, além de identificar a clientela das instituições museológicas (Fujita & Santos, 2016; Rubi, 2012).

Análise do Protocolo Verbal Individual

A primeira categoria de análise foi a Interpretação de informações intrínsecas e extrínsecas, que diz respeito à interpretação de informações relativas ao próprio objeto (informações intrínsecas) ou exteriores a ele (informações extrínsecas) pelo profissional (Padilha, 2014) realizando, assim, a leitura documentária.

Observa-se que o M1 realiza a leitura documentária antes de inserir as informações no seu catálogo:

M1: *“Tem as informações intrínsecas e extrínsecas da obra, então a gente parte primeiro das informações intrínsecas [...] a gente olha para obra, vê as características da obra, da técnica dessas obras”.*

O processo de leitura documentária não é feito pelo sujeito M2, sendo responsabilidade do próprio doador do objeto. No entanto, o M2 revela que, quando necessário, realiza a leitura documentária buscando informações relativas ao próprio objeto para fazer sua descrição em seus catálogos e também informações extrínsecas:

M2: *“[...] a pessoa que doou não sabe o nome da amostra, não é um geólogo, essa amostra não faz parte de uma pesquisa, a pessoa não sabe classificar*

essa amostra, a gente mesmo faz essa classificação. A maioria das amostras macroscopicamente a gente consegue, através de uma descrição simples, a gente já consegue classificar, mas às vezes é necessário também levar para o laboratório e detalhar informações principalmente sobre a composição mineralógica desse material aí que foi doado. Então é uma ferramenta a mais que a gente tem de dispor desses laboratórios que a gente consegue... a gente consegue... caracterizar o nosso material e classificar da melhor forma possível”.

O M3 realiza a leitura documentária do objeto museológico para colocá-lo em coleções específicas dentro do acervo:

M3: “[...] as informações já vão também estar... anotadas a princípio é um caderninho que a gente faz um caderninho de campo então coloca lá no caderninho campanha tal data tudo direitinho, aí animal um daquela campanha tá aí vem todas as informações dela [...] Este caderninho posteriormente que é o que vai ter as informações dele nessa planilha Excel”.

O M4 não realizou a leitura documentária.

A segunda categoria foi a Indexação na documentação museológica, que diz respeito à identificação e seleção de conceitos a partir da análise do objeto para determinar seu assunto (Lancaster, 2004; Reis et al., 2019).

O M1 realizou algumas etapas para a indexação. A primeira foi a leitura documentária do item, na qual identifica os assuntos, a partir da sua leitura imagética e contexto em que foi produzido como pode ser observado em sua fala:

M1: “A descrição da obra a gente tenta fazer a descrição do que nos é representado na imagem [...] a descrição do catálogo é baseada nisso e esse é um primeiro momento também que é essa descrição”.

A segunda foi a identificação e seleção de conceitos, no qual definiu os termos mais específicos para representar o conteúdo daquele objeto em seu catálogo (variável especificidade):

M1: “Aqui usamos termos específicos para as técnicas das obras”.

Estas duas etapas caracterizam o processo de indexação, pois é o principal elemento da política de indexação (Fujita, 2003; Fujita & Santos, 2016). Entretanto, o M1 não demonstrou preocupação com o usuário na escolha desses termos, lembrando que o usuário é um requisito importante para definição dos conceitos (Lancaster, 2004) e para elaboração da política de indexação (Fujita & Santos, 2016).

Os sujeitos M2 e M3 realizaram apenas a leitura documentária do objeto, mas não identificaram e selecionaram conceitos para uma possível recuperação. Apesar de não selecionar conceitos, o M3 agrupa seus objetos por grupos taxonômicos:

M3: “Não, não separamos por assuntos, mas por grupos taxonômicos, como te falei, mamíferos, répteis, insetos, aí isso tá separado”.

O sujeito M4 não realiza nenhuma das etapas para indexação dos objetos.

A terceira categoria foi o Uso de linguagem documentária, a qual considerou a utilização de um ou mais vocabulários controlados na inserção das informações nos catálogos de cada museu.

Observou-se que apenas o M1 utiliza um vocabulário controlado no seu processo de inserção das informações nos catálogos da instituição. O vocabulário que o M1 utilizou é o mesmo empregado no Sistema Donato do Museu Nacional de Belas Artes na documentação do seu acervo, e apesar do sistema ter sido desativado, o vocabulário ainda é utilizado pela instituição, segundo relatou o profissional durante a aplicação do PVI. Assim, verificou-se o uso de linguagem documentária (Kobashi, 2007), considerado um dos elementos da política de indexação (Fujita & Santos, 2016).

Os sujeitos M2, M3 e M4 não utilizam nenhum tipo de linguagem controlada ao fazer seu processo de documentação museológica. O M2 acredita que o uso de termos controlados não é necessário devido ao tipo de acervo que possui:

M2: “E aí como eu comentei anteriormente, o número que ela vai receber vai seguido pelo nome do mineral ou nome da rocha, não tem outros termos... por que a amostra faz parte de um mesmo contexto...”

A última categoria foi a Catalogação do acervo, no qual observou-se a existência de uma padronização da catalogação no processo de documentação em cada museu. Durante o relato do M1 verificou-se a ausência de uma padronização, suas fichas catalográficas variam de acordo com o que o profissional julga necessário:

M1: “[...] então nossos catálogos, eles sempre são... ele sempre tem mudanças, eles nunca seguem o mesmo padrão”.

No M2 esse processo, apesar de padronizado, não é realizado pelo profissional, mas de responsabilidade do doador daquele item para a coleção do museu:

M2: “[...] a ficha que vai ser entregue ao doador, onde ele vai preencher todas as informações referentes àquela mostra que ele está doando. Essas fichas nós já temos impressas”.

No entanto, o M2 realiza o trabalho de catalogação quando julga necessário:

M2: “[...] às vezes o doador ele simplesmente não quer fazer isso. Ele só preenche mesmo, nome, a data de coleta e assina, não tem interesse ou por não saber, ou tá com pressa”.

No M3 verificamos a mesma ausência de padronização das fichas de catalogação observados no M1:

M3: “Agora uma das coisas que eu já percebi que, como sou eu ou algum um aluno que fazem, não tem sistema geral, as minhas tabelas elas estão meio sabe... tem uma tabela que está com as colunas todas de um jeito e outra de eu faço um jeito, não tá padronizado, para cada localidade que eu faço eu faço de um jeito”.

Observou-se, também, que os sujeitos M2 e M3 não possuem a mesma habilidade na realização do processo de catalogação que o M1. O M4 não realiza nenhum tipo de catalogação ou registro para suas peças, porém relatou sentir a necessidade de uma catalogação durante a entrevista:

M4: “Não, catalogação não”.

Dessa forma, entende-se a política de indexação como uma ferramenta que possibilita uma recuperação por assuntos no contexto museológico, visto que permite direcionar a prática da catalogação com vistas para indexação e controle do vocabulário, objetivando o acesso, a recuperação, a localização de forma rápida e fácil das informações dos objetos museológicos por profissionais e usuários.

Conclusões

O problema levantado para o desenvolvimento desta pesquisa se torna ainda mais relevante, pois a ausência de diretrizes dentro dos museus universitários da UFPA acarreta na falta de uniformidade na realização do processo de catalogação do seu acervo, o que resulta em uma lacuna na recuperação da informação por possíveis usuários. Ainda sobre isso, nota-se que os profissionais não possuem total domínio da prática da documentação museológica, sendo pertinente o planejamento de diretrizes, como uma política de indexação, para guiar os afazeres do processo de catalogação das coleções museológicas.

Contudo, apesar da ausência de diretrizes, identificou-se requisitos e elementos de uma política de indexação a partir da análise dos dados obtidos, a saber: cobertura de assunto, linguagem documentária, processo de indexação na documentação museológica, estratégias de busca e a clientela da instituição. Nesse sentido, apesar de detectarmos aspectos importantes de uma política que se assemelha a política de indexação existente em bibliotecas, não há uma sistematização, na forma de um documento dentro da instituição, que dê respaldo para essa atividade.

É importante destacar que a política de indexação pode ser uma ferramenta importante para o tratamento informacional do acervo e, conseqüentemente, para facilitar a posterior recuperação, acesso e uso da informação. Dito isso, mesmo constatada a relevância do tema, ainda se observa a escassez de publicações sobre a temática. Nessa perspectiva, julgamos necessário o aprofundamento das pesquisas sobre o assunto, tendo em vista a construção de uma política de documentação museológica que possa abranger o processo de indexação na catalogação de objetos museológicos.

Referências

Alonso Arroyo, A., Fujita, M. S. L., Gil Leiva, I., & Padiella, A. (2016). Protocolo verbal: análisis de la producción científica, 1941-2013. *Informação & Sociedade*, 26(2), 61-76. <https://digitum.um.es/digitum/bitstream/10201/102026/1/ArroyoSpot..a.pdf>.

- Boccatto, V. R. C. (2012). A linguagem documentária em catálogos on-line para política de indexação. In I. Gil Leiva & M. S. L. Fujita (Eds.), *Política de indexação* (pp. 139-151). Cultura Acadêmica, Oficina Universitária.
- Bottallo, M. (2010). Diretrizes em documentação museológica. In *Documentação e conservação de acervos museológicos: diretrizes* (pp. 48-78). Secretaria de Estado da Cultura de São Paulo, ACAM Portinari.
- Carneiro, M. V. (1985). Diretrizes para uma política de indexação. *Revista da Escola de Biblioteconomia da UFMG*, 14(2), 221-241.
- Carvalho, R. A., Melo D. J., & Linhares, A. P. (2015). Documentação Museológica e Paleontologia: Elaboração de Ficha de catalogação Para o Acervo de Paleoinvertebrados do Museu Paraense Emílio Goeldi. In: *Anais*, 3, 387-406.
- Cintra, A. M. M., Tálamo, M. de. F. G. M., Lara, M. L., G. de., & Kobashi, N. Y. (2002). *Para entender as linguagens documentárias* (2nd ed.). Polis.
- Desvallées, A., & Mairesse, F. (ed). (2013). *Conceitos-Chave de Museologia*. Comitê Internacional para Museologia do ICOM; Comitê Nacional Português do ICOM.
- Dias, E. W., & Naves, M. M. L. (2013). *Análise de assunto: teoria e prática* (2nd ed.) Briquet Lemos.
- Ferrez, H. D. (1994). *Documentação museológica: teoria para uma boa prática*. Minc/Iphan.
- Fujita, M. S. L. (2009). A técnica introspectiva e interativa do Protocolo Verbal para observação do contexto sociocognitivo da indexação na catalogação de livros em bibliotecas universitárias: aplicação e análise. In M. S. L. Fujita (Org.), *A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias. Um estudo de observação do contexto sociocognitivo com protocolos verbais* (pp. 51-79). Unesp, Cultura Acadêmica.
- Fujita, M. S. L., & Santos, L. B. P. (2016). Política de indexação em bibliotecas universitárias: estudo diagnóstico e analítico com pesquisa participante. *Transinformação*, 28(1), 59-76.
- Fujita, M. S. L. (2003). A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação*, 1(1), 60-90.
- Hjorland, B. (2018). Indexing: concepts and theory. *Knowl. Org.* 45(7), 609-639.
- Kobashi, N. Y. (2007). Fundamentos semânticos e pragmáticos da construção de instrumentos de representação de informação. *DataGramaZero*, 8(6).
- Lancaster, F. W. (2004). *Indexação de livros: teoria e prática* (2nd ed.). Briquet de Lemos.
- Mensch, P. V. (1994). *O objeto de estudo da museologia*. UNIRIO/UGF.
- Monteiro, J. (2010). Diretrizes teórico-metodológicas do projeto. In *Documentação e conservação de acervos museológicos: diretrizes* (pp. 30-46). Secretaria de Estado da Cultura de São Paulo, ACAM Portinari.
- Padiilha, R. C. (2014). *Documentação Museológica e Gestão de Acervo*. FCC.
- Redígolo, F. M. (2014). *O processo de análise de assunto na catalogação de livros em bibliotecas universitárias: aplicação do protocolo verbal*. [Tese de Doutorado, Universidade Estadual Paulista].
- Reis, M. S., Morais, J. B. E., Barros, T. H. B., & Santos Júnior, R. L. (2019). Análise documental na museologia: um breve exercício de identificação de conceitos. *Ci. Inf. Rev.*, 6(2), 90-102.
- Rubi, M. P. (2008). *Política de indexação para construção de catálogos coletivos em bibliotecas universitárias*. [Tese de Doutorado, Universidade Estadual Paulista].
- Rubi, M. P. (2012). Política de indexação. In I. Gil Leiva, & M. S. L. Fujita (Eds.), *A política de indexação* (pp. 107-120). Cultura Acadêmica, Oficina Universitária.

(Página deixada propositadamente em branco)

ANÁLISE DA PRODUÇÃO E DA COLABORAÇÃO CIENTÍFICA NA REVISTA KNOWLEDGE ORGANIZATION (2016-2020)

Ana Beatriz Silva¹, Maria Cristina Vieira de Freitas²

¹Doutoranda em Ciência da Informação, Faculdade de Letras, Universidade de Coimbra, silvaanabeatriz024@gmail.com, ORCID iD 0000-0001-5704-4352

²Centro de Estudos Interdisciplinares do Século XX, Faculdade de Letras, Universidade de Coimbra, cristina.freitas@fl.uc.pt, ORCID iD 0000-0002-8849-8792

Resumo

Neste artigo, é realizado um estudo da produção científica em Organização do Conhecimento, no período 2016-2020, divulgada na *Knowledge Organization*, revista de topo da área, intentando identificar os temas mais recorrentes, os/as autores/as mais produtivos/as e mais citados/as, bem como as possíveis redes de coautoria e de colaboração científica. O levantamento e a análise de dados realizam-se manualmente, pela base de dados *Web of Science*[®], ou de forma automatizada, neste caso com o apoio dos *softwares Ucinet*[®] e *Netdraw*[®]. O *corpus* da pesquisa é constituído pelos 175 artigos científicos e pelos 288 autores/as identificados. Os/as 11 autores/as com um maior número de artigos publicados nesta revista, no período de referência, representam 4% do total. Os temas mais frequentes dos artigos não são identificáveis, num nível de granularidade fina, a partir da recolha das palavras-chave ou das palavras relevantes dos resumos ou dos títulos. Quanto às citações, atingiu-se a contagem de 13 autores/as (5%) considerados/as mais citados/as, pelo facto de apresentarem um total de entre sete e 35 citações. Quanto aos/as autores/as mais produtivos/as e mais citados/as, verificou-se que B. Hjørland, R. P. Smiraglia e C. Gnoli estão presentes em ambos os casos, o que indica uma forte influência teórica destes autores nesta revista, no período em estudo. As colaborações entre autores/as foram identificadas e da análise da rede obtida constatou-se que apenas E. Ménard (0,3%) apresentava o maior número de coautorias (11), no período em questão. Conforme se constata, alguns/umas dos/as autores/as mais citados/as do *corpus* de pesquisa são também considerados/as os/as mais produtivos/as do período, não sendo, no entanto, necessariamente os/as mais colaborativos/as, ou seja, os/as que realizaram maiores ligações com outros/as autores/as. Conclui-se que este estudo pode servir como um piloto para futuras investigações, alargando-se a amostra, uma vez que os procedimentos metodológicos nele adotados são transferíveis a outros contextos.

Palavras-chave: Organização do conhecimento, Produção científica, Web of Science, Revistas científicas, Colaboração científica.

Introdução

Num estudo realizado há alguns anos, Guimarães (2014) demonstrou que as publicações veiculadas na revista *Knowledge Organization (KO)* evidenciavam alguma preocupação com a definição desta área de estudo ou domínio científico – Organização do Conhecimento (OC) – quer ao nível epistemológico, quer nos aspetos teórico-metodológicos. Hjørland (2008), cujos estudos sobre essas matérias podem ser acompanhados e tracejados, nas últimas duas décadas, numa publicação realizada nesta mesma revista (*KO*), definiu a OC em dois sentidos: i) estrito, referindo-se, sobretudo, a processos e atividades realizadas na *Web* ou nas tradicionais instituições de informação; ii) amplo, referindo-se a um campo ou domínio científico. Smiraglia (2011), também numa publicação realizada nesta mesma revista (*KO*), percecionou a OC como um subdomínio da Ciência da Informação, relacionado ao processo de construção do conhecimento nos seus aspetos concetuais.

O crescente aumento da produção científica numa dada área pressupõe o aumento da complexidade dos assuntos abordados, tornando pertinente a realização de uma avaliação periódica, de modo a aferir os/as principais autores/as e as contribuições mais significativas (Meschini, & Oliveira, 2017). Assim, a produção científica, enquanto resultado e expressão do trabalho realizado numa dada área, torna-se num importante parâmetro a usar para obter os indicadores que nos informam, direta ou indiretamente, sobre a quantidade e a qualidade do trabalho produzido e que é medido, entre outros aspetos, pela sua maior ou menor visibilidade nos canais usuais de publicação ou pelo seu impacto no meio científico (de que são exemplos o número de citações e a presença em redes de colaboração). Dada a sua evidente visibilidade, um dos meios mais seguros e eficazes para atingir o desejável impacto dos trabalhos produzidos é, pois, a publicação dos resultados em revistas científicas consideradas de topo (Freitas et al., 2014). Não obstante, mapear a produção e a colaboração científica implica selecionar e justificar perspetivas, métodos e técnicas, porque os resultados podem ser diferentes consoante as opções metodológicas adotadas. *Grosso modo*, implica identificar os assuntos subjacentes à área em estudo, bem como os/as autores/as mais produtivos/as e que, em boa medida, são os/as que oferecem um maior contributo para o seu crescimento.

Partindo dos pressupostos prévios, neste artigo são apresentados os resultados de um micro-estudo realizado para realçar aspetos quantitativos e qualitativos da produção científica em OC, recorrendo aos artigos publicados na revista *Knowledge Organization (KO)*, entre os anos de 2016 e 2020. Nesta que é considerada a revista de topo da área, pretende-se identificar os/as autores/as mais produtivos/as, bem como os principais temas abordados nos artigos publicados. Também se pretende identificar os/as autores/as mais citados/as e as coautorias entre autores/as mais produtivos/as.

É de salientar a existência de estudos prévios que também se debruçaram sobre esse tema. Não obstante, conforme se verificou, o recorte temporal de alguns desses estudos limita-se ao ano de 2016. Assim, este artigo se centra na análise de um *corpus* de pesquisa constituído num recorte temporal posterior (2016-2020), o que lhe confere atualidade, originalidade e interesse.

Metodologia

A presente pesquisa é de caráter quantitativo-qualitativo e, dado o seu pendor bibliométrico, apresenta uma inclinação analítico-descritiva, desenvolvendo-se em dois momentos, conforme o que se explica a seguir.

No **primeiro momento**, que correspondeu à fase de construção da problemática e de recolha dos dados, realizou-se, inicialmente, uma pesquisa bibliográfica, com o intuito de identificar estudos prévios sobre o assunto. Posteriormente, e já com uma ideia clara dos objetivos do estudo, efetivou-se um levantamento dos artigos publicados na revista *KO*, nos últimos cinco anos (2016-2020), de modo a propiciar uma análise das tendências a verificar-se no período. Esta revista foi selecionada em função da sua centralidade na abordagem do assunto. As pesquisas foram realizadas em todos os artigos que se encontravam indexados na base de dados *Web of Science® (WoS)*¹. Para atingir estas publicações, usamos os filtros necessários (título da publicação, intervalo cronológico e tipo de documento). As buscas devolveram um total de 176 itens, que foram analisados, tendo em conta os seguintes metadados: título, subtítulo, resumo, palavras-chave e títulos das secções. Após esta análise preliminar, descartamos um dos resultados, considerado redundante, tendo restado 175 artigos, que passaram a constituir o *corpus* da pesquisa. A distribuição dos artigos publicados pelos anos de publicação é a que se expressa na tabela 1.

Tabela 1. Distribuição dos artigos recuperados por ano de publicação

Ano de publicação	Número de artigos	%
2020	22	12,6
2019	33	18,9
2018	30	17,1
2017	40	22,8
2016	50	28,6
n=	175	100,0

Fonte: elaboração nossa, com base no *corpus* de pesquisa.

A seguir às buscas, identificaram-se os principais temas abordados nos 175 artigos, através da recuperação manual das palavras-chave presentes em cada um dos mesmos e, na sua ausência, a partir da seleção das palavras consideradas mais pertinentes, identificadas nos resumos ou, ainda, nos títulos, nos casos em que se verificou a ausência dos resumos. Ao finalizar esse procedimento, chegou-se a uma conta de 19 palavras selecionadas, contabilizando-se as respectivas frequências de ocorrência, que, no total, perfizeram 197².

Ainda, foram identificados/as todos/as os/as autores/as que assinavam os textos, bem como a frequência de publicação, de modo a aferir a sua produtividade, exclusivamente, no contexto deste estudo. Estes/as autores/as foram identificados/as

¹ Todas as consultas e recolhas foram realizadas no período de 06/01/2021 a 22/01/2021.

² Estes dados serão apresentados na secção destinada aos resultados.

usando, primeiramente, o filtro “autores” da base de dados *WoS*[®], cuja distribuição é feita pelo número de publicações (em ordem crescente ou decrescente). Seguidamente, confirmou-se a sua presença, manualmente, em cada um dos 175 artigos que compunham o *corpus* de pesquisa, o que permitiu, inclusive, atingir um número expressivo de 288 autores/as³, tendo sido considerados/as mais produtivos/as, para os efeitos pretendidos, aqueles/as que possuíam entre três e cinco publicações na revista (*KO*), no período de referência, num total de 11(3,8%). Seguindo a mesma lógica, os/as demais autores/as, que perfizeram uma maioria esmagadora (277; 96,2%), foram incluídos/as, consoante as suas frequências de publicação, em dois grupos adicionais (tabela 2). O mesmo filtro foi utilizado para recuperar todos os artigos publicados por esses/as autores/as considerados mais produtivos/as.

Desta forma e dado que o maior número de publicações de um/a só autor/a em todo o período foi de apenas cinco, sendo somente 39 (11,3%) o total de artigos incluídos nesse grupo “seleto” de autores/as (tabela 2), tornou-se exequível, no curso deste estudo, identificar as coautorias e as colaborações entre esses/as autores/as. Este procedimento foi realizado em todos os artigos publicados pelos autores/as considerados/as mais produtivos/as.

Tabela 2. Distribuição dos/as autores/as pelas frequências de publicação

Grupo	N.º de autores/as incluídos/as	%	N.º de publicações no período	Frequência de publicação por grupo	%
1	11	3,8	Entre 3 e 5	39	11,3
2	30	10,4	2	60	17,3
3	247	85,8	1	247	71,4
	288	100,0		346	100,0

Fonte: elaboração nossa, com base no *corpus* de pesquisa.

Para a recolha e a contagem de citações teve-se como base o número fornecido pela *WoS*[®], em cada artigo, nos resultados. Estes dados foram confrontados com outros, provenientes de análise manual, por meio da qual foram sendo verificadas as referências bibliográficas introduzidas em cada um dos 175 artigos incluídos no *corpus* de pesquisa, tendo sido apontadas as menções feitas para cada um/a dos/as autores/as, chegando-se a uma conta de 288. Estes/as autores/as foram divididos em dois grupos, consoante o número de citações apuradas, que variaram entre 0 e 35. Desta forma, ficou-se a saber qual era o número total de citações recebidas por esses/as autores/as, no período, bem como a diferença entre o número de artigos publicados e o número de artigos citados, no mesmo período, tendo sido, a partir desses dados, calculadas as médias de citação dos/as autores/as (cf. tabela 7, adiante).

No **segundo momento**, que correspondeu à etapa de análise dos dados, recorreu-se a programas informáticos. Deste modo, o Microsoft Excel[®] foi utilizado para a elaboração da matriz de colaborações/coautorias e o *Ucinet 6 for Windows: software for Social Network Analysis*[®] (Borgatti et al., 2002) foi utilizado para a mensuração

³ Sumariamente identificados/as no Apêndice I.

das redes de coautoria, tendo servido como ferramenta no cálculo do indicador grau de centralidade das redes de coautoria dos/as autores/as mais produtivos/as, tendo-se verificado, por este procedimento, que oito autores/as ocupavam a posição central. O *NetDraw: Graph Visualization Software*® (Borgatti et al., 2002), incorporado ao *Ucinet*®, foi utilizado para a visualização das redes de colaboração. Neste caso, foram visualizadas, em grafo, nove redes de colaboração (cf. figura 1, adiante). A opção por estas ferramentas prendeu-se com o seu uso em estudos similares (Araújo, & Guimarães, 2017; Amorim Neto et al., 2016; Lima, 2015; Bufrem et al., 2014).

É de referir que os resultados obtidos nas etapas supramencionadas foram plas-mados em tabelas e em figuras (grafos de redes), para a sua ulterior análise e discussão.

Resultados e discussão

1) Estudos prévios

São vários os estudos que nos dão a conhecer aspetos relevantes da produção científica em OC, em língua portuguesa (Souza, 2013; Bufrem et al., 2014; Lima, 2015; Araújo, & Guimarães, 2017; Alves et al., 2019). Souza (2013), por exemplo, realizou uma análise bibliométrica da atividade científica neste campo, a partir de textos de tipologias diversas, indexados na *WoS*® e publicados entre os anos de 1979 e 2006. Bufrem et al. (2014) também realizaram uma análise da produção científica em OC, por meio de artigos selecionados na Base de Dados Referenciais de Artigos de Periódicos de Ciência da Informação (BRAPCI®), tendo como período de referência os anos de 2003 e 2012, organizando os resultados em matrizes de citações e cocitações, apoiando-se no programa Excel®. Lima (2015) efetuou uma análise de domínio em OC, por meio de procedimentos bibliométricos, a partir das publicações indexadas na revista *KO*, recuperadas em *Scopus*®, entre os anos de 2004 e 2013. Também elaborou uma matriz de citações e cocitações, apoiando-se no Excel® e recorre, ainda, ao *software Ucinet*® para auxiliar a construção da rede de cocitações. Araújo e Guimarães (2017), por sua vez, realizam uma análise de citações em artigos científicos que se debruçam sobre a epistemologia da OC e que se encontram indexados na *WoS*® e na *Scopus*®, sem recorrer a filtros temporais. Alves et al. (2019) também analisam artigos científicos publicados na revista *KO*,

Tabela 3. Estudos prévios sobre a produção científica em OC

Autores/as	Procedimentos	Fontes de dados	Períodos cronológicos
Souza (2013)	Análise da atividade científica em OC	<i>WoS</i> ®	1979-2006
Bufrem et al. (2014)	Análise da produção científica em OC	BRAPCI®	2003-2012
Lima (2015)	Análise de domínio em OC	<i>KO</i> / <i>Scopus</i> ®	2004-2013
Araújo e Guimarães (2017)	Análise de citações em OC	<i>WoS</i> ® / <i>Scopus</i> ®	-
Alves et al. (2019)	Análise de artigos científicos (“Elite de pesquisa”)	<i>KO</i> / <i>Scopus</i> ®	1993-2016

Fonte: elaboração nossa, com base nos/as estudos citados.

entre os anos de 1993 a 2016, e indexados em Scopus®, para identificar uma “Elite de Pesquisa”, bem como as suas instituições de origem e mapear as relações de coautoria entre investigadores/as. A modo de resumo, identificamos, na tabela 3, os principais dados desses estudos prévios.

Dos estudos anteriormente descritos (tabela 3), depreende-se que a maioria se centra na análise da produção ou atividade científica, por meio do estudo de citações e cocitações, utilizando como fontes as bases de dados multidisciplinares WoS® e Scopus®, para além do agregador de conteúdos BRAPCI®, juntamente com a revista *KO*. Estes estudos apresentam, pois, o panorama da OC em fontes internacionais, credíveis e especializadas, numa cobertura cronológica de quase 40 anos (entre 1979 e 2016, pelo menos). O nosso estudo pretende focalizar as recolhas e as análises de dados nos/as autores/as, nos temas e nas coautorias e colaborações, num recorte temporal mais recente (2016-2020), circunscrevendo-se a uma única fonte de dados (*KO*) e uma única tipologia documental (artigos científicos), por meio da consulta de uma única base de dados (WoS®).

2) Análise da produção científica identificada na revista *KO* (2016-2020)

2.1) Principais temas abordados nos artigos

Evidentemente, numa revista com o perfil da *KO*, espera-se que as palavras mais utilizadas para sugerir os temas genéricos presentes nos textos recuperados se relacionem diretamente com os processos de classificação (63 ocorrências; 32%), com a própria organização do conhecimento (50 ocorrências; 25%), com as ontologias (19 ocorrências, 10%) e os sistemas de organização do conhecimento (SOC) (8 ocorrências; 4%). Quanto a este último tema, verifica-se, sem grande surpresa, que está presente noutras circunstâncias (por exemplo: tesouros, folksonomias, ontologias, taxonomias), quando é tratado, nos textos, de um modo mais específico. Na tabela 4, apresentam-se estas palavras, recolhidas dos metadados dos artigos, em função da sua frequência de ocorrência.

**Tabela 4. Frequência de ocorrência de palavras nos artigos recuperados.
Período de referência: 2016-2020 (n=175)**

Item	Palavras recuperadas (palavras-chave, resumo ou título)	Ocorrência	%
1	Classificação	63	32
2	Organização do Conhecimento	50	25
3	Ontologias	19	10
4	Sistemas de Organização do Conhecimento	8	4
5	Tesouros	7	4
6	Classificação facetada	5	3
7	Classificação Decimal Universal	5	3
8	Folksonomias	5	3
9	Recuperação da informação	5	3
10	Análise de domínio	4	2
11	Classificação Decimal de Dewey	4	2
12	Epistemologia	4	2
13	Listas de encabeçamento de matérias	3	2

14	Taxonomias	3	2
15	Terminologia	3	2
16	Web semântica	3	2
17	Classificação bibliográfica	2	1
18	Classificação em arquivos	2	1
19	Sistemas de classificação	2	1
n=19		197	100

Fonte: elaboração nossa, com base no *corpus* de pesquisa.

Assim, se nos fixarmos bem nos resultados apresentados na tabela 4, muito facilmente perceberemos que nem as palavras-chave nem as palavras relevantes dos resumos ou dos títulos foram capazes de oferecer, neste caso em particular, pistas dos assuntos abordados nos textos num nível de granularidade fina, pelo que, e para a sua identificação, teria sido necessária a realização de uma análise documental dos conteúdos dos textos, o que fugiria ao objetivo principal deste estudo.

2.2) Autores/as mais produtivos/as

Verificadas as autorias e coautorias, e as respetivas frequências de publicação, constatou-se que dos/as 288 autores/as que publicaram artigos na revista *KO*, no período definido (2016-2020), uma maioria esmagadora contribuiu com apenas uma publicação (247; 71,4%), tendo, ainda, outros/as 30 autores/as (17,3%) contribuído com dois artigos. Apenas uma minoria (11; 11,3%) apresentou três ou mais publicações. Na tabela 5, resumimos esses dados⁴.

Tabela 5. Autores/as: frequência de publicação. Período de referência: 2016-2020. (n=288)

Frequência de publicação	N.º de autores/as envolvidos/as	%
5	2	2,9
4	2	2,3
3	7	6,1
2	30	17,3
1	247	71,4
n=	288	100,0

Fonte: elaboração nossa, com base no *corpus* de pesquisa.

Com base numa leitura direta dessas frequências de publicação, esse grupo minoritário de autores/as, presente no *corpus* de pesquisa com entre três e cinco publicações (tabela 5), foi considerado, no contexto deste estudo, como o grupo dos autores/as mais produtivos/as (tabela 6).

⁴ A tabela completa encontra-se no Apêndice I deste estudo.

Tabela 6. Autores/as mais produtivos/as. Período de referência: 2016-2020 (n=346)

Item	Autores/as (Afiliações, Países)	N.º de artigos publicados	%
1	Hjørland, B. (University of Copenhagen, Denmark)	5	1,4
2	Smiraglia, R.P. (Inst. Knowledge Org. & Struct., USA)	5	1,4
3	Gnoli, C. (University of Pavia, Italy)	4	1,1
4	Martínez-Ávila, D. (Universidade Estadual Paulista, Brazil)	4	1,1
5	Das, S. (University of Trento, Italy)	3	0,9
6	Guimarães, J.A.C. (Universidade Estadual Paulista, Brazil)	3	0,9
7	Hudon, M. (University of Montreal, Canada)	3	0,9
8	Kwasnik, B.H. (Syracuse University, USA)	3	0,9
9	Macfarlane, A. (City University London, Great Britain)	3	0,9
10	Menard, E. (McGill University, Canada)	3	0,9
11	Wei, W. (Zhongnan University of Economics & Law, China)	3	0,9
n=11		n=39	11,3

Fonte: elaboração nossa, com base no *corpus* de pesquisa e na base de dados WoS®.

Conforme se observa na tabela 6, destacam-se como autores com maior produção, no período, por um lado, B. Hjørland e R.P. Smiraglia, com totais individuais de cinco artigos (2,8%) publicados e, por outro, C. Gnoli e D. Martínez-Ávila, com totais individuais de quatro artigos (2,2%) publicados. Relativamente às afiliações recolhidas, estas sugerem uma prevalência, no *corpus* de pesquisa, de autores/as afiliados/as a instituições norte-americanas (4; 36,4%) ou, de igual modo, europeias (4; 36,4%), sendo estes/as seguidos/as por colegas com afiliação a instituições sul-americanas (2; 18,2%). Apenas um (9%) desses/as autores/as produtivos/as tem como afiliação informada uma instituição asiática. Nenhum desses resultados foge à expectativa; ao contrário, coadunam-se, em boa medida, com o perfil da revista (*KO*) e da comunidade discursiva da área que representa (neste caso, OC).

Ainda, nesse mesmo conjunto de autores/as mais produtivos/as no período, verificou-se que uma boa maioria (7; 63,6%) – aqui representada por B. Hjørland, C. Gnoli, M. Hudon, B.H. Kwasnik, R.P. Smiraglia, A. Macfarlane e D. Martínez-Ávila – apresentou artigos em autoria individual – exclusivamente ou inclusivamente, dependendo do caso, – sendo as frequências de ocorrência observadas, respetivamente: 5, 3, 3, 2, 2, 1 e 1.

2.3) Autores/as mais citados/as nos artigos

Dos 288 autores/as presentes no *corpus* de pesquisa, verificou-se, após a contagem manual realizada, que pouco mais da metade (146; 50,7%) apresentava, excluindo-se as autocitações, um número entre uma e 35 citações. Para efeitos de análise, esta discretíssima maioria foi então dividida em dois grupos. Num **primeiro grupo**, foram incluídos/as os/as 133 autores/as (91,1%) que apresentavam, no período em apreço, entre uma e seis citações, distribuídas do seguinte modo: 48 autores/as (36,0%) com apenas uma citação; 38 autores/as (28,6%) com duas citações;

25 autores/as (18,8%) com três citações; três autores/as (2,3%) com quatro citações; 12 autores/as com cinco citações (9,0%) e, finalmente, sete autores/as (5,3%) com seis citações. Num **segundo grupo**, muitíssimo mais discreto, foram incluídos/as os/as 13 autores/as (8,9%) considerados/as mais citados/as, no mesmo período, tendo como critério o facto de apresentarem entre sete e 35 citações. A distribuição dos dados, neste grupo “seleto”, que representa menos de 10% dos/as autores/as citados/as, foi então realizada (tabela 7). É de referir que nem todos os artigos publicados por estes/as autores/as foram citados, razão pela qual optamos por indicar, a título de referência e para o cálculo da média de citações, do conjunto individual dos artigos publicados no período, quantos foram efetivamente citados (tabela 7). Ainda, refira-se, sem grandes surpresas, o que a leis bibliométricas usualmente sugerem, ou seja, a prevalência de muitos/as autores/as com poucas citações e de poucos/as autores/as com muitas citações.

**Tabela 7. Autores/as mais citados/as e distribuição das citações
(período de referência: 2016-2020)**

Item	Autores/as	N.º de citações recebidas	N.º de artigos publicados	N.º de artigos citados	Média de citação por artigo
1	Hjørland, B.	35	5	4	8,8
2	Fox, M.J.	13	2	2	6,5
3	Gnoli, C.	11	4	4	2,8
4	Adler, M.	10	1	1	10
5	Bowker G.C.	10	1	1	10
6	Ibekwe-Sanjuan, F.	10	1	1	10
7	Smiraglia, R.P.	9	5	3	3
8	Wolfram, D.	9	2	2	4,5
9	Zumer, M.	9	1	1	9
10	Campos, M.L.A.	7	1	1	7
11	Fricke, M.	7	2	2	3,5
12	García-Marco, F.J.	7	1	1	7
13	Gomes, H.E.	7	1	1	7

Fonte: elaboração nossa, com base no *corpus* de pesquisa.

De acordo com a tabela 7, desse grupo de primeira grandeza, B. Hjørland foi o autor que, isoladamente, recebeu o maior número de citações (35) nos artigos incluídos no *corpus* de pesquisa, no período analisado (2016-2020), tendo também apresentado um maior número de artigos citados (4), neste caso, juntamente com C. Gnoli (4). Os/as cinco autores/as que aparecem a seguir (M.J. Fox, C. Gnoli, M. Adler, G.C Bowker e F. Ibekwe-Sanjuan) receberam entre 10 e 13 citações, no mesmo período, para um número variável de artigos citados (1, 2 ou 4). Ainda, com totais que variam entre nove e sete citações, encontramos outros/as sete autores/as (pela ordem: R.P. Smiraglia, D. Wolfram, M. Zumer, M. Fricke, M.L.A. Campos, F.J. García-Marco e H. Gomes).

Relativamente ao número de artigos citados, na mesma tabela 7, verificou-se que uma maioria discreta desses/as autores/as (7; 53,8%) recebeu um número razoável de citações (entre 7 e 10), mas concentradas em apenas um artigo publicado no

período, o que, evidentemente, elevou as suas médias de citação (situadas, igualmente, entre 7 e 10). Em situação oposta, observou-se um número um pouco inferior de autores/as (5; 38,5%) que, tendo recebido um número igual ou maior de citações (entre 7 e 13), mas dispersas por vários artigos publicados, viu descer as suas médias de citação (situadas entre 2,8 e 6,5). A exceção a esta regra aplica-se a apenas um autor (7,7%) – B. Hjørland – que, dado o número elevadíssimo de citações recebidas (35) num número comparativamente elevado de artigos publicados (5), obteve uma média elevada de citações (8,8). A importância a atribuir a estas constatações, encontramos-la em Araújo e Guimarães (2017), quando dizem, relativamente aos/as autores/as, que: “a obtenção da maior média de citação indica a sua forte influência teórica nos artigos que os citaram” (Araújo, & Guimarães, 2017, p. 11). Concordando com esta afirmação, confirmou-se, nesse mesmo conjunto dos/as autores/as mais citados/as, que a maioria contribuiu com pesquisas na área da Ciência da Informação/Organização do Conhecimento, com exceção de G.C. Bowker, que desenvolveu os seus estudos na área da Informática, tendo, no entanto, demarcado pelo menos uma colaboração, concretamente, com F. Ibekwe-Sanjuan, no âmbito da Ciência da Informação.

2.4) coautorias entre autores/as mais produtivos/as e redes de colaboração

São objetivos específicos identificar e analisar as situações de coautoria estabelecidas pelos/as 11 autores/as considerados/as mais produtivos/as, segundo os parâmetros adotados por este estudo (cf. tabela 6). Antes de mais, reforça-se a distinção entre coautoria e colaboração científica; conquanto se relacionem entre si, possuem uma demarcação diferente.

Assim, como lembram Alves et al. (2019), a **colaboração científica** é a partilha de experiências e de interações dentro de um domínio científico, envolvendo mais de um/a investigador/a, no sentido de facilitar a realização do projeto de pesquisa e influenciar a produtividade e a visibilidade científica, contribuindo, no seu todo, para o desenvolvimento da ciência. Já a **coautoria**, também segundo Alves et al. (2019), é a participação de dois/duas ou mais investigadores/as, que contribuem com o seu *know-how* intelectual, na concretização de pesquisas científicas. Ainda, referem que

a colaboração científica é essencial para o desenvolvimento do conhecimento e da experiência de pesquisadores/[as] advindos/[as] dos diferentes domínios científicos, ao passo que a coautoria está preocupada com a construção compartilhada do conhecimento científico, como, por exemplo, a elaboração e publicação de um artigo científico, livro, capítulo de livro, ou seja, algo tangível (Alves et al., 2019, p. 140).

Neste trabalho e relativamente às coautorias verificadas nas publicações dos/as 11 autores/as mais produtivos/as, tendo como referência o *corpus* de pesquisa, procedeu-se primeiramente à identificação manual de todas as ocorrências desta natureza (ou seja, as coautorias estabelecidas). Seguidamente, e com o auxílio do Microsoft Excel®, elaborou-se uma matriz com os dados recolhidos, que posteriormente foi inserida no *Ucinet/Netdraw*®, de modo a visualizar automaticamente os dados em redes de coautoria. Assim, os/as autores/as representados/as na cor azul,

num total de 11, são os/as mais produtivos/as. Por sua vez, os/as representados/as na cor vermelha, num total de 43, são os/as respectivos coautores/as, que também fazem parte do mesmo *corpus* de pesquisa (figura 1).

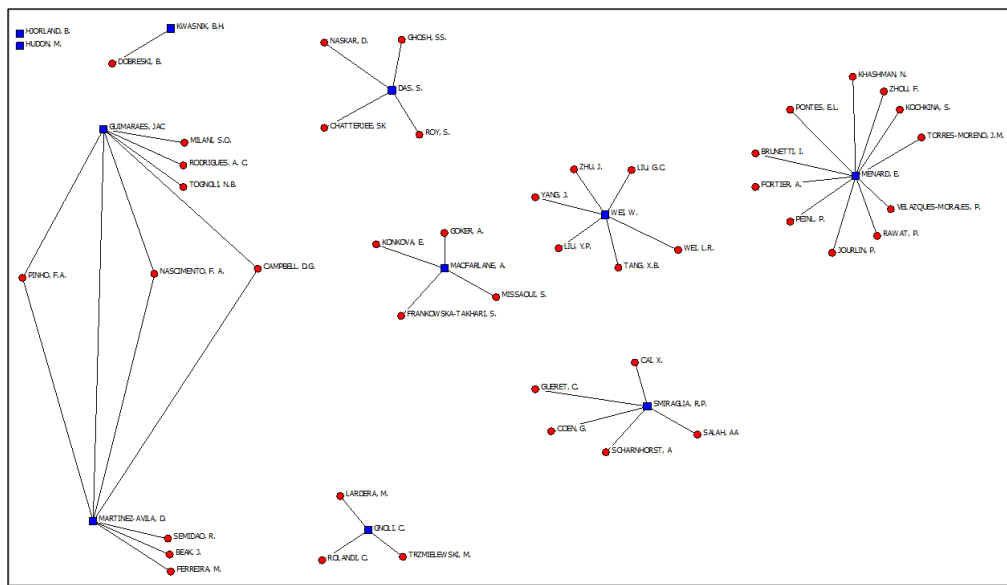


Figura 1. Redes de coautoria dos/as autores/as mais produtivos/as do *corpus* de pesquisa (n=54).

Legenda: ■ Autores/as mais produtivos/as. ● Coautores/as. Uso do NetDraw® (Borgatti et al., 2002).

A partir da análise da figura 1, constata-se que a rede de colaborações visualizada fica composta, primeiramente, por **seis componentes com entre uma e seis ligações**, que representam as coautorias estabelecidas entre os/as seis autores/as mais produtivos/as (R.P. Smiraglia; C. Gnoli; S. Das; B.H. Kwasnik; A. Macfarlane; W. Wei) e outros/as 23 autores/as menos expressivos/as, no âmbito do *corpus* da pesquisa. Seguidamente, observa-se **uma componente maior, composta por nove ligações**, formada a partir das conexões estabelecidas por dois autores mais produtivos (D. Martínez-Ávila e J.A.C. Guimarães), que são coautores entre si e possuem, inclusive, três coautorias comuns, para além de outras três coautorias, individualmente, estabelecendo, deste modo, ligações com nove outros/as autores/as menos expressivos/as no *corpus* da pesquisa. Estas diversas ligações trazem robustez e crescimento à rede, podendo ser interpretadas, neste caso, como um ato de reciprocidade na colaboração entre autores/as. Ainda e em situação oposta, verificam-se **duas componentes isoladas**, ou seja, sem ligações, que representam um autor e uma autora considerados mais produtivos (B. Hjørland e M. Hudon) que, no período sob análise, não apresentaram artigos em coautoria. Ainda, e a contrastar com esta situação, também se constata **uma última componente com 11 ligações**, formada a partir de uma única autora mais produtiva (E. Ménard), que se liga a outros/as 11 autores/as de menor expressão no âmbito do *corpus* de pesquisa, com quem estabelece, certamente, uma rede recíproca de interações.

Através do *software Ucinet*®, foi possível calcular o Grau de Centralidade dos/as autores/as pertencentes à rede de coautorias. Este indicador “consiste no número de atores com os quais um ator está diretamente relacionado” (Velásquez Álvarez, & Aguilar Gallegos, 2005). Neste sentido, E. Ménard apresentou o maior grau de centralidade (11), configurando as mesmas 11 ligações antes mencionadas com outros/as autores/a do *corpus* de pesquisa (A. Fortier; N. Khashman; S. Kochkina; J.M. Torres-Moreno; P. Velázquez-Morales; F. Zhou; P. Jourlin; P. Rawat; P. Peinl; E.L. Pontes; L. Brunetti) (figura 1) e um grau normalizado de 11,3%, também calculado pelo *software*. Este número indica a percentagem de ligações realizadas apenas por esta autora. A secundária estão os autores J.A.C. Guimarães, e D. Martínez-Ávila, que apresentaram um grau de centralidade 7 e um grau normalizado de 6,6%, sendo ambos pertencentes, conforme antes mencionou-se, ao maior componente da rede.

Finalmente, como mostra o grafo resultante desta análise (figura 1), excluindo-se os/as dois autores/as mais produtivos já referidos e que não evidenciaram coautorias, quem apresentou um menor número de ligações com outros/as autores/as da rede foram B.H. Kwasnik, com apenas uma ligação e C. Gnoli, com apenas três ligações.

Conclusão

Este artigo parte de uma análise da produção científica na área da OC, através de um levantamento realizado na base de dados *Web of Science* no período de 2016-2020 e contribui para uma melhor compreensão dos temas mais presentes nos artigos publicados na revista *KO*, dos autores e das autoras mais produtivos e mais citados, bem como das redes de coautoria e das possíveis colaborações entre eles/as estabelecidas.

Quanto aos temas mais presentes nos artigos publicados, concluiu-se que nem as palavras-chave nem as palavras relevantes dos resumos ou dos títulos ofereciam pistas dos assuntos abordados num nível de granularidade fina. Para a sua identificação, teria sido necessária a realização de uma análise documental dos seus conteúdos, o que fugiria ao objetivo principal deste estudo.

Deste modo, chegou-se a um número de 11 autores/as considerados mais produtivos, tendo como referência o *corpus* de pesquisa, constituído por 175 artigos e por 288 autores/as. Estes autores/as assim foram considerados pelo facto de apresentaram entre três e cinco publicações no período.

Quanto às citações, foram 13 os autores/as mais citados/as, com valores situados entre sete e 35 citações, no período. Três desses autores mais citados – B. Hjørland, R.P. Smiraglia e C. Gnoli –, também foram considerados mais produtivos, não se podendo dizer o mesmo relativamente às coautorias ou colaborações, já que dois desses autores – C. Gnoli e B. Hjørland – estabeleceram, respetivamente, poucas (três) ou nenhuma ligações com outros autores/as nos artigos publicados no período. A exceção se faz a R.P. Smiraglia, que apresentou seis ligações relativas a coautorias.

Ainda, no que respeita às coautorias, constatou-se que a autora E. Ménard apresentou um maior número de ligações (11) dessa natureza. Também mereceram

destaque dois outros autores – D. Martínez-Ávila e J.A.C. Guimarães – pelo facto de proporcionarem uma maior robustez à rede, em função das ligações estabelecidas, não apenas entre ambos, como também, e sobretudo com outros/as autores/as.

Estes resultados, ainda que erguidos sobre uma amostra pequena de artigos publicados num único *locus* de análise, de um certo modo, corroboram as afirmações de Souza (2013), para quem o estudo da produção científica em OC é sustentado pela necessidade de se conhecer e avaliar a produtividade, os temas, as colaborações científicas, bem como de mensurar as pesquisas e os/as autores/as nelas envolvidos/as, permitindo a deteção de modelos e de padrões.

As limitações verificadas neste estudo prendem-se, sobretudo, com a natureza da amostra: apenas uma revista científica, ainda que de topo nessas matérias, e analisada num intervalo recente, sendo esta uma limitação relativa, dado haver estudos prévios que remontam ao ano mais recuado deste estudo. Dizer que também não incluímos a perspetiva de género no estudo, dado o carácter da amostra e os objetivos do estudo e também a pouca efetividade da mesma em termos de representação de uma comunidade discursiva, já que não foi esse o desenho original do estudo. Fica esta sugestão para futuras investigações. Também não foram discutidos, por falta de variabilidade e de representatividade, nos dados recolhidos, os temas abordados pelos/as autores/as mais produtivos/as e mais citados/as, individualmente, sugestivos de linhas de investigação, ficando esta proposta igualmente pendente para futuros estudos.

Em pesquisas futuras, será pertinente ampliar o recorte a outras fontes da área ou a outros tipos de documentos (*article, review, proceedings paper, book review, editorial material, letter, biographical item, correction, bibliography*), de forma a obter variabilidade nos dados e enriquecer os resultados e as conclusões, principalmente, no que respeita ao estudo da rede de coautorias, podendo esta densificar-se e potenciar a análise por meio de outros indicadores também fornecidos pelo software *Ucinet*®.

Referências bibliográficas

- Alves, B. H., Dalessandro, R. C., & Santos, F. B. (2019). Colaboração científica no periódico Knowledge Organization: elementos para caracterização de um domínio (pp. 137-144). In *V Congresso Brasileiro em Organização e Representação do Conhecimento*, 8. <http://hdl.handle.net/20.500.11959/brapci/123278>
- Amorim Neto, M. R., & Lima, M. H. T. de F. (2016). O domínio de organização do conhecimento na base BRAPCI: uma análise estatística. *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*, 30(70), 83-104. <https://doi.org/10.1016/j.ibbai.2016.10.005>
- Araújo, P. C. de, & Guimarães, J. A. C. (2017, Outubro 23). Análise de citação da produção científica do domínio de epistemologia da organização do conhecimento. In *XVIII Encontro Nacional de Pesquisa em Ciência da Informação (XVIII ENANCIB)*, 18. <http://hdl.handle.net/20.500.11959/brapci/105227>
- Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *Ucinet 6 for Windows: Software for Social Network Analysis*. Analytic Technologies.
- Bufrem, L. S., Freitas, J. L., & Nascimento, B. S. do. (2014, Maio 14). A pesquisa em organização do conhecimento: uma análise da produção científica na literatura periódica

- em ciência da informação. In *4º EBBBC Encontro Brasileiro de Bibliometria e Cientometria*, 7. <http://hdl.handle.net/20.500.11959/brapci/45576>
- Freitas, M. C. V. de, Borges, M. M., Simões, M. da G. de M., & Cardoso, M. (2014). A visibilidade e a qualidade em revistas de Ciências da Comunicação em acesso aberto no contexto ibero-americano, Portugal e Espanha (pp. 485-499). In M. de L. Martins (Ed.), *II CONFIBERCOM: os desafios da internacionalização*. Centro de Estudos de Comunicação e Sociedade. http://www.google.pt/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CB8QFjAA&url=http%3A%2F%2Fwww.lasics.uminho.pt%2Fconfibercom2014%2Fwp-content%2Fuploads%2F2_Ciencia_Informacao.pdf&ei=JwkPVKSIFNjWasqZgtAG&usq=AFQjCNETnarH60ICEs_ITHUOqctoTX8sdw&bvm=bv.74649129,d.d2s
- Guimarães, J. A. C. (2014). Análise de domínio como perspectiva metodológica em organização da informação. *Ciência da Informação*, 41(1), 13-21
- Hjørland, B. (2008). What is Knowledge Organization (KO)? *Knowledge Organization*, 35(2 e 3), 86-101. <https://doi.org/10.5771/0943-7444-2008-2-3-86>
- Lima, L. S. (2015). *Produção científica em Organização do Conhecimento: Uma análise de domínio via citações de autores* [Mestrado em Ciência da Informação]. Faculdade de Filosofia e Ciências, Universidade Estadual Paulista
- Meschini, F. O., & Oliveira, E. F. T. de. (2017). Produção científica de Birger Hjørland indexada na scopus: artigos e seus citantes. *Revista Guillermo De Ockham*, 15(2), 95-105. <https://doi.org/10.21500/22563202.3055>
- Smiraglia, R. P. (2011). Domain Coherence Within Knowledge Organization: People, Interacting Theoretically, Across Geopolitical and Cultural Boundaries. *Proceedings of the Annual Conference of CAIS / Actes Du Congrès Annuel de l'ACSI*. <https://doi.org/10.29173/cais601>
- Souza, C. (2013). The organization of knowledge: Bibliometric study in database ISI Web of Knowledge. *Biblios: Revista electrónica de bibliotecología, archivología y museología*, 1(51), 20-32. <https://doi.org/10.5195/biblios.2013.108>
- Velásquez Álvarez, O. A., & Aguilar Gallegos, N. (2005). Manual Introdutório à Análise de Redes Sociais. In M. L. L. Aires, J. B. Laranjeiro, & S. C. de A. Silva (Trads.), *Medidas de Centralidade. Exemplos práticos com UCINET 6.109 e NETDRAW 2.28*. https://www2.unicentro.br/lmqqa/files/2016/05/Manualintrodutorio_ex_ucinet.pdf

**APÊNDICE I – DISTRIBUIÇÃO DOS/AS AUTORES/AS RECUPERADOS
NOS TEXTOS PUBLICADOS NA REVISTA KO. PERÍODO
DE REFERÊNCIA: 2016-2020**

Item	Autores/as	Frequência de publicação	%
1	Hjorland, B.	5	1,4
2	Smiraglia, R.P.	5	1,4
3	Gnoli, C.	4	1,1
4	Martinez-Avila, D.	4	1,1
5	Das, S.	3	0,9
6	Guimaraes, J.A.C.	3	0,9
7	Hudon, M.	3	0,9
8	Kwasnik, B.H.	3	0,9
9	Macfarlane, A.	3	0,9
10	Menard, E.	3	0,9
11	Wei, W.	3	0,9
	n=11	39	11,3
12	Barite, M.	2	0,6
13	Broughton, V.	2	0,6
14	Bullard, J.	2	0,6
15	Chen, S.J.	2	0,6
16	Cho, H.	2	0,6
17	Disher, T.	2	0,6
18	Fang, H.	2	0,6
19	Fortier, A.	2	0,6
20	Fox, M.J.	2	0,6
21	Fricke, M.	2	0,6
22	Fujita, M.S.L.	2	0,6
23	Ghosh, S.S.	2	0,6
24	Golub, K.	2	0,6
25	Hider, P.	2	0,6
26	Kaewboonma, N.	2	0,6
27	Keating, S.A.	2	0,6
28	Lee, D.	2	0,6
29	Lee, J.H.	2	0,6
30	Lee, W.C.	2	0,6
31	Liu, G.C.	2	0,6
32	Lykke, M.	2	0,6
33	Marcondes, C.H.	2	0,6
34	Pierozzi, I.	2	0,6
35	Rafferty, P.	2	0,6
36	Solc, R.	2	0,6
37	Souza, R.R.	2	0,6
38	Tang, X.B.	2	0,6
39	Tennis, J.T.	2	0,6
40	Wolfram, D.	2	0,6
41	Zhao, R.Y.	2	0,6
	n=30	60	17,3
42	Abreu, S.C. de	1	0,3
43	Adler, M.	1	0,3
44	Agbaji, D.A.	1	0,3

45	Allar, D.	1	0,3
46	Almeida, C.C. de	1	0,3
47	Alqenaei, Z.M.	1	0,3
48	Andrade, J. de	1	0,3
49	Arboit, A.E.	1	0,3
50	Bak, G.	1	0,3
51	Banerjee, D.	1	0,3
52	Barros, T.H.B.	1	0,3
53	Beak, J.	1	0,3
54	Becnel, K.E.	1	0,3
55	Beheshti, M.S.H.	1	0,3
56	Belarde-Lewis, M.	1	0,3
57	Binding, C.	1	0,3
58	Biniari, K.	1	0,3
59	Bolanos-Mejias, C.	1	0,3
60	Bourcier, P.	1	0,3
61	Bowker, G.C.	1	0,3
62	Brandt, M.B.	1	0,3
63	Brunetti, I.	1	0,3
64	Café, L.M.A.	1	0,3
65	Cai, X.	1	0,3
66	Campbell, D.G.	1	0,3
67	Campos, M.L.D.	1	0,3
68	Castanha, R.C.G.	1	0,3
69	Chansanam, W.	1	0,3
70	Charles, V.	1	0,3
71	Chatterjee, S.K.	1	0,3
72	Chen, L.L.	1	0,3
73	Cheng, Q.K.	1	0,3
74	Cheng, Y.Y.	1	0,3
75	Chhetri, S.	1	0,3
76	Chi, X.H.	1	0,3
77	Choi, I.	1	0,3
78	Choi, N.	1	0,3
79	Clarke, S.G.D.	1	0,3
80	Clavier, V.	1	0,3
81	Cleverley, P.H.	1	0,3
82	Coen, G.	1	0,3
83	Cooper, R.	1	0,3
84	Costa, L.C. da	1	0,3
85	Cruz, M.C.A.E.	1	0,3
86	Dahlberg, I.	1	0,3
87	Davies, S.	1	0,3
88	Day, R.E.	1	0,3
89	Dierking, A.	1	0,3
90	Dillon, A.	1	0,3
91	Dobreski, B.	1	0,3
92	Doggaz, N.	1	0,3
93	Dousa, T.M.	1	0,3
94	Duarte, M.	1	0,3
95	Dunn, H.	1	0,3
96	Dutta, B.	1	0,3
97	Eadon, Y.M.	1	0,3
98	Ejehi, Z.	1	0,3
99	Ejei, F.	1	0,3
100	Farace, D.	1	0,3

101	Ferreira, M.	1	0,3
102	Ferris, S.	1	0,3
103	Foskett, D.J.	1	0,3
104	Frankowska-Takhari, S.	1	0,3
105	Freeman, G.	1	0,3
106	Freitas, M.C.V. de	1	0,3
107	Frota, M.G.D.	1	0,3
108	Gaona-García, P.A.	1	0,3
109	Garcia-Marco, F.J.	1	0,3
110	Gil-Leiva, I.	1	0,3
111	Glushko, R.J.	1	0,3
112	Goker, A.	1	0,3
113	Gomes, H.E.	1	0,3
114	Gomes, P.	1	0,3
115	Green, R.	1	0,3
116	Groth, P.	1	0,3
117	Grundner, A.	1	0,3
118	Guedes, R.D.	1	0,3
119	Gueret, C.	1	0,3
120	Hajibayova, L.	1	0,3
121	Harinarayana, N.S.	1	0,3
122	Hartel, J.	1	0,3
123	Hauser, E.	1	0,3
124	Henttonen, P.	1	0,3
125	Higgins, M.	1	0,3
126	Holanda, M.T. de	1	0,3
127	Hou, Z.P.	1	0,3
128	Humpal, N.	1	0,3
129	Ibekwe-Sanjuan	1	0,3
130	Ishikawa, E.	1	0,3
131	Jansson, I.M.	1	0,3
132	Jantzen, C.	1	0,3
133	Jett, J.	1	0,3
134	Johansson, S.	1	0,3
135	Joo, S.	1	0,3
136	Jourlin, P.	1	0,3
137	Katona, J.M.	1	0,3
138	Keilty, P.	1	0,3
139	Kempf, A.O.	1	0,3
140	Khashman, N.	1	0,3
141	Kochkina, S.	1	0,3
142	Konkova, E.	1	0,3
143	Koopman, R.	1	0,3
144	Kowalska, M.	1	0,3
145	Krishnamurthy, M.	1	0,3
146	Kulthida, T.	1	0,3
147	Lacey, E.	1	0,3
148	Lacruz, M.D.A.	1	0,3
149	Laporte, S.	1	0,3
150	Lara, M.L.G. de	1	0,3
151	Lardera, M.	1	0,3
152	Latham, K.F.	1	0,3
153	Lauruhn, M.	1	0,3
154	Leazer, G.H.	1	0,3
155	Lee, E.	1	0,3
156	Lemos, D.L.D.	1	0,3

157	Leopenwong, S.	1	0,3
158	Li, N.	1	0,3
159	Li, X.	1	0,3
160	Lian, T.H.	1	0,3
161	Lima, G.A.B.D.	1	0,3
162	Lin, F.T.	1	0,3
163	Littletree, S.	1	0,3
164	Liu, X.Z.	1	0,3
165	Liu, Y.P.	1	0,3
166	Liu, Z.F.	1	0,3
167	Lokkegaard, S.	1	0,3
168	Lomas, E.	1	0,3
169	Lu, W.	1	0,3
170	Ludascher, B.	1	0,3
171	Lund, B.D.	1	0,3
172	Lund, H.	1	0,3
173	Ma, S.T.	1	0,3
174	Mai, J.E.	1	0,3
175	Milani, S.O.	1	0,3
176	Missaoui, S.	1	0,3
177	Moeller, R.A.	1	0,3
178	Monarchi, De	1	0,3
179	Mondal, T.K.	1	0,3
180	Moreira, W.	1	0,3
181	Moreiro-Gonzalez, J.A.	1	0,3
182	Moura, M.A.	1	0,3
183	Muir, L.J.	1	0,3
184	Nascimento, F.A.	1	0,3
185	Naskar, D.	1	0,3
186	Netto, C.M.	1	0,3
187	Neubert, J.	1	0,3
188	Oh, D.G.	1	0,3
189	Oliveira, C.E.	1	0,3
190	Ortega, C.D.	1	0,3
191	Osinkska, V.	1	0,3
192	Osinski, Z.	1	0,3
193	Packalen, S.	1	0,3
194	Padmavathi, T.	1	0,3
195	Pando, Da	1	0,3
196	Peinl, P.	1	0,3
197	Pinho, F.A.	1	0,3
198	Piros, A.	1	0,3
199	Poelmans, H.	1	0,3
200	Pontes, E.L.	1	0,3
201	Prost, H.	1	0,3
202	Qiu, J.N.	1	0,3
203	Quan, W.	1	0,3
204	Quinlan, E.	1	0,3
205	Rajabi, T.	1	0,3
206	Rauch, M.	1	0,3
207	Rawat, P.	1	0,3
208	Ribas, R.R.D.	1	0,3
209	Ridenour, L.	1	0,3
210	Ridi, R.	1	0,3
211	Rodrigues, A.C.	1	0,3
212	Rodrigues, F.D.	1	0,3

213	Rodriguez-Bravo, B.	1	0,3
214	Rolandi, C.	1	0,3
215	Roszkowski, M.	1	0,3
216	Roy, S.	1	0,3
217	Salah, A.A.A.	1	0,3
218	Saldanha, G.S.	1	0,3
219	Sanchez-Alonso, S.	1	0,3
220	Sandy, H.M.	1	0,3
221	Sant'ana, R.C.G.	1	0,3
222	Santos, L.B.P. dos	1	0,3
223	Satija, M.P.	1	0,3
224	Scharnhorst, A.	1	0,3
225	Schneider, J.	1	0,3
226	Schopf, J.	1	0,3
227	Semidao, R.	1	0,3
228	Shi, J.S.	1	0,3
229	Silva, C.M.A. da	1	0,3
230	Silveira, N.C.	1	0,3
231	Simões, M.G.M.	1	0,3
232	Sinha, P.K.	1	0,3
233	Skov, M.	1	0,3
234	Slavic, A.	1	0,3
235	Smith, J.D.	1	0,3
236	Soos, C.	1	0,3
237	Sousa, R.T.B. de	1	0,3
238	Souza, R.D.	1	0,3
239	Stoitsis, G.	1	0,3
240	Sun, J.Q.	1	0,3
241	Svenonius, E.	1	0,3
242	Szostak, R.	1	0,3
243	Tai, Y.F.	1	0,3
244	Tang, H.P.	1	0,3
245	Terra, A.L.	1	0,3
246	Tharani, K.	1	0,3
247	Thomer, A.	1	0,3
248	Thorat, R.	1	0,3
249	Tognoli, N.B.	1	0,3
250	Torres-Moreno, J.M.	1	0,3
251	Trzmielewski, M.	1	0,3
252	Tuamsuk, K.	1	0,3
253	Tudhope, D.	1	0,3
254	Twidale, M.	1	0,3
255	Vaidya, P.	1	0,3
256	Van Acker, W.	1	0,3
257	Van Allen, A.	1	0,3
258	Van Brakel, R.	1	0,3
259	Vancauwenbergh, S.	1	0,3
260	Velazquez-Morales, P.	1	0,3
261	Victorino, M.	1	0,3
262	Vieira, R.	1	0,3
263	Viti, E.	1	0,3
264	Wang, S.H.	1	0,3
265	Wang, W.	1	0,3
266	Wang, Y.F.	1	0,3
267	Wei, L.R.	1	0,3
268	Wei, M.K.	1	0,3

269	Wei, X.Q.	1	0,3
270	Wlodarczyk, B.	1	0,3
271	Wu, D.	1	0,3
272	Wu, Y.J.	1	0,3
273	Xu, L.W.	1	0,3
274	Xue, H.L.	1	0,3
275	Yang, J.	1	0,3
276	Yang, L.	1	0,3
277	Yang, Y.F.	1	0,3
278	Yu, C.H.	1	0,3
279	Yuan, Q.J.	1	0,3
280	Zagrouba, E.	1	0,3
281	Zane, A.	1	0,3
282	Zarrad, R.	1	0,3
283	Zhang, C.Z.	1	0,3
284	Zhang, Q.	1	0,3
285	Zhao, H.	1	0,3
286	Zhou, F.	1	0,3
287	Zhu, J.	1	0,3
288	Zumer, M.	1	0,3
n=247		247	71,4
n=288		346	100

Fonte: Dados provenientes da revista *KO*, extraídos através da WoS®, em consultas realizadas entre 06/01/2021 e 22/01/2021.

ELEMENTOS DE INTERDISCIPLINA EN LOS CURSOS DE POSGRADO IBEROAMERICANOS DE ORGANIZACIÓN DEL CONOCIMIENTO: PROTOCOLO DE INVESTIGACIÓN

Mario Barité¹, Varenka Parentelli²

¹Facultad de Información y Comunicación, Universidad de la República,
mario.barite@fic.edu.uy., ORCID iD 0000-0002-2992-6582

²Facultad de Información y Comunicación, Universidad de la República,
varenka.parentelli@fic.edu.uy, ORCID iD 0000-0003-2033-7949

Resumen

La formación de posgrado en Organización de Conocimiento (OC) se pone en tensión cuando, al tiempo de reconocerse como área de saber, comunica sus raíces y su evolución con otros dominios disciplinares, lo que se ha consignado y analizado repetidas veces. El tráfico de ideas entre la OC y sus dominios vinculados considera 150 años de historia acumulada en torno a la representación y recuperación temática de información, y comprende aspectos epistemológicos, teóricos y metodológicos. Algunas teorizaciones e investigaciones realizadas sobre currícula comparada establecen un fundado análisis de las tensiones y los desafíos profesionales en un entorno digital creciente, y se reafirman en la idea de un currículum organizado alrededor de los conceptos clásicos que dan perfil a la profesión, aunque con la posibilidad de integrar herramientas y métodos de otros campos. Para cubrir una laguna identificada en la literatura del área, este trabajo describe el protocolo de la investigación que se realizará sobre la presencia de elementos interdisciplinarios en los contenidos de los cursos de posgrado enfocados en la OC en Iberoamérica. El objetivo es establecer un estado de situación que involucre los alcances y las modalidades en los cuales los estudiantes de posgrado son formados en la interdisciplina en el área, de cara a su desarrollo profesional y a la opción por una carrera orientada a la docencia y la investigación. Se propone la creación de un directorio de cursos universitarios de posgrado en OC y el análisis de sus programas para identificar la presencia o integración de conocimientos interdisciplinarios en los contenidos y bibliografías. Asimismo, se realizará una encuesta a los docentes de los cursos y se triangulará la información recabada. A través de los resultados de esta investigación y su difusión, se aspira a que las instituciones educativas tomen conciencia de la dimensión del factor interdisciplinario y puedan revisar sus currícula en torno al abordaje pedagógico y la estructura curricular. De este modo, podrían expresarse los componentes interdisciplinarios presentes en los cursos de un modo sistemático. Asimismo, los docentes contarán con más elementos para reflexionar sobre el diseño de sus cursos, proyectar su implementación e identificar tendencias en la referencia y uso de textos de autores de otros campos del saber en la comarca iberoamericana. Como avance de la

investigación en curso, se exhiben resultados preliminares a partir de una muestra de programas (con sus respectivas bibliografías), de cursos de cuatro universidades iberoamericanas (México, Brasil, Colombia y Portugal). Para ese fin se establecieron siete categorías de análisis que comprenden la presencia de expresiones propias de lo interdisciplinario, el lugar de los programas en los que se mencionan, la identificación de autores que tienen una procedencia disciplinaria distinta a la ciencia de la información en la bibliografía básica, la identificación de concepciones teóricas subyacentes, las perspectivas didácticas de lo interdisciplinario, la mención explícita de otras disciplinas o de sus vínculos con la OC y aspectos relacionados con habilidades cognitivas o metacognitivas relacionados con la interdisciplina, en relación a objetivos de aprendizaje buscados.

Palabras clave: Organización de conocimiento, Interdisciplina, Abordajes pedagógicos, Currículum, Posgrados.

Introducción

Interdisciplina es un término complejo y polisémico. Para Sutz (2015) implica un proceso de coproducción de conocimiento desde un lenguaje común. Los enfoques sobre interdisciplina de Follari (2007) y Rugarcía (1997) presentan una tensión entre sí, pues el primero centra la posibilidad de la interdisciplina en el surgimiento de una metodología específica, y refiere a la imposibilidad de concebir al sujeto como interdisciplinario. En cambio Rugarcía cree que las personas, tanto en forma individual como colectiva, son capaces de integrar por sí mismas lo que reciben de diversas disciplinas.

Sotolongo y Delgado (2006) sitúan la interdisciplina en un objeto que no puede ser delimitado y abordado por una disciplina, sino por varias. La interdisciplinariedad supone la existencia previa de los campos disciplinarios, no su disolución. Solo se puede construir interdisciplina proyectando y comunicando elementos propios de las disciplinas preexistentes. Desde esta perspectiva no existe oposición ni complementariedad absoluta entre disciplinariedad e interdisciplinariedad (Nieto-Caraveo, 1991, p. 5).

Existen abordajes epistemológicos, metodológicos y teóricos de la cuestión interdisciplinaria. En relación a lo metodológico, para Lenoir (2013) alcanza con la combinación de métodos existentes sin la necesidad de crear algo nuevo. Follari (2007) agrega que los métodos provenientes de diversas disciplinas no deben repetirse en su estado original.

A partir del siglo XIX, en línea con los postulados positivistas dominantes, se establecen en los ámbitos universitarios estructuras organizativas de conocimiento, de las que derivan los modelos tradicionales de enseñanza y de investigación. Estos modelos clásicos enfatizan en lo disciplinario, en la jerarquización y verticalización del currículum, y en la compartimentación de los saberes, provocando diseños tubulares de formación (Omelas Navarro 1982). A medida que la realidad fue exigiendo el necesario vínculo entre disciplinas para resolver problemas comunes, la formación universitaria fue abriéndose a diseños curriculares flexibles para comunicar horizontalmente a las disciplinas, al tiempo de tomar conciencia de las ventajas del aprovechamiento de teorías, métodos y aplicaciones que podían ser útiles fuera

de las estrechas fronteras tradicionales. Las perspectivas de la complejidad y la interdisciplina pudieron desafiar así las estructuras y los modelos tradicionales, ensanchando el horizonte de formación de los estudiantes universitarios (Parentelli 2019a). No obstante, lo anterior se pone en tensión cuando en la organización de las mallas curriculares se continúa con una lógica fragmentada del conocimiento y es a ello que responde la oferta programática de las carreras.

En el caso de la organización del conocimiento (OC), si bien se reconoce a sí misma como un área específica de saber, hunde sus raíces en dominios como la bibliografía, la industria editorial, la bibliotecología y la documentación (Smiraglia, 2012, p. 40). Dialoga además con conocimientos de otras disciplinas, lo que ha sido relevado y reconocido por distintos autores (García Marco, 1995; McIlwaine, 2003; Hjørland, 2008; Szostak, Gnoli & Lopez-Huertas, 2016; Guimarães, 2017; Salaba, 2020). No por azar el encuentro de ISKO-España de 2007 en la ciudad de León, llevó por lema “La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento científico” (Sociedad Internacional para la Organización del Conocimiento (ISKO)-Capítulo Ibérico, 2007).

Los conceptos, técnicas y procedimientos aportados por la inteligencia artificial, el procesamiento del lenguaje natural y las ciencias cognitivas son tres instancias identificadas por McIlwaine (2003), a partir de las cuales la OC construye mejores herramientas teóricas y metodológicas. Hjørland (2008, p. 98-99) señala los beneficios que se obtienen cuando un dominio que comparte con la OC sus bases epistemológicas, desarrolla una teoría sólida. Además, menciona expresamente la contribución de diversas áreas más o menos próximas en el mapamundi del saber: la más cercana de todas, la recuperación de información, y luego la lingüística, la terminología, los estudios cognitivos, la informática y la comunicación. Mai (2000), al utilizar el pensamiento de Pierce para dar fundamento a los procesos de indización, ilumina una perspectiva bastante común y productiva en el área: la apelación a pensadores canónicos (Barthes, Foucault, Bordieu, Eco) para incorporar visiones epistemológicas y teóricas de fundamento común a varias ciencias sociales y humanas. Por último, en la medida en que en el seno de la OC van encontrando un importante espacio las apelaciones a la dimensión cultural del conocimiento, también comienzan a valorarse teorías y métodos tomados de la antropología, la historia social y los estudios culturales.

Quizás la vocación interdisciplinaria de la OC sea consecuencia de que “knowledge organization is of great importance in a global information landscape, and affects all economic, political, social, and educational sectors” (Salaba, 2020, p. 384). El tráfico de ideas y aplicaciones entre la OC y sus dominios vinculados suele ser bidireccional en mayor o menor medida, considerando ciento cincuenta años de desarrollo teórico-práctico en torno a la representación y recuperación temática de información. Puede aventurarse que existen cuatro dimensiones de intercambio de saberes entre la OC y otros dominios: ética, teórica, tecnológica y aplicada, y social.

Ya en 1995, García Marco apuntaba que, probada la aportación de un amplio número de disciplinas al área, la enseñanza universitaria de la OC debía cumplir 3 objetivos, siendo el tercero “capacitar al alumno en el vocabulario y modelos básicos de las ciencias interdisciplinariamente conectadas con ella” (p. 220).

Sin embargo, la literatura muestra que la cuestión de la interdisciplinariedad como área problema en la enseñanza de la OC se ha estudiado de modo esporádico e indirecto. En los primeros Encuentros de Escuelas de Bibliotecología del Mercosur, una serie de trabajos (Barité, 2000a y 2000b; Guimarães, 2000, 2001a y 2001b) identificaron los cuerpos y los enfoques teóricos presentes en los programas del área en esa región con una vocación crítica y no meramente enumerativa: evitar que se redujera la formación a “un conjunto híbrido de nociones y aplicaciones que los educandos no logran interpretar ni utilizar conforme a sus contextos” epistemológicos (Barité, 1999, p. 126).

Algunas teorizaciones (Gorman, 2002; Morgan & Badwen, 2006) e investigaciones realizadas sobre currícula comparada (Al-Ajami & Rehman, 2016; Rehman & Al-Ajami, 2017; Chaudry, 2018; Salaba, 2020), establecen un fundado análisis de las tensiones y los desafíos profesionales en un entorno digital creciente, y se reafirman en la idea de un curriculum organizado en torno a los conceptos clásicos que dan perfil a la profesión, aunque con la posibilidad para integrar herramientas y métodos de otros campos.

Paulatinamente, la interdisciplina adquiere protagonismo y cobra relevancia el modo en el que se expresa en los currícula universitarios, definidos estos como el conjunto de experiencias por la cuales transitan los estudiantes en su proceso formativo (Kelly, 2009) en relación al diseño y su implementación. En lo curricular, la inclusión de la interdisciplina resulta compleja y es frecuente que se involucre en la redacción desde una dimensión discursiva como si de ello naturalmente derivara una interpretación pedagógica (Parentelli, 2019b). Con una mirada benigna podría decirse que esa constatación es al menos el principio de algo nuevo, pero solo se justifica si se logra superar finalmente la frontera del discurso y para ello será necesario reflexionar sobre la interdisciplina misma y tener en cuenta la mirada pedagógica y didáctica entre otros aspectos.

Rugarcía (1997) afirma que un sujeto en solitario es capaz de formarse interdisciplinariamente a través de la integración de conocimientos provenientes de diversas disciplinas. No obstante, Beri y Tello (2006) y Follari (2007) sostienen que la interdisciplina es un efecto de la síntesis colectiva y que por tanto, no pueden existir sujetos interdisciplinarios. Este debate no es menor, pues implica la posibilidad de concebir a los estudiantes como sujetos capaces de integrar, que requerirán -por tanto- dispositivos didácticos específicos para ello.

En este trabajo se describe el protocolo de la investigación que se tiene proyectado realizar en torno a la presencia de elementos interdisciplinarios en los contenidos de los cursos de posgrado universitarios de la región iberoamericana, enfocados en tópicos propios de la OC. Como avance de la investigación en curso, se exhiben resultados preliminares a partir de una muestra de programas de cursos (con sus correspondientes objetivos, contenidos, metodologías y bibliografías), recibidos gracias a la generosa disposición de los responsables de cursos de cuatro universidades iberoamericanas, situadas en México, Brasil, Colombia y Portugal respectivamente. Se espera que el conocimiento previo de este avance, estimule la participación de los docentes de posgrado del resto de las universidades iberoamericanas en la investigación proyectada. Los resultados seguramente permitirán orientar la formación de posgrado en corrientes de pensamiento y tópicos con valor interdisci-

plinario, a partir de la experiencia regional compartida. Asimismo, permitirán visualizar con mayor claridad los dispositivos de orden pedagógico y didáctico que pueden ponerse en juego en la planificación de los cursos, para que los estudiantes de posgrado puedan valorar e incorporar conocimientos provenientes de otros dominios y dialogar creativamente con ellos.

Metodología

La pesquisa planteada se cumplirá en un año y medio, en el marco del programa de investigación TERM-OC, que fue aprobado con financiación para el cuatrienio 2019-2022 por la Comisión Sectorial de Investigación Científica (CSIC) de la Universidad de la República (Uruguay).

La investigación permitirá actualizar datos sobre la cantidad y la ubicación universitaria de los cursos enfocados en tópicos propios de la OC o relacionados con ella, que se llevan adelante en forma regular en la región iberoamericana.

El objetivo general es establecer un estado de situación respecto a la presencia de elementos interdisciplinarios en el universo estudiado, que involucre los alcances y las modalidades en los cuales los estudiantes son formados en la interdisciplina en el área, de cara a su desarrollo profesional y a la opción por una carrera orientada a la docencia y la investigación.

Las preguntas que fundamentan esta investigación son las siguientes: ¿cuáles son los elementos indicadores de aspectos, enfoques y cuestiones de rango interdisciplinario en las carreras de posgrado de OC que se imparten en las universidades iberoamericanas? ¿De qué manera y dónde se mencionan o expresan los aportes de otros dominios en OC, considerando diversas categorías de análisis (por ejemplo: elementos teóricos, metodológicos o didácticos, o implicaciones interdisciplinarias)? ¿Qué consideraciones de índole general pueden realizarse para favorecer, en el futuro, la inserción y organización sistemática de los contenidos de tenor interdisciplinario en los currícula?

Fases metodológicas:

1. Compilación de los programas de cursos universitarios de posgrado en OC de la región iberoamericana.
2. Creación de un directorio con los cursos reunidos.
3. Elaboración de un corpus de programas y bibliografías de los cursos mencionados.
4. Identificación de los contenidos presentes en los programas y bibliografías que puedan involucrar la referencia a, o la presencia o integración de conocimientos interdisciplinarios, en función de criterios de opinión experta a instrumentar.
5. Determinación de las categorías de análisis, a través de las cuales se identificarán contenidos, autores y obras cuyo perfil y producción representativa son ajenos a la OC, las modalidades o tipos de implicación interdisciplinaria que se encuentren, así como aspectos relacionados a la interdisciplina con habilidades cognitivas y las estrategias didácticas para la integración de conocimiento.

6. Diseño e implementación de una encuesta dirigida a los docentes responsables de los cursos analizados.
7. Cruzamiento y triangulación de datos, con el resultado de la encuesta y el marco teórico de la investigación.
8. Publicación de resultados.
9. Disponibilización del directorio generado.

Como se anticipó más arriba, se cumplió con un avance de la investigación proyectada, a partir del análisis de una muestra integrada por los programas y las bibliografías de cuatro cursos de las siguientes universidades iberoamericanas: Universidad Nacional Autónoma de México (UNAM), Universidad Federal Fluminense (UFF) de Brasil, Universidade do Porto (U.Porto) de Portugal, y Universidad de Antioquia (UdeA), de Colombia. Se elaboró una ficha por universidad para recoger los datos. Las cuatro fichas de análisis se presentan en el capítulo de Resultados.

Con el avance realizado se buscó identificar posibles dificultades en la recogida y análisis de los datos a la luz de las categorías predefinidas, poner a prueba algunas de las preguntas de la futura encuesta, y disponer de algunos resultados preliminares.

Resultados

En virtud de que esta ponencia expone, antes que nada, el protocolo básico de la investigación a desarrollarse, en este apartado no se indican los resultados obtenidos por la investigación propiamente dicha, lo que corresponde a una etapa posterior. Solo se anticipan los resultados preliminares del análisis realizado sobre la muestra de cuatro cursos de las universidades arriba mencionadas.

En cuanto a los resultados de la investigación diseñada, solo es posible compartir aquí una proyección de aquellos que se pueden alcanzar, en términos generales, en función del objetivo general y de las preguntas de investigación formuladas. Estos resultados preparatorios se concretan en la creación de distintos productos y/o escenarios. El primero lo constituye el diseño y la puesta a prueba de un método que sea útil para esta investigación, el cual, una vez asegurada su fiabilidad, podría ser agregado a los modelos de análisis de currículas y de experiencias curriculares, registrados en la literatura del área (Aytac et al., 2012; Al-Ajami & Rehman, 2016; Rehman & Al-Ajami, 2017; Braz & Carvalho, 2017; Salaba, 2020).

En segundo lugar, se aportará un directorio actualizado de cursos universitarios de posgrado enfocados en OC, vigentes en las universidades iberoamericanas, del que se proyecta su libre disponibilidad.

A partir del directorio, se estaría en condiciones de formar, a través de los mecanismos establecidos para la colaboración entre universidades, un repositorio regional de los programas y bibliografías de posgrado enfocados en OC, disponibles para facilitar otras investigaciones y el estudio comparado.

Asimismo, esta investigación permitiría sentar las bases para empezar a concretar la aspiración formulada por Chaudry (2018): Que los docentes e investigadores puedan expandir el desarrollo de los cursos de OC a otras áreas de formación, tales como la gestión de información o la arquitectura de información “to respond to the

needs of organization of knowledge resource in the digital environment” (Chaudry, 2018, p. 1385).

En lo que refiere al análisis de la muestra, se visualizan a continuación las cuatro fichas que fueron elaboradas para cada uno de los cursos.

Tabla 1. Ficha UNAM

Parámetros	UNAM – Curso “Seminario de Lenguajes documentales: documentos digitales y multimedia”
Expresiones propias de la interdisciplina	“Analizar el concepto de mediación lingüística en la representación de contenido documental en documentos digitales y multimedia y su relación con los conceptos de información, conocimiento, lenguaje y pensamiento”. “Integrar las teorías de la lingüística textual al análisis de contenido para la organización documental y su nexa con la indización y la clasificación”
Lugar en el programa	Objetivo general. Objetivos específicos. Temario. Referencias bibliográficas
Mención otras disciplinas	Lingüística. Filosofía. Semiótica. Lenguaje Audiovisual
Autores externos a la ciencia de la información	Cabré Castelví, M. Teresa (2017); Gómez Isla, J. (2010); Van Dijk, Teun A. (1992); Castiñeiras González, Manuel Antonio (2007)
Concepciones teóricas subyacentes	Capacidad del sujeto (transversal a la propuesta). Abordaje de objeto complejo
Estrategias didácticas para la interdisciplina	No se observa
Habilidades cognitivas y metacognitivas	Analizar; comprender
Análisis	Inclinación a incorporar aspectos interdisciplinarios y autores de otras áreas. Mención expresa de otras disciplinas de intersección con OC. Se identifican aspectos relacionados con la interdisciplina en la inclusión de otras disciplinas y puede inferirse niveles de abordajes que implican desarrollos cognitivos medio y bajo. No se visualizan en forma explícita dispositivos metodológicos y didácticos para tales fines. No obstante, en el apartado metodológico se expresa: “El método utilizado será el aprendizaje autónomo de los estudiantes a través de las lecturas y de los ejercicios para adquirir la experiencia que complemente las lecturas. La actividad práctica es considerada fundamental en el aprendizaje de la representación de los contenidos documentales”. De lo anterior puede inferirse que la integración de las disciplinas podría darse a partir del ejercicio práctico y autónomo, lo que refuerza la concepción de interdisciplina asociada a la capacidad del sujeto individual.

Tabla 2. Ficha UFF

Parámetros	UFF – Curso “Modelos e teorias de representação de domínios de conhecimentos”
Expresiones propias de la interdisciplina	“Os domínios de conhecimento em Sistemas de Organização do Conhecimento (SOCs) são representados através de estruturas classificatórias, que vêm através dos tempos se adequando a diversas Teorias de Representação. Estas teorias se colocam no espaço da elaboração de Modelos Conceituais independentes de conteúdos específicos de domínios. São meta- representações utilizadas para a organização de conhecimento em domínios diferenciados. Pretende-se, a partir da discussão dos aspectos teóricos e metodológicos relacionados a estas teorias, enfatizar os aspectos relativos à elaboração e análise de modelos de representação para fins de tratamento e recuperação da Informação.”
Lugar en el programa	Resumen del tema, Referencias bibliográficas
Mención otras disciplinas	No explícita
Autores externos a la ciencia de la información	Brachman, R. J. (1979); Le Moigne, Jean-Louis. (1977, 2000); Morin, Edgar (1990); Newell, A. (1982)
Concepciones teóricas subyacentes	Interdisciplina desde un abordaje de la complejidad (por Morin y Le Moigne) Concepción centrada en la integración de aspectos teóricos y metodológicos por lo que se infiere que aborda a la interdisciplina como emergente del objeto de estudio y también desde un enfoque metodológico.
Estrategias didácticas para la interdisciplina	“Pretende-se, a partir da discussão dos aspectos teóricos e metodológicos relacionados a estas teorias, enfatizar os aspectos relativos à elaboração e análise de modelos de representação para fins de tratamento e recuperação da Informação.”
Habilidades cognitivas y metacognitivas	No menciona
Análisis	Hay referencias implícitas a los aspectos interdisciplinarios presentes en los contenidos, y una mención de autores externos a la ciencia de la información. En las estrategias didácticas para la interdisciplina se prevén espacios y tiempos de discusión y método de abordaje para hacer evidente las teorías que según el cuerpo docente “se ubican en el espacio de la elaboración de Modelos Conceptuales independientes de los contenidos específicos del dominio”. En relación a los autores resulta expresiva la mención de autores como Morin y Le Moigne en las referencias bibliográficas, específicamente en relación con textos que abordan el tema de la complejidad. A partir de los dos aspectos anteriores, se infiere que hay un abordaje de la dimensión interdisciplinaria transversal a la propuesta.

Tabla 3. Ficha U.Porto

Parámetros	U.Porto- Curso “Análise de Conteúdo e Indexação”
Expresiones propias de la interdisciplina	No se encuentran
Lugar en el programa	Temario
Mención otras disciplinas	Classificações de I+D. Classificações em contexto empresarial. Classificações em ambiente web
Autores externos a la ciencia de la información	No se mencionan
Concepciones teóricas subyacentes	Se mencionan solo las relacionadas con OC
Estrategias didácticas para la interdisciplina	No se mencionan
Habilidades cognitivas y metacognitivas	No se mencionan
Análisis	Se trata de un curso de corte clásico, enfocado en teorías, métodos y autores propios de la OC. No obstante, al constituir la OC un ámbito interdisciplinario per se el perfil del curso implica cierta atención implícita o indirecta a elementos o componentes interdisciplinarios. Esto se visualiza en la preocupación por el diseño de clasificaciones orientadas a la empresa, o a la ciencia y la tecnología

Referência: Los autores

Tabla 4. Ficha UdeA

Parámetros	UdeA – Curso “Seminario de Sistemas de Organización y Representación del Conocimiento”
Expresiones propias de la interdisciplina	“La Organización y Representación del Conocimiento es uno de los campos de la ciencia de la información en el cual convergen saberes de la lingüística, la computación y las ciencias cognitivas.” También se menciona en una referencia bibliográfica particular, que corresponde a un congreso cuyo nombre contiene las palabras “interdisciplinariedad” y transdisciplinariedad”
Lugar en el programa	Descripción. Objetivo general. Objetivos específicos. Contenidos. Referencias bibliográficas
Mención otras disciplinas	Lingüística. Computación. Ciencias cognitivas. Gestión de los datos. Ética.

Autores externos a la ciencia de la información	Bobrow, DG y Winograd, T. (1977); Cortez Vásquez, A., Vega Huerta, H., Pariona Quispe, J., & Huayna, A. M. (2009); Foucault, Michel (1968); Nitin Induskhya, Fred J. Damerau. (2010)
Concepciones teóricas subyacentes	Hay una propuesta en torno a acercar al estudiante a la temática a través de actividades relacionadas con “estimular”, “familiarizar”, “abordar” y “explorar”, tal como se establece en los objetivos. Luego, en la metodología y evaluación se espera que el estudiante “elabore”. Por lo anterior, podría inferirse que hay una concepción de la interdisciplina relacionada con la capacidad del sujeto.
Estrategias didácticas para la interdisciplina	Podrían observarse estrategias didácticas para la interdisciplina en la metodología, en tanto se prevé que los estudiante practiquen (aplicación de conceptos) y tengan que “elaborar “como etapa final del curso. No se observa ninguna mención específica en relación a la integración de conocimiento.
Habilidades cognitivas y metacognitivas	“Abordar”, “explorar” se relacionan con el objetivo de aprendizaje asociado a “conocer”. “Elaborar” se relaciona con el objetivo de aprendizaje de “aplicación” o bien de “creación”..
Análisis	La apelación a la intersección entre OC y otras disciplinas es explícita y recurrente a lo largo del programa de curso. Se mencionan numerosos autores ajenos al área. En los objetivos se habla de “estimular”, “familiarizar”, “abordar” y “explorar” por lo que se infiere que hay un tratamiento inicial e introductorio que no se propone niveles de desarrollos cognitivos complejos en el proceso de aprendizaje de los estudiantes. No obstante, en el apartado de evaluación se hace referencia a “elaboración de micro tesauros” y “elaboración de ontologías”, y en la metodología “Desarrollo de clases prácticas con la elaboración de micro tesauros, ontologías, sistemas de recuperación, entre otros”. Esto último podría referir a un objetivo de aprendizaje de orden superior en términos de complejidad tal como lo es “crear”. De lo anterior, se observa que la metodología se concibe como un proceso que parte de la “estimulación, familiarización, abordaje y exploración”, transita por una práctica de aplicación y finaliza con la elaboración.

Referencia: Los autores

El análisis de los programas que integraron la muestra permitió valorar positivamente la herramienta que fue diseñada. De todos modos, es posible que se realicen ajustes que permitan un estudio más pormenorizado de los parámetros establecidos, a efectos de que arrojen datos que puedan compararse con mayor detalle.

Corresponde manifestar que la presencia de componentes interdisciplinarios no es una panacea, ni implica una valoración positiva de por sí. Con esto se quiere señalar que un programa de posgrado no es bueno porque incorpore la dimensión interdisciplinaria, ni deja de ser bueno porque no la tome en cuenta. Como se acepta pacíficamente en la literatura del área, los componentes interdisciplinarios están implícitos en la misma base del dominio interdisciplinario de OC.

La investigación proyectada pretende echar luz sobre el mayor o menor reconocimiento de esa evidencia, y las formas en que se da ese reconocimiento, tanto desde la perspectiva de los contenidos propiamente dichos, como desde las herramientas pedagógicas y didácticas que se utilizan para favorecer el aprendizaje.

Conclusiones

El estudio que se proyecta pretende cubrir una laguna identificada en la literatura del área.

La OC es un dominio que ha experimentado en los últimos años un inesperado dinamismo, en virtud del desarrollo de Internet, la web semántica y numerosas aplicaciones relacionadas con la búsqueda y el acceso a la información, a datos y a documentos digitales. Estos cambios han involucrado la generalización de herramientas que permiten un acceso directo, sin mediaciones profesionales, a un universo prácticamente inabordable de documentación. Asimismo, han propiciado la creación de nuevos SOC, como las taxonomías web, las folksonomías y las ontologías, hijas directas de la evolución de la cultura digital. Además, se ha reavivado el viejo debate entre el uso del lenguaje natural o de los vocabularios controlados como herramientas de indización, búsqueda y recuperación de información.

En los programas analizado se observa que una concepción implícita sobre la interdisciplina relacionada a la capacidad del sujeto individual y como emergente del objeto de estudio complejo y no hay una consideración específica de la dimensión didáctica para la enseñanza de la OC.

Este exigente escenario obliga la búsqueda de concepciones y métodos que brinden nueva interpretación a las cuestiones de organización y representación temática. Nada más apropiado, que ir a buscarlos a otros ámbitos del saber donde se ponen a prueba las bases epistemológicas tradicionales (historia de la ciencia, sociología del conocimiento), las formas de dialogar con la información (ciencias cognitivas, lingüística, procesamiento de lenguaje natural) o de organizarla (gestión de contenidos, inteligencia artificial).

Las maneras de seleccionar y utilizar los conocimientos que llegan a la OC luego de cruzar un puente interdisciplinario para potenciar las prácticas propias, impactan en la formación de futuros profesionales e investigadores universitarios, en los contenidos y en las formas de enseñar.

A través de esta investigación se aspira a que las instituciones educativas tomen conciencia de la dimensión del factor interdisciplinario y puedan revisar sus currículas en torno al abordaje pedagógico y la estructura curricular. Los docentes contarán con más elementos para reflexionar sobre el diseño de sus cursos, proyectar su implementación e identificar tendencias en la referencia y uso de textos de autores de otros campos del saber.

Un proyecto de estas características puede contribuir a una concepción común de la OC, más allá de los matices y las diferencias de pensamiento propios de un entorno tan diverso en realidades como el iberoamericano.

Por lo expuesto, se encuentra fundamento suficiente para llevar adelante esta investigación, a partir de un protocolo que respete la secuencia metodológica pre-

sentada. La colaboración de los docentes a nivel regional será condición necesaria para llevar el trabajo a buen puerto.

Nota: Se deja constancia de un especial agradecimiento a los responsables de los cursos estudiados y a sus equipos docentes por su colaboración y su generosidad, al ofrecer material de sus programas de posgrado, para poner a prueba en forma experimental la metodología elegida.

Referencias

- Alajmi, B., & Rehman, S. (2016). Knowledge organization trends in library and information education: Assessment and analysis. *Education for Information*, 32(4), 411-420. doi: 10.3233/EFI-160084
- Aytac, S., Kipp M., Neal, D., & Hsieh-Yee I., (2012). Emerging trends and knowledge organization and information organization courses curriculum. *Proceedings of the American Society for Information Science and Technology*. doi: 10.1002/meet.2011.14504801079.
- Barité, M. (2000a). Referenciales teóricos vigentes en el área de Tratamiento Temático de la Información y su expresión metodológica. En *IV Encuentro de Directores y III de Docentes de Escuelas de Bibliotecología y Ciencia de la Información del Mercosur: 24 al 27 de mayo de 2000, Montevideo, Uruguay. Programas, ponencias, documentos de trabajo, acuerdos y recomendaciones* (pp. 224-233). Montevideo: EUBCA.
- Barité, M. (2000b). Aportaciones teórico-conceptuales para la docencia y la investigación en el área de Tratamiento Temático de la Información. En *EDIBCIC. Actas del V Encuentro* (pp. 54-72). Granada: Universidad de Granada.
- Barité, M. (1999). Propuesta de un marco referencial para la docencia en el área de Procesamiento de la Información. En *Tercer Encuentro de Directores y Segundo de Docentes de las Escuelas de Bibliotecología del Mercosur, Santiago de Chile, 29-31 de octubre de 1998. Formación de recursos humanos en el Área de la Información en el Mercosur: compatibilización curricular: competencias del profesional de la información en el Mercosur. Actas de los Acuerdos, recomendaciones, conferencias y ponencias* (pp. 121-128). Santiago de Chile: UTM; CNLL.
- Beri, C., & Tello, C. (2006). Aproximaciones a la nueva trama entre interdisciplina y formación universitaria. *Questión*, 11, 1-9. Recuperado de <http://sedici.unlp.edu.ar/handle/10915/30134>
- Braz, M. I., & Carvalho, E. S. de (2017). Práticas em tratamento temático da informação: interfaces de ensino e aprendizagem. *Revista Brasileira de Biblioteconomia e Documentação*, 13, 2496-2509. Recuperado de <https://rbbd.febab.org.br/rbbd/article/view/975/951>.
- Chaudhry, A. S. (2018). Expansion and enhancement in Knowledge Organization courses: Response to the Requirements of the Digital Environment. *International Journal of Digital Society*, 9(2), 1382-1386.
- Follari, R. (2007). La interdisciplina en la docencia. *Polis: revista latinoamericana*, (16). Recuperado de <https://polis.ulagos.cl/index.php/polis/article/view/484/883>
- García Marco, F. J. (1995). Los contenidos y la secuencia docente de la organización y representación del conocimiento: una propuesta interdisciplinar. En *Organización del Conocimiento en Sistemas de Información y Documentación*, (Vol. 1, pp. 219-228). Zaragoza: Universidad de Zaragoza. Recuperado de <http://dialnet.unirioja.es/descarga/articulo/2341341.pdf>

- Gorman, M. (2002). Why teach cataloguing and classification? *Cataloging & Classification Quarterly*, 34(1-2), 1-13.
- Guimarães, J. A. Ch. (2017). Organização do conhecimento: passado, presente e futuro sob a perspectiva da ISKO. *Informação & Informação, Londrina*, 22(2), 84-98.
- Guimarães, J. A. Ch. (2001a). Brasil: principales tendencias y enfoques en el área. En *IV Encuentro de Directores y III de Docentes de Escuelas de Bibliotecología y Ciencia de la Información del Mercosur: 24 al 27 de mayo de 2000, Montevideo, Uruguay. Programas, ponencias, documentos de trabajo, acuerdos y recomendaciones* (pp. 336-339). Montevideo: EUBCA.
- Guimarães, J. A. Ch. (2001b). *Ensino de tratamento temático da informação nos cursos de biblioteconomia do Mercosul: análise e perspectivas de un core curriculum à luz dos avanços teóricos da área de organização do conhecimento*. Marília: Faculdade de Filosofia e Ciências da UNESP.
- Guimarães, J. A. Ch. (2000). Perspectivas de ensino e pesquisa em organização do conhecimento em cursos de biblioteconomia do Mercosul: uma reflexão. En *EDIBCIC. Actas del V Encuentro* (pp. 206-216). Granada: Universidad de Granada.
- Hjørland, B. (2008). What is Knowledge Organization? *Knowledge Organization*, 35(2/3), 86-101.
- Kelly, A.V. (2009): *The curriculum: Theory and Practice*. 6th. edition. London: Sage.
- Lenoir, Y. (2013). Interdisciplinariedad en educación: una síntesis de sus especificidades y actualización. *Inter disciplina*, 1(1), 51-86. Recuperado de <http://www.revistas.unam.mx/index.php/inter/article/view/46514>
- Mai, J.-E.. (2000). *The subject indexing process: An investigation into the problems in knowledge representation: Ph.D. dissertation*, University of Texas, Austin.
- McIlwaine, I.C. (2003). Trends in knowledge organization research. *Knowledge Organization*, 30(2), 75-86.
- Morgan, J., & Bawden, D. (2006). Teaching knowledge organization: Educator, employer and professional association perspectives. *Journal of Information Science*, 32(2), 108-115. doi: <https://doi.org/10.1177/0165551506062324>
- Nieto Caraveo, M. L. (1991). Una visión sobre la interdisciplinariedad y su construcción en los Currículos profesionales. *Cuadrante, Revista de Ciencias Sociales y Humanidades, México*, (5-6) (Nueva Época). Recuperado de http://acreditacion.unillanos.edu.co/CapDocentes/contenidos/dis_ambientes_metodos_pedagogicos/Memoria2/vision_sobre.PDF
- Ornelas Navarro, C. (1982) La reforma universitaria y la enseñanza tubular. *Foro universitario*, Época II (19), 29-46.
- Parentelli, V. (2019a). *La interdisciplina en la enseñanza de grado: El caso de la Facultad de Información y Comunicación de la Universidad de la República: Tesis de posgrado presentada en la Universidad Nacional de La Plata para optar al grado de Doctora en Ciencias de la Educación*. Recuperado de <http://www.memoria.fahce.unlp.edu.ar/library?a=d&c=tesis&d=Jte1823>
- Parentelli, V. (2019b). La inclusión de la interdisciplina en los planes de estudios de las carreras de la Facultad de Información y Comunicación. *Informatio*, 24(1), 42-60. Recuperado de <https://informatio.fic.edu.uy/index.php/informatio/article/view/219/216>
- Rehman, S., & Al-Ajami, B. (2017). Knowledge organization content in graduate courses. *Library Review*, 66(1/2). doi: 10.1108/LR-03-2016-0028
- Rugarcía, A. (1997). La interdisciplinariedad: el reino de la confusión. *Revista de la Educación Superior*, 25(98), 1-8. Recuperado de http://publicaciones.anui.es.mx/pdfs/revista/Revista98_S1A4ES.pdf
- Salaba, A. (2020). Knowledge organization requirements in LIS graduate programs. In International Society for Knowledge Organization (ISKO), M. Lykke, T. Svarre, M. Skov

- and D. Martínez-Ávila (Eds.). *Knowledge Organization at the Interface, Proceedings of the Sixteenth International ISKO Conference, 2020 Aalborg, Denmark*, Advances in Knowledge Organization, (17): 384-393 doi: 10.5771/9783956507762-384
- Smiraglia, Richard P. (2012). Knowledge organization: Some trends in an emergent domain. *El profesional de la información*, 21(3), 225-227. Recuperado de <http://dx.doi.org/10.3145/epi.2012.may.01>
- Sociedad Internacional para la Organización del Conocimiento (ISKO)-Capítulo Ibérico (2007). *La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento científico: actas del VIII Congreso ISKO-España. León, 18, 19 y 20 de abril de 2007*: org por Blanca Rodríguez Bravo y María Luisa Alvite Diez. León: Universidad de León.
- Sotolongo, P., & Delgado, C. (2006). *La revolución contemporánea del saber y la complejidad social: hacia unas ciencias sociales de nuevo tipo*. Buenos Aires: CLACSO.
- Sutz, J. (2015) Prólogo. En Espacio Interdisciplinario de la Universidad de la República (2015). *Encuentros sobre la Interdisciplina*. Montevideo: Trilce.
- Szostak, R., Gnoli, C., & Lopez-Huertas, M. (2016). *Interdisciplinary Knowledge Organization*. Switzerland: Springer.

INOVAÇÕES DOS RECURSOS EDUCACIONAIS ABERTOS: O CASO DA UNIVERSIDADE DE SÃO PAULO, BRASIL

Célia Regina de Oliveira Rosa¹

¹Doutoranda em Ciência da Informação pela Universidade de Coimbra,
celia.rosa@student.fl.uc.pt,
ORCID iD 0000-0003-4503-1575

Resumo

Esta comunicação apresenta os recursos educacionais abertos (REA). Expõe as cinco ações para definição do padrão REA estabelecidas pela UNESCO em 2019 visando uma maior adesão mundial. O objetivo geral caracteriza o desenvolvimento da criação dos REA no ensino superior. Para o cumprimento do objetivo geral, identifica-se como objetivo específico verificar se os recursos existentes do sistema e-Aulas concebido pela Universidade de São Paulo (USP), nomeadamente pela Pró-Reitoria de Graduação, atuam como vetores em apoio ao ensino universitário. A metodologia propõe busca a partir do uso de operadores booleanos (AND, OR) nos campos título, assunto e palavras-chave abrangendo o período cronológico de 2016 a 2020 de modo a evidenciar as inovações na criação dos REA nas bases de dados *Library, Information Science & Technology Abstracts* e *Scopus*. Utilizou-se os termos recursos educacionais abertos; REA; ensino superior. O estudo tem características descritivas sugerindo uma abordagem metodológica da pesquisa de cunho qualitativo. Os resultados de pesquisa bibliográfica enfocam a importância do incentivo por parte da universidade de modo a fortalecer e expandir tais recursos, bem como os desafios para implementação. O resultado informa aumento dos recursos utilizados pela comunidade universitária e sustentabilidade com participação de docentes, bibliotecários e alunos. Conclui que o aprimoramento dos recursos entre grupos diversos permitirá a inclusão e equidade referentes à criação, acesso, reutilização, adaptação e redistribuição dos REA.

Palavras-chave: Recursos Educacionais Abertos, REA, Ensino superior, Universidade de São Paulo, Sistema e-Aulas.

Introdução

Somado aos esforços anteriores à década de 70, o Projeto Gutenberg encarregou-se da digitalização e distribuição de obras culturais em formatos abertos. Em 2002, o movimento do Acesso Aberto (AA) é ratificado pela Declaração de Acesso Aberto de Budapeste (BOAI – Budapest Open Access Initiative) permitindo a abertura e acesso à literatura produzida pela academia, consolidando-se mundialmente por meio de diferentes iniciativas de publicação em rede em diversos tipos de materiais como revistas, *proceedings*, *textbooks*, livros digitais abertos.

Neste contexto de disseminação da produção acadêmica, a biblioteca universitária desempenha aplicações inovadoras, auxiliando e trabalhando em parceria com docentes na composição e promoção de materiais de suporte à educação concebidos pelos responsáveis dos seminários para fins de uso e aprendizado educacionais.

Esta comunicação tem como objetivo geral caracterizar o desenvolvimento da criação dos REA no ensino superior. Para o cumprimento do objetivo geral, identifica-se o seguinte objetivo específico: (i) caracterizar os recursos existentes do sistema e-Aulas criado pela Universidade de São Paulo (USP), nomeadamente pela Pró-Reitoria de Graduação, disponíveis em apoio ao ensino universitário.

O conceito referente aos Recursos Educacionais Abertos (REA), cunhado pela UNESCO em 2002, é definido como “materiais de ensino, aprendizagem e pesquisa em qualquer meio, digital ou outro, constituído em domínio público ou lançados sob uma licença aberta que permite acesso gratuito, uso, adaptação e redistribuição por outros com nenhuma ou restrições limitadas”.

Tais recursos educacionais compreendem mapas curriculares, materiais de cursos, livros didáticos, vídeos, multimídia, *podcasts*; qualificados para o ensino e formação acessíveis por educadores e alunos, isentos de custos autorais ou taxas de licença” (FIOCRUZ, 2019).

A Declaração da Cidade do Cabo para Educação Aberta (2007), reitera o conceito de REA relacionado à educação aberta proveniente das práticas pedagógicas abertas pela ampliação da aprendizagem considerando circunstâncias e cenários diferentes combinados por elementos de colaboração e interatividade de modo a prover aos interessados, conhecimento para conteúdos variados.

Wiley (2007), concebeu a estrutura 4Rs de permissões para definição dos REA, adotadas mundialmente por permitir estruturar as atividades a partir dos conceitos de: reutilizar, revisar, remixar e redistribuir. Alguns anos mais tarde o autor propõe o quinto elemento “Retain”: reter, tornando-se os 5Rs denominados de “liberdades mínimas” por gerar um círculo eficaz de comprometimento à medida que os utilizadores usem e adaptem criações de outros autores, compartilhando por sua vez o que criaram para a comunidade novamente (Sebriam; Markun & Gonsales, 2017).

No Brasil, a relevância dos REA atinge governos e a sociedade civil reunidos em esforços para estudo do tema, aplicações e proposta de ações educacionais. Sebriam; Markun & Gonsales (2017) afirmam,

(...) a discussão em torno dos REA foi intensificada a partir de 2008, com a mobilização de uma comunidade eclética, composta por pessoas de distintas áreas do conhecimento e também com a criação do Projeto REA.br – Projeto Brasileiro sobre Recursos Educacionais Abertos: Desafios e Perspectivas. Apoiado pela Open Society Foundation até o ano de 2015, o REA.br gerou uma aproximação com gestores e políticos e também com atores da sociedade civil, por meio de oficinas, debates e eventos.

Em novembro de 2019, este conhecimento acessível aliado ao ensino informal à distância (Santos, 2013) recebeu recomendações relacionadas ao apoio para criação dos REA inclusivos, de qualidade e cooperação internacional por ocasião da 40.ª Conferência Geral da UNESCO.

Particularmente a crise de saúde, sanitária, econômica e até política pelas quais alguns países experimentam em 2020, impõe a chamada “Coalizão Dinâmica REA” almejando buscar e instituir a cooperação global da sociedade.

A recomendação REA da UNESCO pretende o cumprimento de cinco ações impulsionadoras relativas à criação e uso:

- 1) Capacitar interessados na criação, acesso, reutilização, adaptação e redistribuição REA;
- 2) Promover política de apoio para REA;
- 3) Incentivar os REA inclusivos e igualitários;
- 4) Desenvolver modelos de sustentabilidade para REA;
- 5) Favorecer e auxiliar a cooperação internacional em REA.

A execução e implementação das ações para definição do padrão REA contribuirão com o desempenho de parte dos Objetivos de Desenvolvimento Sustentável (ODS) referentes à Agenda 2030 da ONU a saber: ODS 4 (Educação de qualidade), ODS 5 (Igualdade de gênero), ODS 9 (Indústria, inovação e infraestrutura), ODS 10 (Desigualdades reduzidas dentro e entre países), ODS 16 (Paz, justiça e instituições fortes) e ODS 17 (Parcerias para os objetivos). Com isso espera-se um número maior de recursos para uso no ensino-aprendizagem além de docentes orientados para contribuir nesta empreitada deste segundo momento de difusão dos REA.

Decorridos quase 20 anos de existência dos REA, ainda há muito por investir considerando questões como: adoção parcial pelo corpo docente nas universidades; produção descentralizada, falta de conhecimento técnico e fluxos de trabalho para produção prática de materiais, ferramentas de publicação para depósito de uma tipologia de documentos além de vídeos e textos, capacidade e maior participação das bibliotecas universitárias no oferecimento de serviços de apoio para desenvolvimento e uso destes (Thompson & Muir, 2020).

Esta conjuntura demonstra que tais recursos não estão disponíveis a todos os cursos e alguns podem não atender aos padrões de qualidade de seu corpo docente (Stafford, 2020), que por sua vez estão dispostos a incluí-los nos currículos, apesar da escassez de recursos apropriados e do tempo para identificá-los e modificá-los. (Xu, 2018).

Considerando indicadores sobre falta de orientação técnica para produção dos REA (Thompson & Muir, 2020), sugere-se o trabalho conjunto entre biblioteca-docentes (Xu, 2018). Embora docentes se conectem à repositórios de REA, ainda estão afastados e isolados de sua criação (Thompson & Muir, 2020) talvez pelo caráter incomum que tais recursos possam expressar considerando o ambiente tradicional educativo.

Outra adversidade reside na aplicação de licenças e direitos autorais, sendo indispensável o exercício de treinamentos profissionais, reuniões para esclarecimentos sobre a abertura que estes materiais impõem, o acesso aberto (AA) dos recursos fundamentados desde a criação, além do reconhecimento do conceito para os mesmos.

Ao observarmos outros tipos de materiais de apoio ao ensino, nos deparamos com a produção de livros didáticos que poderia ser mais acessível e livre por meio de programas lançados na universidade englobando todas as áreas do conhecimento.

O custo da aquisição deste material pode ser poupado por meio de distribuição gratuita, para que os alunos possam utilizar possíveis valores advindos de bolsas de estudos e verbas em outras despesas. Um número maior de bibliotecas reunidas em esforço colaborativo com docentes autores, pesquisadores, podem alavancar o processo de publicação de livros digitais em auxílio ao ensino de graduação e da pós-graduação nas universidades que ainda não empreenderam esforços nesta matéria.

A disponibilidade em AA desses recursos traz consequências relacionadas à permanência e sucesso do aluno na universidade, ao mesmo tempo que a promoção, adoção e uso dos REA podem ser fortalecidas por palestras rápidas, esclarecedoras sobre demandas de sustentabilidade em programas distintos (Park, Plumer & Deforest, 2018).

Não é qualquer material que pode ser aproveitado como recurso educacional aberto, é fundamental que os docentes sejam ouvidos para exposição de dúvidas em busca de acesso igualitário para formação desta comunidade de educação aberta (Stafford, 2020). Recursos como *Wikipédia* têm uso e importância aumentados, mesmo diante da falta de indicadores ou medidas que comprovem sua contribuição para o desenvolvimento acadêmico; um estudo entre universitários revelou que uso combinado com recursos de aprendizagem convencionais geram uma atuação acadêmica positiva dos alunos (Meseguer-Artola et al., 2020).

Em referência aos bibliotecários, estes se defrontam com inumeráveis desafios técnicos integrantes do cenário de edificação dos REA, relacionados tanto ao alargamento da educação do utilizador quanto à padrões de arquivo aberto, uso de ferramentas em colaboração que permita controlar versões distribuídas (Patterson, 2020), elevando seguramente as capacidades e pontos fortes das Bibliotecas acadêmicas no apoio, à promoção e até mesmo cultivo e criação de recursos educacionais (Okamoto, 2013).

Além de docentes e alunos envolvidos com o processo de criação dos REA, bibliotecários têm participação ao longo do processo de criação destes contribuindo ativamente em diferentes iniciativas pelo estímulo profissional alcançado, decorrentes da relevância do aprendizado, do atendimento aos estudantes e do movimento entre a educação aberta e a biblioteconomia (McLure & Sinkinson, 2020).

A excelência em torno dos REA se dá pela definição, o rigor e qualidade dos recursos, contudo existe um espaço vazio em torno do benefício de tais práticas abertas. Sua praticidade e aplicabilidade sugerem que se tratam de recursos não populares junto à comunidade por implicar em ausência de sentimento de domínio e propriedade (Heck et al., 2020). Alguns docentes inviabilizam o acesso ao REA produzido por comercializarem o conteúdo junto a editoras, perdendo deste modo a autoria sobre a propriedade produzida.

O ensino superior agrega vantagens provenientes da contribuição que pode obter dos alunos que se utilizam de imagens para comunicação (tarefas multimodais) e pelos bibliotecários que prosseguem orientando estes mesmos alunos no uso dos REA, sobre direitos autorais e de propriedade intelectual (Kohout-Taylor & Sheaffer, 2020), implicando em um círculo de sustentação para estes recursos.

Metodologia

O estudo tem características descritivas sugerindo uma abordagem metodológica da pesquisa de cunho qualitativo. A revisão de literatura enfatiza os REA relacionados ao ensino superior.

A escolha das bases de dados *Library, Information Science & Technology Abstracts* (LISA) e SCOPUS foi baseada pela correlação a partir da Biblioteconomia e Ciência da Informação e, pelo caráter geral que proporciona uma visão ampla, respectivamente.

A busca decorreu pelo uso de operadores booleanos (AND, OR) nos campos título, resumo e palavra-chave, abrangendo o período cronológico de 2016 a 2020 nas bases de dados LISA e SCOPUS no sentido de fundamentar a resposta à pergunta de pesquisa “Existe ou não inovações na criação dos REA após quase vinte anos de existência?”

Foram utilizados os termos: recursos educacionais abertos (*open educational resources*), OER, ensino superior (*higher education*) para suporte desse estudo de caso. Na base LISA, a literatura foi limitada ao texto completo em revistas acadêmicas pela expressão de busca “open educational resources OR OER AND higher education”, recuperando-se o total de 39 artigos, com 14 textos selecionados orientados ao ensino superior.

A busca na base de dados SCOPUS obedeceu à mesma expressão de busca informada acima, limitada aos artigos de revistas em acesso aberto nos idiomas inglês e português, recuperando-se o total de 95 documentos com 16 textos selecionados.

Os textos selecionados acentuaram o uso dos livros didáticos abertos (*open textbooks*) como inovação pedagógica aos recursos educacionais abertos; além de abordagens como a implantação de REA por bibliotecas e a importância da adoção desses objetos por docentes; aprendizagem dos educadores orientada aos REA; eficiência de custos de produção e envolvimento positivo dos alunos através do uso dos REA.

Sistema e-aulas da Universidade de São Paulo

Em 2012, a Universidade de São Paulo (USP), nomeadamente a Pró-Reitoria de Graduação, apresenta para fins de aprendizagem, conteúdos educacionais produzidos de mídia digital disponíveis a partir do sistema e-Aulas USP.

O objetivo do serviço é ampliar a disseminação do conhecimento, incentivar e apoiar o corpo docente na criação e acesso de áudios, vídeos, textos, apresentações ligadas a unidades curriculares dos cursos existentes na Universidade.

O gerenciamento do sistema é conduzido pelo Núcleo de Mídias Digitais, o qual ratifica que “os benefícios didáticos observados pelo consumo de recursos educacionais abertos, particularmente vídeos, resultaram em indicativo para o desenvolvimento do sistema”. A linguagem visual e a facilidade de acesso aos conteúdos em rede, resulta em importante fonte de informação de apoio à formação dos alunos e interessados externos à comunidade USP.

A Política de uso do conteúdo se pauta por direitos intelectuais, protegidos de acordo com a Lei 9.610/1998 (Lei de Direitos Autorais ou LDA) e relativo aos direitos patrimoniais pertinentes, às obras derivadas de trabalho individual ou coletivo do corpo docente pertencendo exclusivamente à Universidade de São Paulo.

Metadados asseguram a identificação de autores e descrição das mídias e materiais, caso violem direitos e propriedades intelectuais de terceiros, excluindo-se a divulgação.

O sistema e-Aulas USP permite três modalidades de acesso:

1. Utilizador USP: facultado ao público interno USP com cadastro ativo nos Sistemas da Universidade;
2. Utilizador outros: aberto ao público em geral mediante cadastro prévio no sistema e-Aulas;
3. Livre: aberto ao público em geral sem necessidade de registro.

O sistema propõe tutoriais sobre elaboração de atividades didáticas de graduação online, oficinas para publicação de vídeos, mídias digitais, criação de aulas por meio de webcam/*Powerpoint*.

USP Universidade de São Paulo
e-Aulas: Portal de vídeoaulas

Pesquisar vídeos

SOBRE **COMO FUNCIONA** **POLÍTICA DE USO** **CRÉDITOS** **CONTATO** **AJUDA** **e-Aulas**

Publicando vídeos no E-aulas - USP
FABIO GREGORI E REGINA MELO SILVEIRA
Publicando vídeos no E-aulas

Veja os vídeos disponíveis no e-Aulas em cada nível de ensino:

- Graduação
- Pós-graduação
- Cultura e extensão
- Ensino médio

Pesquisar vídeos por

Áreas de estudo

- Exatas**
- Humanas**
- Biológicas**

Vídeos recentes

- OBSTETRICIA - Práticas em Obstetrícia - Aula 3**
[2340020-322101] Práticas em Obstetrícia Online
Jovane Rodrigues Cruz Moreira
Práticas em Obstetrícia Online - Aula 3 - Ivan Onone Galatin
- [610] Processos de ensino aprendizagem em salas**
Antonio Carlos de Azevedo
Sistemas de Produção Agrícola
- MEDICINA VETERINÁRIA - DNP1100-1) Nutrição Animal**
Paulo Henrique Mazza Rodrigues
Ácido ascórbico (Vitamina C)
- ASTRONOMIA - [AGA215-5] Fundamentos de Astronomia**
Roberto Del'Aglio Dias Da...
Exatas de distância e dimensões em astronomia - Segunda Parte

Vídeos mais vistos

- MATEMÁTICA**
[1600001-1] Tratamento Didático em Matemática para...
João Lopes Lourenço
Quantidade de Divisores Positivos
- MEDICINA VETERINÁRIA**
[Prática Prd-1] Oficina PRG
Fabio Gregori e Regina Melo...
Tutorial sobre criação de vídeos
- MEDICINA VETERINÁRIA**
[PRG-1] Oficina de produção de vídeos
Fabio Gregori e Regina Melo...
Publicando vídeos no E-aulas
- MEDICINA VETERINÁRIA**
[PRG-1] Oficina de produção de vídeos
Fabio Gregori e Regina Melo...
Criando vídeo aulas com Webcam e Powerpoint

Como funciona o e-Aulas?

- Acesse e assista grátis!**
Mais de 4103 horas de vídeos
Tenha acesso a vídeos de diferentes áreas do conhecimento. Acompanhe uma disciplina completa!
[Saiba mais](#)
- Professor**
Disponibilize aulas, palestras e cursos
É professor da USP? Vídeo também é publicação acadêmica! Disponibilize suas aulas! É rápido, é fácil!
[Saiba mais](#)
- Aluno**
Assista, aprenda e organize o conteúdo
Anote, veja e reveja suas aulas! Compartilhe com seus amigos!
[Saiba mais](#)

Pró-Reitoria de Graduação
Telefone: +55 11 3091-9942

Figura 1. Sistema e-Aulas da Universidade de São Paulo

A caracterização dos recursos provenientes do sistema e-Aulas USP desde a criação em 2012 até 2020, apresenta recursos variados como: vídeos, aulas, conferências, palestras, totalizando 105 PDF, 5 ZIP, 2 JPG, 1 TXT; além de 85 links.

Alunos da graduação USP usufruem muito mais dos REA; seguidos por utilizadores “anônimos” sem vínculo USP com acesso a 30% do acervo do e-Aulas USP publicados em acesso aberto sem exigência de cadastro na plataforma.

A figura 2 demonstra queda anual desde 2012 na produção dos REA USP. Percebe-se leve aumento na criação de conteúdos nos anos de 2014 e 2015. Entretanto em 2020, evidencia-se um salto quantitativo de recursos produzidos necessários à manutenção do ensino virtual, justificado pelo efeito da pandemia.

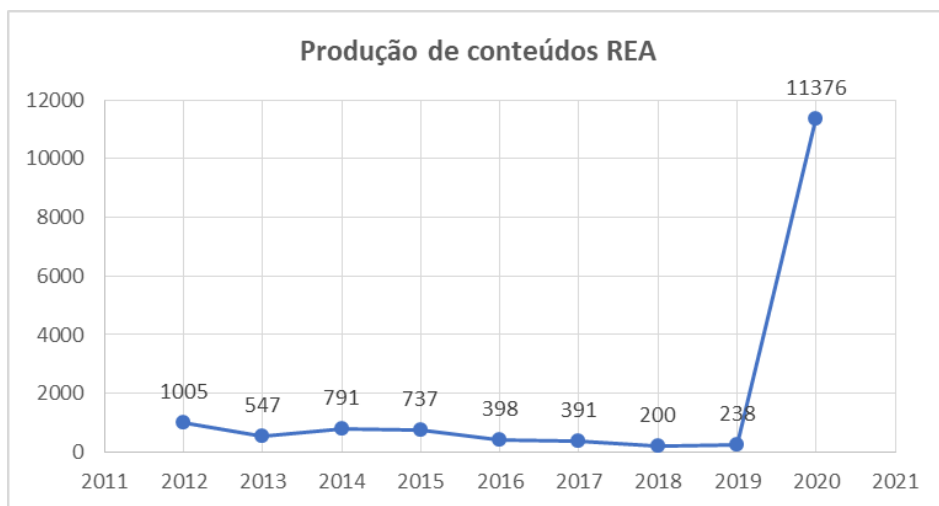


Figura 2. Produção de conteúdos REA (Fonte: Sistema e-Aulas USP)

A comunidade USP ampliou, gradualmente, o interesse pelos REA conferidos pelos acessos registrados. Em 2012, os números atingiram 1083 acessos, retratando aumentos discretos nos anos posteriores de 2019 com 24059 acessos. Contudo, em 2020 houve aumento significativo culminando em 208.977 consultas e uso do sistema.

Os dados coletados do sistema e-Aulas USP demonstram a visibilidade dos recursos compartilhados por docentes, somados à importância das licenças *Creative Commons* e atribuição de acesso gratuito, uso, adaptação e redistribuição dos REA. Canchola (2021), assinala que os REA são criados com a intenção genuína de ajudar outras pessoas pelo “espírito de compartilhamento” existente nesta biblioteca desconhecida ainda por muitos.

Discussão

Diversos níveis do desenvolvimento dos REA pautam a criação nas diferentes instituições de ensino e autores sugerem que integrantes da comunidade universitária se envolvam para geração de mais conteúdos. Programas de pós e graduação que contam com incentivos relacionados ao financiamento e sustentabilidade para REA

(Park, Plumer, & Deforest, 2018) serão extremamente benéficos, uma vez que os alunos poderão participar ativamente do processo de criação destes recursos.

Quanto ao desempenho dos REA pelos alunos USP, vale ressaltar o nível de uso e apreciação destes conteúdos pela facilidade de acesso em rede, esclarecendo que número de downloads não se constituem em dado suficiente para conhecer as preferências dos leitores; além de que a integração dos REA no portal global da universidade pode contribuir com aumento da visibilidade e melhor organização da pesquisa por interessados.

A participação de um número maior de docentes promoverá maior pró-atividade tanto pela produção e melhoria da qualidade dos recursos quanto à promoção de seu uso, elevando a quantidade desses meios na USP e, somando-se à participação de bibliotecários e da biblioteca universitária que não está preocupada somente com questões de armazenamento, gerenciamento e disseminação desta produção específica, mas em extrapolar conhecimento para contribuir com o aumento da criação dos REA.

O trabalho envolve a comunidade universitária, sendo fundamental aliança entre docentes, alunos e biblioteca.

Andone et al. (2020), ressalta que o uso dos recursos educacionais não atingiu sua potencialidade ou capacidade máxima, sugerindo a participação de alunos como “co-criadores” de conteúdos de aprendizagem inéditos; fato que contribui por assimilar os princípios de educação aberta além do aprimoramento de competências de aprendizagem digital. Sugere-se, portanto, que a USP estabeleça diretrizes para envolvimento de alunos acerca da criação dos REA.

O percurso dos recursos educacionais abertos ao longo do tempo tem recebido apoio e questionamentos de seus desenvolvedores e utilizadores a respeito da influência e práticas de capacitação tecnológica. Zhang (2020), sugere incorporação de ferramentas amigáveis para uma “aprendizagem móvel”, o que inclui a utilização de dispositivos móveis para influir positivamente em sua aplicação.

Conclusão

Treinamentos e capacitações dos mais diversos precisam ser extrapolados em benefício desta prática educacional aberta que atinge cidadãos pelos mais diversos interesses.

A coalizão de grupos interessados no desenvolvimento e aprimoramento destes recursos vai garantir fontes inclusivas e igualitárias referentes à criação, acesso, reutilização, adaptação e redistribuição destes conforme a conveniência e necessidade de cada grupo face ao benefício destes recursos serem produzidos a custos baixos.

Neste sentido, a Universidade de São Paulo tem fomentado o uso dos REA para dar cumprimento à ação impulsionadora de capacitação de docentes na criação de novos materiais por meio do acesso, reutilização, adaptação para sua comunidade interna, ao mesmo tempo que oferece e disponibiliza tais recursos à sociedade.

Contudo, mais investimentos dos diferentes setores da universidade são essenciais para composição e geração de novas tipologias de recursos educacionais abertos além dos vídeos, para atingir benefícios e vantagens do uso pela variada comunidade multidisciplinar.

A resposta às inovações dos REA após quase 20 anos de existência, sugere a expansão de muito trabalho participativo para integração e fortalecimento destes recursos no ensino superior.

Referências

- Andone, D., Mihaescu, V., Vert, S., Ternauciuc, A., and Vasii, R. (2020). Students as OERs (Open Educational Resources) co-creators. *IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, Tartu, Estonia, 2020, p. 34-38. <https://doi.org/10.1109/ICALT49669.2020.00017>
- Canchola, A. (2021). We answer your questions about open educational resources. The Observatory of Educational Innovation of Tecnológico de Monterrey. Transcription by Fernanda Ibáñez. <https://observatory.tec.mx/edu-news/we-answer-your-questions-about-oer>
- Declaração de Acesso Aberto de Budapeste. (2002). <https://www.budapestopenaccessinitiative.org/boai-10-translations/portuguese>
- Declaração de Cidade do Cabo para Educação Aberta. (2007). Brazilian Portuguese Translation. <https://www.capetowndeclaration.org/translations/portuguese-translation>
- FIOCRUZ. (2019). *Recursos Educacionais Abertos: Guia completo*. <https://campusvirtual.fiocruz.br/portal/guiarea/OpenEducation.html>
- Heck, T., Peters, I., Mazarakis, A., Scherp, A., Blümel, I., & Bossu, C. (2020). Open science practices in higher education: Discussion of survey results from research and teaching staff in Germany. *Education for Information*, 36(3), 301-323. <https://doi.org/10.3233/EFI-190272>
- Kohout-Taylor, J., & Sheaffer, K.E. (2020). Using open educational resources to empower student creators. *Journal of Electronic Resources Librarianship*, 32(1), 11-18. <https://doi.org/10.1080/1941126X.2020.1709728>
- McLure, M., & Sinkinson, C. (2020). Caring for students in postsecondary open educational resource (OER) and open education initiatives: Inviting student participation and voice. *Reference Services Review*, 48(3), 473-487. <https://doi.org/10.1108/RSR-03-2020-0018>
- Meseguer-Artola, A., Rodríguez-Ardura, I., Ammetler, G. & Rimbau-Gilabert, E. (2020). Academic impact and perceived value of Wikipedia as a primary learning resource in higher education. *El Profesional de La Información*, 29(3), 1-16. <https://doi.org/10.3145/epi.2020.may.29>
- Patterson, J. (2020). User Education and File standards Best Options to Ensure Open Educational Resources are truly Open. *Evidence Based Library & Information Practice*, 15(1), 236-237. <https://doi.org/10.18438/eblip29685>
- Okamoto, k. (2013). Making Higher Education More Affordable, One Course Reading at a Time: Academic Libraries as Key Advocates for Open Access Textbooks and Educational Resources. *Public Services Quarterly*, 9:4, 267-283. <https://doi.org/10.1080/15228959.2013.842397>
- PARK, K., Plumer, D.C., & Deforest, L. (2018). Opening the Door to Open Educational Resources in Higher Education. *Texas Library Journal*, 94(4), 18-19. <https://txla.org/wp-content/uploads/2019/03/TLJ-Winter-2018.pdf>
- Santos, A.I. (2013). *REA no Brasil: O estado da arte, desafios e perspectivas para o desenvolvimento e inovação*. Comitê Gestor da internet no Brasil; 1,6Mb. <https://cetic.br/media/docs/publicacoes/8/rea-andreia-inamorato.pdf>
- Sebriam, D.; Markun, P.; Gonsales, P. (2017). Como implementar uma política de Educação Aberta e Recursos Educacionais Aberto (REA): guia prático para gestores. São Paulo:

- Cereja Editora. https://guiaea.educadigital.org.br/wp-content/uploads/2017/09/Guia_REA_Online.pdf
- Stafford, D. (2020). Promoting Open Educational Resources: A Beginner's Playbook. *Pennsylvania Libraries: research & Practice*, 8(2), 103-114. <https://doi.org/10.5195/palrap.2020.233>
- Thompson, S.D., & Muir, A. (2020). A case study investigation of academic library support for open educational resources in Scottish universities. *Journal of Librarianship & Information Science*, 52(3), 685-693. <https://doi.org/10.1177/0961000619871604>
- UNESCO. (2019). Open Educational Resources (OER). <https://en.unesco.org/themes/building-knowledge-societies/oer>
- UNESCO. (2020). Recommendation on Open Educational Resources (OER). *UNESDOC Digital library*. <https://unesdoc.unesco.org/ark:/48223/pf0000373755?posInSet=1&queryId=4781ded5-9f74-47f9-bdd0-4f137944fdb4>
- UNIVERSIDADE DE SÃO PAULO. Pró-Reitoria de Graduação. Sistema e-Aulas. <http://eaulas.usp.br/portal/home.action;jsessionid=0189176AE8C799C8D752B102F5FC D46B>
- Wiley, D. (2014). The Access Compromise and the 5th R. <https://opencontent.org/blog/archives/3221>
- Xu, Hong. (2018). Obstacles for faculty using open educational resources and solutions. *Texas Library Journal*, 94(3), 85-87.
- Zhang, X., Tlili, A., Huang, R., Chang, T., Burgos, D., Yang, J., & Zhang, J. (2020). A case study of applying open educational practices in higher education during COVID-19: Impacts on learning motivation and perceptions. *Sustainability*, 12(21), 9129. <https://doi.org/10.3390/su12219129>

(Página deixada propositadamente em branco)

ORGANIZACIÓN DEL CONOCIMIENTO EN EL DISEÑO CURRICULAR ACADÉMICO. PERSPECTIVAS DESDE LOS NUEVOS GRADOS EN DOCUMENTACIÓN IMPLANTADOS EN ESPAÑA

María Luisa Alvite Díez

Área de Biblioteconomía y Documentación. Universidad de León,
luisa.alvite@unileon.es, ORCID ID 0000-0003-1490-8936

Resumen

Este trabajo aborda el estudio del currículo relacionado con la Organización del Conocimiento (OC) en los programas de grado propuestos desde el ámbito de la Documentación en las universidades públicas españolas en estos últimos cuatro años con el fin de conocer el alcance de la OC en este perfil curricular más tecnológico. Se ha empleado el análisis de contenido para estudiar los planes de estudio, las guías docentes disponibles y las memorias de verificación de los títulos académicos objeto de análisis. Los resultados se contrastan con trabajos anteriores relacionados con las competencias en OC en los programas académicos y se alinean con las competencias establecidas sobre esta materia en el *Libro blanco* de ANECA, sobre el que se asentaron los grados en Información y Documentación precedentes, así como con las señaladas en el *Libro blanco para el diseño de titulaciones universitarias en el ámbito de la Economía Digital*. Los planes de estudio de estos nuevos grados, ligados a un mercado laboral relacionado con la transformación digital y las necesidades sociales que conlleva, creemos que, en conjunto, fortalecen las competencias en OC y dan muestra de los profundos cambios que se han producido en la disciplina y de su vinculación indiscutible con las tecnologías y la información electrónica.

Palabras-clave: Diseño curricular, Documentación, España, Evolución, Organización del conocimiento, Planes de estudio.

Introducción

Desde que la disciplina de la Documentación (Library and Information Science- LIS-) cuenta con programas formativos de nivel universitario, la Organización del Conocimiento (OC) ha sido tradicionalmente un componente fundamental del contenido curricular de los mismos. Atendiendo a Hjørland (2008), en sentido estricto, la OC trata de actividades como la descripción de documentos, la indexación, la clasificación que se realizan en bibliotecas, bases de datos bibliográficas, archivos y otros tipos de “instituciones de la memoria” por bibliotecarios, archiveros, especialistas en información, especialistas en la materia, así como por algoritmos informáticos y personal no experto. La OC como campo de estudio se preocupa por la natu-

raleza y la calidad de dichos procesos de organización del conocimiento, así como por los sistemas de organización del conocimiento empleados.

Añade el propio Hjørland que la OC se adscribe a la disciplina matriz de la Documentación en sentido riguroso, si bien está seriamente cuestionada por otros campos disciplinares, entre los que subraya la informática. Andersen y Skouvig (2006) describen la OC como un elemento central de la disciplina de la Documentación y para ur Rehman y Alajmi (2017) los cambios constantes y generalizados en los sistemas y herramientas de OC, la aparición de las disciplinas de gestión de la información y del conocimiento han añadido nuevos enfoques de indexación, recuperación y acceso. Asimismo, en las últimas décadas se ha prestado atención a los usuarios finales en el diseño de otras vías para el acceso y uso de la información, particularmente en medios sociales. A todo lo anterior se suma el énfasis, cada vez mayor, en el desarrollo y la disponibilidad de herramientas y servicios específicos para las corporaciones.

Los programas de Documentación (LIS) en los últimos 20 años han abordado cambios notables que responden al proceso vertiginoso de transformación digital en el que nos hallamos inmersos (Moreiro González, 2019; Wyman y Imamverdiyev, 2018). Además, los desarrollos tecnológicos y la explosión de contenidos digitales han dado una nueva orientación y tratamiento a la Organización del Conocimiento, nuevos conceptos, sistemas y herramientas emergentes han de integrarse adecuadamente en los contenidos tradicionales de OC asentados históricamente en los programas formativos de Documentación para que los profesionales de la información cuenten con las capacidades que, en este ámbito concreto, demanda el mercado laboral. El estudio de ur Rehman y Alajmi (2017) aborda precisamente las necesidades de desarrollo de competencias en OC entre los profesionales de la información desde la perspectiva de los hallazgos de un análisis sistemático de los planes de estudio de sesenta y ocho centros educativos. El trabajo subraya la necesidad de incluir en los programas de Documentación temas como taxonomías, folksonomías, ontologías, esquemas de recuperación de información, etc.; herramientas indispensables en los contextos modificados por la expansión del conocimiento, la interdisciplinariedad y el cambio del papel de las instituciones.

En lo que se refiere a la formación universitaria de grado en España, la reordenación de las titulaciones, promovida por la aplicación del Espacio Europeo de Educación Superior, fijó entre sus objetivos adecuar el contenido de los estudios universitarios a las demandas sociales y laborales. El título resultante, Grado en Información y Documentación siguió muy de cerca, en cuanto a objetivos formativos y competencias, las líneas establecidas en el *Libro blanco del Título de Grado en Información y Documentación*, publicado en 2004, recogiendo los consensos alcanzados por las universidades que participaron en su elaboración (Agustín Lacruz y Salvador Olivan, 2016; ANECA, 2004).

El *Libro blanco* situó los contenidos referidos a la OC en la materia troncal “Representación y Recuperación de la información” y recibió una asignación de 36 créditos sobre 240 créditos totales (ANECA, 2004, p. 96). Las competencias relevantes de la materia, señalan García Marco et al. (2006, pp. 25-30), siguen, en buena medida, la formulación de las materias tradicionales del área de Biblioteconomía y Documentación y revelan problemas, entre otros, los relacionados con el solapamiento e indiferenciación de contenidos. Con todo, es esta la materia más represen-

tativa del Grado de Información y Documentación y sobre la que descansan de forma transversal los diferentes perfiles profesionales.

El estudio de Agustín Lacruz, Gómez Díaz y Salvador Oliván (2013) apunta que la materia “Representación y Recuperación de la información” está recogida de manera importante, tanto en número de créditos como de asignaturas, en los planes de estudio de grado en Información y Documentación en las universidades españolas. La mayoría de las universidades ofertan entre 36 y 42 créditos de esta materia, lo que supone un mínimo de 6-7 asignaturas obligatorias y optativas en cada una de las titulaciones. Las denominaciones de las asignaturas evidencian cierta uniformidad. En general, hacen referencia a los procesos documentales (catalogación, indexación, clasificación, resumen, recuperación, etc.), a las herramientas (lenguajes documentales) y a los productos (catálogos).

Los grados en Información y Documentación se pusieron en marcha en las universidades españolas a partir del curso 2009-2010 y a lo largo de esta década se han efectuado reajustes y revisiones que no escapan a sugerencias de cambios más contundentes y adaptaciones efectivas a los entornos profesionales actuales (Abadal, Borrego y Serra Pérez, 2012; Agustín Lacruz y Salvador Oliván, 2016; García González, 2019).

Por el momento, son cuatro las universidades públicas en España que han decidido la supresión del Grado en Información y Documentación y han apostado por la creación de grados vinculados claramente con el sector de la economía digital. Así, en el curso 2017-2018 se puso en marcha el Grado en Gestión de la Información y Contenidos Digitales de la Universidad Carlos III de Madrid (UC3M), en 2019-2020 el Grado en Gestión de Información y Documentación Digital de la Universidad de Barcelona (UB) y el Grado en Gestión Digital de Información y Documentación de la Universidad de La Coruña (UDC) y, finalmente, en el curso académico 2020-2021 se ha iniciado el Grado en Gestión de Información y Contenidos Digitales de la Universidad de Murcia (UM).

Este trabajo examina los planes de estudio de estas nuevas titulaciones con el fin de conocer cómo se han integrado las competencias referidas a la Organización del Conocimiento en estos grados que reflejan singularmente el cambio de paradigma de la disciplina y su vinculación indiscutible con las tecnologías y la información electrónica.

Metodología

El propósito de este estudio es identificar y explorar el currículo relacionado con la Organización del Conocimiento en los programas de grado propuestos desde el ámbito de la Documentación en las universidades públicas españolas en estos últimos cuatro años con el fin de conocer el alcance de la OC en este perfil curricular más tecnológico.

Se ha empleado el análisis de contenido para estudiar los planes de estudio, las guías docentes disponibles y las memorias de verificación de los títulos académicos objeto de análisis. Los resultados se contrastan con trabajos anteriores relacionados con las competencias de OC en los programas académicos y, particularmente, con lo establecido sobre esta materia en el *Libro blanco* (ANECA, 2004) sobre el que se

asentaron los grados en Información y Documentación precedentes. Finalmente, se alinean las competencias en OC con las señaladas en el *Libro blanco para el diseño de titulaciones universitarias en el ámbito de la Economía Digital* (Agenda digital para España, 2015).

Las titulaciones de grado objeto de análisis han sido: Grado en Gestión de Información y Documentación Digital (UB), Grado en Gestión de la Información y Contenidos Digitales (UC3M,) Grado en Gestión Digital de Información y Documentación (UDC) y Grado en Gestión de Información y Contenidos Digitales (UM).

Resultados

Los nuevos planteamientos curriculares están enfocados a la formación de profesionales de la información con competencias para llevar a cabo los procesos inherentes a la documentación en un contexto eminentemente digital en el que coexisten objetos analógicos pero expuestos a tratamientos digitales. Las propias denominaciones de los títulos hacen referencia clara a esta orientación digital.

Se presentan a continuación, ordenados por universidades, los resultados referidos a competencias y asignaturas relacionadas con la OC. Se señala el carácter obligatorio (OB) u optativo (OP) de las asignaturas y los correspondientes créditos ECTS asignados.

El Grado en Gestión de Información y Documentación Digital de la Universidad de Barcelona cuenta con una materia común obligatoria, “Organización y Representación de la información”, de 24 ECTS repartidos en cuatro asignaturas de 6 ECTS (Tabla 1). Además, se ha identificado, dentro de la materia denominada “Comunicación de la información” la asignatura de carácter también obligatorio “Web semántica” y, dentro de la materia optativa etiquetada como “Bibliotecas y archivos”, la asignatura “Catalogación con RDA”. Estas dos últimas asignaturas incluyen amplios contenidos que creemos se hallan englobados en OC.

La competencia asignada en este Grado a la temática que nos ocupa refleja en su redacción la extensión de la OC, teniendo en cuenta la información y los datos de modo abarcador. Igualmente se expresa la amplitud del contexto de aplicación.

Tabla 1. Grado en Gestión de Información y Documentación Digital (UB)

Competencias	Asignaturas	Carácter	Créditos ECTS
Capacidad de definir criterios y aplicar las normativas y otros instrumentos de estructuración, descripción y representación de datos e información en diferentes entornos y contextos.	▪ Representación de la información: metadatos	OB	6
	▪ Análisis de contenido escrito y audiovisual	OB	6
	▪ Representación de la Información: Catalogación	OB	6
	▪ Tratamiento del material audiovisual	OB	6
	▪ Web semántica	OB	6
	▪ Catalogación con RDA	OP	6

La Universidad Carlos III de Madrid es la única institución que ha completado la docencia de los cuatro cursos del título. El plan de estudios del Grado en Gestión de la Información y Contenidos Digitales de esta universidad articula los contenidos relacionados con la OC en una asignatura obligatoria “Organización y Representación del Conocimiento”, adscrita al bloque de “Formación didiciplinar básica” y en dos asignaturas de carácter optativo (Tabla 2).

La formalización de las competencias se fija en remarcar la cohesión de aspectos teóricos e instrumentales tradicionales y actuales. Asimismo, se percibe la conexión estrecha de las competencias centradas en la organización del conocimiento y referidas conjuntamente a la gestión de la información.

Tabla 2. Grado en Gestión de la Información y Contenidos Digitales (UC3M)

Competencias	Asignaturas	Carácter	Créditos ECTS
Conocer las teorías, principios e instrumentos, clásicos y contemporáneos, de la comunicación, la organización y la gestión de la información.	▪ Organización y Representación del Conocimiento	OB	6
	▪ Metadatos avanzados	OP	6
	▪ Web semántica	OP	6
Ser capaz de identificar los procesos y flujos de trabajo en una organización y aplicar las técnicas y herramientas para la descripción, control, gestión y preservación de la documentación, especialmente la electrónica, a lo largo de su ciclo de vida.			

Por su parte, los contenidos sobre OC en el Grado en Gestión Digital de Información y Documentación de la Universidad de la Coruña se distribuyen en tres asignaturas obligatorias adscritas al módulo de “Información y Documentación” y una asignatura optativa perteneciente al módulo de “Tecnologías de la información” denominada “Lenguajes de descripción de información digital”, con contenidos relacionados con la Web semántica, datos abiertos y datos abiertos enlazados.

Tanto los títulos de las asignaturas obligatorias como la expresión de las competencias remarcan la confluencia entre OC y recuperación de información. Además, las competencias acentúan la concatenación de datos, información y conocimiento (Tabla 3).

Tabla 3. Grado en Gestión digital de Información y Documentación (UDC)

Competencias	Asignaturas	Carácter	Créditos ECTS
<p>Dominar los diferentes métodos de representación de los datos, información y el conocimiento que garanticen su recuperación eficiente.</p> <p>Conocer y dominar las técnicas y normativas para la creación y autenticación, reunión, selección, organización, representación, preservación, recuperación, acceso, difusión e intercambio, y evaluación de los recursos y servicios de información.</p>	<ul style="list-style-type: none"> ▪ Organización y Gestión del conocimiento 	OB	6
	<ul style="list-style-type: none"> ▪ Representación y Recuperación digital de la información I 	OB	6
	<ul style="list-style-type: none"> ▪ Representación y Recuperación digital de la información II 	OB	6
	<ul style="list-style-type: none"> ▪ Lenguajes de descripción de información digital 	OP	6

Por último, el Grado en Gestión de Información y Contenidos Digitales de la Universidad de Murcia cuenta con 30 créditos obligatorios asignados a una materia propia denominada “Organización, representación y recuperación de información”, en la que se incluyen cuatro asignaturas específicas de la temática de OC, tal y como figura en la Tabla 4. En la asignatura “Web de datos” se integran contenidos sobre Web semántica, RDF, SKOS, OWL, reutilización de datos semánticos, etc.

Además, el plan de estudios de la UM ofrece dos asignaturas optativas ligadas a la OC, en la vertiente clásica, “Catalogación en bibliotecas”, y en la línea más tecnológica, “Ontologías e ingeniería del conocimiento”.

La competencia más genérica coincide en la expresión con la empleada en la UC3M. En las competencias de carácter más específico se combinan los términos conocimiento, datos e información, mencionando explícitamente la capacidad para transformar información y datos.

Tabla 4. Grado en Gestión de Información y Contenidos Digitales (UM)

Competencias	Asignaturas	Carácter	Créditos ECTS
<p>Conocer las teorías, principios e instrumentos, clásicos y contemporáneos, de la comunicación, la organización y la gestión de la información.</p>	<ul style="list-style-type: none"> ▪ Metadatos y descripción de recursos 	OB	6
	<ul style="list-style-type: none"> ▪ Vocabularios y taxonomías 	OB	6
<p>Capacidad para analizar y transformar información y datos estructurados en cualquier medio.</p>	<ul style="list-style-type: none"> ▪ Tratamiento del contenido textual y audiovisual 	OB	6
	<ul style="list-style-type: none"> ▪ Web de datos 	OB	6
<p>Conocimiento y aplicación de los estándares y pautas para la representación de información, descripción y la organización del conocimiento, a través de esquemas y formatos interoperables.</p>	<ul style="list-style-type: none"> ▪ Catalogación en bibliotecas 	OP	6
	<ul style="list-style-type: none"> ▪ Ontologías e ingeniería del conocimiento 	OP	6

Este trabajo ha identificado las competencias y asignaturas relacionadas con la OC propuestas en estos nuevos estudios de grado, percibiendo aspectos comunes y peculiaridades reseñables.

La asignación de créditos oscila entre 18 y 36 ECTS, sumando las asignaturas obligatorias y optativas. Los planes docentes más proclives a integrar contenidos de la materia que nos ocupa son los de la UB y la UM, siendo, al contrario, la UC3M la que muestra mayor distancia. Precisamente son los grados de la UB y la UM los que disponen, en las Memorias verificadas de los grados correspondientes, de una materia concreta de OC en la estructura de su planificación docente: “Organización y Representación de la información”, de 24 ECTS obligatorios en la UB y, en el caso de la UM, integrando la recuperación de información, “Organización, representación y recuperación de información”, desgranada en 30 ECTS obligatorios.

Parece que la singularidad de la UC3M en la oferta de créditos reacionados con la OC, particularmente en los ECTS de carácter obligatorio, puede guardar relación con un cambio más drástico en la propia concepción del grado, un enfoque que pretende acentuar competencias de gestión tecnológica.

En el trabajo de Agustín Lacruz, Gómez Díaz y Salvador Olivan (2013) referido a los grados en Información y Documentación en España señalan similitudes en los títulos de las asignaturas, eran frecuentes en todos los planes de estudio términos como catalogación, análisis documental, lenguajes documentales e indización.

En este estudio, creemos que, en general, las denominaciones de las asignaturas identifican los contenidos con claridad y son comprensibles para los alumnos. Solo el programa de la UDC ha optado por mantener unidas la representación y la recuperación de información y proponer títulos de asignaturas genéricos con carácter progresivo: “Representación y Recuperación digital de la información I” y “Representación y Recuperación digital de la información II.” En conjunto, la oferta de la Universidad de Murcia es la que cuenta con títulos de asignaturas más distintivos (“Metadatos y descripción de recursos”, “Vocabularios y taxonomías”, “Web de datos”, etc.

Igualmente, se observan diferencias en la especificación de las competencias y preferencias en algunas universidades por interrelacionar organización y gestión de la información. Resulta notoria la inclusión de la representación de datos entre las competencias señaladas. Asimismo, cabe destacar el hecho de que en tres de los grados el término “metadatos” aparezca en el título propuesto de alguna de las asignaturas ofertadas.

La Tabla 5 pretende condensar los cambios significativos observables en las competencias y conocimientos profesionales vinculados a la materia de OC que se perciben en el *Libro blanco del Título de Grado en Información y Documentación* (ANECA, 2004) y aquellos que podemos vincular con la OC en *Libro blanco para el diseño de titulaciones universitarias en el ámbito de la Economía Digital* (Agenda Digital para España, 2015).

El *Libro blanco* (ANECA, 2004) muestra competencias apegadas a las unidades y servicios de información tradicionales y, paralelamente, conocimientos y habilidades profesionales asociados genéricamente a entornos clásicos (García Marco et al., 2006; García González, 2019). Por su parte, las competencias del *Libro blanco* (Agenda Digital para España, 2015) que podemos vincular a la OC presentan un

carácter transversal y se pueden trasladar a otros sectores con desarrollos e implementaciones digitales, todo ello en línea con una de las ideas centrales de este Libro asentada en que los programas formativos universitarios en ámbitos creativos y tecnológicos han de tener una naturaleza multidisciplinar que pueda derivar en perfiles profesionales más especializados.

Tabla 5. Competencias y conocimientos profesionales en OC

	Competencias	Conocimientos profesionales
<i>Libro blanco. Título de Grado en Información y Documentación (ANECA, 2004)</i>	<ul style="list-style-type: none"> ▪ Análisis y representación de la información. ▪ Organización y almacenamiento de la información. 	<ul style="list-style-type: none"> ▪ Catalogación descriptiva. ▪ Indización. ▪ Clasificación. ▪ Resumen. ▪ Desarrollo, mantenimiento y evaluación de instrumentos de organización, control y acceso. ▪ Elaboración, mantenimiento y uso de lenguajes documentales.
<i>Libro blanco para el diseño de titulaciones universitarias en el ámbito de la Economía Digital (Agenda Digital para España, 2015)</i>	<ul style="list-style-type: none"> ▪ Conocer y utilizar los estándares y técnicas de datos abiertos (open data) para su aprovechamiento como usuarios (consumidores) y como proveedores (productores) ▪ Comprender, utilizar y desarrollar la gestión de metadatos en sistemas, bases de datos, documentos, aplicaciones y servicios. ▪ Conocer y utilizar los modelos y estándares del ámbito de los sistemas de grandes volúmenes de datos. 	<ul style="list-style-type: none"> ▪ Fundamentos de Open Data. ▪ Formatos de Open Data. ▪ Linking Open Data. ▪ Reutilización de la Información. ▪ Modelos y estándares de almacenamientos de datos. ▪ Fundamentos de los metadatos. ▪ Usos más comunes de los metadatos.

El *Libro blanco para el diseño de titulaciones universitarias en el ámbito de la Economía Digital* (Agenda Digital para España, 2015) refleja las áreas en las que trabajaron los grupos de expertos relacionados con la economía digital: tecnología de la innovación; diseño visual, arte y creatividad; y estrategia y gestión de negocio. Los perfiles identificados, las competencias y contenidos de las titulaciones, fueron consensuadas por las comisiones con la pretensión final de potenciar y optimizar la inserción laboral de los futuros egresados. El estudio de García González (2019), en el que realiza una aproximación entre el antiguo marco de competencias derivadas del *Libro blanco* (ANECA, 2004), en concreto en el Grado en Información y Documentación de la Universidad de Murcia, y el nuevo escenario desarrollado en el *Libro blanco* (Agenda Digital para España, 2015) indica que debe buscarse un aumento considerable de las competencias para la adquisición de conocimientos profesionales vinculados con la estrategia y gestión de negocio y tecnología de la

innovación. En este proceso de evolución de competencias el objeto de estudio pasa a ser una plataforma global de información, en el que más allá del uso de TIC, las redes y el uso de datos, supone nuevos desarrollos de soluciones emergentes.

Por su parte, el Borrador de las *Guidelines for Professional Library and Information Science (LIS) Education Programmes* (IFLA, 2021), sistematizado alrededor de ocho áreas de conocimiento fundamentales, señala concretamente que el área 5, “Gestión del ciclo de vida de la información”, comprende todas las fases de un ‘documento’, desde su inicio hasta su fin, incluyendo la adquisición, identificación, descripción, organización, descubrimiento y conservación de los documentos, independientemente de su forma, formato, soporte o entorno de información. Esta área fundamental para IFLA requiere el conocimiento y la comprensión de la naturaleza de los documentos, de las necesidades y el comportamiento en la búsqueda de información de los usuarios, así como de los principios de organización e interoperabilidad de la información; de los requisitos funcionales de la organización de documentos, del intercambio y de la presentación de estándares, procedimientos y herramientas. Añade, asimismo, que un profesional de la información debe ser capaz de crear metadatos de calidad para el descubrimiento de recursos; de adoptar, adaptar, planificar, diseñar, desarrollar y/o implementar un sistema de información, herramientas, estándares y servicios de búsqueda de documentos, entre otros.

En los cuatro títulos de grado objeto de análisis en este estudio, se observa un enfoque curricular orientado a una sociedad digital y a entornos profesionales multidisciplinarios. En consonancia, se plantean competencias relacionadas con la OC que reflejan el cambio de paradigma de la disciplina y que se hallan en la línea del borrador de Directrices de la IFLA (IFLA, 2021).

Conclusiones

Los planes de estudio de estos grados, ligados a un mercado laboral relacionado con la transformación digital y las necesidades sociales que conlleva, creemos que, en conjunto, fortalecen las competencias en OC y dan muestra de los profundos cambios que se han producido en la disciplina. Sirven de ejemplo asignaturas ofertadas en estos títulos como: “Ontologías e ingeniería del conocimiento”, “Web semántica” o “Web de datos”.

Se observan soluciones diversas en las distintas universidades y la convivencia de denominaciones clásicas de asignaturas con soluciones más decididas. Los contenidos de las materias propuestas recogen claramente las orientaciones y tratamientos actuales de OC: datos abiertos enlazados, vocabularios semánticos, SKOS, taxonomías, OWL, etc.

Asimismo, las competencias de los grados relacionadas con la OC evidencian la evolución de la disciplina y la perspectiva expandida que integra documentos, información y datos en entornos diversos.

El estadio de implantación de estas titulaciones ha limitado la realización de un nivel de análisis uniforme. La visibilidad de las guías docentes completas desde las sedes web institucionales no siempre ha sido posible en el conjunto de las universidades, atendiendo al curso académico en el que se encuentra el plan de estudios

correspondiente. La labor de revisión de la bibliografía propuesta en cada una de las asignaturas permitirá en un futuro relizar un estudio más profundo sobre la correlación entre la literatura y metodología científica en OC y el contexto docente de la disciplina.

Referencias

- Abadal, E., Borrego, À., & Serra Pérez, R. (2012). Mercado laboral de profesionales de la información: evolución de la oferta y de los perfiles ocupacionales. *BiD: textos universitaris de biblioteconomia i documentació*, 29. <https://dx.doi.org/10.1344/BiD2012.29.8>
- Agenda Digital para España. (2015). *Libro blanco para el diseño de titulaciones universitarias en el ámbito de la Economía Digital*. Ministerio de Industria, Energía y Turismo. <https://www.cci.es/images/ccii/recursos/Libro-Blanco.pdf>
- Agustín Lacruz, M. C., Gómez Díaz, R. & Salvador Oliván, J. A. (2013). Competencias para el análisis, la representación, la organización y la recuperación de la información en los Grados de Información y Documentación en España. En F. Ribeiro & E. Cerveira (org.), *Informação e/ou conhecimento: as duas faces de Jano*. Atas del I Congreso ISKO España e Portugal – XI Congreso ISKO España (pp. 35-48). Cetac.Media.
- Agustín Lacruz, M.C., & Salvador Oliván, J. A. (2016). Análisis y evaluación de las competencias del Grado en Información y Documentación en la Universidad de Zaragoza. *Education in The Knowledge Society (EKS)*, 17 (1), 129-146. <https://doi.org/10.14201/eks2016171129145>
- Andersen, J. and Skouvig, L. (2006), “Knowledge organization: a sociohistorical analysis and critique”, *Library Quarterly*, Vol. 76 No. 3, pp. 300-322.
- ANECA (2004). *Libro blanco. Título de Grado en Información y Documentación*. ANECA. http://www.aneca.es/var/media/150424/libroblanco_jun05_documentacion.pdf
- Borrego Huerta, Ángel (2015). Los estudios de Biblioteconomía y Documentación en Europa: una visión panorámica. *BiD: textos universitaris de biblioteconomia i documentació*, 35 <http://bid.ub.edu/en/35/borrego.htm>
- García González, M. (2019). Adaptación de competencias al marco de la Economía Digital en el grado de Información y Documentación: Universidad de Murcia. *Informatio*, 24(2), 65-90. <https://doi.org/10.35643/Info.24.2.5>
- García Marco, F. J., Agustín Lacruz, M. del C., Pedraza Gracia, M. J., & Salvador Oliván, J. A. (2006). *Diseño y coordinación curricular de las asignaturas de tratamiento y recuperación de la información de la Diplomatura de Biblioteconomía y Documentación*. Prensas Universitarias de Zaragoza.
- Hjørland, B. (2008). What is knowledge organization (KO). *Knowledge Organization*, 35(2/3), 86-101.
- Hudon, M. (2011). Teaching Classification in the 21st Century. *Knowledge Organization*, 38(4), 342–351.
- IFLA (2021). *IFLA Guidelines for Professional Library and Information Science (LIS) Education Programmes DRAFT – February 3, 2021*. <https://lisedu.files.wordpress.com/2021/02/lisepguidelines-consult-draft.pdf>
- Moreiro González, J.A. (2019). Si los nombres denominan a las cosas, esto ya no es lo que era. La profesión. *Anuario ThinkEPI*, 13, e13a01. <https://doi.org/10.3145/thinkepi.2019.e13a01>
- Ortiz-Repiso, V. (2015). Rethinking Library and Information Studies in Spain: Crossing the boundaries”. *BiD: textos universitaris de biblioteconomia i documentació*, (35). <http://bid.ub.edu/en/35/ortiz.htm>

- Ortiz-Repiso, V., Greenberg, J., & Calzada-Prado, J. (2018). A cross-institutional analysis of data-related curricula in information science programmes: A focused look at the iSchools. *Journal of Information Science*, 44(6), 768–784. <https://doi.org/10.1177/0165551517748149>
- Túñez López, M., & Martínez Solana, Y. (2013). El nuevo mapa de los estudios de Documentación y Comunicación en España: oferta de grados, índices de matrícula, tasas de empleo y perfiles de contratación. *Anales de Documentación*, 16 (2). <http://dx.doi.org/10.6018/analesdoc.16.2.168081>
- Universidad de Barcelona (2020). *Memoria verificada del Grado en Gestión de Información y Documentación Digital*. <https://www.ub.edu/portal/web/informacio-mitjans-audiovisuals/verificacio-modificacio>
- Universidad de Barcelona (n.d.). Plan de estudios y guías docentes del Grado en Gestión de Información y Documentación Digital. Retrieved July 7, 2021, from <https://www.ub.edu/portal/web/informacion-medios-audiovisuales/grados/-/ensenyament/detallEnsenyament/6063329/7>
- Universidad Carlos III de Madrid (2016). *Memoria verificada del Grado en Gestión de la Información y Contenidos Digitales*. <https://www.uc3m.es/grado/contenidos-digitales#evaluacionyseguimiento>
- Universidad Carlos III de Madrid (n.d.). Plan de estudios y guías docentes del Grado en Gestión de la Información y Contenidos Digitales. Retrieved July 7, 2021, from <https://www.uc3m.es/grado/contenidos-digitales#programa>
- Universidad de La Coruña (2018). *Memoria verificada del Grado en Gestión Digital de Información y Documentación*. https://humanidades.udc.es/images/stories/documentos/GDID/GDID_MemoriaVerificacionOficial.pdf
- Universidad de La Coruña (n.d.). Plan de estudios y guías docentes del Grado en Gestión Digital de Información y Documentación. Retrieved July 7, 2021, from <https://humanidades.udc.es/estudios/gdid/informaci%C3%B3n-del-t%C3%ADtulo>
- Universidad de Murcia (2020). *Memoria verificada del Grado en Gestión de Información y Contenidos Digitales*. <https://www.um.es/web/estudios/grados/contenidos-digitales/documentacion-del-titulo>
- Universidad de Murcia (n.d.). Plan de estudios y guías docentes Grado en Gestión de Información y Contenidos Digitales. Retrieved July 7, 2021, from: <https://www.um.es/web/comunicacion/grado-en-gestion-de-informacion-y-contenidos-digitales>
- ur Rehman, S., & Alajmi, B. (2017). Knowledge organization content in graduate coursework, *Library Review*, 66(1/2), 90-106. <https://doi.org/10.1108/LR-03-2016-0028>
- Wyman, A., & Imamverdiyev, M. (2018). Global trends and transformations in library science education. *Information and Learning Sciences*, 119 (3/4), 215-225. <https://doi.org/10.1108/ILS-11-2017-0110>

(Página deixada propositadamente em branco)

PERSONAL KNOWLEDGE ORGANIZING THROUGH ONLINE COLLABORATIVE WRITING TOOLS

Dalbert Marques Oliveira¹, Ana Lúcia Terra²

¹Faculdade de Letras da Universidade de Coimbra, dalbertoliveira(a)gmail.com,
ORCID iD 0000-0001-9035-3480

² Faculdade de Letras da Universidade de Coimbra, anatterra(a)fl.uc.pt,
ORCID iD 0000-0003-1292-2849

Abstract

The practice of using online collaborative writing (OCW) tools by students in the three cycles of higher education is a reality in several countries. Otherwise, these tools are also used by researchers in writing scientific papers. It appears that OCW, and the platforms on which they are embedded, offer resources that can be applied to organizing the personal knowledge (PKO) of individuals. When investigating the literature, it is possible to find studies that comment on the use of OCW, especially those that are part of the Google suite, such as Google Docs and Microsoft such as OneDrive. In turn, some studies mention other platforms and tools that allow writing collaboratively. Among these tools are platforms in Wiki format, blogs, EtherPad, DropBox, Manuscripts.io, Overleaf, and Authorea, to mention those mentioned in the recovered literature. During this use, there is the phenomenon of appropriation which, in the case of OCW, consists of the frequent use of the tool, platform, and/or resource to achieve a certain objective or meet some need. Still, on these studies, it is possible to find evidence of the contribution of OCWs to the creation and organization of data and information, not only of the community of users who work in each context but also in the individual use of these OCW. However, there seem to be few works that investigate the appropriation of OCW for the practice of PKO. That said, the objectives of this investigation are: to know the relationship of students in the third cycle of higher education with the OCW; know which are the most popular OCWs by these students and which resources they use the most; realize if there are occurrences of the appropriation of platforms, tools and/or resources for the practice of PKO, even if unconsciously; and if they do not exist, investigate what are the possibilities that OCWs, and their resources, present to users and user communities. These objectives will be achieved through a brief literature review, with which it is intended to visit some of the terms discussed here and bring OCW closer to PKO practices. This approach will be expanded with the application of two research instruments: a questionnaire and a semi-structured interview. These instruments will be applied to students from three doctoral programs. The research instruments were based on another one used in an investigation, at an international level, on OCW. The results of this investigation intend to inform students and researchers about the possibilities of using/appropriating OCW for PKO, thus contributing to expanding the knowledge of these practices.

Keywords: Online collaborative writing platforms; organization of personal knowledge, Ph.D. students.

Introduction

It appears that online writing and collaborative tools (OCW) and the platforms on which they are embedded offer resources that can be applied to the organization of personal knowledge (PKO) of individuals while facilitating the sharing of this knowledge.

The use of online collaborative writing tools (OCW) by higher education students is a reality in several countries, as can be seen in the work of Oliveira and Terra (2021). Among these students are doctoral students, as mentioned by Gallotti (2017). In turn, these tools are also used by researchers in the collaborative writing of scientific papers (Boellstorff et al., 2013).

When investigating the literature, it is possible to find several studies that comment on the use of OCW, especially those that are part of the Google suite, such as Google Docs and Microsoft such as OneDrive. In addition to these OCW, it is possible to find works that mention other platforms and tools that allow collaborative writing online, such as Wiki (Cilliers, 2017), blogs (Garcia et al., 2019), Skype (Akbaba & Başkan, 2017), EtherPad (Olson et al., 2017), DropBox (Hunsinger & Corley, 2012), Manuscripts.io, Overleaf and Authorea (Pepe et al., 2020). These tools are used both for collaborative writing and as a complement to writing itself (cf. Boellstorff et al., 2013).

Although several errors were found in the use of OCW (Ahmed-Nacer et al., 2013; Bescó Baiges & Verdú Surroca, 2014; Olson et al., 2017), the common denominator found by these authors seems to be the ease of interaction between different elements of a group in sharing and editing information to achieve a common goal.

In the literature on the use of OCW, it is possible to find relationships between OCW and the management of personal information, as in the case of the study by Naik, Ajay, and Kolhatkar (2013), who comment that collaborative work tools and applications offer numerous benefits for the efficient management and use of personal information. On the other hand, there are relationships with the management of personal knowledge, as mentioned in the study by Liu et al. (2017), authors who investigate Big Data applied to the integration of knowledge in different areas, seeming to conclude that online collaborative work tools contribute to PKM in the selection of information that will contribute to knowledge.

In these studies, it is possible to find evidence of the contribution of OCW to the creation and organization of knowledge, not only of the community of users who work in each context but also about individuals when they use OCW at an individual level. In turn, OCW platforms have resources that seem to allow the organization and representation of knowledge (cf. Oliveira & Terra, 2021).

Concerning the knowledge organization, Kasten (2007) comments that it is concerned with the creation of mechanisms for classification and search for knowledge. In turn, Botha et al (2014) comment that the organization of knowledge

is supported by classification, mapping, indexing, and categorization tasks that allow the storage, navigation, and retrieval of this knowledge.

On the other hand, there are authors such as Jones (2016) who comments that the only knowledge that can be worked on, which seems to include the organization, is that which was explained in the form of information. As mentioned by Hall (2006), most of the knowledge worked on is still what was codified. In other ways, there are authors such as Cury (2007) who report that tacit knowledge can also be worked through mental techniques of logical reasoning (cf. Morin, 2014).

In turn, as mentioned by Kasten (2007), PKO seems to be an arduous task, since, within an individual, knowledge is shared dynamically, expanding, contracting, being formed, or even disappearing. On the other hand, as already mentioned, OCW has mechanisms that seem to be sufficient to organize and represent knowledge. Among these features is the possibility to rename files and link these files with others through chronological and/or thematic (ontological) and/or hierarchical (taxonomic) relationships, to mention a few possibilities.

Taxonomy relationships work to organize information and/or knowledge, in hierarchical relationships between terms. In turn, the ontology relationships seek to establish semantic relationships between concepts, in the form of conceptual networks (Mazzocchi, 2018).

Another possibility available in OCW is related to the insertion of metadata (see Mayernik, 2021) in the files created or inserted in the platforms. On the other hand, it is possible to create relationships between information using links and sharing options. As an example of sharing options, it is possible to mention the practice of “incorporating” or “embedding” which is related to interconnecting sets of information between documents and web pages (Microsoft, 2021).

That said, this work will investigate the organization of explicit knowledge, that which exists as information with potential meaning. The objectives of this investigation are: to know which are the most popular OCWs by students of the third cycle in the area of communication and information sciences and which are the resources they use the most; realize if there are occurrences of the appropriation of platforms, tools and/or resources for the practice of PKO, even if unconsciously; and if they do not exist, investigate what are the possibilities that OCWs, and their resources, present to users and user communities, in different contexts of use and/or appropriation.

These objectives will be achieved through a brief literature review, with which it is intended to visit the literature referring to the terms referred to here and bring them closer to the practices carried out in OCW. This approach will be expanded with a questionnaire and semi-structured interviews.

The results of this investigation intend to inform students and researchers about the possibilities of using/appropriating OCW for PKO, thus contributing to expand the knowledge of these practices and contribute to the design of interfaces that enhance PKO.

Methodology

To respond to the proposed objectives, this work starts from a constructivist paradigm that, despite accepting the subjectivity inherent to the investigated theme, does not impede its analysis, interpretation, and understanding (Guba & Lincoln, 1994). In line with this paradigm, a qualitative methodology was used, employing the exploration of literature (Gil, 2008; Lakatos & Marconi, 2007).

From the constructivist paradigm and qualitative methodology, this work used the bibliographic method through a brief literature review (Webster & Watson, 2002).

This literature review consisted of a search, for a set of terms and the application of Booleans, in two databases, one of which is more generic, the Web of Science (WS), and the other is more specific, the KO Knowledge Organization. A third database, Google Scholar, was used as a backup, in case the first two did not retrieve enough information for a match.

The terms were chosen to cover the entire investigated question. Among these terms are: “Online Collaborative Writing”, “Personal Knowledge Organization”, “Ontology” and “Taxonomy”. Finally, the retrieved works were read and compared to each other using the interpretive approach (Silva, 2013).

In the design of the questionnaire, closed questions were used to select and group participating students into the categories that we intended to analyze (Gunther & Júnior, 2012). To validate the structure of the questionnaire, it was reviewed and tested by a second investigator (Lakatos & Marconi, 2007). One of the questions in the questionnaire asked the participants to provide their contact details for a possible interview, intending to broaden their perception of the use of OCW for PKO.

The interviewed participants were chosen, initially by drawing, from among the participants who provided contact. After drawing one participant from each course, the others were chosen using the “snowball” technique, where each participant indicated a following until reaching the proportional number of individuals who answered the questionnaire (Tajedini et al., 2018).

The choice of courses was made for convenience, given the proximity of the researcher with the students of these courses. The sample was obtained in a probabilistic way, because it allowed all students of the referred programs to answer the questionnaire (Freitas et al., 2000), which was constructed using Google Forms, and made available online. These students were reached through contact with class representatives, publications in thematic groups on Facebook, and direct contact with students.

The answers obtained were transferred to an Excel sheet (for Mac, version 16.51) and statistically analyzed. Pivot tables were also used, which allow a more sensitive visualization of the relationships obtained between the data.

Results

The responses obtained with the application of the questionnaire totaled 61 participants, 32 belonging to the PDCI, 15 to the ICPD, and 6 to the PPGCOM. Figure 1 illustrates these responses.

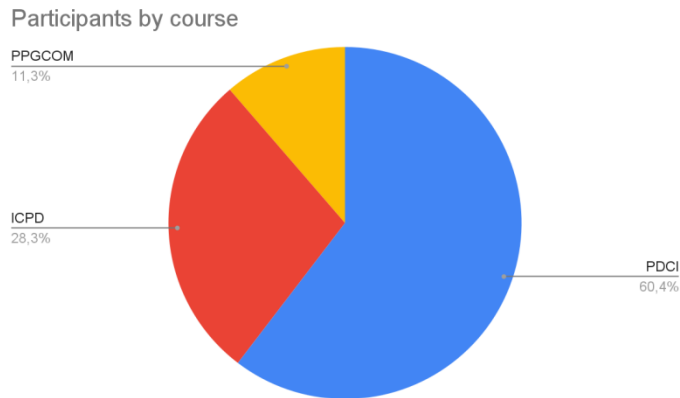


Figure 1. Participants by course.

Source: The authors.

As the intention of this study was to understand PKO through OCW, the participants were asked if they used OCW for PKO and if so, which ones were used. Only 1 respondent claimed not to use OCW for PKO. On the other hand, 39.4% of respondents said they use Google's OCW, 31.5% from DropBox, 27.6% from Microsoft, and 0.8% from Draft or Manuscripts.io. Figure 2 presents these numbers.

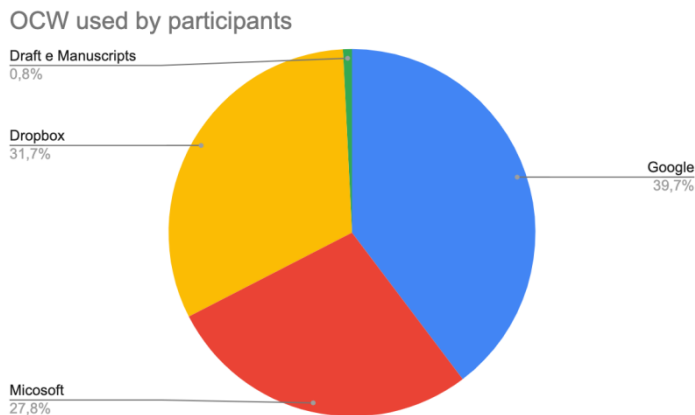


Figure 2. OCW used by participants.

Source: The authors

After confirming that the participants used OCW to practice PKO and verifying which were the most used OCW. One participant from each course was drawn to apply the semi-structured interview. Each of the interviewed participants was asked to recommend another student, from the same program, to be interviewed, except for the randomly chosen participant from the PGGCOM, to maintain proportionality with the questionnaire respondents. Thus, 1 participant from the PGGCOM, 3 from the ICPD, and 5 from the PDCI were interviewed. The purpose of this interview was to understand how these participants practiced PKO in those OCW.

OCW used and its resources for the PKO

The first interview question invited participants to describe how their knowledge was organized in these OCW. The use of folders with nomenclatures proved to be the most used technique, being mentioned by all participants. Although the terms “taxonomy” and “ontologies” were not mentioned, the explanation they provided for the way they organize knowledge seems to relate to taxonomy and ontology (Mazzocchi, 2018).

Otherwise, it was found that some participants used a different OCW for each context and practice, as commented by Participant I:

“I use OneDrive to save old documents, DropBox to share documents with faculty, and Google Drive to work together with colleagues. However, I think I can better organize my knowledge through Google tools, I don’t know if it’s because I don’t know the resources in other tools. Google allows me to organize knowledge in image format through automatic photo recognition, the places I visited through Google Maps functionality, and the timeline of building and editing documents through Google Drive. Google gives me a lot of details about knowledge through context relationships in the information”.

All 9 interviewed participants commented using the folder hierarchy feature to organize personal knowledge (cf. Bergman et al., 2006). Participant D comments that his knowledge is “hierarchically organized by folders”, which seems to be related to the creation of taxonomies, and that he has “shortcuts to these folders organized by subject” (ontology) (Mazzocchi, 2018). Participant A, in turn, uses description resources, present in folders and files, to insert metadata, which are terms through which he can easily access organized knowledge.

Otherwise, participant C uses Google’s Photos platform recognition systems to organize the accumulated knowledge about places he visited, people, objects, and animals he interacted with: “I create automatic themed albums when I’m interested in returning to interact with the knowledge I have accumulated”. This participant also commented that he has a “Thesaurus in the form of a map where [you can] find the keywords related to a certain theme or event”

Participants B and E commented that they use a chronological organization: “I feel better when I have my things organized by year, month and day. In front of that, I insert the theme of information” (Participant B). In turn, Participant E states: “I think better chronologically, and I prefer that knowledge is organized in this way”.

Other participants stated that they employ mixed organizations, which employ chronology, subjects, typology, and events. In this sense, Participant C comments: “organizing my knowledge through the tools present in Google’s suites, such as Drive, Calendar and Photos, is much easier now than before learning about these tools”. Among the resources that Participant C comments on is the possibility of organizing opinions about a certain place (a park or a restaurant, for example), knowing the times of busiest movement in these places and/or the itinerary or the means of transport to get there local. Organizing knowledge by visit points or dates, inserting metadata in the taken photos, grouping similar file types, organizing the week according to scheduled events.

The resources commented on by Participant C are spread across the Google suite, but some of these may be connected, such as Google Photos and Google maps, where the user can keep the representation of the photos he took associated with a particular route.

The use of an automatic organization was mentioned by Participant H. When asked to provide more details, this participant said that he likes the automatic organization mechanisms of OCWs, only occupying themselves with renaming the information set with terms that facilitate their filtering.

Resources for filtering personal knowledge

Since filters are part of the organization, participants were asked to describe how they filtered the accumulated knowledge. The answers seem to suggest that the technological tools present in these OCWs facilitate filtering practices. Information on the same subject can be filtered by entering key terms from these subjects into the OCW search engines. Chronological filters allow for the arrangement and grouping of regular events, such as events that happen at the same time each year, or a particular class that takes place every Monday and Wednesday during the term. Facial recognition allows you to group photos where a certain person appears, and the same applies to objects, environments, or animals.

Otherwise, the use of filters seems to be related to the way knowledge was organized. This relationship can be seen in Participant D’s comment: “I use filters through the directory trees because I have the knowledge organized like that”.

In turn, the knowledge stored in the “use of keywords” was the common denominator of all participants. On the other hand, another commented filter refers to chronological order. Participant B commented: “I try to filter knowledge by year and month, a few times a day [...] I have a good chronological memory, and filters by date support this memory.” On the other hand, Participant E states: “I usually use filters on dates. My knowledge is linked to the time of events and filtering the information by date I can find the context I need to review certain knowledge”.

Otherwise, Participant C commented that he uses filters as needed: “I use a personal thesaurus to filter by terms in topics of interest [...] chronological filter to arrange knowledge by dates and location filters to sort places [...] sometimes I also use time filters when I want to find a specific event within a day full of activities”

In turn, Participant H mentioned that, when he needed to retrieve information about a certain context, he used Google Photos' facial recognition filters: "I needed to know if a certain colleague was present at an event. So, after identifying him in a photo on the Google Photos platform, I asked Google to show me all the photos where he appeared and I was able to retrieve all the photos where he appeared."

As can be seen, Participant H needed to recover knowledge about the presence of a given individual in a given context. For this purpose, this participant inserted metadata in the photos, identifying the information he would like to retrieve and when performing the search, he retrieved the knowledge he intended.

Finally, Participant, I mentioned that he uses different filters on different platforms: "On OneDrive, DropBox and Google Drive, generally filter by keywords and typology; on Google photos filter by theme, I want to retrieve, "houses" for example, on Google maps filter by name of places and by date".

Resources to retrieve personal knowledge

After applying a filter or a set of filters, the study participants were able to access and retrieve the information needed to compose or review the knowledge they wanted.

Some participants, such as B, E, and I, commented that they only needed to visualize the recovered knowledge to perform some action. Participant, I exemplified this visualization: I visualize the knowledge I have about a particular location to "decide what time I will go [at this] location".

On the other hand, other participants reported the need to print the recovered knowledge. This practice of migrating knowledge from digital to physical support was mentioned by participants F and G. Participant F comments that in addition to printing the "generated knowledge", he organizes it in folders. Participant G, in turn, reports: "When I just have to review some knowledge, I just visualize it, but if I have to work with it, I prefer to print it".

Otherwise, the migration between platforms proved to be the option of participants D, H, and I. Participant D comments: "when I recover knowledge [...], I transfer it to some digital document and work on it from there". This participant exemplified: "when I retrieve the set of information contexts on "Document Types", I transfer this set of information to a table where I will work on this knowledge by making new connections". Participant H provides another example: on a "study visit, when I was able to identify the photos where my colleagues appeared, I transferred them to a folder and inserted a unique ID."

Participants A and C comment that they reorganize the knowledge retrieved to facilitate future retrieval. Participant A performs this new organization through folders in the browser's bookmark bar. Otherwise, Participant C stores the retrieved knowledge in relationally organized folders.

Negative points to the use of OCW for the PKO

Despite the ease of organizing, filtering, and recovering accumulated knowledge, the use of OCW, for PKO, seems to be challenging even for individuals who use various resources on these platforms.

The diversity of interfaces seems to be one of the main challenges for PKO as mentioned by Participant F: “I sometimes have difficulty finding myself in the middle of so many screens”. Participant H adds: “I often get lost in the middle of multiple platforms, workspaces, and interfaces”. It seems that these participants refer to the phenomenon of information fragmentation, which was a problem mentioned by Bergman et al. (2006) and more recently by Terra (2017).

On the other hand, lack of resources seems to be another challenge as Participant C reports: “I feel some difficulty in converting the formats. Sometimes I just want to save the metadata of a set of photos, and I don’t know proprietary tools for that”. This student also comments that he experiences the same difficulty in extracting metadata from maps.

Otherwise, some seem to prefer to physically manipulate objects, such as Participant G, who felt it was easier to organize the knowledge he believed he had when he could “manipulate physical objects and interact with these objects in the real world”. Among these objects, the participant mentions “books, notebooks, magazines, booklets, manuals” among others.”

Thinking of individuals like Participant G, the industry has invested in virtual reality technologies that seem to be able to allow the interaction of virtual objects in the real world, for example, through smart glasses and sensors (e.g., Chernyshov et al., 2019).

Finally, the lack of add-ons to the platforms is a challenge commented by Participant A: “I use add-ons in the Google suite [Calendar, Keep, Task], but I miss other add-ons” when asked which other add-ons I missed, the participant mentioned complements that enabled “the creation of mental maps and schedules”. On the other hand, there are add-ons to Google Docs that allow the creation of mind maps and timelines, such as “Mindmeister” and “Creately Diagrams”, present on the Google Workspace Marketplace.

Participant B mentioned the use of similar supplements present in Google Drive “when I’m building a relationship tree with classmates, creating mind maps, for example, I use Google Calendar to retrieve information about class moments and activities in a set, Google Keep as a draft and idea gathering, plus apps allow me to work with Gantt charts”.

Conclusion

In this work, it was possible to verify that the OCW, and their respective platforms, have tools and resources that contribute to the organization of personal knowledge. Although it was not possible to find, in the investigated literature, works that comment on the practice of PKO in OCW, it was possible to retrieve works that reported resources present in OCW that seem to contribute to PKO.

Otherwise, when applying research instruments, namely a questionnaire and a semi-structured interview, to students from the 3rd cycle of the Information-Communication Sciences area, it was possible to verify that these students use resources present in OCW and its platforms, to the practice of PKO.

Among the resources for the practice of PKO, mentioned both in the literature and by the interviewees, there is the possibility of organizing the knowledge explained through automatic organization mechanisms, such as Google Photos pattern recognition and Google Maps location, by inserting nomenclatures and metadata in folders and files, hierarchically, chronologically, by subject, inserting metadata in files and folders and/or the sequential application of these various resources.

On the other hand, the use of filters is carried out using key terms, the choice of dates, pattern recognition, directory hierarchy, and/or the application of sets of these resources. In turn, to recover organized knowledge, some of the participants commented that it was enough to observe the information to import (bring in) that knowledge; otherwise, other students commented that they felt the need to transfer the information to another support and/or transport it to another platform.

However, participants reported some difficulties with PKO practices. Among these, the diversity of interfaces, which can cause a feeling of “being lost”. On the other hand, the lack of proprietary features for metadata migration and add-ons to interconnect tools was also mentioned. Otherwise, the inexistence of the possibility of physically manipulating explicit knowledge in a virtual environment was also referred to as a barrier to the practice of PKO in OCW.

Expressions such as “I feel”, “I prefer”, “decide” and others refer to factors such as the individual’s motivation and feeling. The choice of OCW and/or the way to use OCW seems to be related to these factors. However, this work did not investigate these relationships, which must be considered in the global understanding of the use of OCW for the practice of PKO.

On the other hand, the OCW that seemed only useful for collaborative writing proved to be an asset to the practice of PKO. In turn, references to platforms such as Google’s Photos and Maps, which despite not having been designed as an OCW allow the sharing and editing of the same information and metadata by different users, seems to have demonstrated that the range of OCW for the practice of PKO is even greater than expected at the beginning of the investigations for this work.

That said, the results of this research intend to inform students and researchers about the possibilities of using OCW for PKO, thus contributing to expanding the knowledge of these practices.

Finally, the lack of PKO-oriented supplements appears to be a gap to be filled by the app industry. Otherwise, the apparent confusion between interfaces may lead to discussions that contribute to the design of interfaces that enhance PKO.

References

- Ahmed-Nacer, M., Urso, P., Balegas, V., & Preguiça, N. (2013). Concurrency control and awareness support for multi-synchronous collaborative editing. 9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing, 148-157. <https://doi.org/10.4108/icst.collaboratecom.2013.254113>
- Akbaba, Y., & Başkan, F. (2017). How to merge courses via Skype™? Lessons from an International Blended Learning Project. *Research in Learning Technology*, 25(0). <https://doi.org/10.25304/rlt.v25.1915>
- Bergman, O., Beyth-Marom, R., & Nachmias, R. (2006). The project fragmentation problem in personal information management. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 271-274). Association for Computing Machinery. <https://doi.org/10.1145/1124772.1124813>
- Boellstorff, T., Nardi, B., Pearce, C., & Taylor, T. L. (2013). Words with friends: Writing collaboratively online. *Interactions*, 20(5), 58-61. <https://doi.org/10.1145/2501987>
- Botha, A., Kourie, D., & Snyman, R. (2014). *Coping with Continuous Change in the Business Environment: Knowledge Management and Knowledge Management Technology*. Elsevier.
- Brescó Baiges, E., & Verdú Surroca, N. (2014). Valoración del uso de las herramientas colaborativas Wikispaces y Google Drive, en la educación superior. <https://repositorio.minedu.gob.pe/handle/20.500.12799/3189>
- Chernyshov, G., Kunze, K., Tag, B., Ward, J. A., & Uema, Y. (2019). <http://eyewear.pro>: An open platform to record and analyze large scale data sets from smart eyewear. *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, 264-265. <https://doi.org/10.1145/3341162.3343790>
- Cilliers, L. (2017). Wiki acceptance by university students to improve collaboration in higher education. *Innovations in Education and Teaching International*, 54(5), 485-493. <https://doi.org/10.1080/14703297.2016.1180255>
- Cury, A. (2007). *Inteligência multifocal: Análise da construção dos pensamentos*. Pergaminho.
- Freitas, H., Oliveira, M., Saccol, A. Z., & Moscarola, J. (2000). O método de pesquisa survey. *Revista de Administração da Universidade de São Paulo*, 35(3), 105-112.
- Gallotti, M. M. C. (2017). *Práticas de Comunicação Científica de Doutorandos em Ciência da Informação no Espaço Ibérico e no Brasil: Um estudo exploratório (Repositório Aberto da UP)*. <https://doi.org/10.13140/RG.2.2.11097.47205>
- Garcia, E., Moizer, J., Wilkins, S., & Haddoud, M. Y. (2019). Student learning in higher education through blogging in the classroom. *Computers & Education*, 136, 61-74. <https://doi.org/10.1016/j.compedu.2019.03.011>
- Gil, A. C. (2008). *Métodos e técnicas de pesquisa social (6a)*. Editora Atlas S.A. <https://biblioteca.isced.ac.mz:443/handle/123456789/707>
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In *Handbook of qualitative research* (pp. 105-117). Sage Publications, Inc.
- Gunther, H., & Júnior, J. L. (2012). Perguntas abertas versus perguntas fechadas: Uma comparação empírica. *Psicologia: Teoria e Pesquisa*, 6(2), 203-213.
- Hall, M. (2006). Knowledge management and the limits of knowledge codification. *Journal of Knowledge Management*, 10(3), 117-126. <https://doi.org/10.1108/13673270610670894>
- Hunsinger, D. S., & Corley, J. K. (2012). An Examination of the Factors Influencing Student Usage of DropBox, a File Hosting Service. *Proceedings of the Conference on Information*

- Systems Applied Research, 5, 1-8. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.462.3946&rep=rep1&type=pdf>
- Jones, W. (2016). No knowledge but through information. In *Personal Knowledge Management: Individual, Organizational and Social Perspectives* (2a, pp. 143–166). Routledge. <https://journals3-staging.lib.uic.edu/ojs/index.php/fm/article/view/3062>
- Kasten, J. (2007). Thoughts on the relationship of knowledge organization to knowledge management. *Knowledge Organization*, 34(1), 9-15.
- Lakatos, E. M., & Marconi, M. de A. (2007). *Metodologia do trabalho científico: Procedimentos básicos; pesquisa bibliográfica, projeto e relatório; publicações e trabalhos científicos* (7.^a edição). Atlas.
- Liu, C.-H., Wang, J. S., & Lin, C.-W. (2017). The concepts of big data applied in personal knowledge management. *Journal of Knowledge Management*, 21(1), 213-230. <https://doi.org/10.1108/JKM-07-2015-0298>
- Mayernik, M. S. (2021). Metadata. *Knowledge Organization*, 47(8), 696-713. <https://doi.org/10.5771/0943-7444-2020-8-696>
- Mazzocchi, F. (2018). Knowledge Organization System (KOS): An Introductory Critical Account. *Knowledge Organization*, 45(1), 54-78. <https://doi.org/10.5771/0943-7444-2018-1-54>
- Microsoft (2020) Embed files directly into your website or blog. In Office Support. Disponível em <https://support.microsoft.com/en-us/office/embed-files-directly-into-your-website-or-blog-ed07dd52-8bdb-431d-96a5-cbe8a80b7418>
- Morin, E. (2014). Complex thinking for a complex world—about reductionism, disjunction and systemism. *Systema: Connecting matter, life, culture and technology*, 2(1), 14-22.
- Naik, A., Ajay, A. K., & Kolhatkar, S. S. (2013). Applicability of cloud computing in academia. *Indian Journal of Computer Science and Engineering*, 4(1), 11-15.
- Oliveira, D. M., & Terra, A. L. (2021). Online and Collaborative Tools During Academic and Erasmus Studies. In R. Babo, N. Dey, & A. S. Ashour (Eds.), *Workgroups eAssessment: Planning, Implementing and Analysing Frameworks* (pp. 85-115). Springer. https://doi.org/10.1007/978-981-15-9908-8_4
- Olson, J. S., Wang, D., Olson, G. M., & Zhang, J. (2017). How People Write Together Now: Beginning the Investigation with Advanced Undergraduates in a Project Course. *ACM Transactions on Computer-Human Interaction*, 24(1), 4:1-4:40. <https://doi.org/10.1145/3038919>
- Pepe, A., Cavalleri, M., Best, B. D., Olivotto, V., & Cantiello, M. (2020). Scientific collaboration in the era of COVID-19 [Preprint]. Preprints. <https://doi.org/10.22541/au.158826951.14028294/v2>
- Silva, E. A. da. (2013). As metodologias qualitativas de investigação nas Ciências Sociais. *Revista Angolana de Sociologia*, 12, 77-99. <https://doi.org/10.4000/ras.740>
- Tajedini, O., Sadatmoosavi, A., & Esfandiarpour, Y. (2018). Identification, Extraction, Organization and Sharing of Personal Knowledge: Using Grounded Theory. *International Journal of Information Science and Management*, 16(2), 143-160.
- Terra, A. L. (2017). A gestão da informação pessoal e as competências infocomunicacionais. Atas da X Congresso SOPCOM – Ciências da Comunicação: vinte anos de intervenção em Portugal, 113-123. <http://www.bocc.ubi.pt/pag/sopcom/x-congresso.pdf>
- Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), xiii–xxiii.

INDEXAÇÃO DE IMAGENS FOTOGRÁFICAS RARAS: UM OLHAR SOBRE O ÁLBUM DA CONSTRUÇÃO DO CANAL DO MANGUE NA CIDADE DO RIO DE JANEIRO, BRASIL

Melina de Brito dos Santos¹, Rosali Fernandez de Souza²,
Jeorgina Gentil Rodrigues³

¹Doutoranda PPGCI/IBICT/UFRJ, *melbrisan@yahoo.com.br*,
ORCID iD 0000-0001-5753-4431

²Pesquisadora IBICT, *rosali@ibict.br*, ORCID iD: 0000-0002-0890-7999

³Tecnologista ICICT/FIOCRUZ, *jeorgina.rodrigues@gmail.com*,
ORCID iD:0000-0001-8187-4049

Resumo

O objetivo do estudo é propor o desenvolvimento de práticas de indexação das fotografias do Álbum da construção do Canal do Mangue, Rio de Janeiro, Brasil, pertencente ao acervo iconográfico da Biblioteca de Obras Raras e Antigas do Centro de Tecnologia da Universidade Federal do Rio de Janeiro. Os métodos de análise foram as indexações derivativa e atributiva das fotografias, ressaltando a proveniência, significado, objetivo e história local da construção do Canal do Mangue. Os resultados revelaram elementos essenciais e complementares do conteúdo das imagens como termos de indexação temática. O álbum revelou-se como item documental raro pela sua unicidade e valor histórico e cultural para estudo e pesquisa em temas como urbanização e saneamento básico.

Palavras-chave: *Álbum fotográfico, Fotografia rara, Indexação de imagens, Indexação derivativa, Indexação atributiva.*

Introdução

A construção do Canal do Mangue foi um marco na área da Engenharia nacional realizada no início do século XX, significando uma importante reestruturação do centro da cidade do Rio de Janeiro, modificando não só a arquitetura urbanística como também aumentando a qualidade de vida, principalmente com relação ao saneamento básico e mobilidade urbana. O Canal do Mangue serviu para redirecionar as águas das chuvas para o mar, que anteriormente alagavam a cidade e traziam doenças através das águas contaminadas para dentro de instituições e residências, acarretando na extinção dos alagamentos e uma melhora substancial na saúde coletiva quanto à diminuição de casos de doenças endêmicas na época.

Atualmente, a montagem de um projeto de tratamento de esgoto e despoluição da Baía de Guanabara concretiza uma atualização das obras da construção do Canal do Mangue em âmbito urbanístico, voltada para o equilíbrio dos recursos hídricos e na qualidade de vida da cidade e de sua população. No início de 1990, o Programa de Despoluição da Baía de Guanabara foi concebido no intuito de elevar as condições sanitárias e ambientais da Região Metropolitana do Rio de Janeiro. O projeto fez parte do Programa de Saneamento Ambiental (PSAM), da Secretaria de Estado do Ambiente e Sustentabilidade (Seas), com a redução de 700 litros de esgoto por segundo lançados na Baía de Guanabara, melhorando a água do Canal e despoluindo a Baía e áreas adjacentes. Além disso, a ação da construção da Estação de Tratamento de Esgoto Alegria (ETE Alegria) no Rio de Janeiro (RJ), sendo composta por: Estação de Tratamento de Esgotos da Alegria, Troncos Coletores da Alegria, Tronco Coletor Faria Timbó, Tronco Coletor Manguinhos, Galeria de Cintura da Maré, Tronco Coletor Cidade Nova e subsistemas complementares. Este coletor tronco capta o esgoto sanitário de parte da Cidade Nova, Centro, Catumbi, Rio Comprido, Estácio e Santa Teresa. A instalação de 4,128 km de coletor tronco para a captação do esgoto lançado no Canal do Mangue teve sua destinação adequada na Estação de Tratamento de Esgoto de Alegria (ETE Alegria). As obras quando concretizadas beneficiarão cerca de 2,5 milhões de habitantes, com o tratamento de esgotos oriundos da bacia sanitária Alegria.

O objeto deste estudo é o *Álbum da construção do Canal do Mangue* (s.d.), com 50 fotografias sem tratamento técnico, data provável 1906, localizado no acervo iconográfico da Biblioteca de Obras Raras e Antigas (BOR) do Centro de Tecnologia (CT) da Universidade Federal do Rio de Janeiro (UFRJ). O *Álbum* tem sua devida importância por ser raro e único, ou seja, insubstituível como obra de referência, cujas fotografias são fonte de pesquisa histórica. O objetivo é propor o desenvolvimento de práticas de indexação do conteúdo temático das imagens, visando à recuperação da informação. A revisão da literatura indicou a existência de estudos sobre indexação de fotografias, contudo não foram localizados trabalhos específicos sobre indexação de “álbuns fotográficos”.

A imagem e sua veracidade como fonte de informação parte do uso dentro de um caráter investigativo (Bras et al., 2016), se relacionando às buscas de conhecimento e aos processos de pesquisa. No plano do controle social, a imagem fotográfica foi associada à identificação, passando a figurar, desde o início do século XX, em identidades, passaportes e os mais diferentes tipos de carteiras de reconhecimento social. No âmbito privado, através do retrato de família, a fotografia também serviu de prova (Mauad, 1996). Todavia, para a maioria dos autores que trabalham sob a ótica da imagem como fonte informacional, estes denotam a vertente do uso da fotografia/imagem fotográfica como documento e/ou registro de suma importância para pesquisas em diferentes áreas do conhecimento, sejam elas: administrativa, histórica, jurídica, acadêmica ou de cunho social (Tonello & Madio, 2018). Sendo assim, a criação do processo fotográfico se estende em evoluções perceptivas cumprindo funções de registro, fonte de informação e pesquisa (Mauad, 1996).

A fotografia surge como documento a partir do momento que registra visualmente paisagens, pessoas e eventos, absorvendo as características essenciais dos

documentos de arquivo e conservando no tempo e no espaço as ações e/ou fatos representados por meio da imagem, passando a ser vista como a forma mais adequada de representar, não só as ciências, mas também os acontecimentos históricos, culturais e sociais (Tonello & Madio, 2018). Então, a fotografia é uma forma de expressão cultural da qual foram registrados o tempo/momento nos diversos aspectos como religião, costumes, habitação e acontecimentos sociais de diversas naturezas (Kossov, 2001).

A fotografia se tornou presente em diversas áreas e contextos, dentre elas na cultura, na educação e em nível social proporcionando uma difusão do conhecimento e da informação registrada para que todos obtenham o acesso. Mediante as informações relativas à fotografia visando à utilização como fonte fidedigna de conhecimento, se aplica na “multiplicação dos documentos audiovisuais, em especial a fotografia, exige o estudo de seu significado e de seu conteúdo cultural enquanto registro da história” (Miguel, 1993, p. 122). Então, o uso do documento para registro histórico é digno de ser conservado, transmitido e estudado.

O álbum fotográfico possui um relacionamento particular com o passado como suportes de memórias, tecnologia e intenção de registro das informações (Santos & Albuquerque, 2019). É interpretado como um objeto de domínio constituído de uma maneira que se expande em suas características como um recurso informacional e um domínio do conhecimento (Santos, 2020). O álbum fotográfico revela o papel de narrador e construtor de um vínculo comunicacional, simbólico e afetivo, ao contar um registro ou um evento passado considerado significativo. Então, a produção de álbuns pode ser considerada como uma operação do fazer história, pois há deliberadamente uma intenção de registro de fatos e preservação da memória (Abdala, 2013). Para Santos e Albuquerque, os “álbuns fotográficos são elaborados com uma determinada finalidade que é constituírem uma narrativa de um determinado tema” (Santos & Albuquerque, 2019, p. 177), o caso do álbum fotográfico da construção do Canal do Mangue.

Os procedimentos metodológicos consideraram os processos de indexação derivativa e indexação atributiva. Na indexação derivativa foram selecionadas as palavras-chave atribuídas pelo fotógrafo na descrição temática de cada fotografia. Na indexação atributiva foram conferidas palavras-chave complementando os elementos descritivos. Os dois métodos de análise aumentaram a probabilidade de atender a diferentes interesses de busca de potenciais usuários.

Metodologia

O homem é considerado um ser cognitivo que processa a informação para elaborar conhecimentos com diversas linguagens para a transmissão de ideias e sentimentos e assim realizando o ato da comunicação (Santos, 2020). A determinação do significado das palavras deve passar da fase de indexação para a construção do vocabulário, considerando a operação efetiva de uma informação adequadamente identificada ao entrar no sistema, facilitará que ela seja recuperada (Cesarino & Pinto, 1978). A indexação é uma operação de representação documentária cuja finalidade de proporcionar a recuperação da informação, no ponto de vista evolutivo

do tratamento temático informacional, comportando a análise, a síntese e a representação da informação (Silva & Fujita, 2004).

A indexação de assuntos é uma atividade que visa a atender às necessidades de determinada clientela de usuários de um centro de informação ou de uma publicação específica (Lancaster, 2004). A evolução da prática, da dinâmica na recuperação e do uso da indexação decorre de uma necessidade de informação cada vez mais rápida, precisa e especializada. A construção de índices, metodologias e instrumentos diversificados auxiliam na contextualização de cada tipo de documento (Silva & Fujita, 2004).

Zeng (2007) comenta sobre a importância da construção do diálogo na organização dos registros bibliográficos e no envolvimento de questões como o processo de consenso e controle de informações sistematizadas. O controle de termos é necessário para que a cada conceito não se atribua mais do que um termo (Lancaster, 2004). O processo de descrição parte da análise do assunto reduzindo milhares de palavras a uma frase bem precisa sobre o conteúdo de assunto de um documento (Langridge, 2006).

Os critérios de seleção a serem aplicados à imagem, visando à sua representação, se constituem num problema de delineamento complexo sobre a leitura da imagem, sendo necessário operacionalizá-la com procedimentos que garantam a qualidade e a eficácia da representação da imagem (Smit, 1996). A fotografia de valor histórico auxilia no suporte de pesquisas onde, em muitos casos, culmina no cuidado primário com o tratamento dado às imagens (Santos, 2020).

A descrição de uma imagem nunca é completa tornando a percepção difícil, o que demanda treinamento para obter respostas detalhadas para as perguntas: “quem (seres vivos), onde (ambiente), quando (tempo), onde (espaço), o que (ação) e como (técnica)” (Smit, 1989, p. 109).

A necessidade da imagem ser observada de forma abrangente é compartilhada por vários autores, denotando o envolvimento das colocações idealizadas pelo indexador e os descritores são filtrados e selecionados de acordo com sua relevância para a recuperação, como um desafio constante para o indexador (SMIT, 1989). Sendo assim, a indexação da imagem e/ou da fotografia deve ser flexível, permitindo ao indexador proporcionar a comunicação eficaz com o usuário (Santos, 2020).

Na organização das fotografias, as únicas limitações em termos de representação gráfica seria a condensação das ideias de um conjunto de divisões da classificação, cujo sujeito não se encontra sintetizado na natureza das coisas (Otlet, 1934). Nesse sentido, os álbuns de fotografias podem ser compreendidos como um recurso informacional rico em dados, tendo características próprias que os tornam especiais na sua forma de organização e indexação de fotografias, devendo ter um olhar analítico e sensível perante o conjunto fotográfico (Santos, 2020).

No intuito de tornar as fotografias do *Álbum de Construção do Canal do Manguê* acessíveis, propiciando a visibilidade e recuperação da informação, busca-se um tratamento temático. Qual seria a melhor forma de se indexar as fotografias do *Álbum*?

A natureza da pesquisa é descritiva e exploratória. É descritiva no sentido de tratar as fotografias atribuindo temáticas. É exploratória no sentido de uma investigação empírica com o objetivo de aumentar a familiaridade do pesquisador com o ambiente da construção do Canal do Manguê.

Os procedimentos metodológicos consideraram as características da coleção e os elementos presentes nas 50 fotografias que compõem o referido *Álbum*: séries numéricas e a descrição temática realizada pelo fotógrafo. Além disso, foram atribuídas palavras-chave complementando os elementos descritivos.

A relevância teórica da pesquisa foi explorar os métodos de indexação derivativa e atributiva, de classificação facetada e tesouros para o tratamento das fotografias. A relevância prática do estudo foi a identificação do conteúdo temático das fotografias como subsídios para a construção de cabeçalhos de assuntos e para a hierarquização de conceitos em sistemas de classificação. A metodologia seguida mostrou-se válida para indexar e classificar as fotografias do *Álbum*.

Resultados

Com as descrições temáticas derivativas e atributivas foi possível identificar um conjunto de termos que descrevem as fotografias do *Álbum*. O Quadro 1 apresenta os resultados das descrições do álbum fotográfico. Os exemplos ilustram a fase de escavação do Canal e da sua conclusão e paisagismo.

Quadro 1 – Exemplos de Indexação derivativa e Indexação atributiva

FOTOGRAFIAS	DESCRIÇÃO
	<p>Séries numéricas: 0122 4-2-04</p> <p>Descrição derivativa: Engenheiros J. Marcellino Pinto Alg. Niemeyr Lafayette Pereira Canal do Mangue Excavação da valla com a draga Pristinan para as fundações da muralha de uma das margens do novo Canal no trecho recto. Rio de Janeiro</p> <p>Descrição atributiva: [Elementos]: Vala Escavadeira Obras Figuras humanas Operários [Localização]: Rio de Janeiro</p>

	<p>Séries numéricas: 02-06 18-02-05</p> <p>Descrição derivativa: Canal do Mangue Perspectiva do antigo canal visto da praça depois da sua modificação iniciada em 1903 Rio de Janeiro.</p> <p>Descrição atributiva: [Elementos]: Palmeiras Portão de acesso ao Canal Construções antigas Edificações Antigo Canal</p> <p>[Localização]: Avenida Presidente Vargas Perspectiva da Praça dos Marinheiros (atual Praça da Bandeira).</p>
--	--

Nota. Santos (2020)

A indexação derivativa e a indexação atributiva identificaram termos que descrevem detalhes das fotografias, além de indicarem a relevância histórica do *Álbum*. Esses conjuntos de termos mostraram-se válidos para subsidiar a construção de cabeçalhos de assunto, classificações hierárquicas e tesouros.

A Figura 1 ilustra a complexidade do planejamento do Canal do Mangue em relação ao que já foi construído e as perspectivas de extensão da obra.

Figura 1 – Estação Alegria



Nota. Cedae (s.d.)

A extensão do Canal do Mangue já efetuada (cor verde) e ainda por ser realizada (cor vermelha) evidencia a relevância da obra, para o urbanismo e saneamento da cidade do Rio de Janeiro.

Conclusões

A investigação mostrou que foi possível a descrição das fotografias do *Álbum* pelos métodos de indexação derivativa e atributiva válidos para a indexação de fotografias. Este trabalho oferece uma contribuição para diretrizes de construção de Sistemas de Organização do Conhecimento, tais como: construção de catálogo alfabético de assuntos, taxonomia, classificação facetada e tesouro.

Os registros pós-construção do Canal do Mangue relatam contribuições da obra, ressaltando benefícios para a área da Saúde na diminuição de doenças infecciosas e parasitárias da época. Nesse sentido, a obra de construção do Canal do Mangue continua contribuindo para questões de saneamento, infraestrutura e saúde, em conformidade com os Objetivos de Desenvolvimento Sustentável presentes na Agenda 2030 (Plataforma Agenda 2030, s.d.).

Referências

- Abdala, R. D. (2013). Fotografias escolares: práticas do olhar e representações sociais nos álbuns fotográficos da Escola Caetano de Campos (1895-1966). Tese de doutorado, Universidade de São Paulo, São Paulo, SP, Brasil. Recuperado em 14 julho, 2021, de <https://teses.usp.br/teses/disponiveis/48/48134/tde-04112013-113939/pt-br.php>. [Álbum da construção do Canal do Mangue]. (s.d.). [Manuscrito não publicado].
- Cedae (s. d.). Programa de Despoluição da Baía de Guanabara. Sistema Alegria. Recuperado em 29 julho, 2021, de https://cedae.com.br/despoluicao_baia_guanabara
- Cesarino, M. A. N., & Pinto, M. C. M. F. (1978). Cabeçalho de assunto como linguagem de indexação. *Revista de Biblioteconomia da UFMG*, 7(2), 268-288.
- Figueiredo, A. M. De, & Souza, S. R. G. de. (2011). *Como elaborar projetos, monografias, dissertações e teses: da redação científica à apresentação do texto final*. (4.^a ed). Lumen Juris
- Kossov, B. (2001). *Fotografia & História*. (Edição especial). Ateliê Editorial.
- Lancaster, F. W. (2004). *Indexação e resumos: teoria e prática*. Briquet de Lemos.
- Langridge, D. (1977). *Classificação: abordagem para estudantes de biblioteconomia*. Interciência.
- Mauad, A. M. (1996). Através da imagem: fotografia e história interfaces. *Revista Tempo*, 1(2), 73-98.
- Miguel, M. L. C. (1993). A fotografia como documento: uma investigação à leitura. *Revista Acervo*, 6(1-2), 121-132.
- Otlet, P. (1934). *Traité de Documentation: le livre sur le livre*. Mundaneum.
- Santos, C. R., & Albuquerque, A. C. de. (2019). A caracterização do álbum fotográfico como recurso informacional: elementos para a Organização e representação da Informação. *InCID*, 10 (10), 167-183.
- Santos, M. de B. dos. (2020). *Sistemas de organização do conhecimento e a recuperação de informação: em busca de diretrizes para a indexação de imagens fotográficas raras e*

- antigas*. Dissertação de mestrado, IBICT, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.
- Silva, M. dos R. da, & Fujita, M. S. L. (2004). A prática da indexação: análise da evolução de tendências teóricas e metodológicas. *Transinformação*, 16, (2), 133-161.
- Smit, J. W. (1989). A análise documentária: um primeiro plano. In J. W. Smit (Coord.). *Análise documentária da síntese* (2.^a ed., pp. 101-113). IBICT.
- Smit, J. W. (1996). A representação da imagem. *Informare*, 2 (2), 28-36.
- Plataforma Agenda 2030 (s.d.). Avançando o Desenvolvimento Sustentável. Recuperado em 30 julho, 2021, de <http://www.agenda2030.com.br/sobre>.
- Tonello, I. M. S., & Madio, T. C. de C. (2018). A fotografia como documento: com a palavra Otlet e Briet. *Informação & Informação*, 23 (1), 77-93.
- Zeng, M. L. (2008). Knowledge Organization Systems. *Knowledge Organization*, 35 (2-3), 160-182.

ORGANIZAÇÃO DO CONHECIMENTO EM ARQUIVOS: O USO DA IDENTIFICAÇÃO ARQUIVÍSTICA NA CLASSIFICAÇÃO DE DOCUMENTOS

Natália Bruno Rabelo¹, Clarissa Moreira dos Santos Schmidt²

¹Programa de Pós-graduação em Ciência da Informação da Universidade Federal Fluminense, nataliabruno@id.uff.br, ORCID iD 0000-0002-8910-5516

²Programa de Pós-graduação em Ciência da Informação da Universidade Federal Fluminense, clarissaschmidt@id.uff.br, ORCID iD 0000-0003-1555-4594

Resumo

O propósito deste artigo foi avaliar qualitativa e quantitativamente a produção bibliográfica, do período de 2011 a 2021, que reconhece a identificação arquivística como ferramenta para a classificação de documentos de arquivo. A identificação arquivística foi apontada, a partir da década de 1980, como uma metodologia de produção de instrumentos de gestão de documentos. Sendo assim, nossa investigação foi realizada com base na seguinte pergunta: “com que frequência tem sido utilizada a identificação arquivística como aspecto metodológico da classificação de documentos na literatura arquivística?”. Sustentando-se no método qualitativo e quantitativo, reconheceram-se as relações da organização do conhecimento e a classificação arquivística, contextualizamos a identificação arquivística e seus aspectos teóricos e metodológicos, além de mapearmos a produção bibliográfica sobre o uso da identificação arquivística para a classificação de documentos, em bases de dados selecionadas. Os resultados sugerem que ao mesmo tempo que um evento da área de Arquivologia contém o maior número de publicações sobre a temática, a dominância de discussões teóricas sobre o assunto é encontrada em eventos das áreas de Organização do Conhecimento e Ciência da Informação. Logo, diante dos resultados, infere-se que as temáticas mais frequentes nos eventos das referidas áreas de conhecimento são análises de planos de classificação e aspectos teóricos da identificação arquivística, enquanto a temática presente em ao menos três eventos abarca propostas de planos de classificação.

Palavras-chave: Arquivologia, Ciência da Informação, Organização do Conhecimento, Classificação arquivística, Identificação arquivística.

Introdução

A identificação arquivística consiste em um método fundamentado na Diplomática e na teoria arquivística, que pode ser considerada uma ferramenta significativa para a classificação de documentos de arquivo. A referida função arquivística é a

base para as demais fases da gestão de documento e comumente lhe é incumbida o dever recuperar as relações entre os documentos e as suas respectivas atividades.

Na medida em que a identificação arquivística é importante para a classificação de documentos, implica que a comunidade acadêmica manifeste interesse no tema através de publicações. Em razão disso, é pertinente analisar a produção científica recente da Arquivologia, na qual destacamos o período entre 2011 e 2021, a fim de mapear a quantidade e teor das publicações sobre o assunto e refletir sobre quais as implicações para a área a partir dos resultados obtidos.

Face a isto, cabe destacar que a identificação arquivística é um método utilizado para analisar elementos da forma documental que revelam o vínculo arquivístico, mais explicadamente, entre o documento e seu produtor, e esta ferramenta se torna um meio para a formulação de requisitos para funções arquivísticas, como a classificação e a avaliação (Rodrigues, 2008, p. 223-235).

Metodologia

O presente estudo tem como objetivo avaliar qualitativa e quantitativamente a compreensão da literatura sobre a relevância da identificação arquivística como metodologia na classificação de documentos a partir da seguinte indagação: “com que frequência têm sido utilizada a identificação arquivística como aspecto metodológico da classificação de documentos na literatura arquivística?”. Além disso, desenvolvemos pesquisa bibliográfica e documental para compreender o relacionamento da organização do conhecimento com a classificação de documentos, elucidar a identificação como princípio basilar da classificação e sistematizar a produção bibliográfica produzida nos últimos dez anos.

Nossa pesquisa foi executada através das bases de dados *Scientific Electronic Library* (SCIELO), Base de Dados em Ciência da Informação (BRAPCI), Google Acadêmico e Dialnet, bem como nos anais dos eventos:

- Congresso ISKO Espanha-Portugal, promovido pela *International Society for Knowledge Organization* (ISKO) Espanha-Portugal;
- Congresso Brasileiro em Organização e Representação do Conhecimento realizado pela ISKO Brasil;
- Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), promovido pela Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB) e
- Congresso Nacional de Arquivologia (CNA), organizado pelo Fórum Nacional de Associações de Arquivologia (FNARQ).

Mapeou-se a produção acadêmica sobre a identificação arquivística na classificação de documentos no período de 2011 a 2021 nestes eventos e, posteriormente, analisou-se aspectos teóricos, práticos e históricos no contexto da classificação e da identificação arquivística.

Resultados

A International Society for Knowledge Organization (ISKO), criada em 1989, na Alemanha, podendo ser considerada um marco da solidificação da organização do conhecimento como disciplina científica, tornando-se um espaço de interlocução acadêmica interdisciplinar. Esta área do conhecimento tem por finalidade gerenciar e organizar informações por meio de ferramentas que representem o resultado deste processo, o conhecimento, de modo que seja eficientemente recuperado (Guimarães, 2017).

Tratando-se da Arquivologia, as funções classificação e descrição são exemplificadas como elementos de organização do conhecimento. A classificação arquivística, objeto desta pesquisa, é um processo que se fundamenta no contexto de produção do documento a fim de representar suas respectivas funções ou estruturas relacionadas. Isto é, o conhecimento arquivístico, o objetivo da ciência supracitada, é adquirido por meio da análise dos documentos de um produtor específico que se configuram como informações e é pautado no princípio de respeito aos fundos, no contexto de produção e na forma documental (Tognoli, Rodrigues & Guimarães, 2019, p. 62-71). Portanto, considera-se como um requisito da classificação, o conhecimento prévio da estrutura administrativa da instituição junto as suas funções, das atividades representadas por documentos, da legislação que prevê a produção documental e outros.

Por outro lado, os embasamentos teóricos que constituem o método identificação arquivística, ou identificação da tipologia documental, se originaram da Diplomática quando adaptada à Arquivologia. A Diplomática, em seu sentido mais puro, avaliava o valor jurídico dos documentos, porém, após a necessidade de estender a aplicação da Diplomática como disciplina para também documentos de arquivo que não possuem necessariamente valor jurídico, a Arquivologia adaptou a disciplina em conformidade a essas demandas (Tognoli, 2013, p. 18-78).

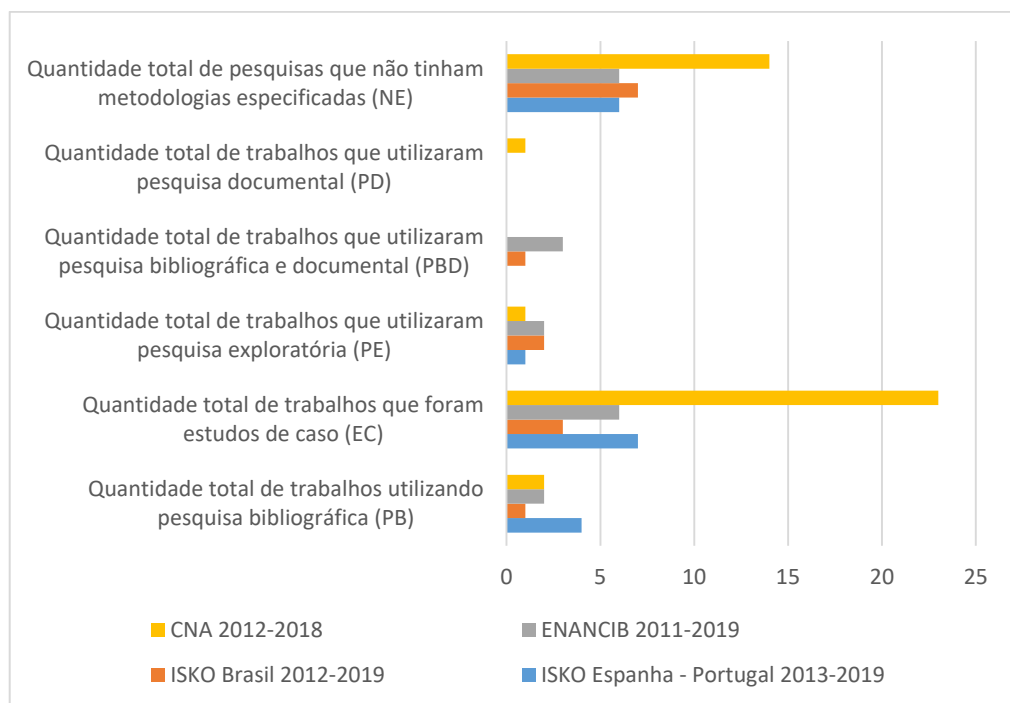
Na medida em que a fundamentação teórica da Diplomática é construída desde meados do século XVII, a identificação arquivística surge como uma ferramenta metodológica utilizada no estágio inicial para a elaboração de um diagnóstico, que pode ser considerado enquanto etapa primeira do processo de montagem de um projeto de arquivo onde se realiza o estudo analítico do documento arquivístico e o seu respectivo vínculo com o órgão produtor. Nesse contexto, a teoria da Diplomática passa a ser aplicada na análise de documentos, passando por uma adequação teórica que culmina na diplomática contemporânea, que contribui oferecendo fundamentos teóricos e metodológicos para a identificação da tipologia documental (Rodrigues, 2008, p. 21-32).

Dessa forma, é relevante analisar a bibliografia recente acerca da temática classificação arquivística que considera o uso da identificação arquivística como uma das metodologias para executar a referida função arquivística. Nesta análise, buscou-se refletir acerca dos eventos (i) Congresso ISKO Espanha-Portugal; (ii) Congresso Brasileiro em Organização e Representação do Conhecimento, promovido pela ISKO Brasil; (iii) ENANCIB; e (iiii) CNA.

Neste sentido, foi realizada a quantificação das publicações e suas respectivas metodologias de pesquisa dos eventos supracitados dentro do recorte temporal pro-

posto, a saber: pesquisa bibliográfica (PB), estudo de caso (EC), pesquisa exploratória (PE), pesquisa bibliográfica e documental (PBD), pesquisa documental (PD), estudo bibliométrico (EB) e metodologia não especificada (NE).

Tabela 1. Dados quantitativos das publicações dos eventos ISKO ESPANHA-PORTUGAL, ISKO BRASIL, ENANCIB e CNA.



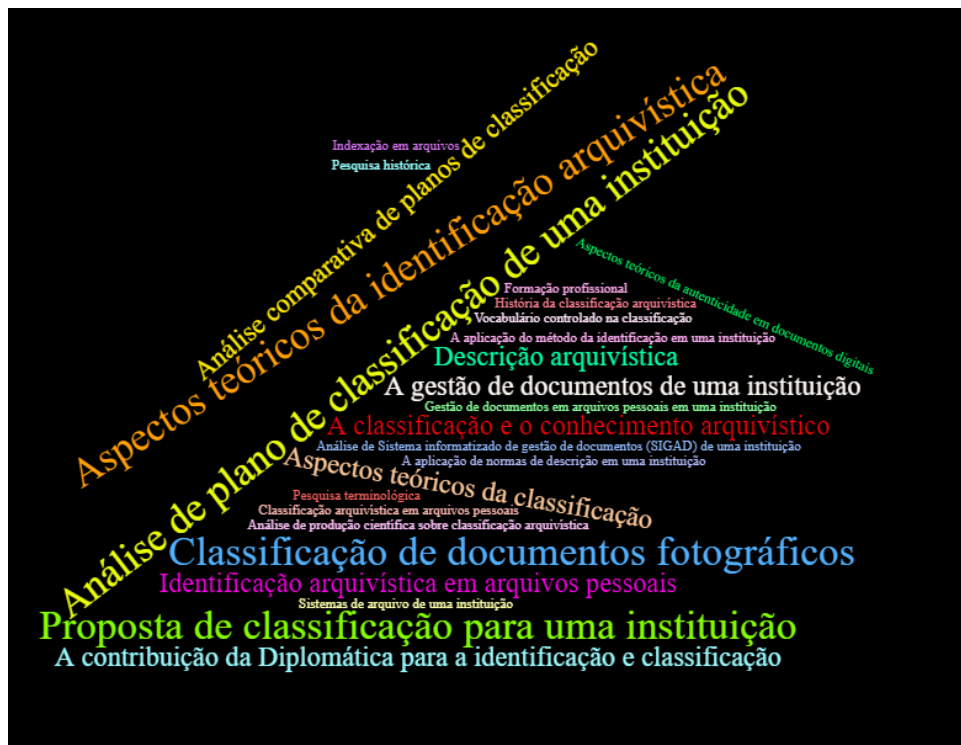
Elaboração das autoras (2021)

Diante dos dados apresentados, conclui-se que, no período analisado, o evento com maior quantidade de trabalhos publicados é o CNA, dispondo de um total de 42 publicações, sendo o segundo lugar ocupados com empate pelo Congresso ISKO Espanha-portugal e ENANCIB, contendo 18 publicações e, sequencialmente, 14 publicações pelo Congresso ISKO Brasil. Além disso, nota-se que o ano que possui maior quantidade de trabalhos publicados foi 2012, onde houve o somatório do CNA e ENANCIB, visto que no ISKO Brasil não foram encontradas publicações com esta temática. Isso nos leva a considerar um possível efeito pós-publicação da Lei n.º 12.527, de novembro de 2011, popularmente conhecida como Lei de acesso à Informação (LAI), que regula o acesso à informação, o qual está previsto na Constituição Federal do Brasil de 1988. A especificação do aporte legal para o acesso à informação no contexto brasileiro pode ter influenciado reflexões sobre a maneira de organizar o conhecimento arquivístico de modo que o mesmo seja acessível para o produtor e para os cidadãos.

Quanto às metodologias de pesquisa majoritariamente utilizadas, repara-se que estudo de caso é o método mais utilizado em todos os eventos estudados, sendo assim, o primeiro lugar nesta quantidade de metodologias é novamente ocupado pelo CNA, contendo um total de 22 estudos de caso. Já o evento com a maior quantidade de pesquisas bibliográficas é o Enancib, um total de cinco – considerando as pesquisas que se autodeclaram como pesquisas bibliográficas e documentais, sendo o segundo lugar ocupado pelo Congresso ISKO Espanha-Portugal. Já quanto os eventos com a maior quantidade de pesquisas exploratórias, destacam-se o Congresso ISKO Brasil e o ENANCIB, ambos com duas publicações utilizando esta metodologia. Entretanto, nota-se que há uma grande quantidade de metodologias não especificadas pelos autores em todos os eventos, totalizando 33 trabalhos não especificados, algo que supera a quantidade de estudos de casos.

Sendo assim, foram identificadas temáticas que envolvem o eixo central da presente revisão de literatura, o uso da identificação arquivística como método na classificação de documentos de arquivo, bem como quantificadas suas ocorrências nos eventos e período estudado nesta pesquisa. Para tanto, estas foram organizadas na nuvem de palavras a seguir onde o maior tamanho das palavras representa a maior quantidade de ocorrências do tema.

Imagem 1. Nuvem de palavras dos assuntos das publicações dos eventos ISKO ESPANHA-PORTUGAL, ISKO BRASIL, ENANCIB e CNA.



Como observado na análise qualitativa das publicações, as abordagens frequentes nos quatro eventos em questão são: a) análises de planos de classificação de uma instituição e b) aspectos teóricos da identificação arquivística. Sequencialmente, os com maior frequência, em ao menos três eventos, são propostas de classificação para uma instituição e classificação de documentos fotográficos. Já a frequência em dois eventos são análise comparativa de planos de classificação, a classificação e o conhecimento arquivístico, a contribuição da Diplomática para a identificação e a classificação, gestão de documentos de uma instituição e identificação arquivística em arquivos pessoais.

Quanto às abordagens com menor frequência nos quatro eventos, considera-se pesquisa terminológica, indexação em arquivos, vocabulário controlado na classificação, história da classificação arquivística, aplicação do método da identificação em uma instituição, aplicação de normas de descrição em uma instituição, gestão de documentos em arquivos pessoais em uma instituição, análise de SIGAD institucional, classificação em arquivos pessoais, formação profissional, pesquisa histórica, sistemas de arquivo de uma instituição, análise de produção científica sobre classificação e aspectos teóricos da autenticidade em documentos digitais.

Conclusões

Deste modo, a análise qualitativa sugere que há um equilíbrio entre discussões teóricas e estudos de casos ao destacar os temas mais frequentes que abrangem os quatro eventos estudados. Entretanto, a maior parte dos temas analisados qualitativamente envolvem um contexto macrossocial, em outras palavras, discussões teóricas que não são realizadas em um contexto institucional. Contudo, de forma oposta, a análise quantitativa sugere que, apesar desta variedade de temas de reflexões teóricas, a maioria das publicações consistem em estudos de caso.

Ao analisar os valores quantitativos como parâmetros desta pesquisa, o fato do CNA ser o evento que detém o maior número de publicações que contenham os temas sobre classificação arquivística e identificação arquivística, alude à percepção de que a Arquivologia é a área que possui, ultimamente, maior aproximação destas discussões em comparação às áreas de Organização do Conhecimento e Ciência da Informação. Entretanto, visto que o CNA é um evento promovido por associações profissionais, é compreensível que haja mais estudos de caso, em outras palavras, experiências profissionais, em comparação às reflexões teóricas que, neste caso, são encontradas em maior quantidade em eventos científicos, como observado a liderança de metodologias utilizadas diferentes de estudos de caso, o que leva a refletir sobre as áreas de Organização do Conhecimento e Ciência da Informação estarem mais próximas deste tipo de discussão.

Referências

- Associação Nacional de Pesquisa e Pós-Graduação (2011). XII Encontro Nacional de Pesquisa em Ciência da Informação, Brasília: Universidade de Brasília.
- Associação Nacional de Pesquisa e Pós-Graduação (2012). XIII Encontro Nacional de Pesquisa em Ciência da Informação, Rio de Janeiro: Fundação Oswaldo Cruz.
- Associação Nacional de Pesquisa e Pós-Graduação (2013). XIV Encontro Nacional de Pesquisa em Ciência da Informação, Santa Catarina: Universidade Federal de Santa Catarina.
- Associação Nacional de Pesquisa e Pós-Graduação (2014). XIV Encontro Nacional de Pesquisa em Ciência da Informação, Rio de Janeiro: Universidade Federal do Rio de Janeiro.
- Associação Nacional de Pesquisa e Pós-Graduação (2015). XVI Encontro Nacional de Pesquisa em Ciência da Informação, João Pessoa: Universidade Federal da Paraíba.
- Associação Nacional de Pesquisa e Pós-Graduação (2016). XVII Encontro Nacional de Pesquisa em Ciência da Informação, Salvador: Universidade Federal da Bahia.
- Associação Nacional de Pesquisa e Pós-Graduação (2017). XVIII Encontro Nacional de Pesquisa em Ciência da Informação, Marília: Universidade Estadual Paulista.
- Associação Nacional de Pesquisa e Pós-Graduação (2018). XIX Encontro Nacional de Pesquisa em Ciência da Informação, Londrina: Universidade Estadual de Londrina.
- Associação Nacional de Pesquisa e Pós-Graduação (2019). XX Encontro Nacional de Pesquisa em Ciência da Informação, Santa Catarina: Universidade Federal de Santa Catarina.
- Fórum Nacional das Associações de Arquivologia (2012). V Congresso Nacional de Arquivologia, Salvador: Universidade Federal da Bahia.
- Fórum Nacional das Associações de Arquivologia (2014). VI Congresso Nacional de Arquivologia, Santa Maria: Universidade Federal de Santa Maria.
- Fórum Nacional das Associações de Arquivologia (2016). VII Congresso Nacional de Arquivologia, Fortaleza: Universidade Federal do Ceará.
- Fórum Nacional das Associações de Arquivologia (2018). VIII Congresso Nacional de Arquivologia, João Pessoa: Universidade Estadual da Paraíba.
- Guimarães, J. A. (2017). Organização do conhecimento: passado, presente e futuro sob a perspectiva da ISKO. *Informação & Informação*, 22(2), 84-98. doi:<http://dx.doi.org/10.5433/1981-8920.2017v22n2p84>
- International Society for Knowledge Organization Brasil (2012). In *Anais*, I Congresso Brasileiro em Organização e Representação do Conhecimento, Marília: Universidade Estadual Paulista Júlio Mesquita.
- International Society for Knowledge Organization Brasil (2013). In *Anais*, II Congresso Brasileiro em Organização e Representação do Conhecimento, Rio de Janeiro: Universidade Estadual Paulista Júlio Mesquita.
- International Society for Knowledge Organization Brasil (2015). In *Anais*, III Congresso Brasileiro em Organização e Representação do Conhecimento, Marília: Universidade Estadual Paulista Júlio Mesquita.
- International Society for Knowledge Organization Brasil (2017). In *Anais*, IV Congresso Brasileiro em Organização e Representação do Conhecimento, Recife: Universidade Federal de Pernambuco.
- International Society for Knowledge Organization Brasil (2019). In *Anais*, V Congresso Brasileiro em Organização e Representação do Conhecimento, Belém: Universidade Federal do Pará.

- International Society for Knowledge Organization España-Portugal (2013). In *Anais, I Congresso ISKO Espanha e Portugal XI Congreso ISKO España*, Porto: Faculdade de Letras da Universidade do Porto.
- International Society for Knowledge Organization España-Portugal (2015). In *Anais, XII Congreso ISKO España II Congreso ISKO España-Portugal*, Murcia: Universidad de Murcia.
- International Society for Knowledge Organization España-Portugal (2017). In *Anais, III Congreso ISKO Espanha-Portugal XIII Congreso ISKO Espanha*, Coimbra: Universidade de Coimbra.
- International Society for Knowledge Organization España-Portugal (2019). In *Anais, IV Congreso ISKO España-Portugal XIV Congreso ISKO España*, Zaragoza: Universitat de Barcelona.
- Rodrigues, A. C. (2008). *Diplomática contemporânea como fundamento metodológico da identificação de tipologia documental em arquivos*. Tese de Doutorado, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo. doi:10.11606/T.8.2008.tde-27112008-151058. Recuperado em 2020-12-28, de www.teses.usp.br
- Tognoli, N. B. (2013). *A construção teórica da Diplomática: em busca de uma sistematização de seus marcos teóricos como subsídio aos estudos arquivísticos*. Tese de Doutorado, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, São Paulo. Recuperado em 2020-12-28, de <https://repositorio.unesp.br/handle/11449/103383>
- Tognoli, N. B., Rodrigues, A. C., & Guimarães, J. A. (2019). Definindo o conhecimento arquivístico: estruturas conceituais. *Informação & Informação*, 24(2), 58-75. doi:<http://dx.doi.org/10.5433/1981-8920.2019v24n2p58>

COLIBRI – ARTES GRÁFICAS

APARTADO 42 001 – 1601-801 LISBOA

TELEFONE | (+351) 21 931 74 99

www.edi-colibri.pt | colibri@edi-colibri.pt

