



UNIVERSIDAD DE CASTILLA-LA MANCHA

Escuela Superior de Informática

Departamento de Tecnologías y Sistemas de Información

TESIS DOCTORAL

**MODELO HÍBRIDO DE RECOMENDACIÓN DE
OBJETOS DE APRENDIZAJE**

Realizada por:

Alfredo Zapata González

Directores:

Dr. D. Manuel Emilio Prieto Méndez

Dr. D. Cristóbal Romero Morales

Ciudad Real, Febrero 2013

AGRADECIMIENTOS

A mis directores de tesis, Dres. Manuel Emilio Prieto Méndez y Cristóbal Romero Morales por el apoyo brindado y por compartir conmigo su experiencia y conocimientos durante la realización de esta tesis doctoral.

A las Mtras. María Elena Barrera y Julieta Guerrero Walker, por la confianza que han depositado en mí y por todas las facilidades administrativas brindadas para la realización del doctorado.

A mi esposa Maritza, quien ha sido mi fuerza durante todos estos años para superar los retos que se fueron presentando.

A mi papá por enseñarme a luchar y alcanzar mis metas, a mi mamá por encausar mi camino y ser siempre mi ejemplo de entrega y superación.

A mi familia, por ser mi motivador y mi apoyo en los momentos más difíciles.

A mis amigos de los grupos de investigación SMLe y MAT, por todos los buenos momentos compartidos dentro y fuera de las paredes del laboratorio.

Este proyecto de investigación fue realizado con el apoyo y financiamiento de la Universidad Autónoma de Yucatán; el Consejo Nacional de Ciencia y Tecnología del gobierno de México y el Consejo de Ciencia y Tecnología del estado de Yucatán a través del “Programa de Becas para Estudios de Posgrado en el Extranjero”.

ÍNDICE DE CONTENIDO

AGRADECIMIENTOS	3
ÍNDICE DE FIGURAS	V
ÍNDICE DE TABLAS	VII
RESUMEN	IX
1 INTRODUCCIÓN	1
1.1 ENTORNO DE LA INVESTIGACIÓN	1
1.2 HIPÓTESIS Y OBJETIVOS	3
1.3 MÉTODO DE TRABAJO	5
1.4 MARCO DE TRABAJO	7
1.4.1 GRUPO DE INVESTIGACIÓN	7
1.4.2 PROYECTOS DE I+D+I	8
1.5 ORGANIZACIÓN DEL DOCUMENTO	11
2 ESTADO DEL ARTE	15
2.1 DEFINICIÓN Y CONCEPTOS DE E-LEARNING	15
2.2 OBJETOS DE APRENDIZAJE Y SU ORGANIZACIÓN	18
2.2.1 DEFINICIÓN	18
2.2.2 CARACTERÍSTICAS	19
2.2.3 METADATOS	20
2.2.4 IEEE-LOM	20
2.2.5 ADL-SCORM	23
2.2.6 OTRAS ESPECIFICACIONES	24
2.2.7 REPOSITARIOS DE OBJETOS DE APRENDIZAJE	25
2.3 BÚSQUEDA Y RECOMENDACIÓN	27
2.3.1 CONCEPTOS Y DEFINICIONES	27
2.3.2 MÉTODOS DE FILTRADO DE INFORMACIÓN	31
2.3.3 ALGORITMOS DE RECOMENDACIÓN	33
2.3.4 APLICACIONES EN ENTORNOS EDUCATIVOS	35
2.4 COLABORACIÓN EN LA BÚSQUEDA Y LA RECOMENDACIÓN	41
2.4.1 COLABORACIÓN	41

2.4.2	BÚSQUEDA SOCIAL COLABORATIVA.....	42
2.4.3	RECOMENDACIONES DE GRUPO.....	44
2.5	DESCUBRIMIENTO DE CONOCIMIENTO Y MINERÍA DE DATOS	48
2.5.1	DEFINICIONES Y TAXONOMÍA.....	48
2.5.2	MINERÍA DE DATOS.....	51
2.5.3	META-APRENDIZAJE	55
2.5.4	MINERÍA DE DATOS EDUCATIVA	58
2.6	CONCLUSIONES	61
3	MODELO DE RECOMENDACIÓN HÍBRIDO DE OBJETOS DE APRENDIZAJE INDIVIDUAL Y COLABORATIVO	63
3.1	BÚSQUEDA INDIVIDUAL.....	63
3.1.1	BÚSQUEDA INDIVIDUAL SIMPLE.....	63
3.1.2	BÚSQUEDA INDIVIDUAL COMPUESTA	65
3.1.3	CALIFICACIÓN Y RECOMENDACIÓN.....	69
3.2	BÚSQUEDA COLABORATIVA	71
3.2.1	GESTIÓN DE GRUPOS.....	72
3.2.2	BÚSQUEDA EN GRUPO	73
3.2.3	VALORACIÓN EN GRUPO	73
3.2.4	RANKING CONSENSUADO.....	74
3.2.5	META-APRENDIZAJE PARA LA SELECCIÓN DEL MÉTODO DE AGREGACIÓN.....	74
3.3	CONCLUSIONES	82
4	DELPHOS. SISTEMA DE RECOMENDACIÓN HÍBRIDO	85
4.1	MARCO ARQUITECTÓNICO	85
4.2	ESPECIFICACIONES TÉCNICAS.....	88
4.2.1	TECNOLOGÍAS DE DESARROLLO	88
4.2.2	ESTRUCTURA DE LA BASE DE DATOS.....	90
4.2.3	REPOSITORIO AGORA	91
4.3	INTERFAZ WEB.....	93
4.4	FUNCIONALIDAD DEL SISTEMA DELPHOS	95
4.4.1	REGISTRO DE USUARIOS.....	95
4.4.2	BÚSQUEDA INDIVIDUAL DE OBJETOS DE APRENDIZAJE.....	96
4.4.3	GESTIÓN DE GRUPOS.....	103
4.4.4	BÚSQUEDA COLABORATIVA DE OBJETOS DE APRENDIZAJE.....	105

4.4.5	ACCIONES DEL GRUPO	108
4.5	CONCLUSIONES	110
5	EXPERIMENTACIÓN	113
5.1	PRUEBA DEL SISTEMA DE RECOMENDACIÓN HÍBRIDO CON PROFESORES	113
5.1.1	RENDIMIENTO	113
5.1.2	USABILIDAD	115
5.2	PRUEBA DEL SISTEMA DE RECOMENDACIÓN HÍBRIDO CON ALUMNOS	117
5.2.1	CRITERIOS DE FILTRADO	117
5.2.2	RENDIMIENTO	122
5.2.3	USABILIDAD	127
5.3	PRUEBA DEL SISTEMA DE RECOMENDACIÓN COLABORATIVO CON GRUPOS	129
5.3.1	FORMACIÓN DE GRUPOS	129
5.3.2	SELECCIÓN DEL MÉTODO DE AGREGACIÓN	133
5.4	CONCLUSIONES	139
6	CONCLUSIONES Y TRABAJO FUTURO.....	141
6.1	ANÁLISIS DE LA CONSECUCIÓN DE OBJETIVOS	141
6.2	ANÁLISIS DE LAS APORTACIONES	144
6.3	DIVULGACIÓN DE RESULTADOS	146
6.3.1	ARTÍCULOS EN REVISTAS INDIZADAS EN EL JOURNAL CITATION REPORTS.....	146
6.3.2	ARTÍCULOS EN OTRAS REVISTAS	146
6.3.3	CAPÍTULOS EN LIBROS INTERNACIONALES	147
6.3.4	CAPÍTULOS EN LIBROS NACIONALES	148
6.3.5	PONENCIAS EN CONGRESOS INTERNACIONALES.....	148
6.3.6	PONENCIAS EN CONGRESOS NACIONALES	148
6.4	TRABAJO A FUTURO	149
6.4.1	MEJORAS EN LAS TÉCNICAS UTILIZADAS	150
6.4.2	MEJORAS EN EL SISTEMA DE RECOMENDACIÓN HÍBRIDO DELPHOS	151
6.4.3	DIVULGACIÓN DEL CONOCIMIENTO ADQUIRIDO	152
	BIBLIOGRAFÍA	155

ÍNDICE DE FIGURAS

Figura 1.1 Esquema general de la tesis.....	13
Figura 2.1 Esquema del estándar IEEE-LOM.....	22
Figura 2.2 Flujo de trabajo del modelo de colaboración 3C	42
Figura 2.3 Transformación de datos en información.....	49
Figura 2.4 Etapas del proceso KDD	50
Figura 2.5 Taxonomía de los modelos de aprendizaje empleados en la minería de datos	52
Figura 2.6 Ciclo de aplicación de minería de datos en entornos educativos	58
Figura 3.1 Método de búsqueda individual simple.....	64
Figura 3.2 Método de búsqueda individual compuesta	66
Figura 3.3 Método de búsqueda colaborativa de objetos de aprendizaje	72
Figura 3.4 Proceso de meta-aprendizaje.....	74
Figura 4.1 Marco arquitectónico del sistema de recomendación híbrido DELPHOS.....	86
Figura 4.2 Interacción de las tecnologías de desarrollo Web	90
Figura 4.3 Esquema de las tablas de la base de datos y sus relaciones.	90
Figura 4.4 Interacción entre el sistema DELPHOS y el repositorio AGORA	92
Figura 4.5 Proceso de diseño centrado en el usuario.....	93
Figura 4.6 Diseño visual de la página Web de entrada a DELPHOS.....	94
Figura 4.7 Búsqueda individual de objetos de aprendizaje	98
Figura 4.8 Panel de los criterios de recomendación	98
Figura 4.9 Ejemplo de listado de objetos de aprendizaje recuperados.....	99
Figura 4.10 Creación de un grupo de búsqueda colaborativa.....	103
Figura 4.11 Listado de grupos a los que pertenece el usuario.....	103
Figura 4.12 Incorporar nuevos miembros al grupo	104
Figura 4.13 Incorporar miembros al grupo.....	105
Figura 4.14 Listado de objetos recomendados en la búsqueda colaborativa.....	105
Figura 4.15 Añadir objetos de aprendizaje a los grupos que pertenece el usuario.....	106
Figura 4.16 Opciones colaborativas de valoración.....	107

Figura 4.17 Panel de las acciones del grupo	109
Figura 4.18 Interfaz del servicio de mensajería para los miembros del grupo	110
Figura 5.1 Resultados del promedio de las evaluaciones llevadas a cabo por los 30 profesores de acuerdo a la encuesta SUS. Los valores tienen un rango entre 1 (muy en desacuerdo) a 5 (muy de acuerdo).	117
Figura 5.2 Resultados del total MRR obtenido de las configuraciones de las siete pruebas.	124
Figura 5.3 Recall de los diferentes top-N.	126
Figura 5.4 Comparativa del Average Recommendation Error obtenido por los algoritmos de clasificación.	138
Figura 5.5 Comparativa del Mean Reciprocal Rank obtenida por los algoritmos de clasificación.	139

ÍNDICE DE TABLAS

Tabla 2.1 Categorías del estándar IEEE-LOM.....	21
Tabla 2.2 Comparación de los sistemas de recomendación	40
Tabla 2.3 Resumen de métodos de agregación.....	46
Tabla 3.1 Ejemplo del método Promedio	76
Tabla 3.2 Ejemplo del método Promedio descartando en valor mínimo.....	76
Tabla 3.3 Ejemplo del método Valor mínimo	77
Tabla 3.4 Ejemplo del método Valor máximo	77
Tabla 3.5 Ejemplo del método Equitativo	78
Tabla 3.6 Ejemplo del método Voto pluralista.....	78
Tabla 3.7 Ejemplo del método Voto aprobado (con umbral 2)	79
Tabla 3.8 Ejemplo del método Voto aprobado (con umbral 3).....	79
Tabla 3.9 Ejemplo del método de Promedio ponderado (más peso al miembro con más experiencia)	80
Tabla 3.10 Ejemplo del método de Promedio ponderado (más peso al miembro más social)	80
Tabla 3.11 Ejemplo del método de Promedio ponderado (más peso al miembro más activo)	81
Tabla 3.12 Ejemplo del método Mediana.....	81
Tabla 4.1 Criterios del perfil de registro de los usuarios.....	96
Tabla 4.2 Descripción de los íconos asociados a los objetos recuperados de la búsqueda individual.....	101
Tabla 4.3 Mecanismo de evaluación pedagógico cualitativo MECOA.....	102
Tabla 4.4 Descripción de los iconos extras que contiene la búsqueda colaborativa	107
Tabla 5.1 Ejemplo de los parámetros de búsqueda en el primer experimento	114
Tabla 5.2 Resultados obtenidos al aplicar diferentes criterios de filtrado.....	115
Tabla 5.3 Preguntas de la Escala de Usabilidad del Sistema (SUS).....	116

Tabla 5.4 Valores de los criterios de recomendación utilizados en cada configuración de prueba.....	118
Tabla 5.5 Ejemplo de parámetros de búsqueda utilizados en el experimento	119
Tabla 5.6 Ranking de los 10 OA mejor posicionados de las siete configuraciones de pruebas	120
Tabla 5.7 Matriz de correlación del rango de Spearman	121
Tabla 5.8 Resultados de los cuestionarios SUS y CSUQ	128
Tabla 5.9 Distribución por áreas de investigación y programas de nivel superior de los usuarios	130
Tabla 5.10 Distribución de los grupos por tamaño	130
Tabla 5.11 Configuración de los 50 grupos conformados	131
Tabla 5.12 Características de los 50 grupos conformados.....	134
Tabla 5.13 Ranking obtenido por los métodos de agregación en el grupo 1	135

RESUMEN

En la actualidad existe una amplia diversidad de repositorios que brindan recursos digitales para la educación en forma de objetos de aprendizaje. Sin embargo, la mayoría de estos sistemas se centran más en la organización e indexado de estos objetos que en los métodos de búsqueda y selección de los objetos. Esto reduce la posibilidad de ofrecer los objetos de aprendizaje más apropiados a las necesidades tanto de usuarios particulares como de grupos de usuarios.

Para solucionar este problema en esta tesis se propone un *modelo híbrido de recomendación de objetos de aprendizaje*, el cual se basa en fuentes de información provenientes de los metadatos asociados a los objetos, los perfiles de registro de los usuarios, así como también, de la interacción que se produce entre los usuarios y los propios objetos.

El modelo desarrollado abarca mecanismos de búsqueda individual y colaborativa. Por un lado, la búsqueda individual permite personalizar la búsqueda utilizando cuatro criterios de recomendación basados en: similitud de contenido, uso, evaluación y similitud de perfil de usuario. Además se utilizan pesos adaptativos y dinámicos, que permiten modificar automáticamente sus valores de acuerdo a la densidad de información disponible sobre cada criterio de filtrado. Por otro lado, en la búsqueda colaborativa se propone facilitar a los miembros de un mismo grupo, poder realizar las actividades más frecuentes como la petición, la recuperación, la incorporación, y valoración de los objetos de acuerdo a los intereses del grupo. Además, en el modelo se implementan estrategias de agregación de votaciones y técnicas de meta-aprendizaje para obtener de forma automática las valoraciones finales sin necesidad de tener que llegar a un consenso.

El modelo se implementó en el *sistema DELPHOS*, con el que se han realizado pruebas experimentales con profesores y alumnos para mostrar la efectividad tanto del modelo híbrido propuesto como de las recomendaciones individuales y de grupo.

1 INTRODUCCIÓN

Este capítulo tiene como propósito ofrecer una visión general del trabajo en esta tesis. Se inicia describiendo el entorno de la investigación, resaltando los problemas relacionados con la recomendación de objetos de aprendizaje que han dado origen a la hipótesis y los objetivos de esta investigación. A continuación, se introduce el método y el marco de trabajo empleado dentro la investigación, así como su relación con proyectos I+D+i desarrollados dentro de los dos grupos de investigación en el que enmarca esta tesis. Finalmente se presenta la estructura organizativa del documento.

1.1 ENTORNO DE LA INVESTIGACIÓN

Un *Objeto de Aprendizaje (OA)*, es una unidad básica de información en sistemas educativos, que combina uno o más recursos didácticos junto con los metadatos que los describen (Wiley, 2002). Esta estructura permite su intercambio y reutilización entre distintos ambientes educativos sin que exista pérdida de su funcionalidad o características.

La proliferación de sitios Web y repositorios dedicados a ofrecer estos recursos destinados a la educación es una prueba del constante desarrollo del e-learning en la actualidad. Es por ello que numerosas instituciones transforman, crean y publican recursos educativos para que sean aprovechados y reciclados. Además, estos espacios congregan a usuarios que comparten muchas veces características e intereses.

Dichos usuarios realizan diversas actividades de gestión de OA para publicar y editar la información que los describe (metadatos), así como también, pueden interactuar y evaluar aquellos recursos que otros usuarios incorporan. La interacción entre los usuarios y los recursos, facilita la compartición, valoración y reutilización de los mismos.

A pesar de estos avances, existen problemas asociados a la localización, selección y reutilización de los OA almacenados en los diversos repositorios existentes.

Por ejemplo, muchos OA son producidos sin que pueda garantizarse su disponibilidad, en ocasiones por falta de metadatos, o bien debido a que no se encuentran organizados y clasificados adecuadamente. Su capacidad de reutilización se ve limitada ya que la información acerca de sus contenidos no está estandarizada y muchas veces se encuentra ausente o incompleta.

Otro inconveniente frecuente en los sistemas y repositorios de recursos educativos, es que se ofrecen a los usuarios recursos que no son apropiados a sus necesidades. Es importante tomar en cuenta que cada usuario posee metas específicas y está dentro de un contexto propio cuando busca información.

De acuerdo con (McLean, 2003), los repositorios actúan como filtros para los recursos referenciados y almacenados, pues la única forma de acceder a ellos es por medio de una búsqueda basada en los metadatos almacenados. Sin embargo, la mayoría de los repositorios tienen motores de búsqueda muchas veces deficientes o confusos para los usuarios.

La generación de metadatos, las evaluaciones y las búsquedas, son actividades frecuentes que pueden aportar datos de interés sobre los recursos, a efectos de decidir su utilización en entornos diferentes a aquellos en los que fueron creados. La personalización de los procesos de filtrado y recuperación de la información puede contribuir a atenuar la sobrecarga de información, ajustando los resultados presentados a las necesidades individuales de cada usuario. Para ello, se requieren herramientas tecnológicas que simplifiquen la localización, reutilización e intercambio de estos recursos.

La tarea de búsqueda normalmente se suele realizar de forma individual, esto se debe a que los principales motores de búsqueda y navegadores Web están diseñados para el uso individual (Morris, 2008). Sin embargo, esta modalidad de búsqueda puede resultar un trabajo extra para los usuarios por el número de consultas y la revisión de páginas Web resultantes.

Existe otro tipo de búsqueda denominada colaborativa, la cual permite que muchas actividades profesionales y ocasionales puedan beneficiarse de las preferencias y experiencias de otras personas con ideas afines (Smyth et al., 2011). Las desventajas de esta

modalidad son las consultas similares que tienden a repetirse y la regularidad de los resultados obtenidos (Preyne, 2004).

Esta investigación se orienta al establecimiento de un modelo de recomendación híbrido de OA, el cual permita ayudar a los profesores y alumnos en la búsqueda de OA tanto de forma individual como colaborativa. La idea es utilizar toda la información disponible sobre los objetos y los usuarios del sistema. De hecho, los repositorios suelen almacenar en bases de datos gran cantidad de información sobre los OA, los usuarios y sus interacciones.

La utilización de toda esta información permitiría mejorar y personalizar las búsquedas, de forma que se puedan ofrecer recomendaciones basadas no sólo en objetos de contenidos similares, sino también en los objetos más utilizados, o en los mejor valorados o incluso en los objetos creados por los usuarios con intereses similares.

Así los objetos finalmente mostrados puedan ordenarse teniendo en cuenta esos criterios de modo que la tarea de búsqueda se convierte en una tarea de recomendación y personalización, que facilita al usuario a encontrar los objetos de mayor interés.

1.2 HIPÓTESIS Y OBJETIVOS

A continuación se presenta la hipótesis, el objetivo general y los objetivos específicos del proyecto. A partir de la pregunta de investigación:

Pregunta de investigación

¿Es posible mejorar los procesos de búsqueda, selección y recuperación de objetos de aprendizaje empleando un modelo de recomendación?

Se pretende probar que la búsqueda, selección y recuperación de OA puede ser más eficiente si se realiza a través de un modelo de recomendación, cuyas recomendaciones se apoyen en diversas fuentes de información asociadas a:

- Los metadatos asociado a los OA
- Las actividades de gestión que han realizado los propios usuarios con los objetos

- Indicadores de calidad obtenidos a través de algunos de los modelos o instrumentos de evaluación implementados.
- Perfiles de registro de los usuarios

Después de un análisis de los problemas asociado a la recuperación eficiente de OA y tomando en cuenta la diversidad de información disponible, se plantea la siguiente hipótesis:

Hipótesis inicial

Los sistemas de recomendación, basados en fuentes de información asociadas a los objetos de Aprendizaje y perfiles de usuario, optimizan la búsqueda, selección y recuperación de dichos objetos de aprendizaje.

A partir de la hipótesis establecida y una vez estudiados los aspectos relacionados con la problemática presentada, se indica el objetivo general de este trabajo:

Objetivo general

Establecer un modelo híbrido de recomendación de objetos de aprendizaje tanto para un sólo usuario individual como para un grupo de usuarios. Se considera un modelo híbrido, en el sentido de utilizar un modelo de conocimiento que permita describir tanto las características de los objetos como los perfiles de los usuarios utilizando simultáneamente diversas técnicas de filtrado y ordenación o ranking.

Este objetivo principal se detalla a través de los siguientes objetivos específicos:

1. Explorar las diversas fuentes de información que almacenan los OA.
2. Implementar un mecanismo de recomendación de OA para el establecimiento de una ordenación o rankings entre los objetos seleccionados.
3. Proponer un método para ayudar a los profesores y alumnos en la búsqueda individual de los OA más apropiados para cada uno.
4. Proponer un método de búsqueda colaborativa de OA que permita ayudar a la creación de grupos de usuarios con intereses comunes así como la recomendación de OA para dichos grupos.

5. Implementar una metodología capaz de adaptarse a diferentes entornos de e-learning, como repositorios o sistemas de gestión del aprendizaje.
6. Definir un conjunto de instrumentos e indicadores para medir la eficiencia y efectividad de la propuesta.
7. Realizar un conjunto de experimentos basados en cantidades suficientes de datos capturados de las actividades de gestión realizada por profesores y alumnos en condiciones reales.

1.3 MÉTODO DE TRABAJO

El método de trabajo utilizado en el desarrollo de esta tesis sigue el enfoque de la Investigación-Acción en Sistemas de Información, el cual se ha convertido en uno de los principales métodos de investigación cualitativa en el campo de los sistemas de información y en la ingeniería del software (Ruiz et al., 2002).

De forma general, se describen a continuación las principales fases consideradas para la realización de la tesis:

- 1. Definición de la problemática.** Se refiere a la determinación de forma eficiente de la problemática asociada a la búsqueda, selección y recuperación de OA. Esto ha originado la necesidad de estudiar las diferentes metodologías que involucran técnicas de sistemas de recomendación y hacen uso de las diversas fuentes de información asociadas a los OA y los perfiles de usuario.
- 2. Análisis del estado del arte.** Donde se ha realizado una revisión sistemática (Biolchini et al, 2005) de las temáticas afines a los OA, repositorios, técnicas de extracción de conocimiento, técnicas de sistemas de recomendación, entre otros. La revisión sistemática contiene una secuencia estricta y bien definida de pasos metodológicos, de acuerdo con un protocolo definido a priori (Kitchenham, 2007).
- 3. Propuesta de un marco metodológico para la recomendación de objetos de aprendizaje.** La propuesta se ha dividido en las siguientes actividades:
 - a. Definición de un modelo de recomendación híbrido de objetos de aprendizaje.
 - b. Establecimiento de las técnicas combinadas de filtrado de información.

- c. Establecimiento de técnicas de recomendación para grupos.
- 4. Implementación de un prototipo.** Esta fase comprende las siguientes actividades:
 - a. Desarrollar una aplicación software a partir del modelo de recomendación híbrido de OA generado.
 - b. Implementar la aplicación software dentro del entorno de una arquitectura abierta y distribuida que fomente la búsqueda y recomendación de OA.
- 5. Experimentación.** Se ha realizado un estudio empírico del método y los sistemas desarrollados para demostrar su validez. Para ello, se han llevado a cabo experimentos con datos reales en cantidades suficientes provenientes de profesores y alumnos de la Universidad Autónoma de Yucatán, México.
- 6. Difusión.** Se han difundido y compartido los resultados obtenidos en las fases anteriores entre la comunidad investigadora mediante la presentación de trabajos en distintos foros nacionales e internacionales, así como en la publicación de un sitio Web donde se informa del proyecto y se promocionan los productos generados. La retroalimentación ha venido tanto en la forma de las revisiones previas a la presentación de las ponencias, así como los comentarios vertidos por los usuarios del sitio Web.

El punto de partida de esta tesis doctoral es el proyecto fin de máster titulado “Extracción de conocimiento a partir de Recursos Digitales para el Aprendizaje” (Zapata, 2009) donde surgieron las primeras ideas de un modelo de recomendación de OA, y la tesis doctoral titulada "Marco conceptual para la gestión integral de Objetos de Aprendizaje" (Menéndez, 2011) que culminó con el desarrollo del repositorio AGORA.

Como recursos de trabajo se han empleado los medios tecnológicos y el material para la realización de la experimentación del grupo de investigación SMILe de la Universidad de Castilla-La Mancha, España, a los que se añaden los recursos bibliográficos físicos o electrónicos que proporciona la propia Universidad.

También constituye una valiosa aportación al proyecto, la participación del autor de esta tesis en foros nacionales e internacionales relacionados con el campo objeto de estudio. De esta manera se contrasta la investigación con el desarrollo del contexto científico en el que está enmarcado.

1.4 MARCO DE TRABAJO

A continuación se presenta el entorno en el que se ha realizado la investigación, así como los proyectos I+D+i en los que se ha contribuido.

1.4.1 GRUPO DE INVESTIGACIÓN

Esta tesis ha sido realizada gracias a la colaboración entre dos grupos de investigación que poseen una amplia trayectoria en temas afines a la investigación desarrollada.

Por un lado el grupo de investigación SMILe (Soft Management of Internet and Learning) de la Universidad de Castilla-La Mancha, España. El cual desarrolla sus líneas de investigación en el ámbito de la aplicación de técnicas de soft computing dentro de la recuperación de información en Internet; de sistemas de aprendizaje y apoyo a la decisión. Resultado de estas contribuciones son las diferentes herramientas de búsqueda y de gestión del conocimiento que se han implementado y puesto en funcionamiento durante los últimos años (SMILe, 2012).

Por otro lado, el grupo de investigación KDIS (Knowledge Discovery and Intelligent Systems) de la Universidad de Córdoba, España. El cual centra su investigación en temas relacionados con las técnicas de soft computing (en especial, los algoritmos bioinspirados), el descubrimiento de conocimiento en bases de datos, el aprendizaje automático y sus aplicaciones a distintos problemas del mundo real: sistemas de recomendación, personalización y mejora de los sistemas educativos, entre otros (KDIS, 2012).

Ambos grupos de investigación han desarrollado diversas líneas de investigación las cuales han sido relevantes para esta tesis:

- Desarrollo de métodos de recuperación de la información para la Web basados en características conceptuales contenidas en los documentos.
- Modelos colaborativos, e-learning y objetos de aprendizaje.
- Desarrollo de modelos de representación de conocimiento.
- Desarrollo de sistemas basados en el conocimiento.
- Modelos de sistemas de recomendación aplicados en entornos educativos.

- Minería de datos en educación.

1.4.2 PROYECTOS DE I+D+I

Esta tesis ha contribuido al cumplimiento de algunos de los objetivos definidos en distintos proyectos de investigación de los grupos SMILe y KDIS. A continuación, se describen los proyectos I+D+i en los cuales este trabajo ha tenido alguna relevancia:

- **Proyecto SCAIWEB-2:** Sistemas de acceso a la información en la Web basados en soft computing. PAC06-0059, PEIC09-0196-3018. Consejería de Ciencia y Tecnología de la Junta de Comunidades de Castilla-La Mancha, España. 2006-2009. La finalidad de este proyecto es la evolución de la plataforma de búsqueda en Internet SCAIWEB hacia la organización y la gestión de los contenidos dinámicos generados por la Web (páginas Web, blogs, correo electrónico). Mediante la utilización de diferentes técnicas basadas en soft computing, principalmente en lógica borrosa, se desarrollan diferentes métodos para organizar automáticamente las colecciones de documentos Web en una estructura jerárquica y borrosa basada en contenidos. Se plantean tareas de revisión de las técnicas hasta ahora aplicadas para esta labor, desarrolla e implanta los complementos necesarios (algoritmos de agrupamiento, retroalimentación por pseudorelevancia, técnicas de visualización avanzada, filtrado de información basado en calidad y en perfiles de usuario) a la plataforma actual para convertirla en una solución integral para la organización, recuperación y gestión de la información Web.
- **ECRT: Metodología y herramienta para la evaluación de la calidad de los recursos para tele-aprendizaje en la formación de profesores.** AA/8172/07, A/016625/08. Agencia Española de Cooperación Internacional para el Desarrollo (AECID), España. 2007-2009. Define un método para medir la calidad de los recursos didácticos de teleaprendizaje utilizados en la formación de profesores (MECOA, Modelo de Evaluación de la Calidad en los Objetos de Aprendizaje). La calidad está basada en las dimensiones del aprendizaje referidas a su aprovechamiento. Este método considera las particularidades de este tipo de recursos, de modo que puedan evaluarse fácilmente por las instituciones educativas y por las personas que desarrollan cursos de formación docente. Se miden, además, las

características informáticas de los recursos, dando una pauta de sus capacidades para utilizarse en las plataformas de instrucción a distancia. El método provee formas de medir la calidad integral de los objetos de aprendizaje. Se realiza una herramienta informática que implementa el método y que permite a las instituciones disponer de un mecanismo sencillo y distribuido para la evaluación de los objetos de aprendizaje en la formación de profesores.

- **PLINIO: Observatorio de efectos del cambio climático basado en la extracción inteligente de información en Internet.** POII10-0133-3516. Consejería de Ciencia y Tecnología de la Junta de Comunidades de Castilla-La Mancha, España. 2009-2010. Existe un gran volumen de información en Internet sobre los efectos del cambio climático. Esta información es proporcionada tanto desde instituciones oficiales como por periódicos, organizaciones no gubernamentales, entre otros. Esta información, que es muy útil con el fin de poseer un conocimiento actualizado sobre el tema, no puede aprovecharse debido a su gran volumen y a la dispersión en las fuentes. Este proyecto está orientado a solucionar este problema mediante la utilización de métodos y tecnologías de última generación relacionadas con la Web semántica como las ontologías, los metabuscadores y los agentes inteligentes con el fin de localizar, analizar, filtrar y categorizar la información que sobre efectos del cambio climático pueda generarse en Internet. Con este sistema se logra poner en disposición del usuario de forma periódica y organizada la información más significativa sobre el tema que haya aparecido en la Web, lo cual puede llegar a ser una ventaja estratégica para la toma de decisiones a corto, medio y largo plazo.
- **FIDELIO: Técnicas de filtrado difuso de la información aplicadas a metabuscadores.** TIN2010-20395. Ministerio de Ciencia e Innovación, España. 2010-2012. El objetivo principal de este trabajo es la aplicación de técnicas borrosas de filtrado de información en metabuscadores con el fin de ofrecer resultados que sean más satisfactorios para el usuario sobre todo en entornos con un alto nivel de dinamismo en sus contenidos (correo electrónico, WWW, blogs etc.). Se plantea por lo tanto el estudio y la aplicación de las técnicas, que dentro del ámbito del soft computing, poseen las características más idóneas para resolver este problema. Además para ello se recurrirá a estrategias como el clustering documental, la

generación automática de ontologías, el filtrado de información basado en perfiles de usuario y la representación del conocimiento mediante prototipos. El trabajo culmina con el desarrollo de una plataforma que implemente los diferentes métodos de organización automática de colecciones dinámicas de documentos en estructuras difusas basadas en contenidos. La evaluación de esta plataforma se realizará mediante su aplicación a herramientas de búsqueda avanzada como los metabuscadores difusos.

- **iNsPIrED:** New Problems in Knowledge Discovery: A Genetic Programming Approach. TIN-2011-22408. Ministerio de Ciencia e Innovación, España 2011-2015. El proyecto iNsPIrED (New Problems In Knowledge Discovery) tiene como principal objetivo el desarrollo de nuevas metodologías de extracción de conocimiento mediante técnicas de programación genética, y su aplicación a distintos problemas reales. El proyecto pretende: desarrollar modelos para la resolución de distintos problemas de extracción de conocimiento, adaptar los modelos anteriores a nuevos problemas con diferentes tipos de datos, aplicar los modelos desarrollados a problemas reales en el ámbito de la minería de datos educativos y de la minería de datos Web, y desarrollar repositorios de datos que permitan a la comunidad científica la comparación de resultados e integración de los modelos desarrollados.
- **ATECSE:** Aplicación de Técnicas de Extracción de Conocimiento en los Sistemas Educativos. P08-TIC-3720. Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía 2008-2012. El proyecto Aplicación de Técnicas de Extracción de Conocimiento a los Sistemas Educativos (ATECSE) se plantea con un doble objetivo. En primer lugar, el desarrollo de nuevas técnicas de extracción de conocimiento interesante en entornos educativos, concretamente, en las áreas de asociación, secuenciación, clasificación y categorización de documentos, así como su aplicación a la resolución de distintos problemas de interés en el ámbito educativo. En segundo lugar, pero no menos importante, se plantea el desarrollo de una herramienta de extracción de conocimiento específica para su aplicación en el contexto de la EDM.

1.5 ORGANIZACIÓN DEL DOCUMENTO

La memoria de esta tesis doctoral se estructura en seis capítulos (ver Figura 1.1). Estos capítulos describen distintos aspectos de la investigación realizada: revisión documental, trabajos relacionados, metodologías, implementaciones y validaciones. A continuación se detallan los contenidos de los distintos capítulos:

- **Capítulo 1: Introducción.** Este capítulo presenta una visión general del trabajo. Se inicia indicando el marco de la investigación, la hipótesis y los objetivos perseguidos en la realización de esta investigación. Se establece el método de trabajo y el espacio donde se enmarca en términos del grupo de investigación y los proyectos relacionados. Por último, se exponen las aportaciones de esta investigación.
- **Capítulo 2: Estado del arte.** En este capítulo se presenta una revisión documental relativa a las temáticas más relacionadas y que resultan primordiales para la investigación desarrollada en este trabajo: conceptos de e-learning, objetos de aprendizaje y su organización, conceptos y procesos involucrados en el descubrimiento de conocimiento, técnicas de minería de datos, definiciones y conceptos sobre búsqueda, recomendación y colaboración.
- **Capítulo 3: Modelo de recomendación híbrido de objetos de aprendizaje individual y colaborativo.** En este capítulo se detalla el modelo de recomendación híbrido de objetos de aprendizaje, el cual incorpora los tipos de búsqueda individual y colaborativa. Para cada uno se describe el marco conceptual y las distintas métricas generadas a partir del marco.
- **Capítulo 4: DELPHOS. Sistema de recomendación híbrido.** Este capítulo presenta la implementación de la metodología descrita en el capítulo previo. Se describe la aplicación software orientada para asistir a los usuarios en las tareas de búsqueda, selección y valoración de objetos de aprendizaje.
- **Capítulo 5: Experimentación.** Este capítulo describe los experimentos realizados para la validación del modelo propuesto y el sistema DELPHOS. Para ello, se establecen medidas que comprueban la efectividad y el rendimiento de las recomendaciones individuales. También se consideran cuestionarios para evaluar la

usabilidad del entorno del sistema. Posteriormente, se seleccionan métodos de agregación y algoritmos de clasificación/predicción para la recomendación colaborativa.

- **Capítulo 6: Conclusiones y trabajo futuro.** En este capítulo se describen las principales aportaciones de esta tesis, el cumplimiento de los objetivos planteados y las principales líneas de trabajo futuro que se abren a partir de esta investigación.

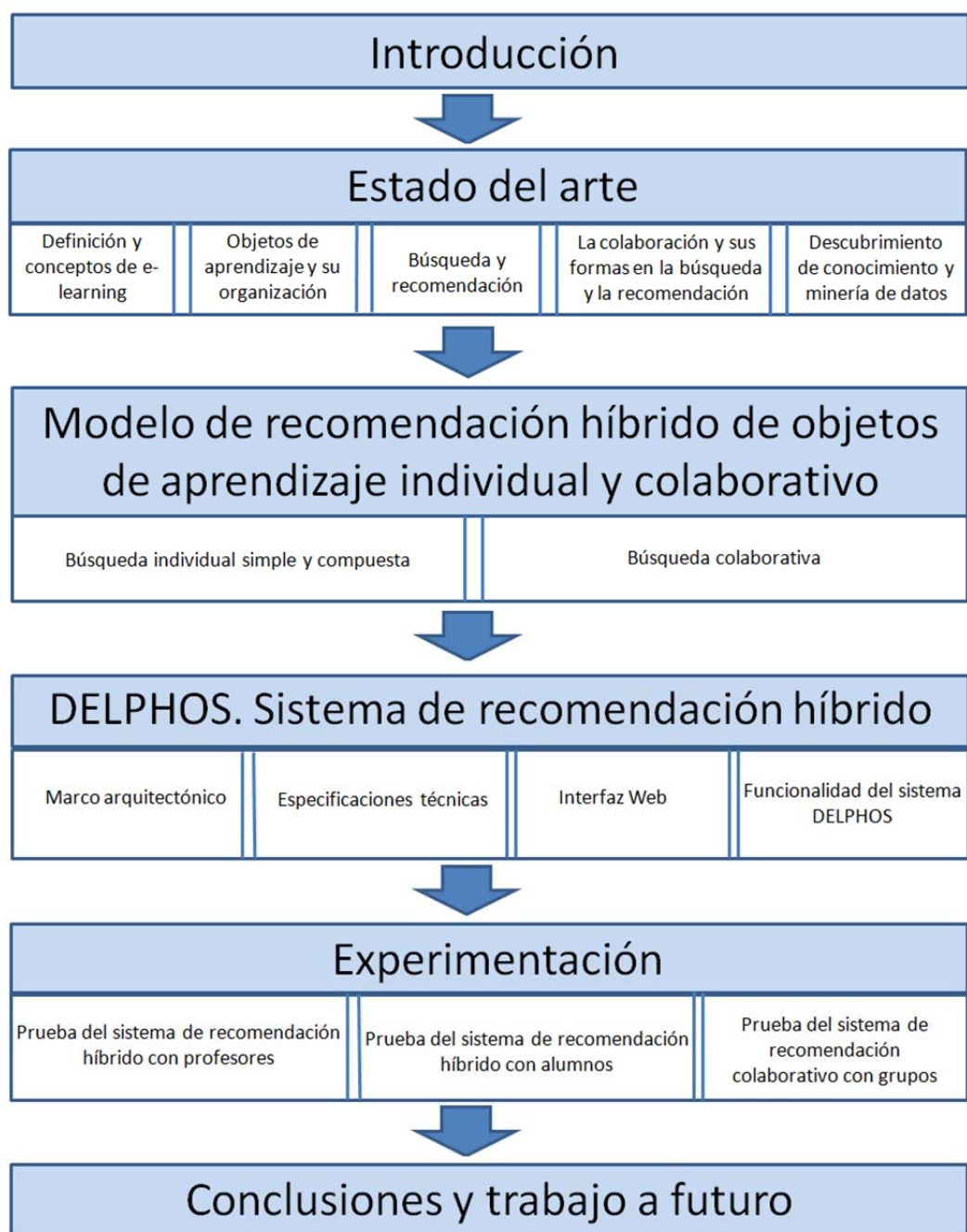


Figura 1.1 Esquema general de la tesis

2 ESTADO DEL ARTE

El objetivo de este capítulo es introducir los antecedentes relacionados con el tema de investigación de esta tesis. Primero se describen los conceptos más importantes relativos al e-learning. Posteriormente se presentan conceptos, características, estándares y especificaciones para la creación y descripción de objetos de aprendizaje, así como las principales metodologías para su desarrollo. En tercer lugar, se aborda el tema de búsqueda y recomendación eje fundamental del trabajo realizado, en el que se incluye una comparativa de los sistemas de recomendación en entornos educativos. Posteriormente, se aborda el tema de la colaboración y sus diferentes modalidades de búsqueda y recomendación que se pueden utilizar en este tipo de sistemas. Finalmente, se expone el estado de la minería de datos Web con énfasis en su aplicación en entornos educativos.

2.1 DEFINICIÓN Y CONCEPTOS DE E-LEARNING

Los avances en las tecnologías de la información junto con los cambios en la sociedad, han dado lugar a la creación de nuevos paradigmas para la educación como el e-learning, en el que la Web juega un papel importante (Khan, 1997). El concepto de e-learning está asociado con actividades de aprendizaje que impliquen el uso de redes y computadores en forma simultánea (Tsai & Machado, 2002). El término e-learning puede definirse de una forma simple como el empleo de Internet para facilitar el aprendizaje. Esto implica utilizar las herramientas disponibles en Internet y más específicamente en la Web, para adaptar el ritmo de aprendizaje al estudiante sin considerar los límites de espacio y tiempo (Brogan, 1999). Existen otras modalidades del e-learning tales como:

- **E-training.** Es aquella opción formativa, compuesta de cursos cortos hasta un nivel de diploma o certificado, tiene como objetivo la actualización de conocimientos o

habilidades en un área sumamente específica, ya sea de un campo disciplinar o bien relacionada con ciertas destrezas y habilidades laborales (Sánchez et al., 2006)

- **Web-based education.** Es la modalidad más extrema del e-learning, debido a que utiliza el streaming de vídeos y las funcionalidades más avanzadas disponibles en el software educativo y donde no hay cara a cara real de contacto entre el profesor y el alumno (Lynch, 2002).
- **Mobile education.** Es una extensión del e-learning que se realiza a través de dispositivos móviles de comunicación, por ejemplo: teléfonos inteligentes, asistentes personales digitales, tabletas, entre otros (Traxler, 2005).

Conviene considerar en e-learning, dos aspectos principales:

- **Educativo.** Está relacionado con la necesidad de contar con un diseño instruccional sólido y acorde con las consideraciones pedagógicas, las características del medio y el comportamiento de los usuarios.
- **Tecnológico.** Se trata de las herramientas que se ofrecen para la presentación de los contenidos; la comunicación entre usuarios o la evaluación, entre otras.

Dado el punto de vista tecnológico, el e-learning ha evolucionado, de simples páginas HTML (HiperText Markup Language) hasta complejos sistemas de gestión del aprendizaje. Se pueden identificar cinco tipos de tecnologías e-learning que pueden ser utilizadas solas o combinadas (Kapp, 2003):

- 1) **Lenguajes de programación.** Permiten tener un completo control y flexibilidad en los mecanismos de interacción del aprendizaje en línea. Sin embargo, una de sus desventajas es que el mantenimiento de tales sistemas es una tarea que requiere contar con personal especializado.
- 2) **Paquetes de autoría.** Emplean elementos, muchas veces visuales, que facilitan la construcción de aplicaciones e-learning sin necesidad de contar con conocimientos de programación. Su principal desventaja es que ofrecen una funcionalidad limitada y generalmente no incluyen los elementos de interacción propios de un sistema e-learning. Además, suelen estar diseñados para un lenguaje de programación específico.

- 3) **Sistemas de Gestión del Contenido (CMS, Content Management System).**
Permiten tener control y orden sobre los recursos educativos debido a que todos los elementos de información están catalogados. Esto permite una búsqueda y exploración del contenido de forma sencilla.
- 4) **Sistemas de Gestión del Aprendizaje (LMS, Learning Management System).**
Están diseñados para asistir al proceso de enseñanza-aprendizaje en un ambiente en línea mediante un conjunto de herramientas. Pueden tener una orientación hacia el alumno o el profesor.
- 5) **Sistemas de Gestión de Contenidos de Aprendizaje (LCMS, Learning Content Management System).** Son sistemas independientes o integrados con un LMS, que gestionan y administran los contenidos de aprendizaje. Una vez que los contenidos están en este sistema, pueden ser combinados, asignados a distintos cursos o descargados, entre muchas otras opciones.

De todas las tecnologías presentadas, las referidas a los LMS son las más empleadas actualmente por numerosas instituciones para organizar y distribuir cursos en línea (Kapp, 2003). Generalmente los LMS son fáciles de usar, se consideran flexibles en términos pedagógicos y eficientes con relación a los costos. Además de permitir organizar los contenidos educativos a la vez ofrece una serie de funcionalidades de comunicación (Downes, 2005). Algunos ejemplos de LMS son:

- **MOODLE (Modular Object Oriented Distance Learning Environment).** Es un sistema de gestión de cursos de aprendizaje en línea de distribución libre. Su diseño está basado del enfoque pedagógico constructivista social. Posee una arquitectura modular, lo que permite incorporar una gran diversidad de componentes y funcionalidades. Implementa una interfaz de navegador de tecnología sencilla, ligera, eficiente y compatible (Dougiamas & Taylor, 2003).
- **Dokeos.** Es un sistema de aprendizaje en línea basado en software libre. Posee una interfaz intuitiva que facilita la creación y organización de contenidos interactivos y ejercicios, así como también, incorpora herramientas de seguimiento e informes sobre el desempeño de los alumnos en el curso. Su arquitectura está basada en módulos lo que facilita la incorporación de nuevas funciones (De Praetere, 2010).

- **ILIAS (Integrated Learning, Information and cooperAtion System).** Es un sistema de gestión del aprendizaje de código abierto. Posee un ambiente de aprendizaje flexible, el cual permite la reutilización y el intercambio de materiales de instrucción entre profesores y alumnos. Destaca su sistema de control de acceso basado en roles y la facilidad para la creación de cuestionarios (University of cologne, 2004).
- **ATutor.** Es un sistema de código abierto que se utiliza para la impartición de cursos en línea. Su instalación es rápida y pueden desarrollarse temas para cambiar la apariencia de la plataforma así como instalar nuevos módulos para ampliar las funcionalidades de la misma. Los profesores pueden crear, configurar y empaquetar cursos para su redistribución en la red y los estudiantes aprenden en un entorno adaptativo y social (Toronto University, 2002).

2.2 OBJETOS DE APRENDIZAJE Y SU ORGANIZACIÓN

Los objetos de aprendizaje tienen gran importancia en la actualidad. Se consideran piezas básicas en las modalidades de enseñanza basadas en la computación y las comunicaciones; siendo útiles en todos los niveles y formas de organización de la actividad docente. Son muy utilizados en los sistemas e-learning. Además constituyen medios auxiliares para la capacitación en las empresas e incluso en diversas formas de instrucción no reglada.

El propósito de los objetos de aprendizaje es proporcionar colecciones de unidades de información basadas en estándares que proporcionen flexibilidad, portabilidad y reutilización de los contenidos instruccionales, propiciando un alto grado de control a los profesores y los estudiantes (Wiley, 2002).

2.2.1 DEFINICIÓN

El concepto de “objeto de aprendizaje” no es sencillo de definir, ya que después de varios años de haberse acuñado el término, no existe un consenso sobre su definición. En consecuencia se tienen diversas concepciones sobre lo que debe ser un objeto de aprendizaje (Knolmayer, 2003) (Mohan, 2004).

Una de las definiciones más generales es la que propone el estándar IEEE-LOM (IEEE-Learning Object Metadata) (IEEE-LTSC, 2002), el cual considera un OA, como cualquier entidad digital o no digital, que puede ser utilizada, reutilizada o referenciada durante el aprendizaje asistido por la tecnología, como sistemas de instrucción basados en computador, sistemas de educación a distancia, tutores inteligentes, entre otros.

Otra definición más enfocada en los entornos virtuales educativos es la de McGreal (McGreal, 2004), que define los OA como cualquier recurso digital reusable que tiene encapsulada una lección o ensamblado un grupo de lecciones en unidades, módulos, cursos e incluso programas.

Ante esta falta de consenso y para efectos de este trabajo, un OA se define como un elemento reciclable con contenido multimedia que tiene un propósito instructivo y cumple con alguna especificación tecnológica. Esto engloba a cualquier entidad digital desarrollada siguiendo un diseño instruccional (Merrill, 1996), que puede utilizarse, reutilizarse o referenciarse para el aprendizaje.

2.2.2 CARACTERÍSTICAS

Si existe el consenso de que, idealmente, un OA debe tener las siguientes características básicas (Wiley, 2002), (ADL, 2004):

- **Granularidad.** Los OA son definidos como unidades finas o "granos" que pueden combinarse e incorporarse de múltiples maneras.
- **Autocontenido.** Por sí solo el objeto debe ser capaz de cumplir un objetivo establecido.
- **Reutilización.** Una vez creados, los OA deben funcionar en diferentes contextos de aprendizaje.
- **Interoperabilidad.** El objeto debe basarse en estándares que aseguren su utilización en distintos sistemas de e-learning.
- **Accesibilidad.** Los OA deben etiquetarse para que sean identificados dentro de un conjunto y facilitar su acceso y gestión.
- **Escalabilidad.** Los OA deben ser duraderos y actualizables en el tiempo, para que sus componentes puedan ser mejorados.

- **Organización.** Los OA deben ser clasificables y secuenciales en un mismo entorno de aprendizaje.

Los OA están compuestos de dos partes principales: por un lado, el contenido del objeto en sí mismo (elementos multimedia o bien otros OA) y por otro, las etiquetas (también denominadas metadatos) que describen el contenido (Wiley, 2002). Un OA también puede incluir un conjunto de descriptores de comportamiento o los requerimientos para su utilización y composición, es decir, un contrato (Sánchez-Alonso, 2005).

2.2.3 METADATOS

Los metadatos contienen información primordial y objetiva acerca de los OA. Representan un modo de caracterización de análisis introspectivo de los objetos. Estos pueden concebirse como un conjunto de atributos o elementos estructurados para describir la naturaleza de un recurso, tal como sus requerimientos, sus características, entre otros, lo que permite su gestión, localización y recuperación (Al-Khalifa & Davis, 2006). Por lo tanto, son uno de los principales factores que influyen en la reusabilidad del objeto (Wiley, 2002).

El almacenamiento de los metadatos se realiza a través de una estructura compacta normalmente codificada en un formato basado en XML (eXtensible Markup Language), que se anexa al objeto como una capa extra de información que a su vez se adapta a especificaciones y estándares (Decker et al., 2000).

Los estándares y especificaciones cubren varios aspectos tales como el empaquetamiento, el etiquetado y la secuenciación. Actualmente existen propuestas de estándares que cubren los aspectos mencionados, como son IEEE-LOM, ADL-SCORM, entre otros.

2.2.4 IEEE-LOM

Es uno de los estándares más completos, debido a que establece la sintaxis y la semántica de los metadatos del objeto de aprendizaje como atributos requeridos para describir de forma adecuada y completa a un objeto (Fischer, 2001). Es el principal estándar de catalogación de OA.

IEEE-LOM nace como producto del trabajo cooperativo de grupos relacionados con el desarrollo de estándares y la gestión de objetos de aprendizaje tales como: ARIADNE

(ARIADNE Foundation, 2001), IMS GLC (IMS Global Learning Consortium) (IMS Global Learning Consortium, 1997), DCMI (Dublin Core Metadata Initiative) (Dublin Core Metadata Initiative, 1995) y IEEE-LTSC (IEEE Learning Technology Standard Committee) (IEEE-LTSC, 2002), siendo este último su principal promotor.

El estándar IEEE-LOM establece un esquema conceptual para la representación de metadatos de un objeto de aprendizaje. Define una estructura jerárquica formada por 9 categorías (ver Tabla 2.1) y 68 elementos. Cada categoría contiene elementos que pueden almacenar otros elementos o valores, como analogía a las ramas y hojas de un árbol (ver Figura 2.1). Su estructura es lo suficientemente flexible como para incorporar nuevos metadatos así como definir vocabularios controlados para sus valores.

Tabla 2.1 Categorías del estándar IEEE-LOM

Categoría	Descripción
General	Proporciona información general sobre el OA. Sus valores están referidos al objeto como un todo. Contiene 10 subelementos.
LifeCycle	Agrupar todas las características y datos relacionados al proceso de desarrollo del OA hasta su estado actual, así como a los participantes en ese proceso. Contiene 6 subelementos.
MetaMetadata	Provee información sobre los metadatos definidos para la instancia, quién desarrolló la instancia, cuándo, qué referencias utilizó. Contiene 9 subelementos.
technical	Describe los requerimientos técnicos y las características tecnológicas del recurso. Contiene 12 subelementos.
Educational	Describe al OA en términos instruccionales y pedagógicos. Puede haber varias instancias de esta categoría. Contiene 11 subelementos.
Rights	Describe los derechos de propiedad intelectual, así como las condiciones de uso del OA. Contiene 3 subelementos.
Relation	Agrupar elementos que establecen las relaciones de un OA con otros objetos. Define 12 tipos de relaciones. Pueden existir varias instancias de esta categoría para definir múltiples relaciones. Contiene 6 subelementos.
Annotation	Proporciona comentarios sobre el OA, principalmente del tipo educativo, así como quién y cuándo se realizó la anotación. Pueden existir numerosas anotaciones para un mismo objeto. Contiene 3 subelementos.
Classification	Describe al OA con respecto a un esquema de clasificación. Al haber múltiples instancias de esta categoría, es posible clasificar un mismo objeto utilizando distintos esquemas. Contiene 8 subelementos.

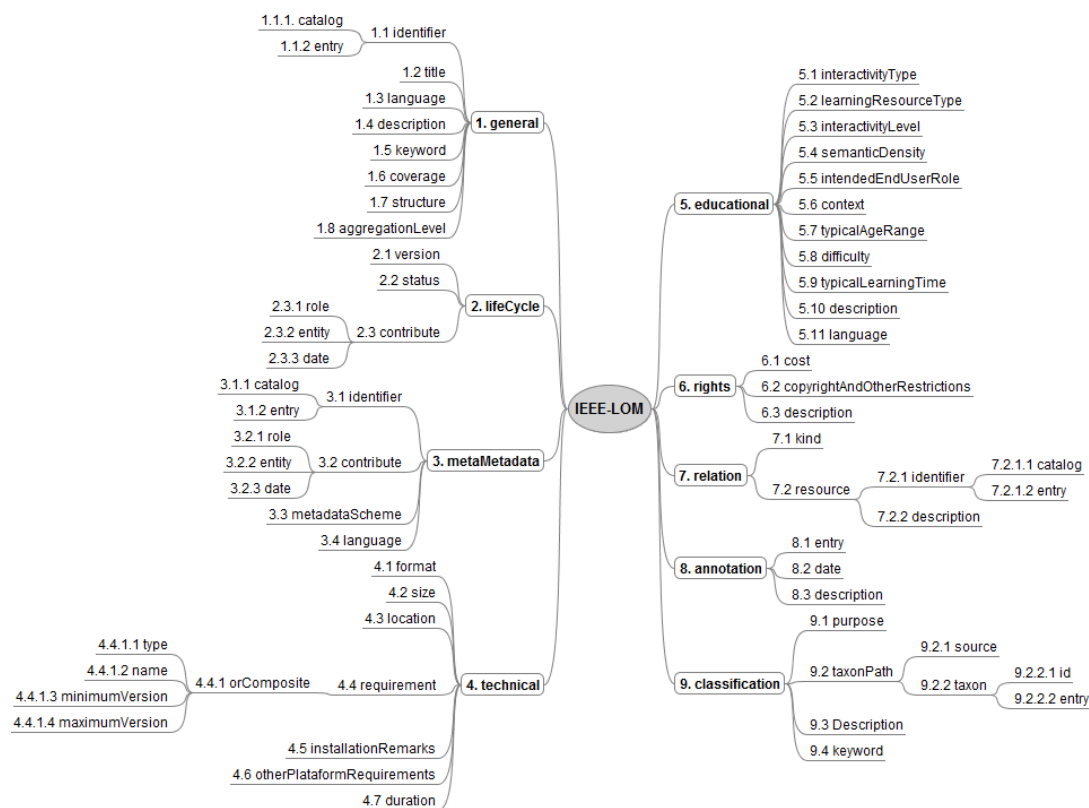


Figura 2.1 Esquema del estándar IEEE-LOM

EL estándar IEEE-LOM permite que las etiquetas puedan rellenarse con dos tipos de valores: los correspondientes a vocabularios controlados y los valores de texto libre. Las etiquetas se formalizan en un multiesquema XML que implementa la especificación, de forma que los metadatos de un OA se asocian creando una instancia XML del multiesquema definido. También proporcionan un mecanismo de adaptación de la especificación que se denomina “perfiles de aplicación” el cual debe cumplir las dos siguientes restricciones:

- Debe mantener los tipos de datos y espacios de valores de los elementos del esquema base.
- No puede definir nuevos tipos de datos ni espacios de valores para los elementos agregados.

Hoy en día se han desarrollado diversos perfiles de aplicación, en particular para países de Hispanoamérica destacando la extensión del estándar IEEE-LOM al idioma castellano

denominado LOM-ES (Blanco et al., 2008), en ella se mantienen las 9 categorías pero establece nuevos vocabularios controlados, traduce vocabularios ya existentes al castellano y hacen obligatorios campos de metadatos que en el estándar original fueron propuestos como opcionales.

De entre el resto de países, una de las iniciativas más activas es CanCore Learning Resource Metadata Initiative (CanCore LRM Initiative, 2000), que es gestionada por varias organizaciones de Canadá y se centra principalmente en establecer simplificaciones al estándar IEEE-LOM para facilitar las implementaciones en función del contexto de uso de los OA. También proporciona recomendaciones y directrices para facilitar la aplicación y la interoperabilidad de la iniciativa con otras especificaciones existentes. Un aspecto relevante de este perfil es que define los criterios para los cuales ciertos metadatos son más útiles que otros según la función realizada (búsqueda, visualización y catalogación automática o manual).

Otros ejemplos de estos perfiles de aplicación que incorporan nuevos metadatos o extienden los vocabularios existentes para adecuarse a las necesidades de países específicos son: *UK-LOM* (United Kingdom Learning Object Metadata Core) (CETIS, 2008), *KEM* (Korea Educational Metadata) (KERIS, 2004), *LRE* (European Schoolnet Learning Resource Exchange Metadata Application Profile) (EUN, 2007) y *ANZ-LOM* (Australia New Zeland Learning Object Metadata) (Education Services Australia, 2008).

2.2.5 ADL-SCORM

SCORM (Sharable Content Object Reference Mode) (ADL, 2004), es un estándar de la iniciativa ADL (Advanced Distributed Learning), orientado a compartir contenido de aprendizaje de una manera estandarizada. Su principal objetivo es facilitar la portabilidad de contenidos de aprendizaje de un LMS a otro, así como contribuir con la reusabilidad de los SCO (Sharable Content Object), denominación usada por SCORM para los OA.

Los SCO, cumplen con esta portabilidad entre distintos LMS a través de interfaces estandarizadas que proporciona SCORM. Se caracterizan por poseer un bajo nivel de granularidad, por lo que pueden ser administrados de mejor manera por los LMS, además

debido a su independencia del contexto de aprendizaje, pueden ser reusables en diferentes situaciones de aprendizaje (Bohl et al., 2002).

El estándar SCORM cuenta con tres componentes (ADL, 2004):

- 1) **Modelo de agregación de contenidos.** Proporciona métodos coherentes en materia de almacenamiento, identificación, condicionamiento de intercambios y recuperación de contenidos.
- 2) **Entorno de ejecución.** Describe las especificaciones que un sistema de gestión del aprendizaje debe implementar para administrar el entorno de ejecución con el contenido SCORM.
- 3) **Modelo de secuenciación y de navegación.** Permite una presentación dinámica del contenido. Describe cómo el sistema interpreta las reglas de secuenciación introducidas por un desarrollador de contenidos, así como los eventos de navegación realizados por el estudiante o por el sistema.

SCORM emplea al estándar IEEE-LOM concretamente incorpora 43 elementos para describir a los objetos de aprendizaje y proporciona especificaciones orientadas a la interoperabilidad entre distintas soluciones e-learning.

2.2.6 OTRAS ESPECIFICACIONES

Existen otros estándares, de entre los que cabe destacar el IMS-LD (IMS-Learning Design) (IMS Global Learning Consortium, 2003a), que es una especificación del IMS Global Consortium que apoya el uso de un amplio rango de pedagogías usadas en e-learning. Proporciona un lenguaje para representar la pedagogía usada en los cursos, apoya el cómo enseñar más que el qué enseñar. Facilita la reutilización de las prácticas pedagógicas más que el uso de contenidos. Se menciona a IMS-LD como la especificación que permite la incorporación de manera más efectiva de los aspectos instruccionales (Harper et al., 2005). Los aspectos importantes de IMS-LD son los siguientes:

- Ofrece soporte para múltiples alumnos, y contempla la comunicación entre ellos.
- Modela el rol del profesor.

- Permite combinar recursos educativos con actividades pedagógicas, y con las interacciones entre personas con diferentes roles.

Esta especificación es una de las pocas que considera el aspecto pedagógico en los OA, razón por la cual puede vincularse con otras especificaciones como *IMS-Content Packaging* (IMS Global Learning Consortium, 2001), la cual describe las estructuras de datos que pueden utilizarse para intercambiar datos entre sistemas que deseen importar, exportar, agregar y desagregar los paquetes de contenido; *IMS Question & Test Interoperability Specification* (IMS Global Learning Consortium, 2003b), permite representar preguntas y evaluaciones de los estudiantes e *IMS Reusable Definition of Competency or Educational Objective Specification* (IMS Global Learning Consortium, 2002), se enfoca en la representación de competencias educativas.

2.2.7 REPOSITORIOS DE OBJETOS DE APRENDIZAJE

La proliferación de sitios y repositorios dedicados a ofrecer recursos destinados a la educación es una prueba del constante desarrollo del e-learning (UNESCO, 2010). Es por ello que numerosas instituciones transforman y publican recursos educativos para que sean aprovechados y reciclados. Por consiguiente, estos espacios congregan a usuarios que comparten características e intereses.

La especificación IMS-DRI (IMS Digital Repositories Interoperability) (IMS Global Learning Consortium, 2003a), define a los repositorios digitales como "cualquier colección de recursos que son accesibles a través de una red sin el conocimiento previo de la estructura de la colección".

En los repositorios, los OA son agrupados y almacenados bajo dos planteamientos (IMS Global Learning Consortium, 2003a): los que contienen los objetos y los que solo contienen sus metadatos. En los primeros, los objetos y sus descriptores se encuentran dentro de un mismo sistema e incluso dentro de un mismo servidor. Por otra parte, se encuentran aquellos que contienen sólo los metadatos o los descriptores y se accede al objeto a través de una referencia a su ubicación física que se encuentra en otro sistema o repositorio de objetos.

Actualmente algunos de los principales repositorios de objetos de aprendizaje son:

- **MERLOT (Multimedia Educacional Resource for Learning and Online Teaching)** (Shell, 2002). Desarrollado y mantenido por la California State University Center for Distributed Learning (CSU-CDL). Es un repositorio de materiales de aprendizaje en línea centralizado que contiene sólo los metadatos y apunta a los objetos ubicados en sitios remotos. Ofrece un mecanismo de búsqueda avanzada de OA basada en: palabras clave, título, idioma, categoría, tipo de material, entre otros. También permite la búsqueda de usuarios con intereses comunes para establecer contactos. Con respecto a la recuperación de los resultados los ordena por diversos filtrados de información, tales como: relevancia (basada en valoraciones de los usuarios), título, autor, tipo de material, fecha de publicación y fecha de edición.
- **ARIADNE (Alliance of Remote Instructional Authoring & Distribution Networks for Europe)** (Duval et al., 2001). Es un proyecto de European Commission's Telematics for Education and Training Program, está orientado a la producción de contenido reutilizable para el aprendizaje. Contempla la descripción, almacenamiento, recuperación y uso en cursos estructurados. Incluye una colección de servicios que interactúan entre sí para ofrecer una red de repositorios que soportan búsquedas federadas de OA. Sin embargo, posee una búsqueda simple de objetos y limitaciones en cuanto a la comunicación y colaboración entre los usuarios del sistema. En cuanto a la recuperación de resultados, estos son ordenados por los siguientes filtrados de información: proveedor, idioma, formato, contexto y tipo de recurso.
- **MACE (Metadata for Architectural Contents in Europe)** (Stefaner et al., 2007). Es una iniciativa europea destinada a integrar repositorios de objetos de aprendizaje distribuidos por Europa. Posee una arquitectura que permite el acceso a los contenidos ofreciendo distintos tipos de metadatos y estructuras de clasificación. Ofrece componentes, llamados widgets, que permiten ejecutar un recurso de forma independiente o bien incrustarlo en una aplicación o un sistema Web. Incorpora una búsqueda simple y ordena los resultados obtenidos por diversos filtrados de información, tales como: fecha de publicación, más populares, orden alfabético y categoría.

- **AGORA (Ayuda para la Gestión de Objetos Reusables de Aprendizaje)** (Prieto, 2009). Es un proyecto hispanoamericano desarrollado por parte del grupo de investigación SMILe de la Universidad de Castilla-La Mancha, España. Permite almacenar los metadatos y los archivos de los recursos didácticos empleando un editor de metadatos y un asistente de generación de OA. Para aquellos recursos publicados en internet es posible importar los archivos directamente en el repositorio o bien mantener solo su referencia externa. Los recursos almacenados son catalogados utilizando el estándar IEEE-LOM y publicados para su descarga en el formato original como OA en formato SCORM o un enlace. La búsqueda y recuperación de objetos se realiza a través de tres etapas iterativas. Primero, la búsqueda se determina por medio de palabras clave en un campo de texto. Posteriormente, se establece un grado de similitud de metadatos con el patrón de la búsqueda. Finalmente, se presenta un listado de OA ordenados de acuerdo a su grado de similitud.

Después de una revisión general de todas las propuestas anteriores, cabe destacar que la mayoría de los repositorios de objetos de aprendizaje descritos se centran en la organización, indizado y métodos simples para la recuperación y ranking.

2.3 BÚSQUEDA Y RECOMENDACIÓN

Esta sección tiene como objetivo describir los diferentes conceptos, características y métodos que engloban la búsqueda y la recomendación de información en sistemas basados en la Web, con una particular orientación de su aplicación en entornos educativos.

2.3.1 CONCEPTOS Y DEFINICIONES

En los últimos años, la cantidad de información disponible en Internet ha crecido tan rápidamente que supera la capacidad de procesamiento de un humano (Zafra et al., 2009). La obtención de información en Internet es una actividad compleja. Encontrar la información apropiada, requerida por cada usuario, en la Web no es algo sencilla (Porcel et al., 2009). Es por ello, que el desarrollo de sistemas que permiten obtener información de Internet se ha incrementado de manera sustancial desde hace algunos años.

Las dos actividades principales que resuelven esta problemática son la *búsqueda* y la *recomendación*. Estas dos actividades se complementan, debido a que una es subsecuente de la otra. En este contexto, se considera a la *búsqueda*, aquella actividad donde un usuario a nivel particular busca, selecciona, descarga, califica y comenta los ítems que cree interesantes para él. De esta forma, los ítems finalmente mostrados pueden ordenarse teniendo en cuenta estos aspectos y convertir una tarea de búsqueda en una tarea de recomendación y personalización, que puede ayudar al usuario a encontrar los ítems que realmente más le puede interesar.

Por otra parte, la *recomendación* puede tener distintos objetivos, siendo la tarea más común la de recomendar al usuario un solo ítem de entre todos los ítems disponibles, además existen otras posibilidades más complejas tales como:

- Encontrar algunos ítems interesantes para el usuario, donde el objetivo es recomendar una lista ordenada de ítems.
- Encontrar todos los ítems interesantes, donde el objetivo es recomendar todos los ítems que puedan satisfacer al usuario.
- Realizar anotaciones en la lista de ítems, donde el objetivo es enfatizar algunos ítems con información del contexto.
- Proporcionar credibilidad, donde el objetivo es proporcionar explicaciones sobre por qué se le han recomendado esos ítems específicos.

Los *Sistemas de Recomendación* (SR), nacieron de los *Sistemas de Recuperación de Información* (SRI), pero se diferencian de estos en su personalización. Habitualmente, un SRI es definido como el proceso que trata la representación, almacenamiento, organización y acceso de elementos de información (Salton, 1986). Es decir, es un sistema capaz de almacenar, recuperar y mantener información (Kowalski, 1997). Para un usuario, el proceso de recuperación de información consiste en realizar una pregunta al sistema y obtener como respuesta un conjunto de documentos ordenados. Pero en todo sistema de recuperación de información es necesario la realización de una serie de pasos previos y diferenciados para poder llegar a sus respuestas, los más relevantes son (Olivas, 2011):

- 1) **Indexación.** El SRI crea un índice que contiene los términos que el sistema considera importantes (después de un preprocesado de cada documento) y su ubicación en los documentos.
- 2) **Consulta.** El usuario formula una pregunta al sistema, en un lenguaje procesable por éste (formalismo).
- 3) **Evaluación.** El sistema devuelve los resultados (documentos que satisfacen en cierto grado la demanda de información del usuario), ordenados según su relevancia con respecto a la consulta formulada.
- 4) **Retroalimentación del usuario.** El sistema opcionalmente aprende de las diferentes consultas de un usuario, focalizando la recuperación según este conocimiento adquirido.

Por otro lado, los sistemas de recomendación, tienen diversas definiciones. A continuación se presentan las más relevantes:

- Los SR, proporcionan sugerencias personalizadas sobre temas que los usuarios encontrarán interesantes. Habitualmente los SR requieren una interfaz de usuario que pueda inteligentemente determinar el interés de este y utilizar la información para hacer sugerencias (Claypool et al., 2001).
- Un SR, es aquel sistema que tiene como principal tarea seleccionar ciertos objetos de acuerdo a los requerimientos del usuario, dado que estos objetos están almacenados y caracterizados, con base en sus atributos (Wang, 1998).
- Los SR, se definen como aplicaciones que exploran las fuentes de información para sugerir productos y proporcionar a los usuarios información para facilitar sus procesos de toma de decisiones (Schafer, 2001).

Los SR tienen algunas características importantes que conviene resaltar, ya que los diferencia claramente de los sistemas de recuperación de la información (Hanani et al., 2001):

- Se pueden aplicar sobre datos sin estructurar o semi-estructurados, por ejemplo documentos Web o mensajes de correo electrónico.

- Están basados en perfiles de usuario, en lugar de que los usuarios expresen sus necesidades mediante consultas.
- Gestionan grandes cantidades de información.
- Trabajan fundamentalmente con información en modo texto.
- Eliminan información irrelevante del flujo de entrada.

Por otra parte, los SR están compuestos de (Burke, 2002):

- **Datos de los ítems.** Es la información que el sistema tiene antes de que el proceso de recomendación comience.
- **Datos del usuario.** Es la información que el usuario debe comunicar al sistema para generar una recomendación.
- **Algoritmos de recomendación.** Combinan los datos de los ítems y del usuario para hacer recomendaciones.

Todo sistema de recomendación contiene necesariamente dos elementos: el *usuario*, al que se le denomina usuario activo, y los *productos* que recomienda, a los cuales se les refiere de modo genérico como ítems. Los ítems que maneja el sistema deben almacenarse de algún modo para que éste pueda extraer las características de cada uno de ellos y utilizarlas en su algoritmo de recomendación. A continuación se describen tres formas de representar estos datos:

- **Estructurados.** El ejemplo más común para representar datos estructurados es en una base de datos relacional, que contiene atributos como: número de identificación, nombre y tipo, entre otros.
- **No estructurados.** Son aquellos datos sobre los cuales se ignora sobre su naturaleza a priori y que, en general, están escritos en lenguaje natural. Algunos ejemplos sobre datos no estructurados pueden ser las noticias que se han publicado hoy en un periódico, los comentarios de un usuario en una página Web, entre otros.
- **Semi-estructurados.** Este tipo de representación puede tener un conjunto de atributos que admiten valores booleanos (verdadero o falso), o incluso atributos que admiten texto libre en lenguaje natural. Un ejemplo podría ser un correo electrónico.

2.3.2 MÉTODOS DE FILTRADO DE INFORMACIÓN

Considerando los elementos de que disponen los sistemas de recomendación (datos de los ítems, datos del usuario y algoritmo de recomendación) se pueden distinguir varios métodos de filtrado de información (Burke, 2002):

- **Filtrado colaborativo.** Utiliza información de preferencias o valoraciones, de muchos usuarios para filtrar y recomendar productos a un usuario, ignorando la representación de los ítems. En el caso más simple, estos sistemas predicen las preferencias de un usuario como una suma ponderada de las preferencias de otros, en los cuales los pesos son proporcionales a las correlaciones sobre el conjunto de ítems comunes evaluados por dos personas (Goldberg et al., 1992). Para utilizar este filtrado se ha utilizado una variedad de técnicas de aprendizaje, incluidas las redes neuronales, indexación semántica latente y las redes bayesianas. Sin embargo, estos sistemas presentan algunos problemas, los cuales se describen a continuación:
 - *Primer evaluador (First-Rater).* Ocurre cuando un nuevo ítem es introducido en el catálogo y no se tiene ningún tipo de información acerca de él. Este ítem no podrá ser recomendado hasta que algún usuario activo especifique algunas de sus características o un conjunto de ítems similares.
 - *Arranque frío (Cold-Star).* Sucede cuando un usuario se da de alta en el sistema, y, como éste no conoce nada acerca de sus gustos e intereses, no se puede hacer ningún tipo de recomendación. Existe un problema similar cuando un usuario tiene intereses que no son comunes entre el resto de usuarios del sistema.
- **Filtrado basado en contenido.** Este tipo de filtrado analiza la descripción de los productos para identificar los elementos que son de interés particular para el usuario (Pazzani & Billsus, 2007). A pesar de ser una estrategia muy precisa, presenta varias limitaciones reconocidas que se describen a continuación:
 - Esta técnica suele recomendar productos excesivamente parecidos a los que el usuario activo ya conoce, e incluso, demasiado similares entre sí. En este último caso, si alguno de los productos sugeridos no interesa al usuario su

confianza en el sistema se verá comprometida, debido a que el resto de las sugerencias ofrecidas también se parecen a dicho producto.

- o Otra desventaja tiene su origen en la especialización de los atributos de los productos considerados, ya que es una tarea costosa que en algunos casos requiere la participación de un experto de conocimiento, capaz de describir el dominio del sistema de forma precisa.
- o La llegada de nuevos usuarios al sistema. En este escenario, el sistema de recomendación dispone, típicamente, de muy poco conocimiento sobre sus preferencias personales, de ahí la baja precisión de las recomendaciones ofrecidas.
- **Filtrado demográfico.** Clasifica a los usuarios en grupos demográficos basándose en atributos personales (edad, sexo, estado civil, ocupación profesional, aficiones, entre otros) y les proporcionan recomendaciones potencialmente interesantes para cualquier persona perteneciente a dicho grupo demográfico (Krulwich, 1997). El principal beneficio es que no requiere de un historial de valoraciones de los usuarios. Por el contrario, este tipo de recomendación tiene dos limitaciones principales:
 - o Puede conducir a recomendaciones demasiado generales e imprecisas para los usuarios, por considerar únicamente sus características demográficas.
 - o Además, este método no permite que las sugerencias ofrecidas se adapten a posibles cambios en las preferencias de los usuarios, dado que sus datos personales suelen permanecer invariables a lo largo del tiempo.
- **Filtrado basado en conocimiento.** Realiza recomendaciones partiendo del conocimiento que proporciona el propio usuario sobre sus necesidades, y del conocimiento de los productos a recomendar, buscando los que mejor se adapten a las necesidades de los usuarios (Burke, 1999). El perfil puede ser cualquier estructura de datos que soporte esta relación. Algunas de sus desventajas son:
 - o Aunque requiera de poca cantidad de información para proporcionar recomendaciones sobre un producto en particular al usuario, para la extracción del conocimiento es necesario implementar un sistema basado en reglas de decisión.

- o Las recomendaciones no son personalizadas, es decir, son del tipo estático para los usuarios que posean las mismas necesidades se proporcionarán las mismas recomendaciones.
- **Filtrado híbrido.** Este tipo de filtrado surgió con el objetivo de solventar algunos problemas presentados por los filtrados anteriores ante algunas situaciones. Por lo tanto, un sistema de recomendación se considera híbrido, si combina al menos dos filtrados de los mencionados anteriormente (Basu et al., 1998). El enfoque más frecuente es combinar filtrado colaborativo con el basado en contenido, a fin de destacar sus puntos fuertes. De acuerdo con Burke (Burke, 2002), se propone una clasificación analítica de sistemas híbridos el cual combina pares de algoritmos de recomendación:
 - o *Ponderado (Weighted)*. La puntuación (o votos) de un ítem recomendado se calcula con los resultados de todas las técnicas de recomendación presentes en el sistema. La combinación más simple híbrida sería una combinación lineal de recomendación de calificaciones.
 - o *Conmutación (Switching)*. Una conmutación híbrida utiliza un criterio para cambiar entre técnicas de recomendación. La conmutación híbrida introduce complejidad adicional dentro del proceso de recomendación desde los criterios de conmutación que debe determinar, y esto introduce otro nivel de parametrización.
 - o *Mixta (Mixed)*. Las recomendaciones de diferentes sistemas son presentados al mismo tiempo. Se utiliza, cuando es práctico hacer un gran número de recomendaciones de forma simultánea.

2.3.3 ALGORITMOS DE RECOMENDACIÓN

Existe una variedad de algoritmos que se han aplicado a los sistemas de recomendación. Cada uno tiene sus rasgos específicos, ya sea para implementarse a un determinado enfoque o para manejar una cantidad determinada de información de usuarios e ítems. En general, los algoritmos de recomendación deben encontrar la preferencia del usuario, que puede expresarse de dos formas (Sarwar et al., 2001):

- **Predicción.** Consiste en un valor numérico $p(a, j)$ que predice la preferencia del usuario activo (U_a) hacia el ítem (I_j) no usado aun por U_a .
- **Recomendación.** Contiene una lista de los N primeros ítems, que el usuario U_a posiblemente prefiere, no usados aun por el usuario U_a , (los top- N).

Algunos algoritmos están basados en el usuario y usan estadísticas para buscar el conjunto de los usuarios más similares al U_a , posteriormente, combinan las preferencias de esos vecinos para producir una predicción o recomendación (top- N) para U_a . Otros algoritmos están basados en los ítems y suponen que un usuario estaría interesado en ítems similares a los que le interesaron antes a ese mismo usuario, y buscan entre el conjunto de ítems que el usuario ha usado y jerarquizado, calculan la similaridad con el ítem objetivo, y seleccionan los k más similares. También se les llama “basados en el modelo”, pues desarrollan primero un modelo de las opiniones de los usuarios, usando técnicas de aprendizaje automático.

A continuación se describen algunos de los algoritmos de código abierto más populares en los sistemas de recomendación:

- **Ripper.** Es un algoritmo de inducción de reglas que trabaja de forma similar a los algoritmos que construyen árboles de decisión: “particionando” recursivamente el conjunto de datos. Su eficacia se debe a un algoritmo de poda que optimiza el conjunto de reglas que se obtienen a partir de un determinado conjunto de datos. Además, soporta atributos multi-valuados. Esta característica puede ser muy interesante cuando se trabaja con textos semi-estructurados (Cohen, 1995).
- **Rocchio.** Este algoritmo implementa un enfoque basado en búsquedas y tiene como objetivo optimizar las consultas de manera incremental. Para ello, el sistema permite al usuario que valore los ítems que el sistema le devuelve. Esta valoración se realiza en términos de lo interesantes que le resultan estos ítems al usuario en referencia a la consulta que hizo. Gracias a esta información suministrada por los usuarios, el algoritmo puede modificar la consulta inicial de las búsquedas ya sea estableciendo pesos a los términos, o añadiendo y eliminando términos de la búsqueda para ofrecer ítems más interesantes (Rocchio, 1971).
- **Slope One.** Este algoritmo es de los más utilizados, gracias a su eficiencia y facilidad de implementación. Combina las opiniones y gustos de los distintos

usuarios para ofrecer recomendaciones personalizadas al usuario activo. Existen dos variantes: basado en usuario, el cual calcula la similitud entre usuarios; y basado en ítems, el cual opera sobre valoraciones de los ítems y estadísticas de uso (Lemire, 2005) .

- **Easyrec.** Proporciona recomendaciones personalizadas usando el servicio Web RESTful para integrarse dentro de aplicaciones Web. Está escrito en Java, utiliza una base de datos MySQL y viene con una herramienta de administración. A diferencia de los algoritmos anteriores, es una aplicación lista para usar, incluyendo funciones de integración y administración (Studio Smart Agent Technologies, 2010).

2.3.4 APLICACIONES EN ENTORNOS EDUCATIVOS

Los sistemas de recomendación se han aplicado con éxito en diferentes dominios, tales como entretenimiento, contenidos, servicios, comercio electrónico o e-learning (Ricci et al., 2011). En e-learning por ejemplo los sistemas de recomendación se han utilizado para personalizar el proceso de aprendizaje y ajuste a los conocimientos previos del usuario, capacidades y preferencias (Manouselis et al., 2011). Otros ejemplos de las tareas de personalización en los sistemas de e-learning mediante sistemas de recomendación son la interacción adaptativa, la ejecución de cursos adaptativos, el descubrimiento de contenidos personalizados, el soporte colaborativo adaptativo y la búsqueda de recursos de aprendizaje.

En los últimos años, los sistemas de recomendación también se han utilizado para ayudar en la búsqueda de objetos de aprendizaje en los repositorios, puesto que son capaces de seleccionar, de forma automática y personalizada, el recurso que mejor se adapta a las preferencias o necesidades de cada usuario (Casali et al., 2011). A continuación se describen algunos trabajos sobre esta línea de investigación:

Uno de los primeros intentos para desarrollar un sistema de recomendación orientado a objetos de aprendizaje, es la propuesta *RACOFI (Rule-Appling Collaborative Filtering)* (Anderson et al., 2003). Esta aplicación es el resultado de dos sistemas integrados: COFI (Collaborative Filtering) y RALOCA (Rule Applying Learning Object Comparison Object). RACOFI combina dos enfoques de recomendación mediante la integración de un motor de filtrado colaborativo que trabaja con las calificaciones que los usuarios proporcionan a los

OA. Posteriormente a través de técnicas de minería de datos se generan reglas de asociación que determinan las futuras recomendaciones a los usuarios.

Otra propuesta interesante es *RecoSearch* (Fiaidhi, 2004), es un modelo que combina las técnicas de filtrado colaborativo y por contenido. Aplica una infraestructura colaborativa para la publicación, búsqueda, recomendación y presentación de objetos de aprendizaje enfocados a la enseñanza del lenguaje de programación en Java. Para este proceso intervienen dos motores de filtrado especializados que trabajan simultáneamente denominados CollabroSearch y CollabroRecommender.

El sistema *Altered Vista* (Walker et al., 2004), propone un modelo para la recomendación de objetos de aprendizaje basados en filtrado colaborativo. Este sistema está especialmente dirigido a profesores y estudiantes de cursos que se imparten totalmente en línea. El objetivo principal es analizar las evaluaciones de los OA realizadas por los alumnos para posteriormente realizar las recomendaciones.

Otro sistema colaborativo es *CYCLADES* (Avancini & Straccia, 2005), este sistema ofrece una amplia gama de funciones enfocadas para los usuarios que desean buscar y navegar a través de bibliotecas digitales, favorece la creación de comunidades agrupándolos de acuerdo a sus intereses similares. Su metodología de recomendación incorpora algoritmos de agrupamiento (Expectation Maximisation) y de clasificación (redes bayesianas).

El sistema *QSLA* (*Questions Sharing and Interactive Assignments*) (Rafaeli et al., 2004), se enfoca en el contexto de las comunidades en línea con el fin de aprovechar el punto de vista social en el aprendizaje y promover la colaboración, la recomendación en línea y la posterior formación de las comunidades de aprendices. Su arquitectura incorpora el filtrado colaborativo y está basada en actividades relacionadas con el conocimiento tales como: generación, compartir, evaluación y gestión.

Un enfoque similar es el propuesto por *Tang* (Tang & McCalla, 2005), es un sistema que agrupa a sus usuarios de acuerdo a sus intereses comunes, por otra parte, las recomendaciones de los recursos educativos se basan en función de las consultas realizadas en el pasado. La arquitectura de este sistema incorpora diversos módulos tales como: filtrado

colaborativo, agrupamiento, mantenimiento de los recursos educativos y modelado de los estudiantes.

El sistema *LORM (Learning Object Recommendation Model)* (Tsai et al., 2006), fue desarrollado para recuperar y recomendar objetos de aprendizaje para los estudiantes. Adopta un enfoque ontológico para realizar descubrimientos semánticos basados en las preferencias de los alumnos expresadas hacia los objetos. El mecanismo de esta herramienta es un método híbrido de recomendación que conjuga dos algoritmos, el primero calcula las calificaciones del alumno hacia los objetos, el segundo establece una correlación de experiencias que han expresado los alumnos con determinados objetos. Por último, las dos puntuaciones se sumarán a la puntuación de recomendación.

Una propuesta que explora una nueva forma de obtener una calificación de la calidad de los OA es la desarrollada por *Kumar* (KUMAR et al., 2007). Este sistema utiliza las redes bayesianas de confianza para superar la insuficiencia y la falta de críticas realizadas a los objetos. Contiene un enfoque híbrido (filtrados de contenido y colaborativo) que implementa un modelo basado en las cadenas de Markov, para verificar las reglas vigentes sobre la calidad de los OA y determina si se han alcanzado todas las variables en el modelo de evaluación de la calidad.

La propuesta de *Al-Khalifa* (Al-Khalifa, 2008), se basa en un repositorio de objetos de aprendizaje árabe con capacidades de recomendación denominado “Marifah”. Dicho repositorio ha integrado características avanzadas de búsqueda que no suelen incluir los motores de búsqueda convencionales. Su entorno permite a los usuarios realizar actividades de gestión con los recursos tales como: publicar, editar y eliminar. También permite evaluar la calidad de los recursos y consultar las estadísticas que acumulan cada uno.

Ruiz-Iniesta (Ruiz-Iniesta et al., 2009), propone un enfoque proactivo para la recomendación de OA que se adapta al perfil del estudiante. Para ello, implementa una ontología basada en temáticas de lenguaje de programación, la cual sirve como índice para organizar los objetos en el repositorio. Por otra parte, el perfil de los estudiantes almacena información sobre su historial de navegación. El método incorpora un enfoque híbrido que combina aspectos sociales y de contenido.

La propuesta de *Bozo* (Bozo et al., 2010), presenta un enfoque de recomendación para la búsqueda de objetos de aprendizaje orientada a los profesores. La metodología incorpora dos métodos de filtrados de información (colaborativo y por contenido) y un modelo conceptual del plan de estudios utilizado en el sistema educativo chileno. Como fuentes de información incorpora: información asociada a los OA (metadatos, uso, evaluaciones y estadísticas) y los perfiles de los profesores.

Manouselis (Manouselis et al., 2010), describe un estudio piloto europeo denominado *CELEBRATE*, cuyo objetivo es permitir a los profesores y alumnos localizar fácilmente, utilizar y reutilizar tanto contenido de código abierto, así como el contenido de proveedores comerciales. Para ello, el entorno de simulación CollaFiS, permite parametrizar, ejecutar y evaluar todas las variaciones consideradas del algoritmo empleado. Esta propuesta muestra una visión general de los sistemas de recomendación seleccionados para el aprendizaje de recursos y estudios relacionados con las evaluaciones.

El sistema *LRMDCR* (*Learner's Role based Multidimensional Collaborative Recommendation*) (Wan & Okamoto, 2011). Utiliza un modelo basado en las cadenas de Markov, para dividir el grupo de alumnos avanzados y principiantes mediante el uso de las actividades y los procesos de aprendizaje. Implementa un esquema de cálculo basado en el filtrado colaborativo multidimensional para decidir sobre los objetos de aprendizaje recomendadas para cada alumno del grupo.

Finalmente, *Casali* (Casali et al., 2011), propone un sistema para ayudar a los usuarios a encontrar los recursos educativos más apropiados de acuerdo a sus necesidades y preferencias. Para ello, utiliza una búsqueda federada de repositorios y la recomendación utiliza los metadatos asociados a los OA y el perfil del usuario que realiza la consulta. Incorpora una arquitectura multiagente que incluye varios tipos de agentes con diferentes funcionalidades.

Después de una revisión general de todas las propuestas anteriores, se destacan los siguientes puntos:

- Menos de la mitad de las propuestas (seis de catorce) son sistemas completos: Altered Vista, CYCLADES, QSIA, Casali, Tang y Manouselis.

- Sólo seis propuestas fueron evaluadas mediante el uso de sus propios repositorios o bases de datos: RACOFI, Altered Vista, CYCLADES, Tang, Al-Khalifa y Manouselis.
- Ocho propuestas fueron experiencias piloto con usuarios humanos: Altered Vista, CYCLADES, QSIA, Tang, LRMDCR, Casali, Ruiz-Iniesta y Manouselis.

Por otro lado, un análisis más profundo sobre estas propuestas revela las limitaciones de cada una posee. La Tabla 2.2 proporciona una visión general de las características principales de los trabajos anteriores. Las características de la tabla comparativa se describen a continuación:

- 1) **Enfoque híbrido.** Combina al menos dos enfoques, tales como las recomendaciones de colaboración, basadas en contenido, demográficas, basadas en conocimiento, entre otros.
- 2) **Búsqueda avanzada.** Permite no sólo el uso de palabras clave, sino también valores de los metadatos para la búsqueda de los OA.
- 3) **Criterios de filtrado.** Contiene algunos de los criterios de filtrado para la búsqueda de los OA. Estos filtros permiten priorizar la búsqueda de los usuarios de acuerdo a sus intereses particulares.
- 4) **Calificación de los objetos de aprendizaje.** Muestra la calificación final obtenida para cada una de los OA recomendados.
- 5) **Estadísticas de los objetos de aprendizaje.** Muestra algunas estadísticas asociadas a los OA recomendados, como el número de descargas, el número de evaluaciones, la evaluación promedio, entre otros.
- 6) **Explicación.** Muestra una explicación acerca de por qué ese objeto de aprendizaje específico ha sido recomendado y no otros.
- 7) **Evaluación.** Proporciona un instrumento para evaluar los objetos de aprendizaje.

Tabla 2.2 Comparación de los sistemas de recomendación en educación

Sistemas	1	2	3	4	5	6	7
RACOFI	No	SI	SI	SI	No	No	SI
RecoSearch	SI	SI	No	SI	No	No	SI
Altered Vista	No	SI	No	No	SI	No	SI
CYCLADES	No	SI	SI	SI	No	No	SI
QSIA	No	SI	No	SI	No	No	SI
Tang	No	No	No	No	No	No	SI
LORM	SI	No	No	SI	No	No	SI
Kumar	SI	No	No	SI	No	No	SI
Al-Khalifa	No	No	No	SI	SI	No	SI
LRMDCR	SI	No	SI	SI	No	No	SI
Ruiz-Iniesta	SI	No	No	No	No	No	SI
Bozo	SI	SI	SI	SI	No	No	SI
Manouselis	No	SI	SI	SI	SI	No	SI
Casali	SI	SI	SI	No	No	No	SI

Como se muestra en la Tabla 2.2, las características que la mayoría de los sistemas comparados contienen son: búsqueda avanzada, calificaciones y evaluación de los OA. Por otra parte, más de la mitad de ellos (ocho) utiliza un enfoque híbrido. Además, ninguno cumple con todas las características, que será uno de los objetivos de esta tesis.

2.4 COLABORACIÓN EN LA BÚSQUEDA Y LA RECOMENDACIÓN

La búsqueda en la Web suele ser vista como una actividad solitaria, a pesar de que millones de usuarios utilizan servicios como Google, Yahoo, entre otros. Esta tendencia ha empezado a cambiar en años recientes, con propuestas de modelos de búsqueda de colaborativa en el que grupos de usuarios pueden cooperar para buscar con mayor eficacia. En esta sección se describen conceptos y ejemplos de sistemas que se basan en la colaboración para la mejora en la búsqueda y recomendación.

2.4.1 COLABORACIÓN

La *colaboración* puede ser vista como la combinación de la cooperación, comunicación y la coordinación. Donde, la *cooperación* es la acción conjunta durante una sesión dentro de un espacio de trabajo compartido, donde los miembros de un grupo producen y organizan la información tales como documentos, gráficos, entre otros. Por otra parte, la *comunicación* está relacionada con el intercambio de mensajes e información entre las personas. Finalmente, la *coordinación* está relacionada con la gestión de las personas, sus actividades y recursos y puede ser vista como la unión de la comunicación y la cooperación (Fuks et al., 2005). Estos conceptos conforman el modelo 3C (Ellis et al., 1991), cuya filosofía inspira a la metodología implementada en la *Búsqueda colaborativa de objetos de aprendizaje* que se va a proponer en esta tesis y se explica a detalle en el epígrafe 3.2.

El flujo de trabajo del modelo 3C muestra la naturaleza iterativa de la colaboración, donde los miembros de un grupo obtienen una retroalimentación de sus propias acciones y las de sus compañeros a través de una base de conocimiento relacionado entre los participantes (ver Figura 2.2):

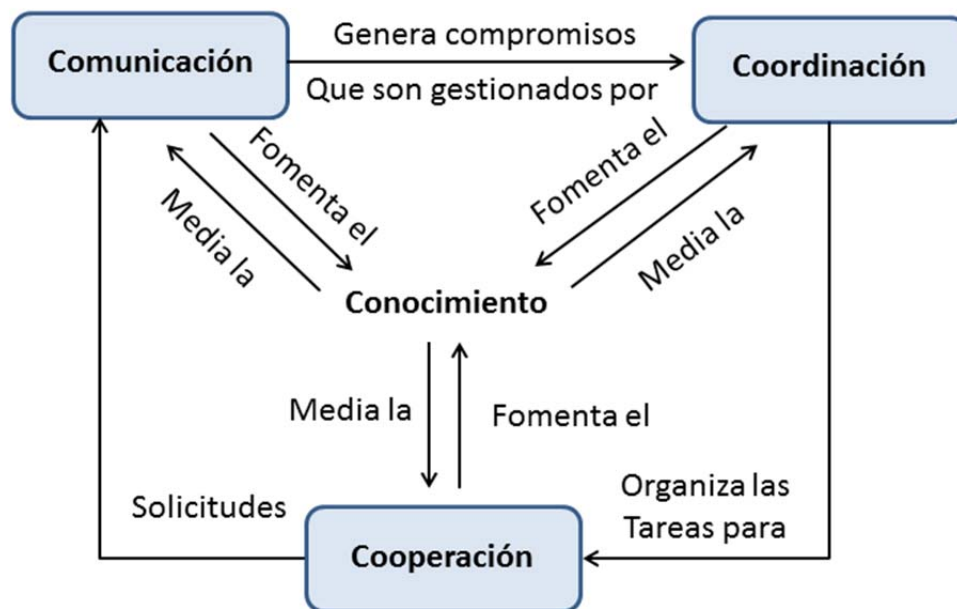


Figura 2.2 Flujo de trabajo del modelo de colaboración 3C

2.4.2 BÚSQUEDA SOCIAL COLABORATIVA

La *Búsqueda Social Colaborativa* (BSC) está enfocada en facilitar a los miembros de un mismo grupo realizar las actividades de búsqueda, selección, valoración, etiquetado y comentario de los ítems de su interés a través de una interfaz común al grupo. Es un enfoque basado en una comunidad de búsqueda en la Web que apoya el intercambio de resultados de las selecciones anteriores entre un grupo de investigadores relacionados con el fin de personalizar las listas de resultados y reflejar las preferencias de la comunidad en su conjunto (Coyle & Smyth, 2008).

Actualmente la BSC se ha implementado en los motores de búsqueda genéricos, y se dividen en dos tipos de colaboración: implícita y explícita. A continuación se describen y se presentan ejemplos de estos sistemas:

- **Los motores de búsqueda implícitos.** Se caracterizan por inferir necesidades similares de información. Dos ejemplos de estos sistemas, ambos de distribución libre son:

- o *Jumper 2.0* (Jumper Networks Inc., 2011). Orientado a los usuarios que comparten una base de datos, permite realizar búsquedas de forma colaborativa a través de marcadores y etiquetas de conocimiento.
 - o *Seeks* (Seeks Project, 2011). Implementa una arquitectura descentralizada peer-to-peer permitiendo a los usuarios personalizar y compartir sus resultados preferidos en una búsqueda de forma privada.
- **Los motores de búsqueda explícitos.** Se caracterizan porque los usuarios comparten un sistema de información y han acordado que necesitan trabajar juntos para este objetivo. Algunos ejemplos de estos sistemas son:
 - o *SearchTogether* (Morris & Horvitz, 2007). Es un plug-in para Internet Explorer creado por Microsoft, el cual permite realizar búsquedas colaborativas en dicho navegador. Para ello, requiere iniciar sesión con Windows Live y crear una “sesión de búsqueda” invitando a nuestros contactos de Windows Live Messenger. Este plug-in permite interactuar con las personas que se haya invitado directamente en el navegador. Otra característica es que las búsquedas pueden ser realizadas a través de varios buscadores, compartiendo los resultados entre los miembros del grupo. Además, muestra las búsquedas previas realizadas permitiendo marcar un resultado para que lo vea el resto, e incluso comentar y evaluar dichos resultados.
 - o *Fischlár-DiamondTouch* (Smeaton et al., 2005). Este motor de búsqueda utiliza una combinación de búsqueda de texto, que opera sobre el texto reconocido automáticamente del habla, y una búsqueda de imágenes que utiliza las características de bajo nivel de la imagen que se compara con los fotogramas clave del vídeo. Los dos aspectos más novedosos de esta propuesta son: que implementa una evaluación colaborativa basada en la detección de grupos de usuarios que trabajan juntos y el diseño de una interfaz adaptada a dispositivos móviles.
 - o *Cerchiamo* (Golovchinsky et al., 2008). Permite a grupos de usuarios explorar y evaluar documentos de forma asíncrona. Su motor de búsqueda implementa tres interfaces independientes: usuario, reguladora y del

algoritmo. Los componentes del sistema se comunican a través de la interfaz de usuario que posee diferentes configuraciones de visualización (pantalla completa y compartida).

- o CoSearch (Amershi & Morris, 2008). Es un sistema orientado a las búsquedas colaborativas que realizan los usuarios a través de dispositivos móviles. En cuanto a la interfaz de usuario incorpora diversos elementos para que los miembros del grupo se comuniquen entre sí, tales como: etiquetas de colores para categorizar las búsquedas realizadas, un área para redactar comentarios y una sección de notas que se adjuntan al documento localizado.
- o *Heystaks* (Smyth et al., 2009). Es un plug-in que se integra al motor de búsqueda de Google. Su idea se enfoca en cómo la reutilización de la experiencia puede apoyar a la búsqueda colaborativa de información. El escenario que plantea para que los miembros del grupo se comuniquen es a través de los resultados que devuelve Google. A los que se añaden opciones colaborativas representadas a través de iconos. Entonces, el usuario puede evaluar, comentar y etiquetar los resultados. Posteriormente acceden a un sitio Web privado donde podrán visualizar la actividad de los otros miembros del grupo.

Cabe destacar que no se han encontrado sistemas enfocados en facilitar la búsqueda social colaborativa de objetos de aprendizaje dentro de un entorno educativo.

2.4.3 RECOMENDACIONES DE GRUPO

El objetivo de la recomendación de grupo consiste en calcular una puntuación de recomendación para cada ítem que refleje los intereses y preferencias de todos los miembros del grupo. En general, los miembros del grupo no siempre tienen los mismos gustos y por consenso una puntuación para cada ítem debe ser cuidadosamente diseñado (Amer-yahia et al., 2009). Es por ello, que recomendar a los grupos es aún más complicado que recomendar individualmente (Masthoff, 2011).

Los grupos de usuarios normalmente surgen por una necesidad compartida, sin embargo no siempre se conforman por iniciativa propia. Actualmente existen herramientas que facilitan

y promueven su conformación, y que identifican al menos a cuatro tipos de grupos (Boratto & Carta, 2010):

- 1) **Establecido.** Un número de personas que expresamente optan por ser parte de un grupo, debido a que comparten intereses a largo plazo.
- 2) **Ocasional.** Un número de personas que hacen algo de vez en cuando juntos. Sus miembros tienen un objetivo común en un momento determinado.
- 3) **Aleatorio.** Un número de personas que comparten un entorno en un determinado momento, sin intereses explícitos que los vinculan.
- 4) **Automáticamente identificado.** Son grupos detectados automáticamente teniendo en cuenta las preferencias de los usuarios y /o los recursos disponibles.

Los componentes clave para una interfaz de recomendación del grupo son: una sección para solicitar ítems específicos y otra, para la presentación de las recomendaciones devueltas. Se proponen tres modelos de organización de la información disponible cuando se muestran las recomendaciones (Connor et al., 2001):

- 1) **De solo grupo.** Estas interfaces evitan información que revela las preferencias de los otros miembros del grupo y previene que se comparen calificaciones.
- 2) **Compuesta.** Despliega una lista de ítems recomendados por los miembros del grupo y usuarios particulares. Permite comparar sus opiniones para determinar los ítems más adecuados para los intereses del grupo.
- 3) **Individual centrada.** Muestra las preferencias de los otros miembros del grupo y permite comparar las calificaciones realizadas a los ítems para facilitar la toma de decisiones.

En la actualidad existen una serie de métodos de decisión en grupo o de agregación de votaciones (Masthoff, 2011) que establecen diferentes formas de cómo un grupo de personas pueden llegar a un consenso (ver Tabla 2.3).

Tabla 2.3 Resumen de métodos de agregación

Método	Descripción
Voto pluralista (Plurality Voting)	Es un sistema de valoraciones de un solo ganador, donde el ítem con mayor número de valoraciones es elegido.
Promedio (Average)	Promedio de las valoraciones individuales.
Multiplicativo (Multiplicative)	Multiplica las valoraciones individuales.
Cuenta Borda (Borda Count)	Cuentan los puntos de los ítems rankeados en las listas de preferencias individuales con la opción inferior recibiendo 0 puntos, el siguiente recibe 1 punto y así sucesivamente.
Regla Copeland (Copeland Rule)	Contabiliza la frecuencia de un ítem cuando se asocia con otros ítems (utilizando la mayoría de las valoraciones) y el número de veces que se pierde.
Votación por aprobación (Approval Voting)	Contabiliza las valoraciones individuales para un ítem por encima de un umbral de aprobación.
Valor mínimo (Least Misery)	Selecciona el mínimo de las valoraciones individuales.
Mejor valorado (Most Pleasure)	Selecciona la máxima de las valoraciones individuales.
Promedio descartando el valor mínimo (Average without Misery)	Promedio de las valoraciones individuales, después de excluir los ítems con valoraciones individuales por debajo de un cierto umbral.
Equitativo (Fairness)	Los ítems son rankeados como si los miembros los escogieran a cada uno de ellos.
La persona más respetada (Most respected Person)	Utiliza la valoración del miembro más respetado.

De acuerdo con Gartrell (Gartrell et al., 2010), se han identificado una serie de descriptores asociados a los grupos que ayudan a inferir el proceso de la toma de decisiones, los cuales son:

- **El nivel del contacto social.** Representa el nivel de contacto social que mantiene con los miembros que conforman el grupo.
- **Experiencia.** Representa el nivel de experiencia de los miembros del grupo.
- **Disimilaridad.** Este valor representa la diferencia de calificaciones que emiten a un mismo ítem entre los miembros del grupo.

A continuación se describen algunos sistemas de recomendaciones de grupo desarrollados para diferentes entornos:

El sistema propuesto por Ankur (Ankur et al., 2004), implementa un algoritmo de agrupamiento para la identificación automática de grupos sociales de profesores y alumnos denominado Group-Place Identification. Para ello, utiliza los rastros de movilidad de la comunidad a través de cualquier tecnología de localización proporcionada a través de un dispositivo móvil.

El método desarrollado por Yu (Yu et al., 2005), propone una recomendación basada en el contenido multimedia a grupos de usuarios que viajan en autobuses, trenes, aviones, entre otros. Implementa un modelo de espacio vectorial para calcular el grado de relevancia entre el material multimedia y los perfiles de usuario.

El sistema de Campos (Campos et al., 2008), está enfocado para la recomendación automática de una lista clasificada de nuevos ítems a un grupo de usuarios. Para ello, implementa una metodología basada en las redes bayesianas para modelar las incertidumbres que aparecen tanto en las interacciones entre los miembros del grupo y los procesos que conducen a la elección final o recomendación.

Chen (Chen et al., 2008), propone un sistema que incorpora un enfoque colaborativo, el cual implementa un algoritmo genético para predecir las posibles interacciones entre los miembros del grupo, con el objetivo de calcular correctamente la calificación que puedan otorgarle a un ítem.

La propuesta de Boratto (Boratto et al., 2009), implementa un algoritmo que detecta los grupos de usuarios cuyas preferencias de películas son similares y predice las recomendaciones para tales grupos. El algoritmo trabaja en cuatro pasos: detecta la similitud

de preferencia entre dos usuarios, detecta comunidades de usuarios, predice las calificaciones de los ítems que recomienda a los usuarios y finalmente extiende las predicciones a otros ítems similares.

Gartrell (Gartrell et al., 2010), desarrolla un sistema que predice las preferencias de películas para un grupo de usuarios. Para ello, basa su arquitectura en tres componentes del grupo: sus descriptores, las estrategias de decisión de los usuarios y las reglas de asociación basadas de las calificaciones emitidas a los ítems.

Después de una revisión general de todas las iniciativas anteriores, es importante destacar la ausencia de propuestas orientadas a la recomendación de objetos de aprendizaje para grupos de usuarios dentro de entornos educativos.

2.5 DESCUBRIMIENTO DE CONOCIMIENTO Y MINERÍA DE DATOS

El objetivo de esta sección es describir los diferentes conceptos y procesos involucrados en el descubrimiento de conocimiento, así como también, abordar temas relacionados de la minería de datos y su aplicación en entornos educativos, específicamente en repositorios de objetos de aprendizaje.

2.5.1 DEFINICIONES Y TAXONOMÍA

El descubrimiento de conocimiento implica diversos conceptos con los que está relacionado, tales como *los Datos*, *la Información*, *la Gestión de la información* y *la gestión del conocimiento*. *Los datos* consisten en hechos, imágenes, o sonidos. Cuando se combinan con la interpretación y el significado, se convierte en información. Por lo tanto, *la información* consiste en datos que han sido formateados, filtrados, y resumidos (Chen, 2001). El proceso de la transformación de datos en información se representa en el siguiente esquema (Itmazi, 2005):

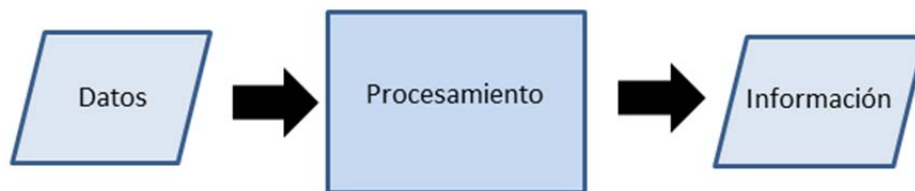


Figura 2.3 Transformación de datos en información.

Una vez establecidos los conceptos datos e información, se puede definir del término de *Gestión de la Información* (GI), como “el proceso que utiliza herramientas y técnicas para capturar, recoger, analizar, organizar, manipular y recuperar la información. De la misma manera, controla y explota los recursos de la información de una organización” (Itmazi, 2005).

Por otro lado, la *Gestión del Conocimiento* (GC), se define como una disciplina emergente cuyo objetivo es generar, compartir y utilizar el conocimiento tácito y explícito existente en un determinado espacio para dar respuesta a las necesidades de los individuos y de las comunidades en su desarrollo (Alavi et al., 2001).

Finalmente el *Descubrimiento del Conocimiento en Bases de Datos* (en inglés, Knowledge Discovery in Databases, KDD), se define como “el proceso no trivial de identificar patrones válidos, novedosos y potencialmente útiles y en última instancia, comprensible a partir de los datos” (Fayyad et al., 1996).

El KDD es un proceso iterativo e interactivo y de acuerdo con Hernández (Hernández et al., 2004) su taxonomía se puede organizar en cinco fases (ver Figura 2.4):

- **Integración y recopilación.** En esta actividad se determinan las fuentes de información que pueden ser útiles y dónde conseguirlas. Posteriormente, se transforman los datos a un formato común. Dado que los datos provienen de diferentes fuentes, pueden contener valores erróneos o faltantes.
- **Selección, limpieza y transformación.** Se eliminan o corrigen los datos incorrectos y se decide la estrategia a seguir con los datos incompletos. Además, se proyectan los datos para conseguir únicamente aquellas variables o atributos que van a ser

relevantes, con el objetivo de hacer más fácil la tarea propia de minería y para que los resultados de la misma sean más útiles.

- **Minería de datos.** Se decide cuál es la tarea a realizar (clasificar, agrupar, entre otras) y se elige el método a utilizar.
- **Evaluación e interpretación.** Los expertos evalúan y se analizan los patrones, y si es necesario se vuelve a las fases anteriores para una nueva iteración. Esto incluye resolver posibles conflictos con el conocimiento que se disponía anteriormente.
- **Difusión y uso.** Se hace uso del nuevo conocimiento y se hace partícipe de él a todos los posibles usuarios. Este conocimiento se suele utilizar en procesos de toma de decisiones.

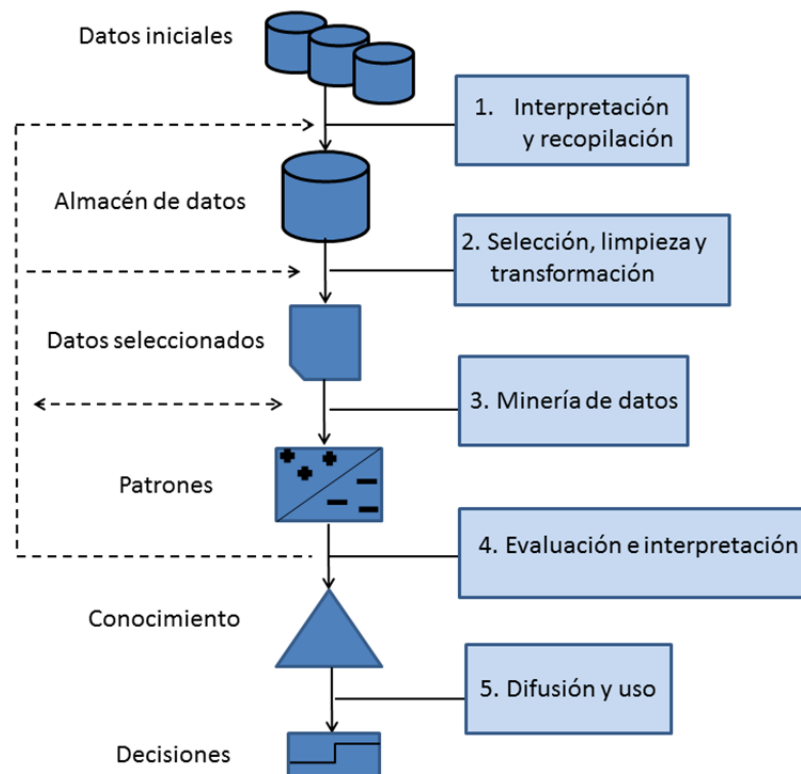


Figura 2.4 Etapas del proceso KDD

Como se observa en la Figura 2.4 la *minería de datos* es una etapa crucial de las técnicas incluidas en el proceso KDD. Se define como “el proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos” (Fayyad et al., 1996).

2.5.2 MINERÍA DE DATOS

La minería de datos, trabaja con grandes volúmenes de datos, procedentes en su mayoría de sistemas de información, con los problemas que ello conlleva (ruido, datos ausentes, intratabilidad, volatilidad de los datos, entre otros), y aplica técnicas adecuadas para analizar estos datos y extraer conocimiento novedoso y útil (Hernández et al., 2004).

De acuerdo con (Hernández et al., 2004), los modelos de aprendizaje empleados en la minería de datos pueden dividirse en dos categorías principales: predictivos y descriptivos. Cada una de estas categorías contiene una taxonomía la cual se describe (ver Figura 2.5):

- **Modelos predictivos o aprendizaje supervisado.** Predicen el valor del atributo de un conjunto de datos a partir de información previamente conocida. El objetivo de este modelo es proporcionar características que describan las relaciones entre los datos. Los algoritmos aplicados en los modelos predictivos se catalogan en:
 - *Clasificación.* Son utilizados para clasificar datos, con la finalidad de predecir clases de objetos cuyas categorizaciones no se han definido.
 - *Regresión.* Predicen una o más variables continuas, como por ejemplo las pérdidas o los beneficios, basándose en otros atributos del conjunto de datos.
 - *Predicción.* Son aquellos que pueden ser usados para predecir tendencias de los valores en los datos o de las clases.
- **Modelos descriptivos o aprendizaje no supervisado.** Descubren patrones y tendencias en los datos. Sirven para llevar a cabo acciones y obtener un beneficio o conocimiento de ellas. Estos algoritmos se clasifican en:
 - *Agrupamiento.* Son aquellos que no requieren una clasificación predefinida para particionar los datos obteniendo el conocimiento de acuerdo a las características de los mismos.
 - *Asociación.* Se basan en el descubrimiento de reglas de asociación que muestran condiciones en los valores de los atributos que ocurren simultáneamente de forma frecuente en un determinado conjunto de datos.
 - *Correlación y dependencias.* Establecen un patrón en el que uno o más atributos determinan el valor de otro. Su función principal es predecir

diferentes valores de los datos, como ejemplo: ganancias, ventas, tasas, entre otros.

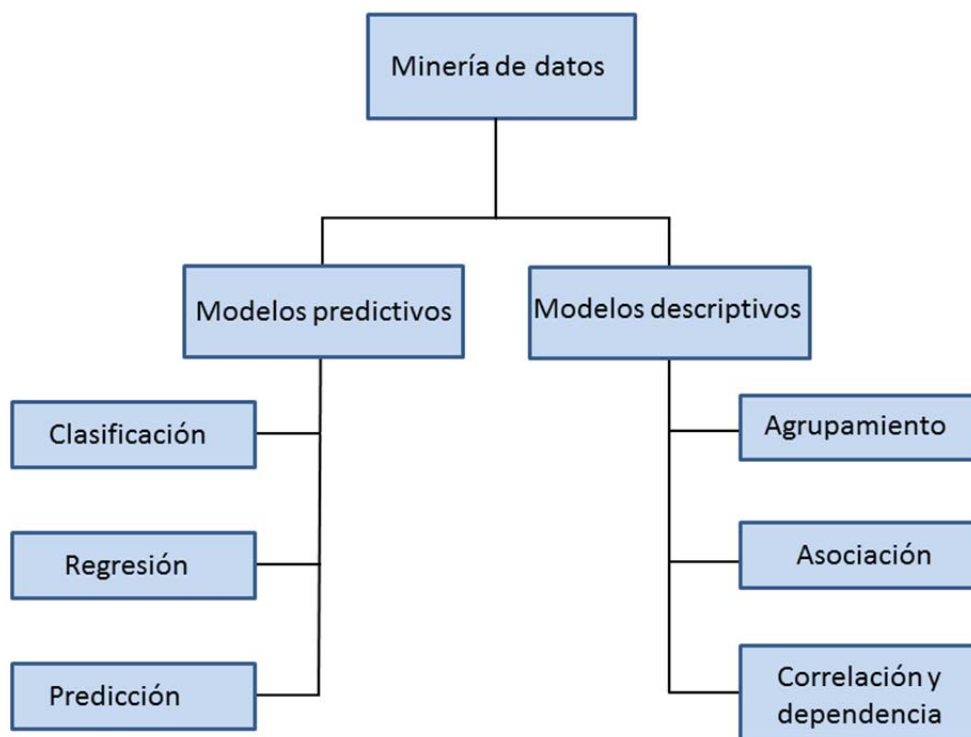


Figura 2.5 Taxonomía de los modelos de aprendizaje empleados en la minería de datos

En la minería de datos existen multitud de técnicas que pueden resolver una gran diversidad de problemas. A continuación se presenta un resumen de las técnicas más populares:

- **Algebraicas y estadísticas.** Se basan, generalmente, en expresar modelos y patrones mediante fórmulas algebraicas, funciones lineales, funciones no lineales, distribuciones o valores agregados estadísticos tales como medias, varianzas, correlaciones, entre otras. Algunos de los algoritmos más conocidos dentro de este grupo de técnicas son la regresión lineal (Edwards, 1976) y la regresión logística (Hosmer, 2000).
- **Redes bayesianas.** Se basan en estimar la probabilidad de pertenencia (a una clase o grupo), mediante la estimación de las probabilidades condicionales inversas o a priori, utilizando el teorema de Bayes. Poseen varias ventajas, una de ellas es que pueden tratar muchos atributos y son muy robustas al ruido que pueden provocar los datos. Por el contrario, la expresividad es limitada y depende de la discretización.

Algunos algoritmos populares son EM (Dempster et al., 1977) y Naive Bayes (H. Zhang, 2004).

- **Algoritmos evolutivos.** Son métodos sistemáticos para la resolución de problemas de búsqueda y optimización, a los cuales se aplican los mismos métodos de la evolución biológica, tales como: selección basada en la población, reproducción sexual y mutación (Eiben & Smith, 2003).
- **Conteos de frecuencias y tablas de contingencia.** Se basan en contar la frecuencia en la que dos o más sucesos se presenten conjuntamente. Ejemplos de estos algoritmos son Apriori (Agrawal et al., 1993) y Predictive A priori (Scheffer, 2004).
- **Árboles de decisión.** Está técnica se encuentra dentro de una metodología de aprendizaje supervisado. Su representación es en forma de árbol, en donde cada nodo es una decisión, los cuales a su vez generan reglas para la clasificación de un conjunto de datos. Entre sus ventajas se encuentran: que admiten atributos discretos y continuos, también se desempeñan de forma eficiente con los atributos no significativos. Entre sus desventajas, se encuentran la sobreadaptación y el ruido generado por los datos. Algunos ejemplos de estos algoritmos son ID3 (Quinlan, 2007), J48 (Ye, 2006) y C4.5 (Quinlan, 1993).
- **Relacionales, declarativas y estructurales.** La característica principal de este conjunto de técnicas es que representan los modelos mediante lenguajes declarativos, como los lenguajes lógicos, funcionales o lógico-funcionales. Las técnicas de ILP (programación lógica inductiva) son las más representativas y las que han dado nombre a un conjunto de técnicas denominadas minería de datos relacional.
- **Redes neuronales artificiales.** Se utilizan de manera frecuente para detectar categorías comunes en los datos, debido a que detectan y aprenden complejos patrones de comportamiento. Sus principales ventajas son: cuando están bien ajustadas, obtienen precisiones muy altas y se puede aplicar tanto en modelos predictivos como descriptivos. Su principal desventaja son los datos incompletos. Algunos ejemplos son las redes de kohonen (Kohonen, 1988) y el modelo de Retro-Propagación (Rumelhart et al., 1986).

- **Basadas en núcleo y máquinas de soporte vectorial.** Se trata de técnicas que intentan maximizar el margen entre los grupos o las clases formadas. Existen muchas variantes, dependiendo del núcleo utilizado y la manera de trabajar con el margen.
- **Estocásticas y borrosas.** Este conjunto de técnicas contienen características que se describen por separado. Por una parte, las técnicas estocásticas se basan en las teorías del análisis de probabilidades. Contienen datos con alguna función de densidad de probabilidad asociada. Por otra parte, las técnicas borrosas, aunque también se basan en la asunción de un conocimiento impreciso, son distintas de las estocásticas. En vez de tener una función de densidad de probabilidad asociada, se caracterizan por una función de pertenencia que expresa el grado en que un elemento pertenece a un conjunto. Estas técnicas junto con las redes neuronales y los algoritmos evolutivos, forman lo que se denomina computación flexible.
- **Basadas en casos, en densidad o distancia.** Son métodos que se basan en distancias al resto de elementos, ya sea directamente, como los vecinos más próximos o mediante la estimación de funciones de densidad. Algunos algoritmos conocidos son el K-medias (Macqueen, 1967) y COBWEB (Fisher, 1987).

Además de todo lo anterior existen diversas combinaciones que dificultan aún más realizar una taxonomía razonable que contemple a todos los grupos de técnicas.

Por otro lado, existe una variedad de paquetes software para realizar minería de datos. Cada uno de ellos posee características apropiadas para realizar determinadas tareas o para analizar cierto tipo de datos. A continuación se describen algunos ejemplos:

- **WEKA (Waikato Enviroment for Knowledge Analysis)** (Witten, 2011), es una herramienta visual de libre distribución desarrollada por los investigadores de la Universidad de Waikato en Nueva Zelanda. Está implementado en el lenguaje de programación Java y opera en diferentes entornos de sistemas operativos. Algunas de sus principales características son: acceso a datos, preprocesado de datos, modelos de aprendizaje y contiene cuatro entornos.
- **KEEL (Knowledge Extraction based on Evolutionary Learning)** (Alcalá-Fdez et al., 2008), fue desarrollado por 5 grupos de Universidades españolas, permite utilizar y construir diferentes modelos para minería de datos y tiene como característica

importante la inclusión de una librería de algoritmos de aprendizaje evolutivo con código abierto en Java. Las características principales son: algoritmos de preprocesamiento, librería de algoritmos de extracción de conocimiento, librería de herramientas estadísticas para el análisis de los algoritmos y aplicación para su uso vía Web; además contiene cuatro entornos.

- **DBMiner** (DBMiner Technology Inc., 2003), es un software de distribución libre desarrollado por la Universidad de Simon Fraser, Canadá. Está concebido para la extracción del conocimiento en bases de datos relacionales, almacenes de datos y Web, su arquitectura de diseño incorpora OLAP (online analytic processing) y OLAM (online analytic mining).
- **RapidMiner** (Universidad de Dortmund, 2001). Este sistema antes era conocido como YALE (Yet Another Learning Environment), está desarrollado en Java y es de código abierto. Cuenta con una interfaz gráfica fácil de utilizar y además integra un mecanismo sencillo para desarrollar extensiones que hacen posible integrar nuevos operadores y con ello adaptar el software para los requerimientos personales.

2.5.3 META-APRENDIZAJE

Los métodos del meta-aprendizaje están dirigidos a la detección automática de modelos de datos interesantes. Estos métodos pertenecen a una rama del aprendizaje automático que intenta sustituir a los expertos humanos involucrados en el proceso de minería de datos para crear varios modelos computacionales de aprendizaje a partir de los datos. Los nuevos datos y la descripción de los objetivos de los sistemas de meta-aprendizaje deben apoyar la toma de decisiones en la clasificación, regresión, las tareas de asociación y/o proporcionar modelos de datos comprensibles (Duch, 2011).

Los métodos del meta-aprendizaje suelen implementarse para seleccionar un modelo o algoritmo de entre un conjunto para un problema concreto, también se han empleado, para la selección de los parámetros de un algoritmo que mejor funciona con un determinado tipo de datos.

Para la comprensión de los datos, se requiere de los siguientes pasos: preprocesamiento, selección de la información relevante, transformaciones y métodos de clasificación. Sin un

conocimiento a priori sobre un determinado problema para encontrar una secuencia óptima de las transformaciones representa un gran desafío. De acuerdo al teorema “No-free lunch” (Duda et al., 2001), ningún método puede superar a todos los demás en todas las situaciones. La mejor recomendación es probar diferentes enfoques y transferir los conocimientos obtenidos mediante la resolución de otros problemas.

Las técnicas de meta-aprendizaje ayudan a seleccionar o crear modelos de predicción óptimos y reutilizan la experiencia previa desde el análisis de otros problemas, facilitan a los humanos la mayoría de los trabajos y a lograr el objetivo de los programas informáticos que mejoran con la experiencia (Brazdil et al., 2009) (Jankowski et al., 2011). Estos métodos están diseñados para automatizar las decisiones necesarias para la aplicación de técnicas de aprendizaje computacional.

El objetivo del meta-aprendizaje es recomendar el mejor método para determinados datos, que constan de elementos descritos por algunas características. Esto requiere la caracterización del conjunto de datos por meta-características, capacidad para estimar la similitud de los datos nuevos a los ya analizados del conjunto de datos, y una base de datos de los resultados obtenidos por diversos métodos sobre estos conjuntos de datos.

El papel de los sistemas de meta-aprendizaje en los problemas de aprendizaje supervisado, es recomendar el uso de métodos que ofrezcan el mejor resultado en un conjunto de datos similar. Este enfoque se ha utilizado en varios proyectos (Brazdil et al., 2009), algunos ejemplos son:

- **StatLog** (Michie et al., 1994), compara el rendimiento de un conjunto fijo de algoritmos sobre algunas bases de datos. El objetivo es la generación automática de reglas basadas en las características de los datos que guíen la selección del clasificador.
- **METAL** (Giraud-Carrier et al., 2004), desarrolla un prototipo de sistema que ayuda a los usuarios con la selección del modelo y el método de combinación, también los guía a través de los experimentos. Esta orientación no se limita solo a la selección de un método apropiado, sino que también sugiere pasos de transformación de datos que a menudo son cruciales para obtener buenos resultados.

- **MLC++** (Kohavi & Karl, 1994), es una librería enfocada a la selección de modelos. El objetivo es aplicar todos los algoritmos posibles al problema actual, y la selección final se hará tomando en cuenta el rendimiento de cada uno de los algoritmos.

El gran desafío en la minería de datos es crear sistemas flexibles de aprendizaje que pueden extraer información relevante de datos, transferencia de conocimiento de la experiencia pasada, y volverlos a configurar para encontrar diferentes soluciones interesantes para una tarea determinada (Duch, 2007). En lugar de un algoritmo de aprendizaje único diseñado para resolver el problema especializado, deben establecerse prioridades para definir lo que hace que una solución sea interesante, y la búsqueda de configuraciones de módulos computacionales que crean automáticamente los algoritmos que respondan a una demanda específica.

Esta búsqueda de todos los posibles modelos puede verse limitada por las prioridades de los usuarios y debe guiarse por la experiencia basados en la solución de problemas de índole similar. Sin conocimiento previo acerca de un problema dado, encontrar una secuencia óptima de transformaciones puede no ser posible, tiene que ser aprendida mediante el análisis de patrones de flujos de trabajo que han tenido éxito para muchos tipos de modelos en varios tipos de datos.

De acuerdo con (Duch, 2011), para el meta-aprendizaje no es importante que una combinación específica de características demuestre tener éxito en alguna tarea, pero es importante que una transformación específica de un subconjunto de características sea de utilidad para los datos de este tipo, o que la distribución de los patrones en el espacio tenga algunas características que puedan ser descritas por cierto modelo de datos específico y por lo tanto sea fácil de adaptar a un nuevo conjunto de datos de tipo similar. Dicha información permite la generalización de los conocimientos en el ámbito de la búsqueda de patrones para una nueva composición de transformaciones, facilitando la transferencia de conocimientos entre diferentes tareas.

La aplicación de los métodos de meta-aprendizaje va a cambiar el futuro de la minería de datos, permitiendo a los usuarios con poca experiencia en los métodos de cálculo, realizar un análisis profundo de los datos que pueda conducir a importantes descubrimientos.

2.5.4 MINERÍA DE DATOS EDUCATIVA

La Minería de Datos en Educación o Minería de Datos Educativa (Educational Data Mining, EDM), es la aplicación de técnicas de minería de datos a información generada en los entornos educativos. La EDM se define como “ el proceso de transformar los datos en bruto recopilados por los sistemas de enseñanza en información útil que pueda utilizarse para tomar decisiones informadas y responder preguntas de investigación” (Heiner et al., 2006).

La aplicación de minería de datos en sistemas educativos es un ciclo iterativo de formación de hipótesis, pruebas y refinamiento. El conocimiento descubierto, una vez filtrado, sirve de guía, facilita y mejora el proceso de aprendizaje a través de la toma de decisiones. Como se observa en la Figura 2.6, los educadores, investigadores y responsables académicos son los encargados de diseñar, planificar, crear y mantener los sistemas educacionales y los alumnos usan e interactúan con él (Romero & Ventura, 2007).

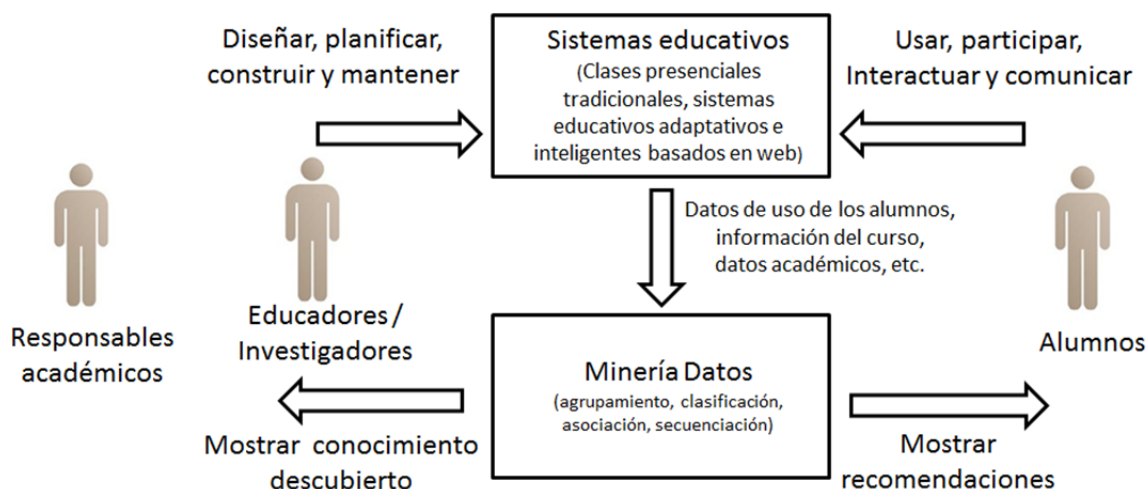


Figura 2.6 Ciclo de aplicación de minería de datos en entornos educativos

La aplicación de técnicas de EDM se puede ver desde distintos puntos de vista dependiendo del tipo de usuario (Romero & Ventura, 2010) (Romero & Ventura, 2013):

- **Del estudiante.** El objetivo es ayudar o realizar recomendaciones a los alumnos durante su interacción con el sistema de e-learning para mejorar su aprendizaje. Sus principales aplicaciones son: sugerir buenas experiencias de aprendizaje a los estudiantes, adaptar el curso según el progreso del aprendiz, ayudar a los estudiantes

dando sugerencias y atajos, recomendar caminos más cortos y personalizados, entre otros.

- **Del profesor.** El objetivo es apoyar a los profesores de los sistemas de e-learning para que puedan mejorar el funcionamiento o rendimiento de estos sistemas a partir de la información generada por los alumnos al utilizar los cursos. Sus principales aplicaciones son: obtener una mayor realimentación de la enseñanza, conocer más sobre cómo los estudiantes aprenden en la Web, evaluar a los estudiantes por sus patrones de navegación, reestructurar los contenidos en el sitio Web para personalizar los cursos, clasificar a los estudiantes en grupos, entre otros.
- **De los investigadores.** El objetivo es comparar técnicas de minería de datos que permitan recomendar la más útil para una tarea específica educativa o problema, para evaluar la efectividad del aprendizaje utilizando diferentes configuraciones y métodos, y así sucesivamente.
- **De las instituciones educativas.** El objetivo es obtener información útil que permita mejorar la eficiencia del sitio y adaptarlo a la conducta de los usuarios (tamaño óptimo del servidor, distribución del tráfico de red, entre otros). Otros objetivos son adquirir medidas que permitan organizar mejor los recursos institucionales (humanos y materiales) y determinar la eficiencia de los programas educativos del aprendizaje a distancia soportado por ordenador.

A continuación se describen algunos ejemplos de trabajos donde se han aplicado con éxito técnicas de EDM en sistemas de recomendación:

Zhang (Zhang et al., 2008), propone un Sistema de Instrucción Personalizada de Recomendación (PIRS, en inglés), el cual está diseñado para el aprendizaje basado en Web. Este sistema reconoce los diferentes patrones de estilo de aprendizaje y hábitos de uso de Internet a través de pruebas aplicadas a los estudiantes. Para ello, implementan reglas de asociación a través del algoritmo Apriori a sus registros de navegación Web.

Un enfoque similar es el propuesto por Ba-Omar (Ba-omar et al., 2007), es un sistema centrado en la identificación de patrones de aprendizaje de los estudiantes y la secuencia de elección de los recursos educativos (audio, texto, entre otros) en relación con sus estilos de aprendizaje. Para detectar estos patrones utilizaron el algoritmo Apriori.

Romero (Romero et al., 2009), propone una arquitectura avanzada cuyo objetivo es realizar recomendaciones a los estudiantes los sitios o páginas Web más apropiados cuando visitan un sistema educativo adaptativo basado en Web. Estas recomendaciones se basan de las reglas descubiertas por diversos algoritmos de minería Web.

La propuesta de Rištofič (Rištofič, 2005), se enfoca a la recomendación de lecciones (objetos de aprendizaje o conceptos) que los estudiantes deben estudiar la próxima vez que utilicen un sistema educativo adaptativo basado en Web. Para descubrir los patrones de comportamiento se implementaron reglas de asociación.

García (García et al., 2007), propone un sistema de recomendación basado en técnicas de minería de datos que ayude al profesor a detectar posibles problemas en la estructura y contenidos de los cursos, basado en la propia información que brindan los datos de utilización de los cursos por los alumnos.

La propuesta desarrollada por Wang (Wang & Shao, 2004), establece un modelo de recomendación para estudiantes, el cual está basado en intervalos de tiempo específico de las sesiones de navegación realizadas por cada usuario. Posteriormente, se integran grupos de usuarios con perfiles de navegación similares y se ejecutan técnicas de agrupamiento y reglas de asociación.

Vialardi (Vialardi et al., 2009), propone un sistema para predecir qué tan conveniente es para un estudiante matricularse a un curso específico. Para ello, se utilizan los resultados de estudiantes con un perfil similar que han llevado el curso. Para descubrir los patrones de comportamiento se implementó el algoritmo C4.5.

Finalmente, la propuesta de Guo (Guo & Zhang, 2009), es un sistema de recomendación orientado al aprendizaje personalizado y adaptativo de los estudiantes basado en el desempeño de las actividades realizadas en el pasado. Para ello, utilizan redes neuronales y árboles de decisión.

Después de una revisión general de todas las propuestas anteriores, se destaca que ninguna de ellas ha aplicado meta-aprendizaje dentro de un sistema de recomendación.

2.6 CONCLUSIONES

En este capítulo se presentaron los temas que sustentan la línea de investigación de esta tesis. Se ha llevado a cabo una revisión del estado del arte de temas afines a la línea de investigación y que destacan por su aportación al desarrollo de sistemas de recomendación orientados a entornos educativos.

Un tema central de esta tesis es la descripción de las técnicas de filtrado de información y su taxonomía. Destaca la comparativa realizada entre 14 sistemas de recomendación de objetos de aprendizaje (Tabla 2.2), en la cual se muestra que ninguno presenta todas las características comparadas. También se ha estudiado el enfoque de la *Búsqueda social colaborativa* para la recomendación a un grupo de usuarios. Cabe destacar que no se han encontrado sistemas de recomendación colaborativos y de grupo para los objetos de aprendizaje. Finalmente, con respecto a la aplicación de EDM a los sistemas de recomendación, cabe destacar que no se han utilizado métodos de meta-aprendizaje.

En el siguiente capítulo se describe una propuesta de *modelo de recomendación híbrido para Objetos de Aprendizaje* orientado a la búsqueda individual y social colaborativa.

3 MODELO DE RECOMENDACIÓN HÍBRIDO DE OBJETOS DE APRENDIZAJE INDIVIDUAL Y COLABORATIVO

En este capítulo se propone un *Modelo de Recomendación Híbrido de Objetos de Aprendizaje Individual y Colaborativo (MRH-OA-IC)* que utiliza los filtrados de información por contenido, colaborativo y demográfico para mejorar los mecanismos de selección de los objetos recuperados y facilita las actividades de recomendación, colaboración y etiquetado social. De hecho, el MRH-OA-IC se orienta tanto una *búsqueda individual* de un sólo usuario como una *búsqueda colaborativa* de un grupo de usuarios. Para cada una, se ha establecido una concepción diferente de su funcionalidad que se describe de forma detallada en este capítulo.

3.1 BÚSQUEDA INDIVIDUAL

La Búsqueda individual es aquella en la que un usuario a nivel particular realiza la petición, selecciona, descarga y evalúa los OA que cree interesantes para él. En la realización de la *búsqueda individual* se proponen dos modalidades: *simple* y *compuesta*. Cada una se realiza de forma independiente y a través de métodos que involucran métricas y filtrados de información que se ajusten a las necesidades del solicitante.

3.1.1 BÚSQUEDA INDIVIDUAL SIMPLE

Esta búsqueda se basa en el contenido de los OA y los metadatos que pueden suministrar durante las consultas. La *búsqueda individual simple* considera tres grupos de procesos principales: *gestión del objeto patrón*, *búsqueda-recuperación* y finalmente el *ranking* (ver Figura 3.1). El objetivo es recuperar un conjunto de OA ordenado según una medida de relevancia.

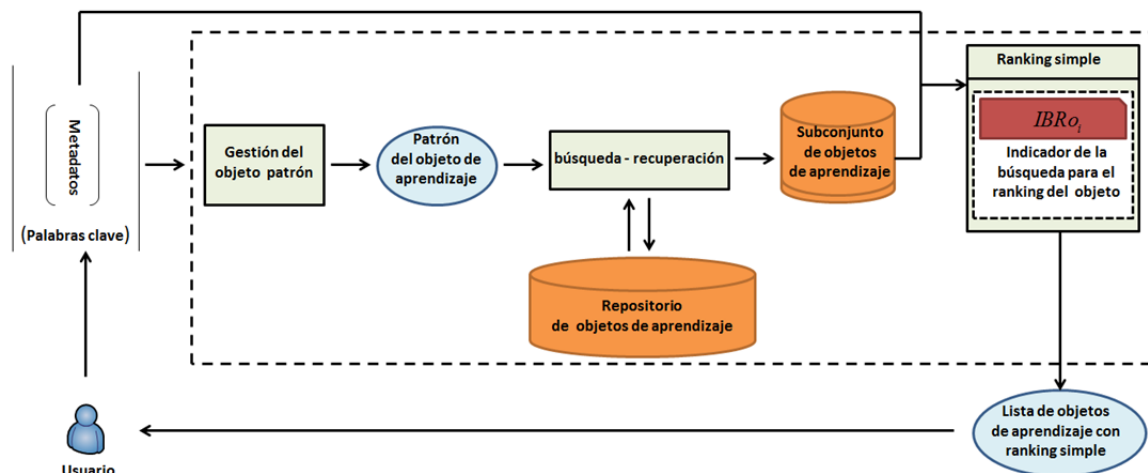


Figura 3.1 Método de búsqueda individual simple

Como se observa en la Figura 3.1, el área punteada comprende los tres procesos de la *búsqueda individual simple*, que también forman parte del método de la *búsqueda individual compuesta* que será explicado en el epígrafe 3.1.2.

La generación del *objeto patrón* es un proceso en el que un usuario proporciona palabras clave y los valores operacionales de varios metadatos de los 68 definidos en el estándar IEEE-LOM (IEEE-LTSC, 2002). A partir de esos valores, el proceso inicial debe ocuparse de generar el *objeto patrón* ya con la estructura de metadatos IEEE-LOM que es la misma que tienen los OA del repositorio.

El proceso de *búsqueda-recuperación* en el repositorio, proporciona un *subconjunto de objetos de aprendizaje* que coincide con la especificación del *objeto patrón*. Por último, el proceso de *ranking* utiliza una medida para dar orden en el subconjunto recuperado de OA.

A los efectos de este trabajo, se ha adoptado una medida concreta inspirada en el índice de Jaccard (Hamers et al., 1989) y modificada por Menéndez (Menéndez & Zapata et al., 2011). Esta métrica calcula el grado de similitud ($\text{SimMeta}_{\text{texto}}$) entre los metadatos de tipo cadena o LangString (16 de los 68 definidos en el estándar) asociados a cada uno de los elementos del *subconjunto de objetos de aprendizaje* (O_x) y las palabras clave contenidas en el *objeto patrón* (O_y). El objetivo es establecer la proporción de términos relevantes que son comunes entre ambos objetos. Esta medida se muestra en la Fórmula 3.1 con la que se obtienen los indicadores de similitud entre las palabras clave y los textos contenidos en cada

metadato:

$$\text{SimMeta}_{\text{texto}}(O_x, O_y) = \frac{\sum_{k=1}^n \sum_{l=1}^m (\text{Sim}(v_{xk}, v_{yl}) \cdot (\frac{f_{xk}}{f_x} \cdot \frac{f_{yl}}{f_y}))}{n \cdot m} \quad (3.1)$$

Donde:

- n, m = Representan la cantidad de términos relevantes que tienen los textos v_x y v_y .
- v_{xk}, v_{yl} = Son términos relevantes de v_x y v_y .
- $\text{sim}(v_{xk}, v_{yl})$ = Devuelve un valor entre $[0,1]$ que representa la similitud semántica de dos términos.
- f_{xk}, f_{yl} = Son la frecuencia de aparición de los términos v_{xk} y v_{yl} en v_x y v_y .
- f_x, f_y = Son la frecuencia total de términos en v_x y v_y .

A partir de esos indicadores $\text{SimMeta}_{\text{texto}}(O_x, O_y)$ se obtiene el *Indicador de la Búsqueda para el Ranking del Objeto (IBRo_i)* cuyo valor varía de $[0,1]$ y se calcula a través de la Fórmula 3.2:

$$\text{IBRo}_i = \frac{\sum [\text{SimMeta}_{\text{texto}}(O_x, O_y)]}{M} \quad (3.2)$$

Donde:

- M = Número total de metadatos a comparar.
- $\text{SimMeta}_{\text{texto}}(O_x, O_y)$ = Porcentaje de similitud obtenido de un metadato.

Finalmente, se puede mostrar al usuario una *lista de objetos de aprendizaje* ordenados según sus grados de similitud sintáctica respecto de los términos de la petición.

3.1.2 BÚSQUEDA INDIVIDUAL COMPUESTA

Se basa en los filtrados de información por contenido, colaborativo y demográfico, lo que significa que incorpora un enfoque de recomendación híbrido. La *búsqueda individual compuesta* al igual que la simple comprende tres procesos: *gestión del objeto patrón*, *búsqueda-recuperación* y la conformación del *ranking* (ver Figura 3.2). Por lo que permite

gestionar una petición pero basada en palabras clave, valores de metadatos y criterios de filtrado de la información, donde cada criterio viene acompañado de un peso ponderado, cuya función es potenciar o disminuir las preferencias del solicitante según sea el caso. De esta forma los objetos de aprendizaje recuperados pueden ordenarse teniendo en cuenta diferentes aspectos y transformar el proceso de la *búsqueda-recuperación* en un proceso de *recomendación híbrido*.

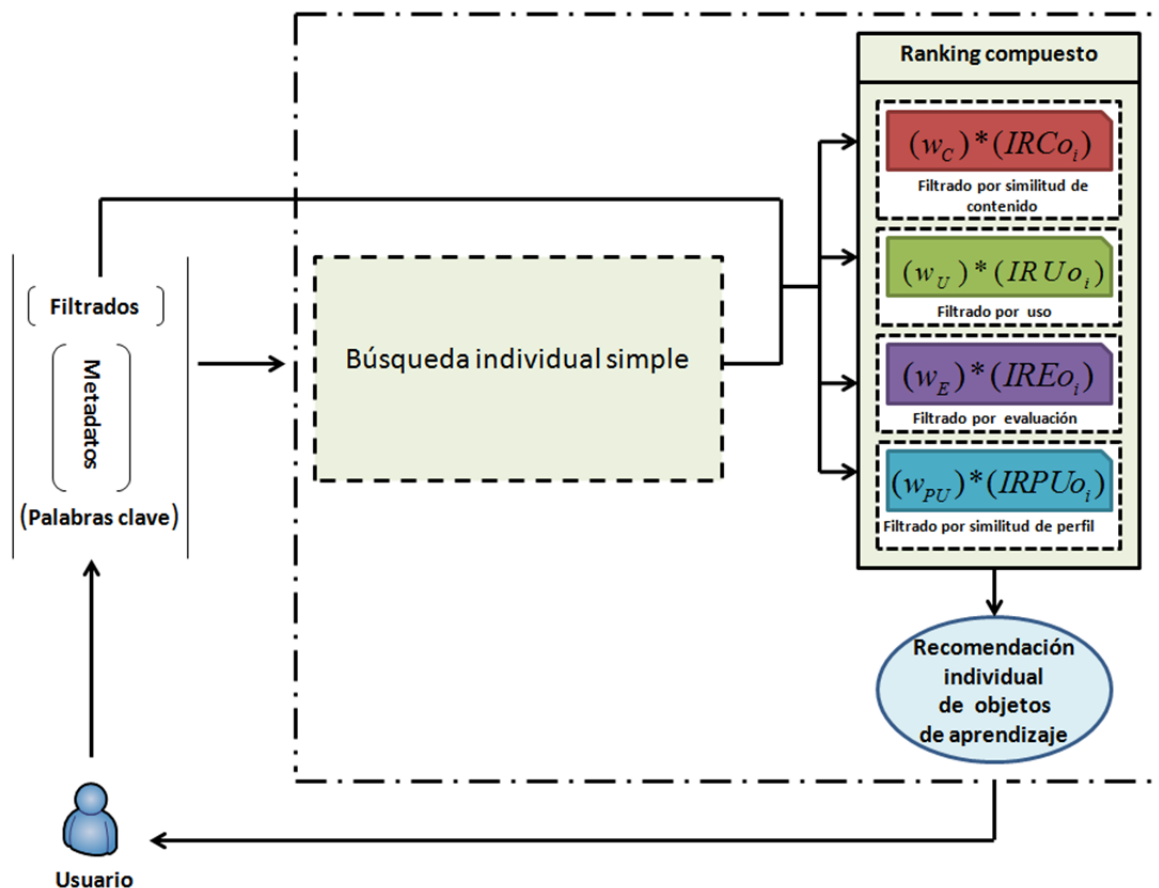


Figura 3.2 Método de búsqueda individual compuesta

A los efectos de esta tesis y como puede observarse en la Figura 3.2 se han propuesto cuatro criterios de filtrado basados en: similitud de contenido, uso, evaluaciones y similitud de perfil (Zapata et al., 2013) . A continuación se describen:

Filtrado por similitud de contenido. En este filtrado se utiliza el *Indicador de Ranking de Contenido para un objeto (IRC_{o_i})* y es coincidente con la *Fórmula 3.2* (descrita en la sección 3.1.1). Su valor varía de [0,1] y se calcula a través de la *Fórmula 3.3*:

$$IRCo_i = \frac{\sum \left[SimMeta_{\text{texto}} (O_x, O_y) \right]}{M} \quad (3.3)$$

Donde:

- M = Número total de metadatos a comparar.
- $SimMeta_{\text{texto}} (O_x, O_y)$ = Porcentaje de similitud obtenido de un metadato.

Filtrado por uso. Este filtrado utiliza un enfoque colaborativo y su fuente de información está basada de las actividades de uso asociadas a los objetos de aprendizaje. A los efectos de este trabajo, el objetivo particular es proporcionar una puntuación superior a los objetos de aprendizaje más descargados, es decir a los más utilizados. El *Indicador de Ranking de Uso para un objeto* ($IRUo_i$) posee un rango $[0,1]$ y se calcula a través de la Fórmula 3.4:

$$IRUo_i = \frac{\sum_{I=1}^N DOx_i}{MaxDOy} \quad (3.4)$$

Donde:

- DOx_i = Número de descargas de un objeto de aprendizaje (Ox_i).
- $MaxDOy$ = Número máximo de descargas que tiene asociado un objeto de aprendizaje (Oy).

Filtrado por evaluación. En este tipo de filtrado utiliza también un enfoque colaborativo y su fuente de información proviene de las evaluaciones asociadas a los objetos de aprendizaje. Su objetivo es proporcionar una puntuación superior a los objetos mejor evaluados por los usuarios. Esta evaluación se debe de llevar a cabo por los usuarios a través de algún instrumento específico.

El *Indicador de Ranking de Evaluación para un objeto* ($IREo_i$) cuyo valor se encuentra en el intervalo $[0,1]$ se calcula a través de la Fórmula 3.5:

$$IREO_i = \frac{\left[\sum_{I=1}^N \sum_{J=1}^M \alpha_{IJ} \right]}{\sum_{K=1}^M Max_K} \quad (3.5)$$

Donde:

- α_{IJ} = Puntuación media de las evaluaciones de un objeto.
- N = Número total de usuarios que han evaluado un objeto.
- Max_K = Valor máximo de la evaluación de un objeto.
- M = Número de reactivos en el instrumento de evaluación.

Filtrado por similitud de perfil. En este filtrado se implementa el enfoque demográfico y a los efectos de este trabajo se ha definido una métrica para calcular el grado de similitud entre dos perfiles de usuarios. Su objetivo es proporcionar una puntuación superior a aquellos objetos de aprendizaje que han sido publicados por usuarios que poseen un perfil similar a quien realiza la consulta.

El *Indicador de Ranking de Perfil de Usuario para un objeto* ($IRPUo_i$) cuyo valor $[0,1]$ se calcula a través de la Fórmula 3.6:

$$IRPUo_i = Sim(Pr u_x, Pr u_y) = \frac{\sum_{a \in A} [SimAtributo(a_x, a_y)]}{A} \quad (3.6)$$

Donde:

- $SimAtributo(a_x, a_y)$ = Distancia entre los atributos correspondientes al perfil de registro de los usuarios que han publicado el OA ($Pr u_x$) y el que ha realizado la búsqueda ($Pr u_y$).
- A = Número total de atributos del perfil de registro a comparar.

3.1.3 CALIFICACIÓN Y RECOMENDACIÓN

El objetivo de este proceso es obtener el valor de la calificación final para de cada uno de los objetos de aprendizaje recuperados de la *búsqueda individual compuesta*. Para ello, se utiliza una combinación lineal ponderada de los valores obtenidos por cada uno de los criterios de filtrado, los cuales son multiplicados de forma correspondiente por un peso que representa la densidad de los valores asociados a ese filtrado. El valor de la calificación final tiene un rango de $[0,1]$ y se obtiene a través de la Fórmula 3.7:

$$\mathcal{R}o_i = \frac{(w_C * IRCo_i + w_U * IRUo_i + w_E * IREo_i + w_{PU} * IRPUo_i)}{N} \quad (3.7)$$

Donde:

- W_C = Peso para la similitud de contenido.
- $IRCo_i$ = Indicador de Ranking de Contenido para un objeto.
- W_U = Peso para el uso.
- $IRUo_i$ = Indicador de Ranking de Uso para un objeto.
- W_E = Peso para la evaluación.
- $IREo_i$ = Indicador de Ranking de Evaluación para un objeto.
- W_{PU} = Peso para la similitud de perfil.
- $IRPUo_i$ = Indicador de Ranking de Perfil de Usuario para un objeto.
- N = Es el número de criterios de filtrado habilitados en la petición y que tienen asociado un peso mayor que cero. El valor de N se encuentra ubicado en el rango $[1,4]$.

Aunque los valores estáticos pueden ser asignados a los parámetros (W_C , W_U , W_E y W_{PU}), es mejor ajustar las proporciones óptimas para los pesos dinámicamente. Con el fin de resolver este problema, en esta tesis se propone la obtención automática de estos pesos de manera adaptativa y dinámica. Es decir, el valor de estos pesos se ajusta o adapta a la densidad de información disponible sobre cada criterio de filtrado (Zapata et al., 2013). A continuación, las cuatro ecuaciones propuestas para volver a calcular automáticamente las ponderaciones utilizadas se describen:

Peso adaptativo para la similitud por contenido (W_C). Es un factor de corrección para el *filtrado por similitud de contenido* cuyo valor está representado por la *Fórmula 3.3*. Este peso cambia de acuerdo a la densidad de metadatos valorados en el repositorio. Su valor $[0,1]$ se calcula a través de la *Fórmula 3.8*:

$$w_C = \frac{\sum NMV_{Rep}}{LOM_{Metadato} \cdot n} \quad (3.8)$$

Donde:

- $\frac{\sum NMV_{Rep}}{n}$ = Promedio del número de metadatos valorados en el repositorio.
- $LOM_{Metadato} = 68$ (equivale al número total máximo de los metadatos definidos en el estándar IEEE-LOM).

Peso adaptativo para el uso (W_U). Es un factor de corrección para el *filtrado por uso* el cual está representado por la *Fórmula 3.4*. Este peso varía de acuerdo al ratio de descargas totales asociadas a los objetos de aprendizaje en el repositorio. Su valor $[0,1]$ se calcula a través de la *Fórmula 3.9*:

$$w_U = \frac{NOD_{Rep}}{OA_{Publicados}} \quad (3.9)$$

Donde:

- NOD_{Rep} = Número total de objetos de aprendizaje descargados en el repositorio.
- $OA_{publicados}$ = Número total de objetos de aprendizaje publicados en un repositorio.

Peso adaptativo para la evaluación (W_E). Es un factor de corrección para el *filtrado por evaluación* cuyo valor está representado por la *Fórmula 3.5*. Este peso cambia de acuerdo al ratio de objetos de aprendizaje evaluados en el repositorio. Su valor $[0,1]$ se calcula utilizando la *Fórmula 3.10*:

$$w_E = \frac{NOE_{Rep}}{OA_{Publicados}} \quad (3.10)$$

Donde:

- NOE_{Rep} = Número total de objetos de aprendizaje evaluados en el repositorio.
- $OA_{Publicados}$ = Número total de objetos de aprendizaje publicados en el repositorio.

Peso adaptativo para la similitud de perfil (W_{PU}). Es un factor de corrección para el *filtrado por similitud de perfil* el cual está representado por la *Fórmula 3.6*. Este peso cambia de acuerdo a la densidad de atributos del perfil de registro valorados en el repositorio. Su valor $[0,1]$ se calcula a través de la *Fórmula 3.11*:

$$w_{PU} = \frac{\sum APV_{Rep}}{AP_{Total}} \quad (3.11)$$

Donde:

- $\frac{\sum APV_{Rep}}{n}$ = Promedio del número de atributos valorados en los perfiles de registro en el repositorio.
- AP_{Total} = Número total de atributos que contiene el perfil de registro.

Como resultado final se obtiene un *ranking* u ordenación de los objetos de aprendizaje de acuerdo a la ponderación de las calificaciones obtenidas de los distintos criterios de filtrado. De esta forma, el mecanismo propuesto aporta una lista de recomendación de los objetos mejor ajustados a las características de la petición, la relevancia de los objetos y los perfiles de los usuarios.

3.2 BÚSQUEDA COLABORATIVA

La búsqueda colaborativa está enfocada a facilitar a los miembros de un mismo grupo de usuarios realizar las actividades de petición, recuperación e incorporación de los OA que cree conveniente a los distintos grupos que pertenece. Posteriormente se les podrá asignar

una calificación de acuerdo a los intereses del grupo.

Para ello se propone otro nivel de generalización en el MRH-OA-IC, incorporando un enfoque de *Búsqueda Colaborativa* a la *Búsqueda Individual* antes expuesta. El objetivo es conformar grupos con intereses comunes y recomendaciones de objetos de aprendizaje para los miembros de dichos grupos. La búsqueda colaborativa consta de cuatro procesos principales (ver Figura 3.3): *Gestión de grupos*, *Búsqueda en grupo*, *Valoración en grupo*, y la generación del *Ranking de grupo* (Zapata et al., 2012).

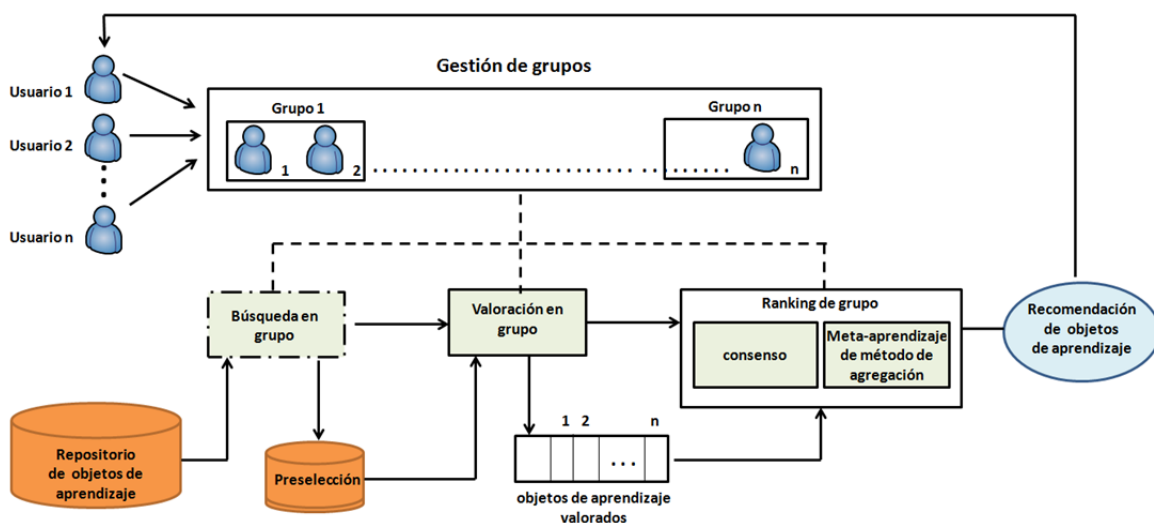


Figura 3.3 Método de búsqueda colaborativa de objetos de aprendizaje

A continuación se describen a detalle estos procesos que aparecen resumidos en la Figura 3.3.

3.2.1 GESTIÓN DE GRUPOS

La *gestión de grupos* es un proceso transversal dentro del *Método de Búsqueda Colaborativa*, ya que facilita todas las actividades de gestión en todos los procesos principales durante un período de tiempo previamente definido.

La gestión comienza en el momento que cualquier usuario decide crear un grupo de búsqueda colaborativa de objetos de aprendizaje. Una vez conformado, es posible el envío de peticiones de incorporación a otros usuarios. Ambos subprocesos se describen:

- **Creación de un grupo.** Se pueden crear grupos con el propósito de definir una temática específica para la búsqueda en común de objetos de aprendizaje. El creador del grupo se convierte en el mediador del mismo, lo cual le permite definir una serie de descriptores de forma explícita, por ejemplo: título, objetivos, descripción general, entre otros.
- **Petición de incorporación a un grupo.** Una vez conformado el grupo, el mediador puede enviar peticiones de incorporación a otros usuarios para formar parte del mismo. Esta petición de incorporación se expresa de forma explícita por parte del mediador a través de alguna herramienta de comunicación. Estos grupos se suelen formar entre profesores de la misma asignatura, área de conocimiento, trabajo conjunto en diseño de objetos de aprendizaje, entre otros.

3.2.2 BÚSQUEDA EN GRUPO

Durante este proceso los miembros del grupo siguen realizando las búsquedas a través de palabras clave, valores de metadatos y criterios de filtrado de la información de forma síncrona o asíncrona. Pero ahora, el mediador del grupo puede establecer un período de tiempo y cada usuario puede visualizar las búsquedas realizadas por todos los miembros de los grupos a los que pertenece.

De este modo se hace posible añadir a los grupos correspondientes los objetos de aprendizaje que se consideren pertinentes o interesantes para el grupo de acuerdo a la temática previamente definida. Para facilitar estas decisiones y orientar a otros miembros, se puede incorporar un componente social para el conjunto de objetos preseleccionados que permita a los miembros la posibilidad de agregar etiquetas y comentarios.

3.2.3 VALORACIÓN EN GRUPO

Una vez finalizadas las búsquedas de los miembros del grupo, pueden asignar valoraciones a los objetos de aprendizaje añadidos, las cuales se dan en una escala [1,5] donde el valor más alto se considera como el OA que más se ajusta los intereses del grupo y el valor más bajo equivale al menos adecuado.

También se ha considerado como estrategia de valoración, que el mediador del grupo pueda

definir un tiempo límite para que todos los miembros las asignen (de forma síncrona o asíncrona). Las valoraciones emitidas pueden visualizarse por todos los miembros de tal forma que se permite la comparación de opiniones.

3.2.4 RANKING CONSENSUADO

Una vez valorados todos los OA hay que establecer una única valoración final del grupo para todos los objetos de aprendizaje. El problema es que una técnica manual requiere de la intervención de todos los miembros del grupo. Lo ideal es obtener el valor final real mediante una técnica manual democrática, lo cual significa que todos los miembros del grupo deciden de forma consensuada la calificación final de cada uno de los OA. De esta forma, se obtiene una recomendación de objetos de aprendizaje, en la que los mejores posicionados resultan los más relevantes para los intereses planteados en el grupo.

Por otro lado, existen técnicas automáticas de valoración en las que se utiliza un método preestablecido para agregar todas las valoraciones de cada uno los miembros a cada uno de los OA. Algunos ejemplos son: el promedio, la mejor puntuación, la valoración anárquica, la peor puntuación, entre otras. Estas formas de valoración se describen brevemente en la *Tabla 2.3* (epígrafe 2.4.3).

3.2.5 META-APRENDIZAJE PARA LA SELECCIÓN DEL MÉTODO DE AGREGACIÓN

Con el objetivo de poder determinar que técnica de agregación de valoraciones funciona mejor para cada tipo de grupo, se propone utilizar un proceso de meta-aprendizaje (Figura 3.4).

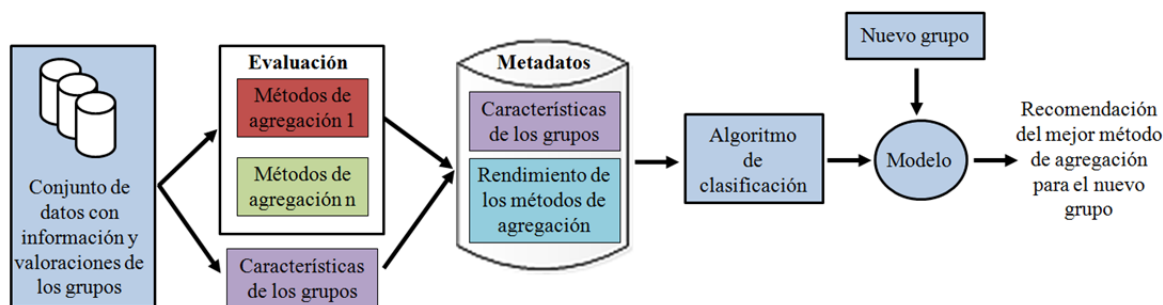


Figura 3.4 Proceso de meta-aprendizaje

Como se observa en la Figura 3.4, el proceso de meta-aprendizaje parte de un *conjunto de datos* que contiene información sobre los grupos y sus valoraciones. A partir de la información de los grupos se definen sus características, las cuales son:

- **Tamaño de grupo.** Representa el número de miembros que conforman el grupo.
- **Nivel del contacto social.** Representa el nivel de contacto social entre los miembros del grupo desde el punto de vista profesional o académico. Este valor se obtiene de forma explícita y con el fin de cuantificarla se propone categorizarlo en cinco niveles (0: Nula, 1: Baja, 2: Normal, 3: Alta, 4: Muy alta). El valor de este descriptor para un grupo es el promedio y la varianza de los valores indicados por todos los miembros del grupo.
- **Nivel de experiencia.** Representa el nivel de experiencia basado en cuatro atributos del perfil de registro de los miembros del grupo, los cuales son: años de docencia, tecnología ofimática, didáctica y diseño instruccional. El valor de los atributos se ha categorizado en cuatro niveles (0: ninguna, 1: inicial, 2: media, 3: avanzada) y el valor final de este descriptor para el grupo es el promedio y la varianza de los valores de todos los miembros del grupo.
- **Nivel de disimilaridad.** Representa el desacuerdo entre los miembros del grupo con respecto a las valoraciones de un objeto de aprendizaje. El valor de este descriptor se basa en el promedio y varianza de la diferencia de las valoraciones emitidas a los objetos de aprendizaje.
- **Nivel de actividad.** Representa el nivel de actividad de descargas y publicaciones de objetos de aprendizaje realizado por los miembros del grupo. Para ambas actividades se ha propuesto categorizarlas en cinco niveles (0: Nula, 1: Baja, 2: Normal, 3: Alta, 4: Muy alta). El valor de este descriptor es el promedio y varianza de las actividades de todos los miembros.

Por otro lado, a partir de las valoraciones individuales de cada uno de los miembros de un grupo se pueden utilizar diversos métodos de agregación para obtener una valoración final del grupo para cada OA:

- 1) Promedio (Average).** En este método el OA más influyente del grupo, es aquel que obtiene, el mayor valor resultante entre la suma de todas las valoraciones y el número de miembros del grupo (Tabla 3.1).

Tabla 3.1 Ejemplo del método Promedio

Miembros	OA									
	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M₁	10	4	3	6	10	9	6	8	10	8
M₂	1	9	8	9	7	9	6	9	3	8
M₃	10	5	2	7	9	8	5	6	7	6
grupo	7	6	4.3	7.3	8.7	8.7	5.7	7.7	6.7	7.3

- 2) Promedio descartando el valor mínimo (Average without misery).** Es igual que en el *método 1*, la diferencia radica en aquellos OA que tienen una valoración por debajo de un umbral determinado (por ejemplo 3), los cuales no serán considerados en las recomendaciones del grupo (Tabla 3.2).

Tabla 3.2 Ejemplo del método Promedio descartando en valor mínimo

Miembros	OA									
	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M₁	10	4	3	6	10	9	6	8	10	8
M₂	1	9	8	9	7	9	6	9	3	8
M₃	10	5	2	7	9	8	5	6	7	6
grupo	-	6	-	7.3	8.7	8.7	5.7	7.7	-	7.3

- 3) Valor mínimo (Least misery).** Se considera al OA con la valoración más baja como el más influyente del grupo. De esta forma, el OA mejor valorado de los valores mínimos representa la valoración del grupo (Tabla 3.3).

Tabla 3.3 Ejemplo del método Valor mínimo

	OA									
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M ₁	10	4	3	6	10	9	6	8	10	8
M ₂	1	9	8	9	7	9	6	9	3	8
M ₃	10	5	2	7	9	8	5	6	7	6
grupo	1	4	2	6	7	8	5	6	3	6

- 4) **El más votado (Most pleasure)**. Es similar al *método 3*, pero en lugar de considerar la valoración más baja de los miembros, selecciona la más alta. El OA con la valoración más alta es la más influyente para el grupo (Tabla 3.4).

Tabla 3.4 Ejemplo del método Valor máximo

	OA									
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M ₁	10	4	3	6	10	9	6	8	10	8
M ₂	1	9	8	9	7	9	6	9	3	8
M ₃	10	5	2	7	9	8	5	6	7	6
grupo	10	9	8	9	10	9	6	9	10	8

- 5) **Equitativo (Fairness)**. Este método se describe mejor a través de un ejemplo (ver Tabla 3.5): se selecciona de forma aleatoria un miembro, por ejemplo M₁, cuyos objetos mejores valorados son OA₁, OA₅ y OA₉. A partir de estos tres objetos se visualizan los valores mínimos que poseen y se selecciona el mayor de ellos. Por lo tanto, OA₅ tiene una valoración de 10. El proceso continua de la misma forma teniendo en cuenta el resto de los objetos disminuyendo uniformemente de a 1.

Tabla 3.5 Ejemplo del método Equitativo

	OA									
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M ₁	10	4	3	6	10	9	6	8	10	8
M ₂	1	9	8	9	7	9	6	9	3	8
M ₃	10	5	2	7	9	8	5	6	7	6
grupo	4	3	1	8	10	9	5	7	2	6

- 6) **Voto pluralista (Plurality voting).** Es similar al *método 5*, pero en lugar de seleccionar un número definido de los objetos mejor valorados considerando sus valores mínimos, se elige la alternativa que más valoraciones haya obtenido (Tabla 3.6).

Tabla 3.6 Ejemplo del método Voto pluralista

	OA									
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M ₁	10	4	3	6	10	9	6	8	10	8
M ₂	1	9	8	9	7	9	6	9	3	8
M ₃	10	5	2	7	9	8	5	6	7	6
grupo	5	3	1	8	10	9	2	7	4	6

- 7) **Votación por aprobación (Approval voting) (con umbral 2).** En este método se toman en cuenta solo las valoraciones de los miembros que sean mayores al umbral 2. Por cada miembro que asigne al OA una valoración que supere este umbral, se contabiliza como 1. El OA más influyente para el grupo es aquel que acumule un número mayor de valoraciones (Tabla 3.7).

Tabla 3.7 Ejemplo del método Voto aprobado (con umbral 2)

OA										
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M ₁	10	4	3	6	10	9	6	8	10	8
M ₂	1	9	8	9	7	9	6	9	3	8
M ₃	10	5	2	7	9	8	5	6	7	6

 Umbral = 2

OA										
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M ₁	1	1	1	1	1	1	1	1	1	1
M ₂		1	1	1	1	1	1	1	1	1
M ₃	1	1		1	1	1	1	1	1	1
grupo	2	3	2	3	3	3	3	3	3	3

- 8) **Votación por aprobación (Approval voting) (con umbral 3).** Es similar al *método* 7, pero considerando las valoraciones de los miembros que sean mayores al umbral 3 (Tabla 3.8).

Tabla 3.8 Ejemplo del método Voto aprobado (con umbral 3)

OA										
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M ₁	10	4	3	6	10	9	6	8	10	8
M ₂	1	9	8	9	7	9	6	9	3	8
M ₃	10	5	2	7	9	8	5	6	7	6

 Umbral = 3

OA										
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M ₁	1	1	1	1	1	1	1	1	1	1
M ₂		1	1	1	1	1	1	1		1
M ₃	1	1		1	1	1	1	1	1	1
grupo	2	3	2	3	3	3	3	3	2	3

- 9) Promedio ponderado (weighted Average) (más peso al miembro con más experiencia).** Es similar al *método 1*. La diferencia radica en que se establece un mayor peso a las valoraciones de un determinado miembro del grupo. Ese mismo peso asignado se toma en cuenta en el denominador al momento de calcular el promedio. Para este método, se considera un peso doble a las valoraciones del miembro con más experiencia del grupo (M_1), de acuerdo a su perfil de registro (Tabla 3.9).

Tabla 3.9 Ejemplo del método de Promedio ponderado (más peso al miembro con más experiencia)

	OA									
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M₁	20	8	6	12	20	18	12	16	20	16
M₂	1	9	8	9	7	9	6	9	3	8
M₃	10	5	2	7	9	8	5	6	7	6
grupo	7.75	5.5	4	7	9	8.75	5.75	7.75	7.5	7.5

- 10) Promedio ponderado (más peso al miembro más social).** Es similar al *método 9*, pero se caracteriza por proporcionar un peso doble a las valoraciones del miembro más social del grupo (M_2), es decir, aquel que mantiene las relaciones más estrechas desde un punto de vista profesional o académico con el resto de los miembros (Tabla 3.10).

Tabla 3.10 Ejemplo del método de Promedio ponderado (más peso al miembro más social)

	OA									
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M₁	10	4	3	6	10	9	6	8	10	8
M₂	2	18	16	18	14	18	12	18	6	16
M₃	10	5	2	7	9	8	5	6	7	6
grupo	5.5	6.75	5.25	7.75	8.25	8.75	5.75	8	5.75	7.5

- 11) Promedio ponderado (más peso al miembro más activo).** Es similar a los *métodos 9 y 10*, pero se diferencia de ambos por asignar un peso doble a las valoraciones del

miembro más activo del grupo (M_3), de acuerdo al número de descargas y publicaciones realizadas de objetos de aprendizaje (Tabla 3.11).

Tabla 3.11 Ejemplo del método de Promedio ponderado (más peso al miembro más activo)

	OA									
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M₁	10	4	3	6	10	9	6	8	10	8
M₂	1	9	8	9	7	9	6	9	3	8
M₃	20	10	4	14	18	16	10	12	14	12
grupo	7.75	5.75	3.75	7.25	8.75	8.5	5.5	7.25	6.75	7

12) Mediana (Median). Consiste en establecer el valor medio de las valoraciones realizadas por los miembros del grupo a un OA específico, ordenados de menor a mayor (Tabla 3.12).

Tabla 3.12 Ejemplo del método Mediana

	OA									
Miembros	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆	OA ₇	OA ₈	OA ₉	OA ₁₀
M₁	10	4	3	6	10	9	6	8	10	8
M₂	1	9	8	9	7	9	6	9	3	8
M₃	10	5	2	7	9	8	5	6	7	6
grupo	10	5	3	7	9	9	6	8	7	8

Posteriormente, se realizará una *evaluación* para determinar el método de agregación que obtenga un menor error con respecto a la valoración final consensuada de los miembros del grupo para todos los OA, concretamente, se va a seleccionar aquel método que obtenga el valor más bajo de RMSE (Root-Mean-Square Error). Esta es una medida ampliamente utilizada para calcular las diferencias entre los valores pronosticados por un modelo o un estimador y los valores realmente observados (Hyndman & Koehler, 2006). Su valor es de 0 a infinito, donde los valores más pequeños indican una mejor precisión y se calcula a través de la Fórmula 3.12:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{1,i} - X_{2,i})^2}{n}} \quad (3.12)$$

Donde:

X_1 = Representa el valor real obtenido democráticamente por el grupo.

X_2 = Representa el valor predicho por el método de agregación.

n = Es el número de decisiones/valoraciones del grupo.

Una vez finalizada esta fase de evaluación de los métodos de agregación, se procede a la conformación de un nuevo conjunto de datos o metadatos, que contiene tanto las características de cada uno de los grupos así como también el método de agregación que presentó el mejor rendimiento para el grupo.

A partir de este conjunto de metadatos se puede predecir que método de agregación es el más apropiado para un nuevo grupo dada sus características. Esta es una tarea de clasificación, donde el valor a predecir es el nombre del método de agregación. El problema es que existen multitud de técnicas de clasificación, por lo que habrá que seleccionar un número representativo de algoritmos de clasificación para y comparar su rendimiento al utilizar nuestro conjunto de metadatos y poder así determinar el algoritmo que ofrezca un mejor rendimiento y utilizarlo para predecir el método de agregación más apropiado. De forma que utilizando el modelo de clasificación obtenido por el algoritmo elegido se podrá predecir en tiempo real cuál es el mejor método de agregación para un nuevo grupo en función de las características de los miembros que lo componen y sus valoraciones individuales.

3.3 CONCLUSIONES

En este capítulo se describió el diseño conceptual del modelo de recomendación híbrido de objetos de aprendizaje individual y colaborativo, cuyo objetivo es mejorar los mecanismos de recuperación y clasificación de objetos de aprendizaje en repositorios. Para ello, se

establecieron filtros de información por contenido, colaborativo y demográfico, razón por la cual se considera un enfoque de recomendación híbrido.

El MRH-OA-IC, fue diseñado para implementarse en entornos reales de repositorios de objetos de aprendizaje que se gestionan bajo estándares internacionales, tales como: MERLOT, ARIADNE, MACE, entre otros.

En el siguiente capítulo, se describirá la implementación del MRH-OA-IC que se ha realizado utilizando la *plataforma AGORA*, un sistema de gestión integral de objetos de aprendizaje. Esta implementación se ha denominado *DELPHOS* (Zapata et al., 2012).

4 DELPHOS. SISTEMA DE RECOMENDACIÓN HÍBRIDO

En este capítulo se describe el sistema de recomendación híbrido DELPHOS que es el resultado de la implementación del *MRH-OA-IC* descrito en el capítulo anterior. Se inicia presentando las características principales del marco arquitectónico. Posteriormente, se describen las especificaciones técnicas que involucran las tecnologías utilizadas para el desarrollo, el esquema de la base de datos y la forma como interactúa con el repositorio AGORA. También, se detallan las características del diseño de la interfaz Web para realizar diversas actividades tales como: gestionar el perfil del usuario, realizar la búsqueda de objetos de aprendizaje, realizar la gestión de grupos y las acciones colaborativas entre los miembros del grupo. Finalmente, se exponen las principales funcionalidades del sistema.

4.1 MARCO ARQUITECTÓNICO

DELPHOS emplea una arquitectura abierta y modular basada en un diseño de tres capas (Valipour et al., 2009), el cual separa la interfaz de uso, el comportamiento del sistema y las operaciones para la gestión de la información; con el fin de simplificar la implementación y mejorar la reutilización.

En la arquitectura (ver Figura 4.1), cada una de las capas está constituida por una colección de servicios modelados como componentes, los cuales, al ser ejecutados en forma individual o conjunta permiten la realización de los procesos necesarios para la búsqueda y recomendación de objetos de aprendizaje.

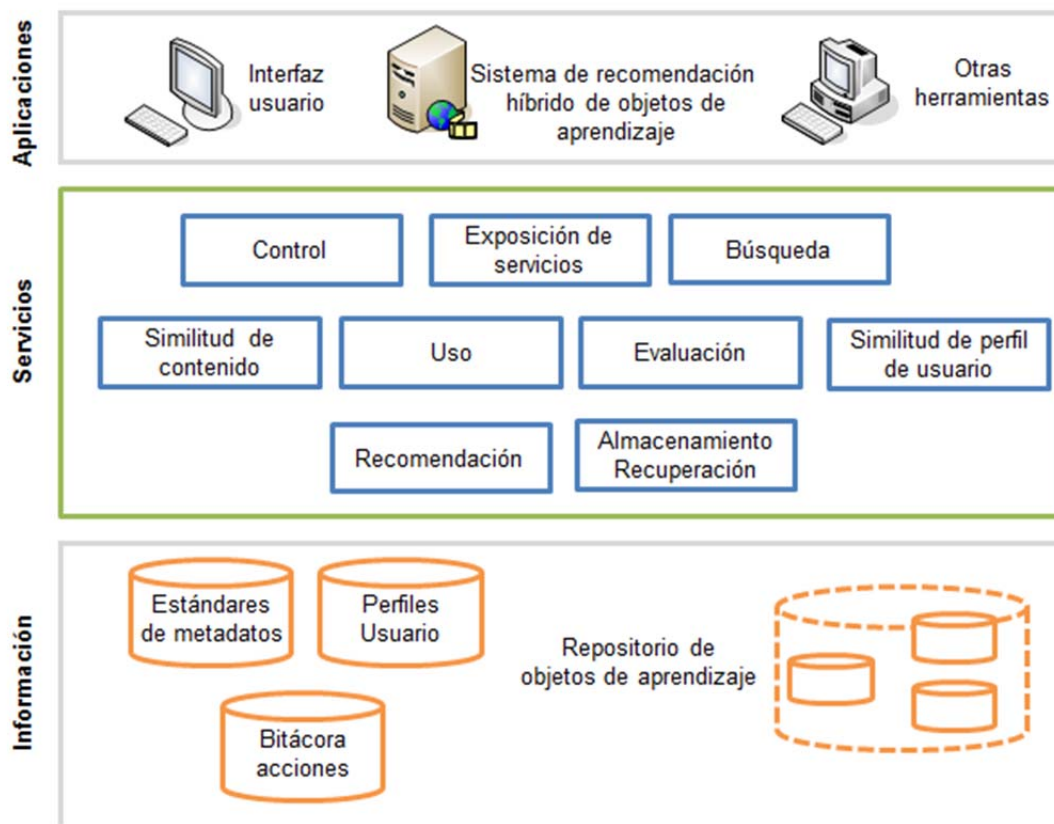


Figura 4.1 Marco arquitectónico del sistema de recomendación híbrido DELPHOS.

Como se observa en la Figura 4.1, cada una de las capas está especializada en realizar determinadas funciones, las cuales se describen:

- 1) **Aplicaciones.** En esta capa aparecen referidos los usuarios que consultan y descargan los objetos de aprendizaje. Cualquier herramienta o sistema e-learning que soporte un estándar de distribución o una interfaz de interoperabilidad, podrá utilizar la información y las funcionalidades expuestas por el sistema.
- 2) **Servicios.** En la capa intermedia se describen los servicios integrados que dan soporte a la recomendación híbrida de objetos de aprendizaje de acuerdo a un conjunto de especificaciones definidas. Estos servicios pueden agruparse para ofrecer funcionalidades más complejas o ejecutarse de forma individual con el fin de facilitar las actividades de búsqueda, recomendación y almacenamiento. Se describen sus componentes:

- a. *Control*. Monitorea y registra toda la actividad realizada dentro de la arquitectura, con la finalidad de determinar el grado de calidad de los distintos procesos y los productos generados, así como generar información para detectar y solucionar situaciones problemáticas.
- b. *Exposición de servicios*. Este servicio tiene como finalidad favorecer la interconexión de la arquitectura con otras aplicaciones o sistemas externos. Esto se realiza mediante interfaces que facilitan el acceso a los servicios que constituyen la arquitectura.
- c. *Búsqueda*. Este servicio se centra en la búsqueda de objetos de aprendizaje almacenados en un repositorio o bien en sistemas gestores del aprendizaje. La consulta implica el uso de palabras clave, metadatos y considera los criterios por similitud de contenido, uso, evaluación y similitud de perfil.
- d. *Similitud de contenido*. Tiene como finalidad establecer la semejanza entre colecciones de objetos de aprendizaje. Implementa la métrica de un modelo de similitud que considera la semántica de los metadatos que conforman a un OA para establecer la proximidad entre grupos de objetos.
- e. *Uso*. Este servicio implementa una métrica para calcular la relevancia basada en las actividades de descargas asociadas a los objetos de aprendizaje.
- f. *Evaluación*. Este servicio implementa una métrica para calcular la relevancia basada en las evaluaciones asociadas a los objetos de aprendizaje.
- g. *Similitud de perfil*. Calcula la semejanza entre dos perfiles de usuario. Implementa una métrica que compara los campos del perfil de registro, para establecer la proximidad entre dos usuarios.
- h. *Recomendación*. Es el servicio más importante de la arquitectura, debido a que se encarga de proporcionar los objetos de aprendizaje de acuerdo a los criterios considerados en la búsqueda.
- i. *Almacenamiento y recuperación*. Realiza todas las operaciones de acceso al repositorio. Guarda y recupera los recursos y metadatos que constituyen al

OA sin importar su ubicación y formato de almacenamiento, lo que permite la distribución de la información. Es el responsable de la actualización adecuada de los recursos y metadatos de los objetos.

3) Información. En la capa de información aparece el repositorio de objetos de aprendizaje, el cual almacena información útil para la arquitectura. Este repositorio puede funcionar como:

- a. *Contenedor de metadatos.* El repositorio contiene sólo los metadatos y se accede al OA a través de una referencia a su ubicación física que se encuentra en otro sistema o repositorio de objetos. Esto es útil cuando el recurso referenciado no puede ser descargado bien por su estructura distribuida o por las políticas de uso.
- b. *Contenedor de objetos de aprendizaje.* Extiende el concepto anterior para guardar los contenidos que constituyen al objeto.

4.2 ESPECIFICACIONES TÉCNICAS

En esta sección se describen los aspectos técnicos relacionados con la arquitectura del sistema DELPHOS. Esta información permite una mejor comprensión de los distintos elementos que intervienen en el funcionamiento del sistema. Primero, se describen las tecnologías que intervienen en el desarrollo y las características del servidor Web. Posteriormente, se detalla la estructura de la base de datos y finalmente se explica la forma como interactúa el sistema DELPHOS con el repositorio AGORA.

4.2.1 TECNOLOGÍAS DE DESARROLLO

Las tecnologías involucradas en el desarrollo del sistema constituyen parte de los requerimientos mínimos para su ejecución. A continuación se describe el papel que desempeña cada una:

- **Lenguaje HTML (HyperText Markup Language).** Se implementa para la elaboración de las páginas Web que conforman la interfaz visible del sistema DELPHOS. Establece la estructura y contenido de tipo texto o imágenes de las

páginas Web.

- **Lenguaje CSS (Cascading Style Sheet).** Define la presentación de una página Web en términos de un formato y tipografía específica, independiente de la implementación. Los archivos CSS se les denomina hojas de estilo.
- **Lenguaje JavaScript.** Se implementa para manipular dinámicamente los distintos elementos de la interfaz, lo que permite establecer nuevos estilos de interacción. Generalmente los archivos JavaScript constituyen componentes interactivos (widgets).
- **Lenguaje PHP (Hypertext Pre-processor).** Se utiliza para el desarrollo de componentes software, funciones y bibliotecas de DELPHOS. Es el lenguaje fundamental del sistema y su intérprete puede integrarse de manera eficiente en el servidor Web Apache.
- **Lenguaje SQL (Structured Query Language).** Manipula la información de la base de datos. Si bien existen diversos gestores de bases de datos que pueden utilizarse, para efectos de este sistema se implementa el gestor de base de datos MySQL.

Adicionalmente, para ejecutar o interpretar el código desarrollado para DELPHOS se ha utilizado el servidor de aplicaciones Web XAMPP (Apache Friends Project, 2002), una aplicación software de libre distribución que facilita la instalación de un servidor Web (Apache), un gestor de bases de datos (MySQL) y dos intérpretes de comandos (PHP y PERL). La Figura 4.2 presenta el esquema de interacción de todas las tecnologías involucradas en la ejecución de DELPHOS.

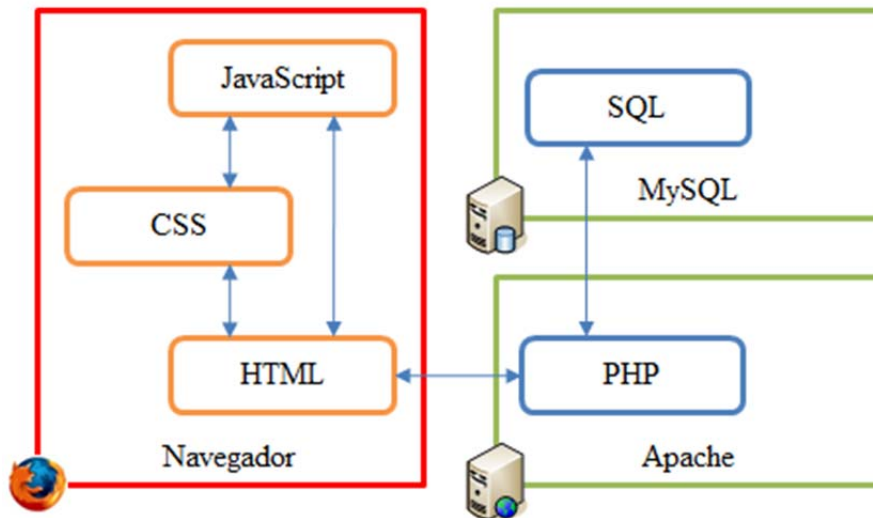


Figura 4.2 Interacción de las tecnologías de desarrollo Web

4.2.2 ESTRUCTURA DE LA BASE DE DATOS

El sistema DELPHOS registra toda la información que se genera de los OA, así como los usuarios y las relaciones que se generan entre ambos. Estos registros se almacenan en 8 tablas que conforman la base de datos. Las estructuras de estos datos almacenados se encuentran relacionadas (Figura 4. 3).

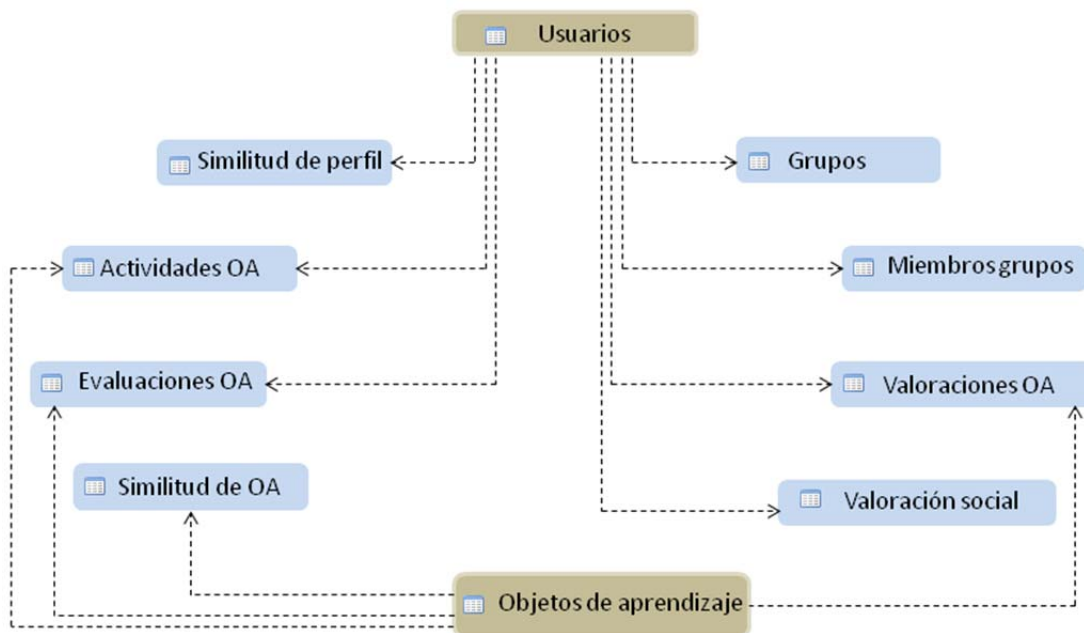


Figura 4.3 Esquema de las tablas de la base de datos y sus relaciones.

El contenido de cada una de las tablas que conforma la base de datos se describe a continuación:

- 1) **Similitud de perfil.** Almacena el ranking de similitud de los perfiles de usuario.
- 2) **Actividades OA.** Almacena el ranking de las actividades de descargas asociadas a los objetos de aprendizaje.
- 3) **Evaluaciones OA.** Contiene el ranking de las evaluaciones pedagógicas asociadas a los objetos de aprendizaje.
- 4) **Similitud de OA.** Almacena el ranking de los objetos de aprendizaje similares de acuerdo a sus metadatos.
- 5) **Grupos.** Contiene las características asociadas de los grupos de búsqueda colaborativa de objetos de aprendizaje.
- 6) **Miembros grupos.** Almacena a los miembros que conforman los grupos de búsqueda colaborativa de OA.
- 7) **Valoraciones OA.** Recopila las valoraciones realizadas a los objetos de aprendizaje por los miembros de los grupos de búsqueda colaborativa.
- 8) **Valoración social.** Contiene las valoraciones del nivel social que existe entre los miembros de los grupos.

Todas las tablas anteriores tienen una configuración distinta con el paso del tiempo, es por ello que son dinámicas. Debido a esta situación se cuenta con un sistema de actualización inmediata de la base de datos. Pero además, se han programado diversos umbrales de actualización automática cada vez que se acumulan las siguientes nuevas acciones en el sistema: 10 OA publicados, 5 evaluaciones pedagógicas, 20 descargas y 5 usuarios nuevos.

4.2.3 REPOSITORIO AGORA

El sistema DELPHOS ha sido desarrollado para recuperar los OA que son más relevantes para las necesidades del usuario, utilizando el repositorio AGORA (Menéndez, 2011) que es un sistema de gestión integral de OA. AGORA comparte aproximadamente unos 800 OA

que han sido publicados por cerca de 300 usuarios de diferentes universidades españolas y latinoamericanas. AGORA es una plataforma orientada a facilitar al usuario el proceso de construcción de OA, conforme a sus necesidades instruccionales. El repositorio almacena objetos de aprendizaje, incluyendo metadatos y sus recursos asociados. Entre sus principales funcionalidades ofrece un gestor de repositorios distribuido y la generación de metadatos conforme al estándar IEEE-LOM. AGORA permite además acceder a sus funcionalidades a través de un conjunto de servicios y componentes que facilitan su extensibilidad y la interoperabilidad con otras aplicaciones como DELPHOS. Una descripción general de las funciones e interacciones con AGORA se muestra en la Figura 4.4.

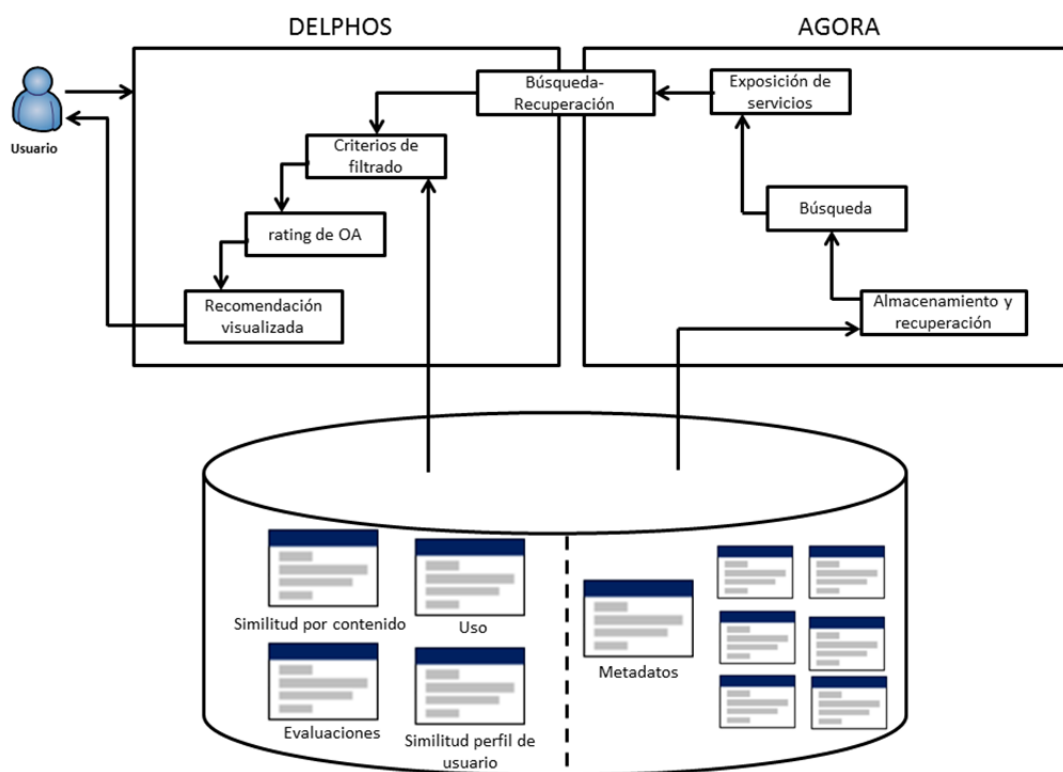


Figura 4.4 Interacción entre el sistema DELPHOS y el repositorio AGORA

Como se observa en la Figura 4.4, estas dos aplicaciones se comunican cuando un usuario utiliza la interfaz gráfica para definir un objeto de aprendizaje y devuelve resultados a la consulta, lo que se denomina el proceso de búsqueda-recuperación. La interfaz del sistema DELPHOS emplea un conjunto de componentes Web para permitir una operación fácil e interactiva. Los parámetros de consulta se utilizan para recuperar una colección de objetos

de aprendizaje almacenados en el repositorio AGORA, así como otra información importante, como metadatos, perfiles de usuario y los registros de la actividad asociada a los OA.

4.3 INTERFAZ WEB

En el diseño de la interfaz Web de DELPHOS se consideraron algunos tópicos de la arquitectura de la información (Rosenfeld, 2002), como son la distribución del contenido textual, la cantidad de elementos gráficos, la presencia de menús descriptivos, el uso de enlaces, entre otros. El propósito era crear un sistema funcional y estético en términos de la interfaz de usuario. Para ello, se ha optado por un diseño centrado en el usuario (Jacko & Sears, 2003), cuyo propósito es mejorar su experiencia mediante su caracterización, la identificación de sus necesidades y el análisis de su comportamiento (ver Figura 4.5).



Figura 4.5 Proceso de diseño centrado en el usuario

La interfaz Web se desarrolló con el lenguaje CSS para una fácil modificación y aplicación de las nuevas características en el futuro. Para el diseño visual se han considerado tonos que ofrecen suficiente contraste entre el texto y el fondo, lo que facilita la lectura y evita la distracción del usuario. En consecuencia, se pretende que su atención se centre en la información que se le presenta (ver Figura 4.6):

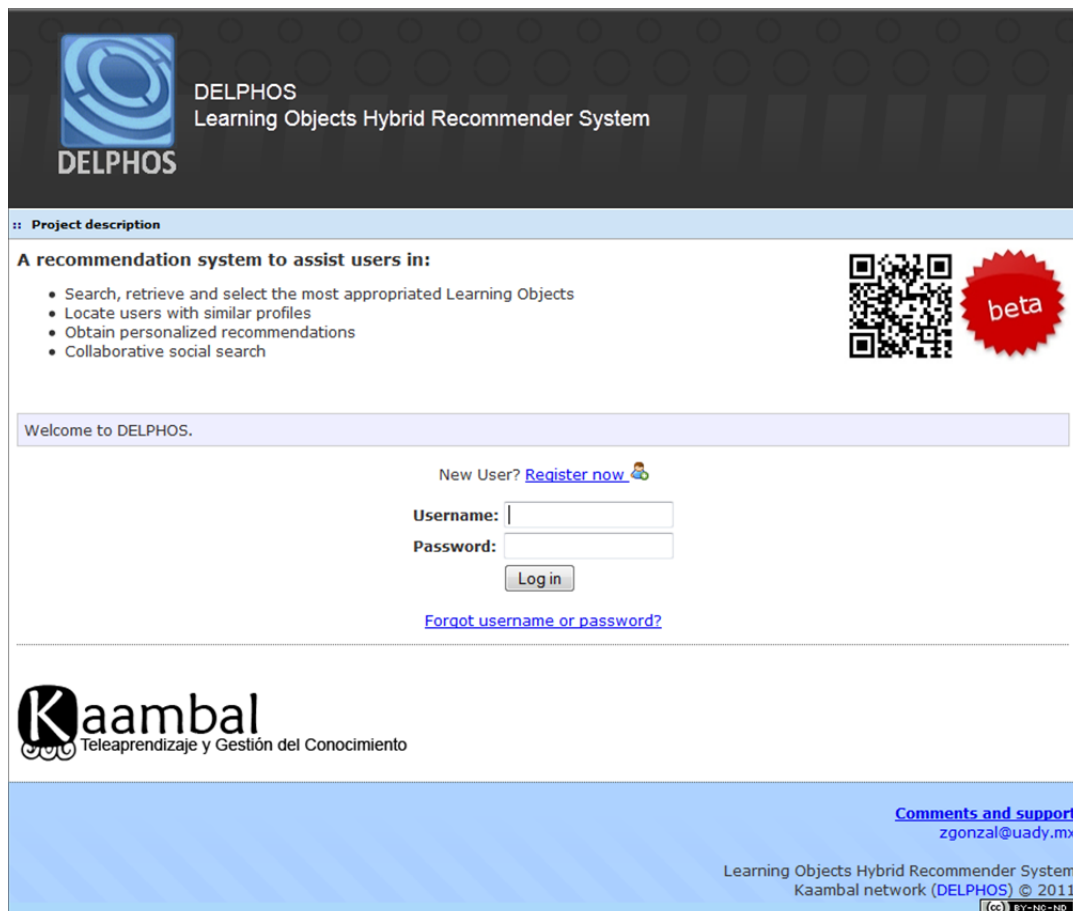


Figura 4.6 Diseño visual de la página Web de entrada a DELPHOS

Como se observa en la Figura 4.6, la interfaz Web está conformada por varias secciones las cuales comparten la misma estructura de diseño, que se describen a continuación:

- Un encabezado donde se presentan las generalidades del proyecto (logotipo, título entre otros).
- La barra de navegación presenta un enlace a la descripción del proyecto.
- El panel de título describe el contenido de la página.
- El panel de contenido presenta formularios de acción o el resultado de una solicitud de información. En él se incluye el título del panel y una breve descripción de su contenido. Opcionalmente puede presentar la iconografía utilizada. Algunas veces el panel de contenido incluye un panel de avisos y notificaciones.

- Una barra de botones de acción.
- El pie de página, que presenta medios de contacto y soporte.

Esta estructura permite organizar todo el conjunto con una consistencia visual y fomenta la sensación minimalista y de estabilidad dentro de la aplicación. Esto trae como consecuencia una mayor facilidad en la navegación y exploración por parte de los nuevos usuarios, agilizando la ejecución de las tareas.

Cabe destacar que los contenidos y descripciones han sido estructurados para una fácil lectura y comprensión. Las imágenes utilizadas han sido limitadas tanto en cantidad como en variedad. La resolución y tamaño elegidos permiten una disminución del tiempo de espera con respecto a la descarga de la página sin restar su significado. También, se han aprovechado las características de las tecnologías utilizadas (componentes, plantillas, caché, compresión, simplificación de código) para optimizar la descarga y el despliegue de las páginas Web.

4.4 FUNCIONALIDAD DEL SISTEMA DELPHOS

En esta sección se describen las funcionalidades que integra DELPHOS, tales como: el registro de usuarios, la búsqueda individual de objetos de aprendizaje, la gestión de grupos, la búsqueda colaborativa de OA y las acciones del grupo.

4.4.1 REGISTRO DE USUARIOS

Para utilizar los servicios y herramientas que ofrece el sistema DELPHOS se requiere previamente rellenar un formulario de registro, el cual podrá ser editado en cualquier momento. Es muy importante para el sistema que los usuarios proporcionen información sobre sí mismos, completando todos los campos del formulario de registro. Utilizando esta información DELPHOS puede mejorar el criterio de recomendación por similitud de perfil de usuario.

En el perfil de usuario existen dos tipos de información. Por un lado, al igual que la mayoría de los sistemas, los usuarios deben proporcionar datos generales y personales (algunos de los cuales son opcionales). Por otro lado, también se solicita información más orientada al

perfil académico tales como: nivel de experiencia como profesor, disciplina de estudios, nivel de enseñanza en el que imparte clases, entre otros (ver Tabla 4.1).

Tabla 4.1 Criterios del perfil de registro de los usuarios

Atributo	Valores asociados				
Experiencia como profesor	0-5 Años	6-10 Años	11-15 Años	16-20 Años	Más de 20 Años
Experiencia en el uso de tecnología ofimática	Ninguna	Inicial	Media	Avanzada	
Experiencia en temas de didáctica	Ninguna	Inicial	Media	Avanzada	
Experiencia en diseño instruccional	Ninguna	Inicial	Media	Avanzada	
Disciplina de estudios	Ciencias sociales y administrativas		Educación, Humanidades y Arte		
	Ciencias de la Salud	Ciencias naturales y exactas		Ingeniería y Tecnología	
Disciplina de la asignatura	Ciencias sociales y Administrativas		Educación, Humanidades y Arte		
	Ciencias de la Salud	Ciencias naturales y exactas		Ingeniería y Tecnología	
Nivel de enseñanza en el que imparte clases	Educación Infantil		Educación Primaria	Educación Secundaria Obligatoria	
	Bachillerato	Formación Profesional		Estudios de posgrado	
Lugar de clases de la asignatura	Aula	Laboratorio	Entorno Real	Domicilio	Mixto
Modalidad de la asignatura	Presencial		Semipresencial		A distancia
Rol desempeñado en la asignatura	Docente		Tutor		Familia
	Compañero		Independiente		Mixta
Estrategia de agrupamiento en la asignatura	Individual		Grupo		Mixto
Plataforma de educación en línea	Ninguna		Moodle		Dokeos
	Claroline		Desarrollo propio		Otra
Editor de objetos de aprendizaje que utiliza	Ninguno		Reload		WebCT
	SCORM editor		Desarrollo propio		Otro
Repositorio de objetos de aprendizaje que utiliza	Ninguno		AGREGA		ARIADNE
	MERLOT		CAREO		Otro

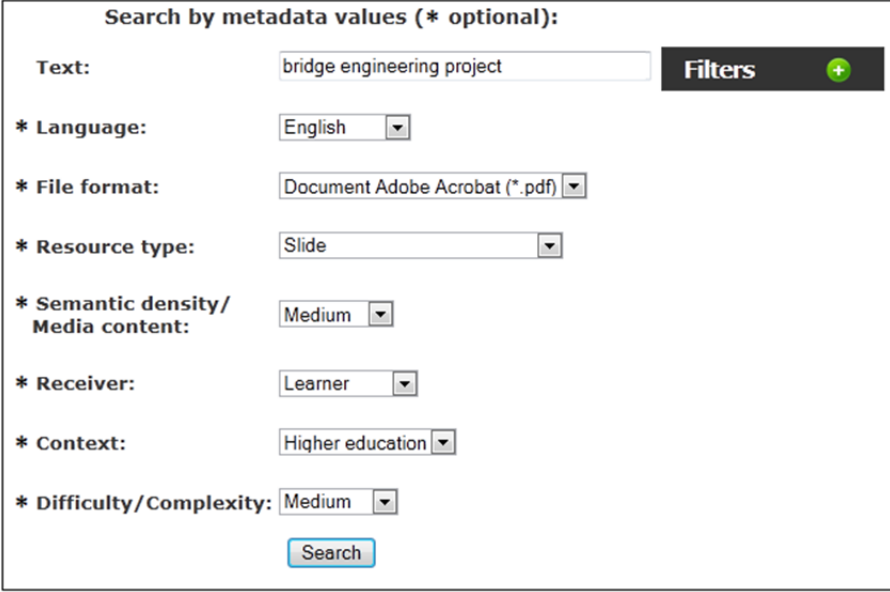
4.4.2 BÚSQUEDA INDIVIDUAL DE OBJETOS DE APRENDIZAJE

En correspondencia a la propuesta del epígrafe 3.1.2, la búsqueda individual de objetos de aprendizaje está diseñada para ser muy flexible, permitiendo tanto el uso de palabras clave (campo obligatorio Text de la Figura 4.7) como algunos valores de metadatos de acuerdo al estándar IEEE-LOM (campos opcionales de la Figura 4.7), los cuales fueron seleccionados

en base a la recomendación CanCore-LRM (Friesen, 2004) que define los metadatos de mayor incidencia en opciones de búsqueda, y que se describen a continuación:

- **Idioma.** El idioma o idiomas predominantes en cada OA para la comunicación con el usuario (Español, Inglés, Francés, Portugués, Italiano).
- **Formato del archivo.** El formato o extensión del archivo del OA (doc, pdf, html, txt, xls, ppt, swf, mid, mp3, wav, ra, bmp, gif, jpg, png, avi, mpg, mov, rv, asf, wmv, flv).
- **Tipo de recurso.** El uso o el tipo de recurso que contiene el OA (Ejercicio, Cuestionario, Figura, Gráfica, diapositiva, Tabla, Examen, Experimento, Lectura, Fotografía, Vídeo, Música).
- **Densidad semántica.** La densidad semántica de un OA puede ser estimada en función la cantidad de información que contiene (Muy bajo, Bajo, Medio, Alto, Muy alto).
- **Destinatario.** El usuario(s) principal(es) para el que ha sido diseñado este OA (Profesor, Estudiante, Administrador, Profesional).
- **Contexto.** El entorno principal en el que se utilizará este OA (Escuela, Educación superior, Entrenamiento, Otro).
- **Dificultad.** Este elemento describe lo difícil o complejo que puede resultar a los destinatarios típicos trabajar con este OA (Muy fácil, Fácil, Medio, Difícil, Muy difícil).

Todos estos elementos aparecen en la interfaz según se aprecia en la Figura 4.7.



Search by metadata values (* optional):

Text: **Filters** +

*** Language:**

*** File format:**

*** Resource type:**

*** Semantic density/
Media content:**

*** Receiver:**

*** Context:**

*** Difficulty/Complexity:**

Figura 4.7 Búsqueda individual de objetos de aprendizaje

La búsqueda individual de OA no solamente implica definir los parámetros de la búsqueda basados en palabras clave y metadatos, sino que también, ofrece la posibilidad de considerar los diferentes filtrados de información o criterios de recomendación basados en: similitud de contenido, uso, evaluación y similitud de perfil de usuario (Figura 4.8).

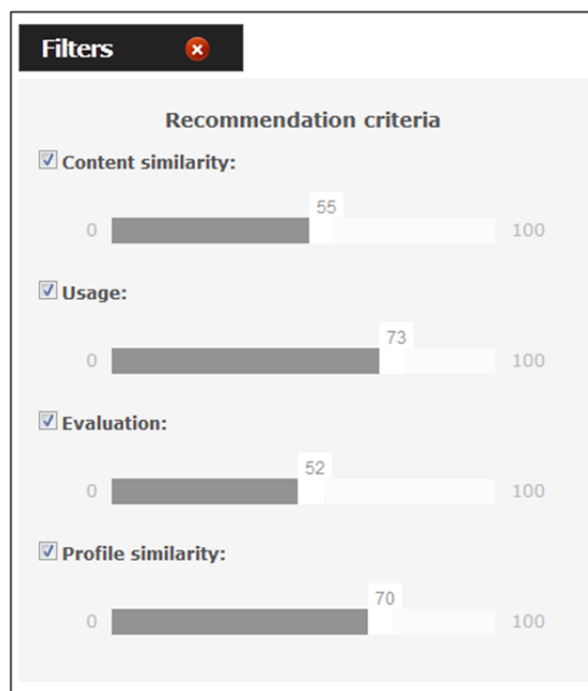


Figura 4.8 Panel de los criterios de recomendación

Como se observa en la Figura 4.8, el panel de los criterios de recomendación permite a los usuarios modificar los diferentes valores del peso de cada criterio de recomendación. Los usuarios pueden asignar nuevos valores (en un rango de 0 a 100%) mediante el uso de una barra deslizante, y también se puede activar o desactivar todos los criterios de recomendación. Sin embargo, al menos un criterio debe permanecer activo a fin de ser capaz de calcular el puntaje asociado al OA recomendado. Es importante notar que el sistema DELPHOS proporciona al usuario unos valores recomendados del peso de cada criterio o valores por defecto. Estos valores se calculan de forma adaptativa y se consideran dinámicos debido a que se actualizan de forma periódica. La modificación de estos valores es una acción opcional y queda a voluntad de los usuarios.

Finalmente, la búsqueda individual devuelve al usuario un listado de objetos de aprendizaje que cumplen con los parámetros introducidos. Además junto a cada objeto el sistema proporciona una serie de información adicional (ver Figura 4.9). Esta información puede ser muy útil para los usuarios, ya que les ayuda a seleccionar los mejores y más interesantes OA. El usuario podrá descargar y utilizar ninguno, uno o varios OA de la lista obtenida.

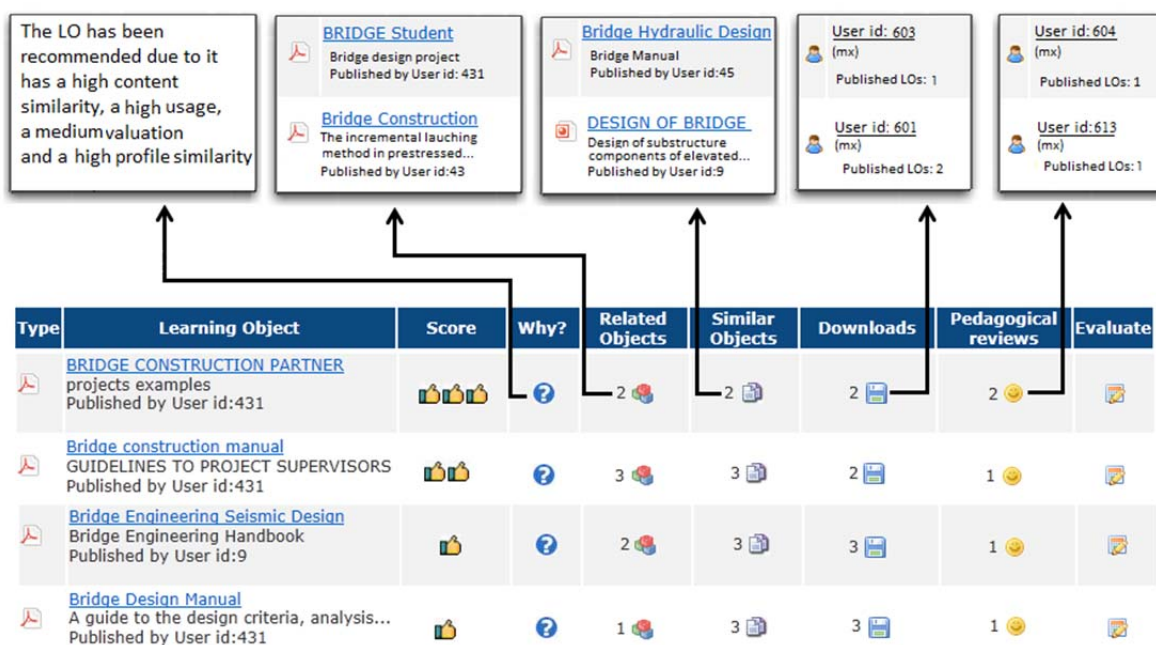


Figura 4.9 Ejemplo de listado de objetos de aprendizaje recuperados

Como se observa en la Figura 4.9, cada OA recomendado aparece acompañado de una calificación calculada automáticamente mediante el sistema de calificación y recomendación híbrido basado en pesos, objetos relacionados y similares, así como las referencias de los usuarios que lo han descargado y evaluado, descrito en el epígrafe 3.1.3. Para facilitar su interpretabilidad al usuario, se ha transformado el valor real entre $[0,1]$ de la calificación en un icono gráfico con 5 posibles valores (ver columna Score de la Figura 4.9) mediante una discretización en anchura: cinco iconos (valor > 0.8 y ≤ 1), cuatro iconos (valor ≤ 0.8 y > 0.6), tres iconos (valor ≤ 0.6 y ≥ 0.4), dos iconos (valor < 0.4 y ≥ 0.2) y un icono (valor < 0.2 y ≥ 0).

Cabe destacar que el sistema DELPHOS ofrece una explicación breve o sentencia del porqué este OA en particular ha sido recomendado (ver columna Why? de la Figura 4.9). La sentencia se forma mostrando un texto y los valores obtenidos por cada criterio de recomendación del OA, con el fin de mostrar al usuario una explicación humana real. Nuevamente se han discretizado los valores de cada criterio de filtrado en tres etiquetas de la siguiente forma: Alto (valor > 0.7 y ≤ 1), Medio (valor ≤ 0.7 y ≥ 0.3) y Bajo (valor < 0.3 y ≥ 0).

Además DELPHOS proporciona información estadística de las descargas de los OA y objetos relacionados y similares también representados por íconos, que se describen en la Tabla 4.2:

Tabla 4.2 Descripción de los íconos asociados a los objetos recuperados de la búsqueda individual

Iconos	Descripción	
	Score	Cálculo basado en una escala de 5 pulgares.
	Why?	Se muestra información relativa a la justificación de su posición en el ranking.
	Related Objects	Se muestra una lista de los objetos relacionados con este OA.
	Similar Objects	Se muestra una lista de los objetos más similares de acuerdo al estándar IEEE-LOM de metadatos.
	Downloads	Se muestra cuántos usuarios han descargado el OA.
	Pedagogical reviews	Se muestra cuántos usuarios han evaluado el OA.
	Evaluate	El usuario puede realizar una evaluación pedagógica del OA.

En la versión actual de DELPHOS, es posible realizar una evaluación pedagógica del objeto. Durante el desarrollo de esta tesis se adoptó un mecanismo de evaluación pedagógica cualitativa propuesta y testada, en el que participaron docentes e investigadores con especialidades de varias áreas de conocimiento y varios países a través de un proyecto liderado por miembros del grupo de investigación SMILe. Esta iniciativa se denomina *Modelo para la Evaluación de la Calidad de Objetos de Aprendizaje (MECOA)* (Zapata et al., 2009), consta de seis categorías a evaluar: Contenido, Representación, Competencias, Autogestión, Significación y Creatividad. Estas seis categorías contienen 12 ítems (ver Tabla 4.3).

Tabla 4.3 Mecanismo de evaluación pedagógico cualitativo MECOA

Categoría a evaluar	Atributo	Valores asociados		
Contenido del objeto de aprendizaje	¿En qué grado el objeto de aprendizaje especifica lo que el alumno aprenderá?	Nulo Más que suficiente	Escaso Suficiente-adecuado Excesivo	
	¿Qué tipo de aprendizaje se propone que el estudiante adquiriera por medio del objeto?	Conceptual	Procesual	Actitudinal
	¿Qué cantidad de medios (gráficos, visuales, auditivos, animación, etc.) contiene el objeto de aprendizaje?	Nulo Adecuado	Deficiente Saturado	Regular
Representación del objeto de aprendizaje	¿Cómo se integran los elementos del objeto de aprendizaje?	Aislados Redundantes	Secuenciados Saturados	Integrados
Competencias logradas a partir de la interacción con el objeto de aprendizaje	¿Qué adquiere el estudiante por medio del uso de este objeto de aprendizaje?	Conocimientos Valores	Habilidades Mixto	Actitudes
	¿Qué categorías de aprendizaje logra el alumno?	Acceso información Aplicación	Meta-cognición	Organización Procesamiento
Autogestión en la interacción con el objeto de aprendizaje	¿De qué forma el diseño ofrece al aprendiz el control?	Trayectoria de navegación Términos de problemas o casos	Respuestas programadas Diseño cerrado del objeto	Solución de problemas No aplica
Significación del objeto de aprendizaje	¿Qué tipo de tarea propone en el objeto de aprendizaje?	Interrogación Introducción a problemática	Caso a resolver	Prueba diagnóstica Ninguno
	¿Especifica algún tipo de organizador conceptual?	Mapa conceptual Gráfico o figura	Índice o tabla de contenido Ninguno	Numeración de ordenamiento
	¿Cómo se da la transferencia de conocimientos?	Ejemplos Interpretaciones	Demostraciones Ninguno	Aplicaciones
Creatividad que desarrolla el objeto de aprendizaje.	¿Qué opciones plantea el objeto de aprendizaje para generar el conocimiento propio?	Reconocimiento de intereses Plantea retos	Generación de nuevas preguntas No contiene	Líneas derivadas de aplicación
	¿Qué alternativas brinda para la solución de problemas?	Permite plantear términos de problema Ofrece diversas alternativas de solución	Provee alternativas de solución No contiene planteamiento problemático	Permite planteamiento de soluciones

4.4.3 GESTIÓN DE GRUPOS

En esta sección se describen las funcionalidades que facilitan la gestión de los grupos de búsqueda colaborativa de objetos de aprendizaje, las cuales se detallan:

Creación de grupos. Cualquier persona registrada puede crear un grupo de búsqueda colaborativa de OA. El objetivo de esta funcionalidad es definir un título y una descripción. Quién crea el grupo pasa a convertirse automáticamente en el mediador del mismo (Figura 4.10).

Figura 4.10 Creación de un grupo de búsqueda colaborativa

Visualización de los grupos. Los miembros de un grupo pueden visualizar sus características o eliminarse del mismo. Adicionalmente, el mediador puede acceder a las funciones de edición y panel de los miembros (Figura 4.11).

Edit |
 Panel's members |
 Eliminate

Number: 3

id	Group name	Description	Role	Options
176	Conceptos matemáticos	Este grupo está enfocado a la búsqueda de objetos de aprendizaje orientados a la enseñanza de las matemáticas de conceptos específicos.	Mediator	
213	Diseño de cursos en línea	Este grupo tiene como objetivo principal la búsqueda de recursos educativos para el diseño de cursos en línea.	Mediator	
214	El uso de las TICs en el aula	El objetivo es obtener recursos educativos sobre herramientas tecnológicas que se emplean en el aula.	Member	

Figura 4.11 Listado de grupos a los que pertenece el usuario

El panel de los miembros, es donde el mediador puede realizar diversas actividades de gestión, tales como: la incorporación de nuevos miembros, asignar cambios del rol que desempeñan, así como también, eliminarlos del grupo.

Incorporación de nuevos miembros al grupo. Esta actividad se realiza a través de un buscador por medio del nombre de usuario, apellidos o cuenta de correo. Como resultado se obtiene un listado de las personas coincidentes con los parámetros de la búsqueda. Es importante notar, que cada persona listada tiene asociado los objetos que han publicado, las asignaturas que han registrado y los grupos a los que pertenece. El mediador puede incorporar uno o varios miembros al grupo de forma automática a través de la validación o puede enviarle una invitación por correo electrónico (Figura 4.12).

Add members | **Members' management**

Add members:
For search of users please enter name, username or e-mail address .

Name:

☐ Objects published |
 ☐ Subjects |
 ☐ Groups |
 ☐ Send invitation

Users list: 5

Name	Username	Options	Add members
Zapata Alfonso	zgonzal		☐
Zapata Gonzalez Alfredo Autonomous University of Yucatan	alfredo		☐
Sánchez Zapata Carolina Elizabeth	carolina		☐
Zapata Monica	mzapata		☐
Zapata Gonzalez Margarita	alumno31		☐

Figura 4.12 Incorporar nuevos miembros al grupo

Gestión de los miembros del grupo. Una vez incorporado un nuevo miembro al grupo, el mediador puede realizar las siguientes actividades de gestión (ver Figura 4.13):

- o Cambios de rol de miembro a mediador y viceversa las veces que se requiera.
- o Eliminarlo del grupo y posteriormente si se desea volver a incorporarlo.

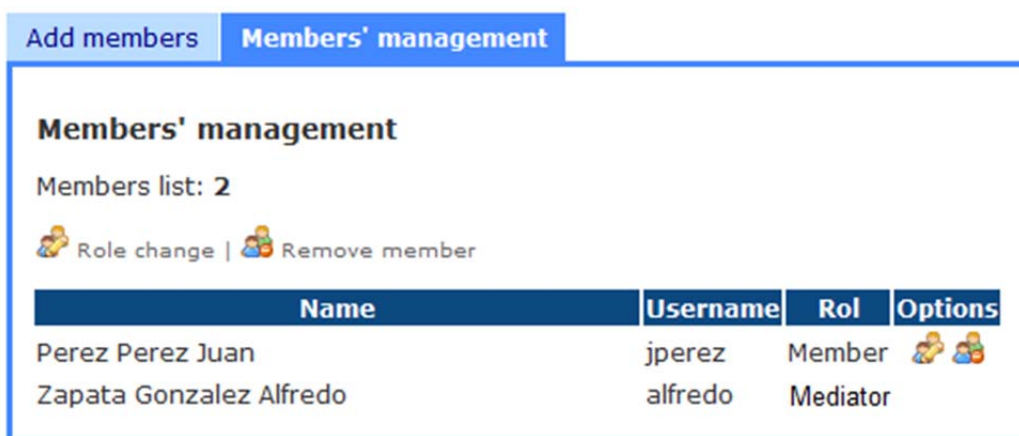


Figura 4.13 Incorporar miembros al grupo

4.4.4 BÚSQUEDA COLABORATIVA DE OBJETOS DE APRENDIZAJE

Acorde a la propuesta planteada en el epígrafe 3.2, en la búsqueda colaborativa, al igual que en la búsqueda individual se consideran parámetros de búsqueda basados en palabras clave, valores de metadatos y criterios de recomendación. Adicionalmente, se han incorporado nuevos elementos al panel de los OA, que están orientados al trabajo colaborativo (ver botón *Add to group* y columna *Collaboration options* de la Figura 4.14).

Score Why? Related LOs Similar LOs Downloads Pedagogical revs. Evaluate Rate Tag Comment									
Learning Objects retrieved list: 3									
Type	Learning Object	Score	Why?	Related Objects	Similar Objects	Downloads	Pedagogical reviews	Evaluate	Collaboration options
	Learning Objects película en flash sobre objetos de aprendizaje Published by User id:43 Add to group			3	3	7	6		
	Un modelo blended Learning características de la modelo Blended learning Published by User id:45 Add to group			1	3	2	1		
	Learning Archivo de vídeo como recurso en formato flash Published by User id:83 Add to group			2	3	2	1		

Figura 4.14 Listado de objetos recomendados en la búsqueda colaborativa

Como se observa en la figura 4.14, cada OA recomendado además de contener la misma información que en la búsqueda individual, incorpora: un nuevo botón *Add to group* que permite añadir o recomendar el objeto a un grupo (Figura 4.15).

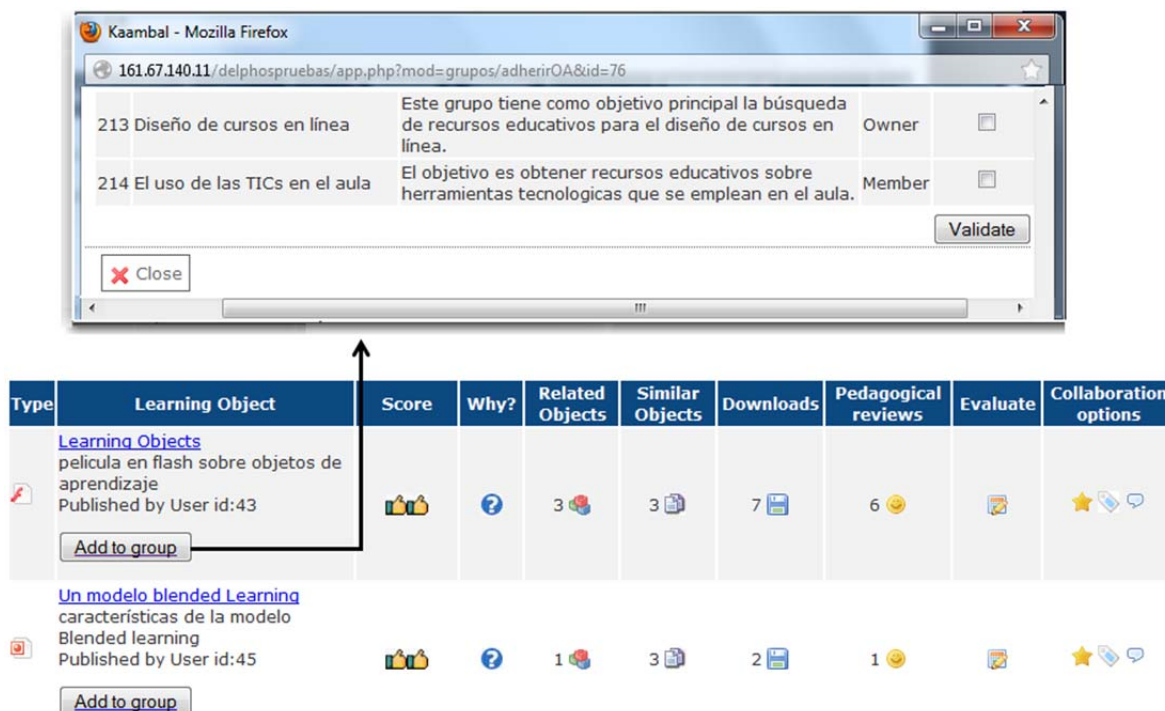


Figura 4.15 Añadir objetos de aprendizaje a los grupos que pertenece el usuario

También integra una nueva columna de iconos (la más a la derecha) con nuevas opciones colaborativas para que el usuario pueda realizar las siguientes actividades:

- **Valorar.** Los miembros pueden valorar a los OA de acuerdo a los intereses y necesidades del grupo (ver Figura 4.16). Esta valoración se representa con una escala de iconos con cinco niveles (★ : No recomendable, ★★ : Poco Recomendable, ★★★ : Recomendable, ★★★★ : Muy Recomendable, ★★★★★ : Muy altamente recomendable).

Recommend if this Learning Object is adjusted to interests and needs of group.

☐ ★★★★★ = Extremely good recommended
☐ ★★★★ = Good recommended
☒ ★★★ = Recommended
☐ ★★ = Low recommended
☐ ★ = Not recommended

Rate global ★★☆☆ (3.0) **Based on 2 rating(s)**

Alfonso Zapata	★★★★
Alfredo Zapata Gonzalez	★★★★
Eyra Gutierrez	★★★★

Group name: Conceptos matemáticos **Rate global** ★★☆☆ (3.0) **Based on 2 rating(s)**

Alfonso Zapata	★★★★
Alfredo Zapata Gonzalez	★★★★

Figura 4.16 Opciones colaborativas de valoración

- **Etiquetar y comentar el objeto.** Los miembros del grupo pueden agregar una o más etiquetas a los objetos de aprendizaje y añadir comentarios para recomendar, criticar o proporcionar información adicional.

Toda esta información extra de la búsqueda colaborativa está disponible desde varios íconos, los cuales se describen en la Tabla 4.4:

Tabla 4.4 Descripción de los iconos extras que contiene la búsqueda colaborativa

Iconos	Descripción	
 Add to group	Add to group	Despliega una ventana donde el usuario puede adicionar el objeto a los distintos grupos que pertenece.
 Rate	Rate	Permite asignarle una valoración en una escala de 5 estrellas.
 Tag	Tag	Permite añadir información a los metadatos.
 Comment	Comment	Permite añadir comentarios.

4.4.5 ACCIONES DEL GRUPO

Cada uno de los grupos contiene a un panel donde los miembros pueden realizar las actividades colaborativas de valoración, las cuales se describen (ver Figura 4.17):

- **Visualización de los OA.** Los miembros pueden visualizar todos los objetos de aprendizaje que han sido adicionados al grupo.
- **Valoraciones individuales.** Los miembros del grupo pueden valorar a los objetos de aprendizaje en este panel, en caso de no haberlo realizado anteriormente. Esta valoración está representada con la misma escala y pueden ser visualizadas por todos los miembros.
- **Valoración final consensuada.** Esta valoración tiene como objetivo involucrar a todos los miembros para tomar una decisión final sobre los OA más acordes a los intereses del grupo. El mediador es el encargado coordinar esta actividad de forma o bien consensuada/democrática (mediante comunicación entre todos los miembros) o bien semiautomática (mediante la utilización de métodos de agregación de votos) y de asignar la valoración final, la cual se representa con una escala de cinco niveles (★ : No interesante, ★ ★ : Poco interesante, ★ ★ ★ : Interesante, ★ ★ ★ ★ : Muy interesante, ★ ★ ★ ★ ★ : Muy altamente interesante). Las decisiones finales pueden ser visualizadas por todos los miembros y pueden modificarse en cualquier momento.
- **Nivel social de los miembros del grupo.** Los miembros del grupo pueden valorar el nivel de relación social que mantienen. Esta valoración se asigna en una escala de cinco niveles (👤 : Nula, 👤 👤 : Baja, 👤 👤 👤 : Normal, 👤 👤 👤 👤 : Alta, 👤 👤 👤 👤 👤 : Muy alta).

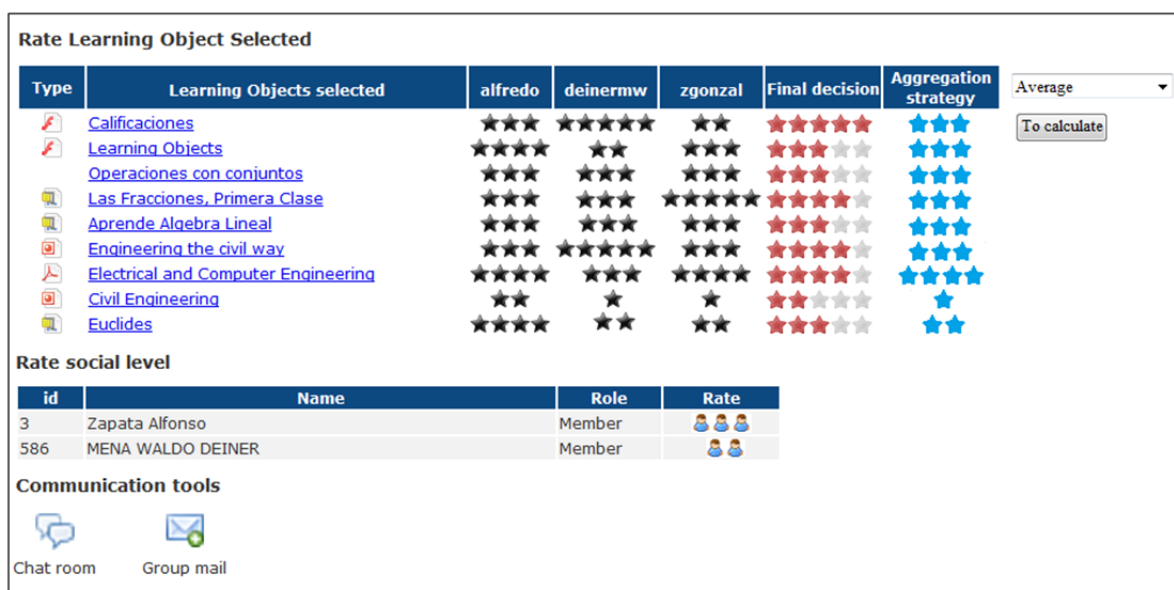


Figura 4.17 Panel de las acciones del grupo

Como se observa en la Figura 4.17, el panel de las acciones del grupo permite al mediador del grupo simular que valoración se obtendría dependiendo del método de agregación que seleccione de entre los 12 disponibles (ver columna *Aggregation Strategy* de la Figura 4.17), la cual se representa también con una escala de cinco niveles.

El mediador del grupo también puede utilizar la opinión de cada uno de los miembros del grupo para llegar a una decisión consensuada sobre la puntuación de cada OA. Para ello, el panel de las acciones del grupo integra herramientas de comunicación exclusivas para sus miembros (ver sección *Communication tools* de la Figura 4.17), las cuales funcionan de forma síncrona como el servicio de mensajería instantánea o chat (ver Figura 4.18) y asíncrona como el correo electrónico, ambas facilitan el intercambio de opiniones para el consenso de la decisión final.

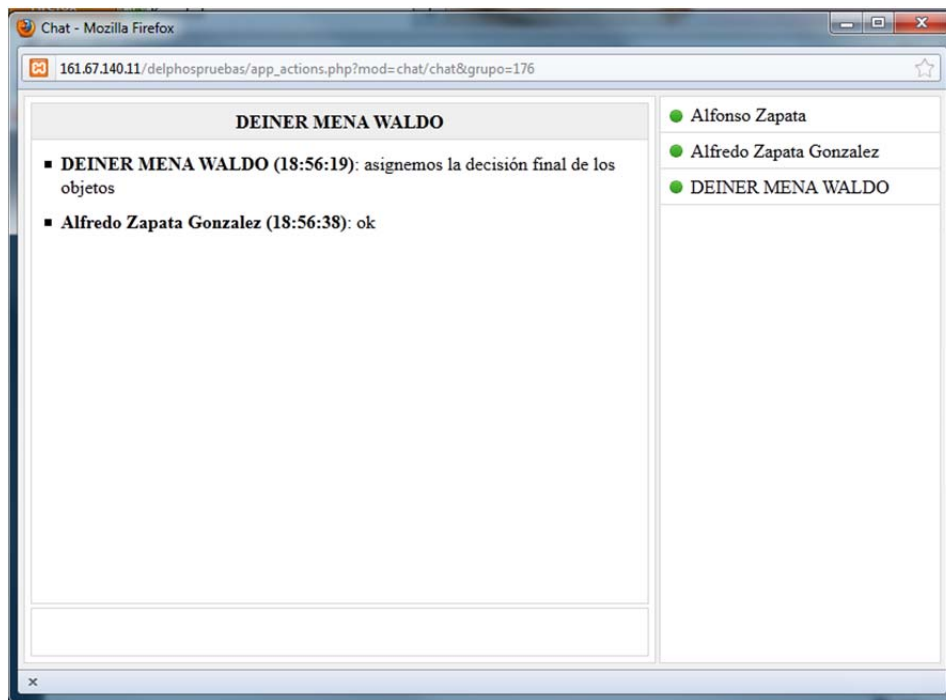


Figura 4.18 Interfaz del servicio de mensajería para los miembros del grupo

4.5 CONCLUSIONES

En este capítulo se ha presentado una descripción completa del *sistema de recomendación híbrido DELPHOS*. Para ello, se ha detallado cada uno de los componentes, tales como: el marco arquitectónico, las especificaciones técnicas y la descripción del entorno. Actualmente el sistema está disponible para su acceso a través de la dirección <http://smile.esi.uclm.es/delphospruebas/>.

Se destaca en forma de resumen que el sistema DELPHOS posee las siguientes funcionalidades:

- Dispone de una interfaz basada en servicios para la interconexión con otras aplicaciones e incorporación de nuevas funcionalidades.
- Está basado en tecnologías y protocolos estándares de libre distribución.
- Proporciona un gran número de características en comparación con otros sistemas similares.

- Se basa en un enfoque de recomendación ponderado híbrido que utiliza diferentes filtrados o criterios recomendación.
- Los valores de los pesos que acompañan a los criterios de recomendación pueden ser asignados por el usuario o pueden ser calculados automáticamente por el sistema.
- Permite la conformación de grupos de búsqueda colaborativa.
- Proporciona un panel de actividades que facilita las valoraciones colaborativas de objetos de aprendizaje y del nivel social de los miembros del grupo.
- Provee mecanismos de comunicación entre los usuarios para las actividades colaborativas.

En el siguiente capítulo, se describen los diferentes experimentos realizados con profesores y alumnos que han utilizado el sistema DELPHOS y han permitido validar la efectividad de la metodología propuesta y la usabilidad de la interfaz Web.

5 EXPERIMENTACIÓN

En este capítulo se describen los experimentos que se han realizado para evaluar el sistema de recomendación híbrido DELPHOS, que implementa el modelo propuesto MRH-OA-IC. Para ello, se ha involucrado tanto a profesores como a alumnos de la Universidad Autónoma de Yucatán. Se han realizado tres pruebas diferentes: la primera utilizando profesores para una validación sobre el correcto funcionamiento del sistema, la segunda utilizando alumnos, para evaluar la recomendación híbrida, y la última utilizando ambos tipos de usuarios para formar grupos y evaluar la recomendación colaborativa.

5.1 PRUEBA DEL SISTEMA DE RECOMENDACIÓN HÍBRIDO CON PROFESORES

Para hacer una primera validación del correcto funcionamiento del sistema se han utilizado un grupo de profesores que han actuado de evaluadores iniciales del software o beta-testers. Para ello, se invitaron a 30 profesores de nivel superior de la Universidad Autónoma de Yucatán, México, los cuales fueron seleccionados con base en su perfil de registro destacando su experiencia en el uso de tecnología ofimática, temas de didáctica y diseño instruccional (Zapata et al., 2011).

5.1.1 RENDIMIENTO

Los profesores dispusieron de una semana para realizar las búsquedas cuyos parámetros incluyeran palabras clave sobre un tema específico de alguna de las asignaturas que imparten, también debían incluir valores de metadatos, sin embargo tenían que deshabilitar los cuatro criterios de recomendación. La búsqueda tenía que devolver resultados para que se considerara como válida para el experimento.

Una vez finalizado el período de búsqueda de los profesores, se procedió a comprobar cuál era el comportamiento de los objetos de aprendizaje recuperados en el ranking aplicando

distintas configuraciones de los criterios de filtrado. A modo de ejemplo, nos basaremos en una búsqueda cuyos parámetros incluyen texto y valores de metadatos, los cuales se muestran en la Tabla 5.1.

Tabla 5.1 Ejemplo de los parámetros de búsqueda en el primer experimento

Parámetros	Valores
Palabra clave	Social network
Idioma	All
Formato de archivo	All
Tipo de recurso	All
Densidad semántica	High
Receptor	All
Contexto	Higher education
Dificultad	Medium

Como resultado de la búsqueda se obtuvo un subconjunto de 6 OA cuyos números de identificación son: 35, 41, 66, 71, 82 y 96 (Tabla 5.2, parte izquierda). Posteriormente, se utilizó sólo el criterio correspondiente aplicándole un peso (W) del 100%, mientras que los otros tres criterios permanecían con un peso de 0% (Tabla 5.2, parte central). Es interesante observar en la Tabla 5.2 como los anteriores 6 OA recuperados poseen una ordenación diferente en cada filtrado. Por tanto, como era de suponer, el tipo de filtrado afecta directamente en el orden de presentación de los objetos. Finalmente, los cuatro criterios de filtrado se aplicaron de forma homogénea, es decir, recibieron un mismo peso del 25% (Tabla 5.2, parte derecha).

Tabla 5.2 Resultados obtenidos al aplicar diferentes criterios de filtrado

OA recuperados		Criterios de filtrado				Ranking OA
Sin aplicar criterios de filtrado		Similitud de contenido Pesos: W1:100%, W2=W3=W4=0%	<u>Uso</u> Pesos: W2:100%, W1=W3=W4=0%	Evaluation Pesos: W3:100%, W1=W2=W4=0%	<u>Similitud de perfil</u> Pesos: W4:100%, W1=W2=W3=0%	Los cuatro criterios de filtrado aplicados weights: W1=W2=W3=W4=25%
1	96	35	41	71	35	35
2	82	82	66	66	41	71
3	71	71	82	41	66	41
4	66	41	96	35	71	82
5	41	66	71	82	82	66
6	35	96	35	96	96	96

Como se observa en la Tabla 5.2, los objetos 35 y 96 han sido ordenados el primero y el último respectivamente, tanto por dos filtrados como en el ranking híbrido final, y de forma inversa cuando no se aplica ningún criterio. Evidentemente, si se aplicara otros pesos distintos, el ranking final sería distinto. Por lo tanto, esta primera validación del funcionamiento del sistema confirmó que puede adaptarse bien a cualquier búsqueda, sin embargo detectó la necesidad de realizar otro estudio experimental para determinar los valores óptimos de los pesos, el cual se describirá en el epígrafe 5.2.

5.1.2 USABILIDAD

Los mismos 30 profesores fueron invitados para dar su opinión acerca de la utilidad del sistema DELPHOS, mediante una encuesta anónima. Para ello, se utilizó el cuestionario denominado *Escala de Usabilidad del Sistema* o SUS (*System Usability Scale*) (Brooke, 1996), la cual provee una escala simple de 10 ítems o preguntas que dan una visión global de las evaluaciones subjetivas sobre la usabilidad (ver Tabla 5.3). Todas las preguntas de este cuestionario requieren una respuesta en una escala Likert de 1 (muy en desacuerdo) a 5 (muy de acuerdo). Como resultado de la validación, se obtuvo el 74.58% de satisfacción sobre la base de la opinión del usuario acerca de las funcionalidades proporcionadas por el sistema.

Tabla 5.3 Preguntas de la Escala de Usabilidad del Sistema (SUS)

Pregunta	Cuestionario SUS
1	Volveré a usar esta aplicación.
2	La aplicación es compleja.
3	La aplicación es fácil de usar.
4	Es necesario el apoyo de un experto para usar la aplicación.
5	Las opciones de la aplicación están bien integradas.
6	La aplicación es inconsistente.
7	Aprendí rápidamente a usar la aplicación.
8	Es incómodo usar la aplicación.
9	Me gustó usar la aplicación.
10	Necesito aprender muchas cosas antes de usar la aplicación.

Los resultados del análisis (véase la Figura 5.1) muestran que los usuarios perciben que el sistema es fácil de usar, y facilita en gran medida las acciones de búsqueda y recuperación de objetos de aprendizaje para adaptarse a sus necesidades. La encuesta incluyó un campo de texto en el que los usuarios pueden expresar sus comentarios y sugerencias; hubo 17 observaciones que dieron lugar a las siguientes mejoras:

- Incluir un componente social que permitiese la posibilidad de valorar, etiquetar y comentar los OA.
- Dotar al entorno del sistema con algunas herramientas de comunicación como el chat y correo electrónico.
- Incluir nuevas funcionalidades orientadas a la búsqueda colaborativa.

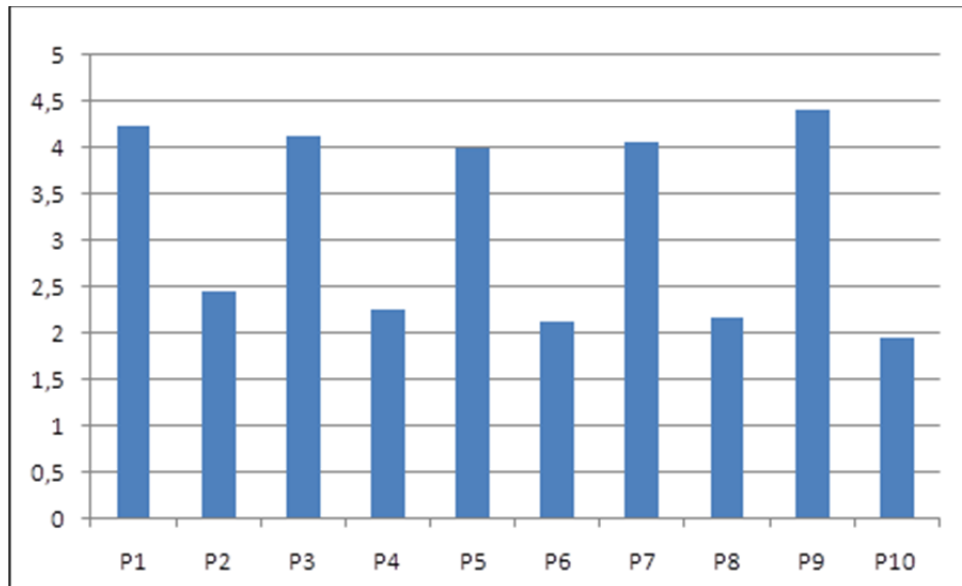


Figura 5.1 Resultados del promedio de las evaluaciones llevadas a cabo por los 30 profesores de acuerdo a la encuesta SUS. Los valores tienen un rango entre 1 (muy en desacuerdo) a 5 (muy de acuerdo).

5.2 PRUEBA DEL SISTEMA DE RECOMENDACIÓN HÍBRIDO CON ALUMNOS

A los efectos de esta prueba, se contó con la participación de un grupo de 24 estudiantes de primer año de la licenciatura en ingeniería civil de la Universidad Autónoma de Yucatán, México. Los experimentos fueron realizados durante las clases prácticas de la asignatura de *Introducción al desarrollo de aplicaciones informáticas* a finales del segundo semestre del año 2012. El profesor de la asignatura introdujo los sistemas DELPHOS y AGORA a los estudiantes en una sesión y en la siguiente los estudiantes los utilizaron para la búsqueda específica de los objetos de aprendizaje, siguiendo las instrucciones proporcionadas por el profesor. Cabe destacar, que actualmente el repositorio AGORA cuenta con aproximadamente 70 objetos de aprendizaje relacionados con el área de la ingeniería eléctrica, civil y ambiental (Zapata et al., 2013).

5.2.1 CRITERIOS DE FILTRADO

En este segundo experimento se muestra cómo el sistema DELPHOS trabaja con diferentes criterios de filtrado. El objetivo fue realizar una evaluación del sistema con los 24 estudiantes de licenciatura y probar el comportamiento en el ranking de recomendación de

los objetos de aprendizaje con diferentes pesos.

Se compararon un total de siete configuraciones de prueba, en cada una de ellas se utilizó la misma búsqueda, pero con diferentes valores para cada uno de los cuatro criterios de recomendación (ver Tabla 5.4). El objetivo de cada prueba se resume a continuación:

- **Prueba 1.** Valida el uso de los pesos adaptativos que se proponen como configuración predeterminada de cada criterio de recomendación y que se calculan automáticamente por el sistema.
- **Pruebas 2, 3, 4 y 5.** Validan el uso de cada uno de los criterios de recomendación de forma independiente o separada, es decir, uno a la vez, de forma que un criterio tiene un peso de 100% cuando los otros tres criterios tienen un peso de 0%.
- **Pruebas 6 y 7.** Validan el uso de valores aleatorios en los pesos de los cuatro criterios de recomendación.

Tabla 5.4 Valores de los criterios de recomendación utilizados en cada configuración de prueba

Criterios de Recomendación	Prueba 1 (%)	Prueba 2 (%)	Prueba 3 (%)	Prueba 4 (%)	Prueba 5 (%)	Prueba 6 (%)	Prueba 7 (%)
Similitud de contenido	55	100	0	0	0	80	10
Uso	73	0	100	0	0	50	90
Evaluación	52	0	0	100	0	65	30
Similitud de perfil	70	0	0	0	100	10	15

En este experimento, todos los estudiantes llevaron a cabo la consulta que querían; es decir, cada estudiante utilizó una búsqueda diferente. En la Tabla 5.5, se muestra un ejemplo de búsqueda utilizada por uno de los estudiantes durante este experimento.

Tabla 5.5 Ejemplo de parámetros de búsqueda utilizados en el experimento

Parámetros	Valores
Palabra clave	Engineering
Idioma	English
Formato de archivo	All
Tipo de recurso	All
Densidad semántica	Medium
Receptor	Learner
Contexto	Higher education
Dificultad	Medium

Por ejemplo, cuando se ejecutó la búsqueda anterior (ver Tabla 5.5) el sistema recuperó un subconjunto de 19 objetos de aprendizaje que se ajustaban a los parámetros declarados. Posteriormente, se aplicaron de forma automática los criterios de filtrado a este subconjunto acorde a todas las configuraciones declaradas de la Tabla 5.4, a fin de obtener el ranking final en cada una de las siete configuraciones de pruebas. Finalmente, en la Tabla 5.6 se muestra el ranking obtenido por los 10 primeros OA, es decir, los 10 OA mejor posicionados en cada prueba.

Tabla 5.6 Ranking de los 10 OA mejor posicionados de las siete configuraciones de pruebas

Prueba 1	Prueba 2	Prueba 3	Prueba 4	Prueba 5	Prueba 6	Prueba 7
OA684	OA666	OA684	OA666	OA689	OA666	OA684
OA666	OA677	OA670	OA670	OA663	OA684	OA670
OA670	OA674	OA666	OA663	OA665	OA670	OA666
OA677	OA668	OA668	OA677	OA669	OA677	OA677
OA668	OA681	OA677	OA681	OA674	OA668	OA668
OA663	OA690	OA679	OA684	OA677	OA679	OA679
OA679	OA679	OA663	OA690	OA680	OA663	OA663
OA674	OA663	OA678	OA679	OA682	OA690	OA674
OA690	OA665	OA674	OA683	OA684	OA681	OA690
OA665	OA684	OA690	OA669	OA683	OA674	OA665

Como puede verse en la Tabla 5.6, todas las configuraciones de prueba muestran diferentes rankings; es decir, la lista de los OA recomendados tiene un orden diferente en cada prueba. Cada OA tiene una posición diferente en el ranking de acuerdo con los criterios de recomendación específicos y sus valores de peso. A modo de ejemplo de cómo las posiciones de los objetos de aprendizaje cambian en cada prueba, se analizó el comportamiento del OA con el número de identificación 684 (seguir las flechas en la Tabla 5.6).

Podemos ver que el OA con el número de identificación 684, se encuentra en la primera o segunda posición en el ranking en las pruebas 1, 3, 6 y 7. Sin embargo, en las pruebas 2, 4 y 5, baja varias posiciones. Tomando en cuenta, que cada estudiante utilizó una búsqueda diferente, se implementó el coeficiente de correlación de Spearman para comprobar si existían relaciones estadísticas entre todos los rankings (diferentes órdenes de OA) obtenidos cuando se realizaron las 24 búsquedas de los estudiantes. El coeficiente de correlación de Spearman o rho de Spearman, es una medida no paramétrica de dependencia estadística entre dos rangos (Fieller et al., 2008). En él se evalúa el nivel de las relaciones entre dos variables o rangos que pueden ser descritos utilizando una función monótona y se interpreta exactamente igual que el coeficiente de correlación de Pearson entre dos rangos de

variables. Con el fin de obtener la matriz de correlación de Spearman, se ha calculado el coeficiente rho de Spearman entre cada par de configuraciones o pruebas (ver Tabla 5.7).

Tabla 5.7 Matriz de correlación del rango de Spearman

	Prueba 1	Prueba 2	Prueba 3	Prueba 4	Prueba 5	Prueba 6	Prueba 7
Prueba 1	1	0.73978*	0.82837**	0.73336*	0.66792*	0.96694**	0.95804**
Prueba 2	0.73978*	1	0.65203*	0.51602	0.41860	0.61851*	0.54475*
Prueba 3	0.82837**	0.65203*	1	0.64226*	0.49907*	0.83584**	0.86936**
Prueba 4	0.73336*	0.51602	0.64226*	1	0.72225**	0.70002**	0.73058**
Prueba 5	0.66792*	0.41860	0.49907*	0.72225**	1	0.61519*	0.54195
Prueba 6	0.96694**	0.61851*	0.83584**	0.70002**	0.61519*	1	0.95796**
Prueba 7	0.95804**	0.54475*	0.86936**	0.73058**	0.54195	0.95796	1

La matriz de correlación de Spearman, al igual que la de Pearson, es simétrica porque la correlación entre dos variables X e Y es la misma que la correlación entre Y y X. Además se realizó la prueba *t de Student* con un valor de $P(p)$ que muestra si las diferencias entre las dos variables puede ser considerada estadísticamente significativa con un nivel de confianza del 99% (** $p < 0,01$) o 95% (* $p < 0,05$).

Como se observa en la Tabla 5.7, todas las pruebas poseen una correlación positiva en el ranking de los OA. Por una parte, la correlación más alta (mayor que 0.9) está entre las configuraciones o pruebas 1, 6 y 7. Se esperaba que esta correlación fuera alta, debido a que las tres configuraciones utilizan valores en los cuatro pesos, es decir, utilizan la información de los cuatro criterios de filtrado. Por otra parte, de entre las otras cuatro configuraciones que sólo utilizan un criterio simultáneamente, destaca la prueba 3 (que utiliza sólo la información de uso), la cual también muestra una correlación alta (mayor que 0,8) con las pruebas 1, 6 y 7. Esto muestra que, en este caso, la información de uso parece ser la información individual más importante de entre toda la información disponible (contenido, uso, evaluación y perfil de usuario) para los propósitos de recomendación.

En conclusión, este experimento muestra cómo el sistema DELPHOS obtiene diferentes rankings en los OA de la misma búsqueda, en función del criterio de recomendación utilizado y sus valores de peso. Además, el sistema recomienda de forma automática el uso

de unos valores de pesos por defecto que el usuario puede utilizar o bien puede priorizar qué criterio de recomendación es el más interesante y en qué porcentaje.

5.2.2 RENDIMIENTO

Para evaluar el rendimiento del sistema de recomendación híbrido se ha realizado un experimento donde se analiza cuáles son los OA más interesantes para cada usuario en particular, y también se determina cuál es la mejor configuración de las 7 configuraciones de pruebas anteriores; es decir, la configuración que devuelve estos OA en las primeras posiciones del ranking de OA. Para determinar el interés del usuario se ha utilizado una puntuación implícita a partir de los *click data* o datos de pulsaciones que almacena el sistema (Jung et al., 2007). Concretamente se van a utilizar las acciones de seleccionar/descargar un OA (del ranking o lista de OA devueltos tras una búsqueda) por parte de los usuarios para conocer los intereses reales de los usuarios, asumiendo que si un usuario pulsa en un OA para descargarlo significa que está interesado en dicho OA.

En el entorno del sistema DELPHOS, los usuarios pueden seleccionar la descarga de uno, varios o ninguno de los OA recomendados y se asume que están interesados en estos OA que se han descargado. El orden o la posición del OA que ha sido descargado, es registrada y almacenada en una tabla o bitácora que contiene todos los OA recomendados en cada consulta realizada por el usuario.

El objetivo de este experimento es comparar y encontrar cuál o cuáles de las configuraciones previamente propuestas de los criterios de recomendación o valores de peso (ver Tabla 5.4) obtiene los mejores resultados en el top del ranking de los OA, es decir, cuando el usuario selecciona o descarga el OA o los OA que se han posicionado más alto o en las primeras posiciones de la lista de OA recomendados.

En este experimento, se les solicitó a los 24 estudiantes buscar objetos relacionados con la ingeniería civil. Concretamente, el tema específico propuesto por el profesor fue el diseño y construcción de puentes. Cada estudiante realizó tres búsquedas diferentes con el objetivo de seleccionar/descargar sólo los OA más interesantes o más relacionados con el tema señalado. Algunos ejemplos de las frases o palabras clave más utilizadas por los alumnos

durante las búsquedas fueron: proyectos de puentes, construcción de puentes, diseño de puentes y obra civil de puentes.

En resumen, los 24 estudiantes ejecutaron un total de 72 búsquedas, utilizando la misma configuración nula. Esta configuración nula significaba que no se han utilizado ninguno de los criterios de filtrado, es decir, los cuatro valores de peso se fijaron en 0%. De esta manera, la lista de OA no se mostraba ordenada de acuerdo a ninguno de los criterios de filtrado. Los estudiantes tuvieron que basarse únicamente en la información descriptiva que acompaña a los OA recuperados para su selección o descarga, debido a que se les indicó que la posición que ocupaban los OA en la lista no era relevante. Para cada búsqueda, se almacenó en una tabla o bitácora tanto la lista de OA mostrada al usuario como los OA seleccionados/descargados por el usuario. Posteriormente y en modo fuera de línea, las siete configuraciones de prueba (ver Tabla 5.4) se calcularon automáticamente a partir de las 72 búsquedas, con el objetivo de obtener la posición del OA o aquellos OA seleccionados/descargados por cada usuario en cada uno de los siete rankings. Finalmente, para evaluar el rendimiento de cada una de la configuración de las pruebas, se han utilizado dos indicadores o medidas ampliamente utilizadas en los sistemas de recomendación de listas de elementos como son la tasa media de éxito recíproco (*average reciprocal hit rate*) y la cobertura (*recall*).

En primer lugar, se calculó la tasa media de éxito recíproco (*ARHR*, *Average Reciprocal Hit Rate*) también conocida como *Mean Reciprocal Rank (MRR)*, la cual se ha utilizado con el fin de comparar la posición de los primeros OA descargados en las siete configuraciones de las pruebas. El *MRR* de cada búsqueda o consulta es el recíproco del rango o posición del primer OA descargado que aparece en la lista o ranking de OA recomendados, o cero, en el caso de que ninguno de los OA recomendados haya sido seleccionado/descargado. La puntuación correspondiente para un conjunto de búsquedas es la media de las posiciones recíprocas de cada una de las búsquedas (Bian et al., 2008), tal como se expresa en la Fórmula 5.1:

$$MRR = \frac{1}{|B|} \sum_{i=1}^{|B|} \frac{1}{rank_i} \quad (5.1)$$

Donde:

- B = El número de búsquedas.
- $Rank_i$ = El rango o posición del primer OA descargado para la búsqueda i .

Los valores más altos de ARHR o MRR indican que la lista o ranking correspondiente contiene los OA seleccionados/descargados en las posiciones más altas o primeras posiciones de la lista, y por tanto, para nuestro problema de recomendación estos valores más altos de ARHR o MRR muestran los mejores rankings.

Los resultados obtenidos en este experimento, se han representado mediante un diagrama de barras verticales, donde se muestra el valor total del MRR obtenido a partir de las 72 búsquedas de los 24 estudiantes con cada una de las siete configuraciones de prueba (ver Figura 5.2).

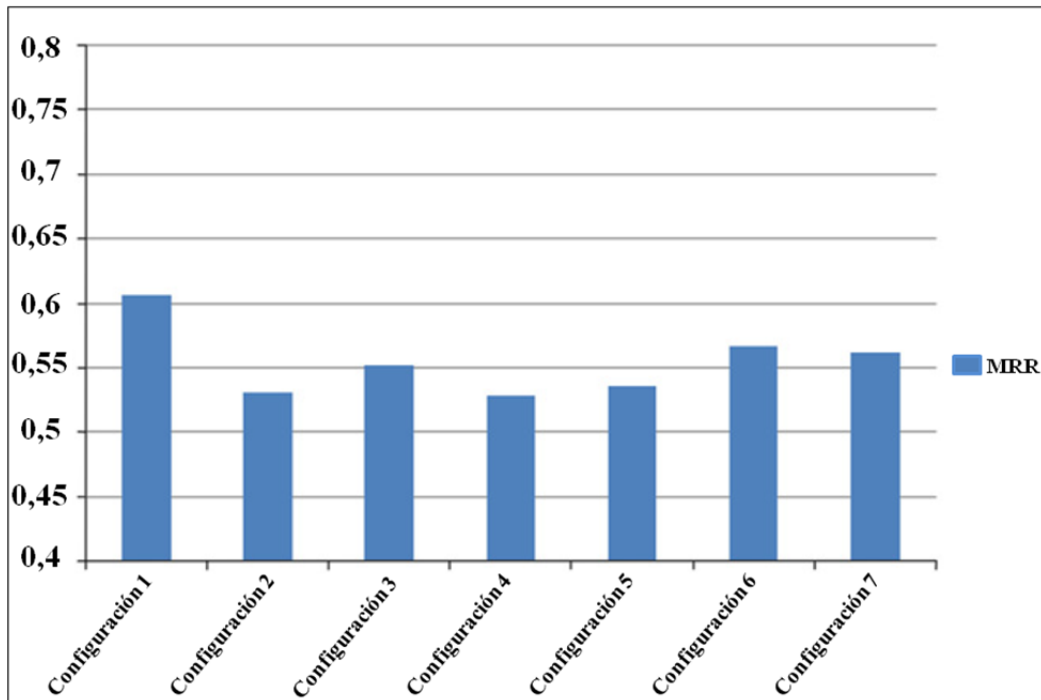


Figura 5.2 Resultados del total MRR obtenido de las configuraciones de las siete pruebas.

Como se observa en la Figura 5.2, la configuración 1 (que es la propuesta de configuración por defecto) obtuvo el resultado más alto del *MRR*, comparado con las otras seis configuraciones que obtuvieron valores muy similares entre ellas. Las configuraciones 6 y 7 obtuvieron la segunda y tercera posición del *MRR*, seguidos por la configuración 3.

En segundo lugar, se calculó el valor del *recall* para los OA que se ubicaban en las posiciones *top-N* (los *N* primeros lugares del ranking), debido a que esta medida se ha utilizado de forma efectiva como una métrica de cuantos resultados relevantes devuelve un sistema de recomendación (Cremonesi et al., 2010). El *recall* para una búsqueda a un nivel *N* se le asocia el valor de 1 si el usuario realiza al menos una selección/descarga del OA en una posición $\leq N$, o cero si el usuario no selecciona/descarga ningún OA de entre los *N* primeros OA. Finalmente, el *recall* total para cada nivel *N* se calcula como el promedio sobre todas las búsquedas y se expresa a través de la Fórmula 5.2:

$$Recall(N) = \frac{\#selecciones}{B} \quad (5.2)$$

Donde:

- *#selecciones* = El número de las selecciones/descargas.
- *B* = El número de búsquedas.
- *N* = La posición o el nivel en la lista de los OA.

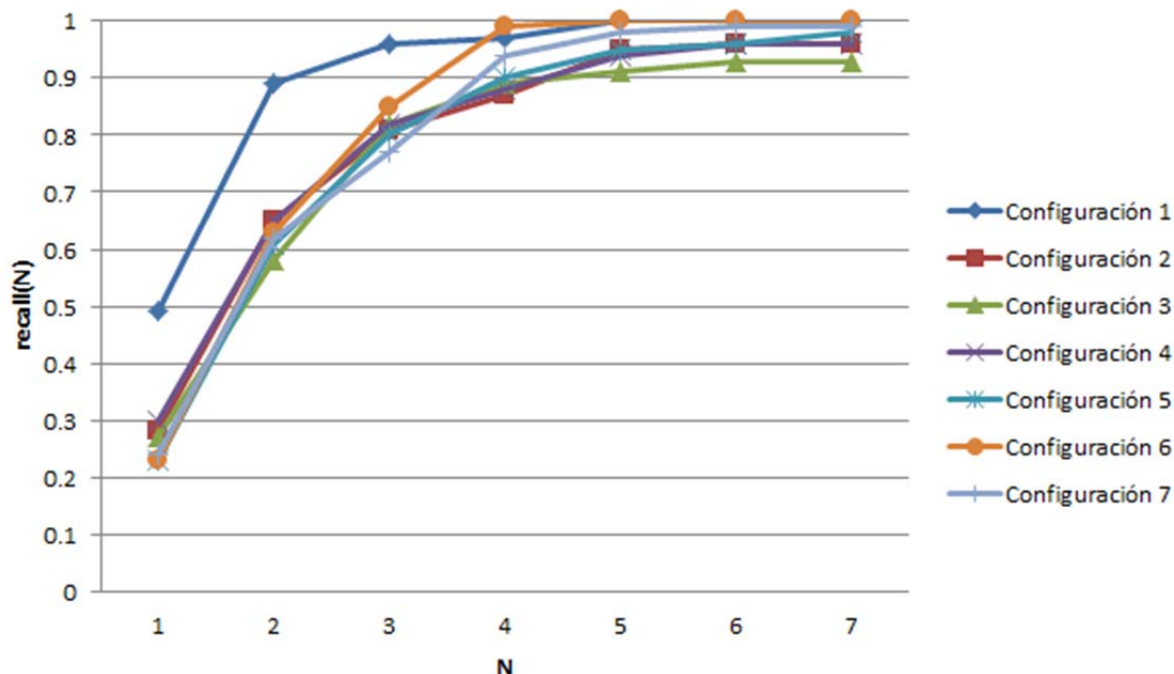


Figura 5.3 Recall de los diferentes top-N.

Como se observa en la Figura 5.3, se ha realizado una comparación entre las siete configuraciones con respecto al *recall* a diferentes *top-N* (se tomaron en cuenta las 7 primeras posiciones consecutivamente). En general, el *recall* incrementó muy rápido al aumentar *N*. De hecho, todas las configuraciones presentaron un comportamiento similar al obtener un valor muy elevado de *recall* (cercano a 1) a partir de *N*=5. Los valores más altos de *recall* se han obtenido de nuevo por la configuración 1, seguido por las configuraciones de 6, 7 y 3.

En conclusión, este experimento ha mostrado que la configuración 1 se comporta mejor que las otras, ya que obtiene el mayor valor de *MRR* y *recall*. Es decir, los valores por defecto de los pesos recomendados (calculados automáticamente por el sistema) han demostrado que puede adaptarse bien a cualquier búsqueda. Esto es muy importante para el sistema DELPHOS, ya que facilita la utilización del sistema de recomendación híbrido mediante el uso de los valores por defecto, siendo de esta forma innecesaria la solicitud del sistema al usuario de los valores de los pesos o filtrados específicos para realizar la búsqueda.

5.2.3 USABILIDAD

Para validar la usabilidad del sistema DELPHOS se han utilizado dos cuestionarios. Para ello, los mismos 24 estudiantes que participaron en los experimentos de los epígrafes 5.2.1 y 5.2.2 fueron invitados a completar ambos cuestionarios, donde debían responder a una serie de preguntas acerca de la usabilidad del sistema (ver Tabla 5.8).

Por un lado, se aplicó de nuevo el cuestionario *SUS* (utilizado en el epígrafe 5.1.2) para obtener una visión global de la usabilidad. Por otro lado, se implementó el *Cuestionario de Usabilidad del Sistema Informático* o CSUQ (*Computer System Usability Questionnaire* (Lewis, 1995), el cual ha sido desarrollado por IBM y se compone de una escala de 19 ítems o preguntas. Se califica por medio de una escala de siete puntos que va desde 1 (muy en desacuerdo) a 7 (muy de acuerdo) y No aplicable (N/A). Finalmente, el grado general de la usabilidad del sistema en cada cuestionario se obtiene promediando las respuestas de todos los estudiantes en un único valor entre 0% y 100%.

Tabla 5.8 Resultados de los cuestionarios SUS y CSUQ

Cuestionario SUS	Promedio
1. Volveré a usar esta aplicación.	4.45
2. La aplicación es compleja.	2.37
3. La aplicación es fácil de usar.	4.16
4. Es necesario el apoyo de un experto para usar la aplicación.	2.08
5. Las opciones de la aplicación están bien integradas.	4.08
6. La aplicación es inconsistente.	1.83
7. Aprendí rápidamente a usar la aplicación.	4.12
8. Es incómodo usar la aplicación.	1.87
9. Me gustó usar la aplicación.	3.95
10. Necesito aprender muchas cosas antes de usar la aplicación.	2.04
Usabilidad	76.46%
Cuestionario CSUQ	Promedio
1. En general, estoy satisfecho con lo fácil que es utilizar el sistema.	4.14
2. Era fácil de usar sistema.	2
3. Pude completar efectivamente mi trabajo usando el sistema.	3.71
4. Soy capaz de completar mi trabajo de forma rápida utilizando el sistema.	4.03
5. Soy capaz de completar mi trabajo de manera eficiente utilizando el sistema.	2.74
6. Me siento cómodo utilizando el sistema.	3.65
7. Era fácil de aprender a utilizar el sistema.	3.8
8. Creo que me convertí rápidamente productivo utilizando el sistema.	2.28
9. El sistema proporciona mensajes de error que claramente me indican cómo solucionar el problema.	3.78
10. Cada vez que cometo un error usando el sistema, puedo recuperarme de forma fácil y rápidamente.	2.31
11. La información (como la ayuda en línea, los mensajes que aparecen en pantalla, y otros documentos) proporcionada por el sistema es clara.	1.69
12. Es fácil encontrar la información que necesitaba.	3.57
13. La información proporcionada por el sistema es fácil de entender.	3.66
14. La información es eficaz para ayudar a completar las tareas y escenarios	2.29
15. La organización de la información en las pantallas del sistema es clara.	3.72
16. La interfaz del sistema es agradable.	2.02
17. Me gusta usar la interfaz de sistema.	2.48
18. El sistema cuenta con todas las funciones y capacidades que espero que tenga.	1.6
19. En general, estoy satisfecho con el sistema.	3.84
Usabilidad	74.25%

Los resultados de los cuestionarios SUS y CSUQ (ver Tabla 5.8), muestran que los usuarios tienen una buena opinión acerca de las funcionalidades proporcionadas por el sistema DELPHOS, al obtener un valor de 76.46% (SUS) y 74.25% (CSUQ). En general, se considera que el sistema es fácil de usar, y facilita en gran medida las acciones de búsqueda y recuperación de OA para satisfacer sus necesidades específicas. Ambos cuestionarios también incluyen un campo de entrada de texto libre donde los usuarios pueden escribir sus comentarios y sugerencias; hubo 11 observaciones las cuales confirmaron la incorporación de nuevas funcionalidades orientadas para la búsqueda colaborativa y relacionadas con proporcionar componentes sociales (etiquetas y comentarios) y herramientas de comunicación (chat y correo electrónico). A partir de estas observaciones, se desarrollaron e implementaron al entorno del sistema: la modalidad de la búsqueda colaborativa, el módulo de gestión de grupos y el panel de actividades del grupo.

5.3 PRUEBA DEL SISTEMA DE RECOMENDACIÓN COLABORATIVO CON GRUPOS

Con el fin de validar el mecanismo de la recomendación colaborativa para grupos, se ha llevado a cabo un experimento con alumnos y profesores de la Universidad de Yucatán. El objetivo de dicho experimento es doble. Por un lado, se pretende validar la funcionalidad de la búsqueda colaborativa. Por otro lado, se pretende predecir cuál es el mejor método de agregación para calcular de forma automática las puntuaciones de cada grupo para cada uno de los OA. De esta forma el sistema DELPHOS podrá recomendar a cada nuevo grupo de reciente creación cuál es el método de agregación más idóneo de acuerdo a las características del grupo y de los miembros que lo conforman.

5.3.1 FORMACIÓN DE GRUPOS

En este experimento se contó con la participación de 71 usuarios (23 profesores y 48 alumnos) de la Facultad de Educación de la Universidad Autónoma de Yucatán, México. El desglose por áreas de investigación y programas de nivel superior a los que pertenecen los profesores y estudiantes respectivamente se muestran en la Tabla 5.9.

Tabla 5.9 Distribución por áreas de investigación y programas de nivel superior de los usuarios

Área de investigación	Número de profesores
Currículum e instrucción	10
Administración y políticas educativas	8
Psicología y educación	5
Programa de nivel superior	Número de estudiantes
Licenciatura en educación (modalidad presencial)	14
Maestría en innovación educativa (modalidad presencial)	10
Maestría en innovación educativa (modalidad en línea)	24

Como se observa en la Tabla 5.9, se contó con la participación de 23 profesores que pertenecen a tres áreas de investigación y 48 estudiantes matriculados en tres programas de nivel superior. Todos ellos eran usuarios de AGORA y se les ha impartido un pequeño curso introductorio sobre DELPHOS.

Para realizar este experimento, se crearon un total de 50 grupos diferentes con el objetivo de realizar una prueba de búsqueda colaborativa de OA con una gran variedad de tipologías diferentes de grupos. Para ello, primero se definieron diversos tamaños de grupos de 2, 3, 4 y 5 miembros (ver Tabla 5.10) y se agruparon a los usuarios de diferentes formas, de manera que se pudiera tener una gran diversidad de tipos distintos de usuarios en los grupos. La distribución final de los 71 usuarios por áreas y programas en los 50 grupos se muestra en la Tabla 5.11.

Tabla 5.10 Distribución de los grupos por tamaño

Tamaño de grupo	Número de grupos
2	20
3	15
4	10
5	5

Tabla 5.11 Configuración de los 50 grupos conformados

Áreas y programas de los miembros de los grupos	ID grupo	Número de miembros
Profesores de una misma área de investigación.	1	5
	2	4
	3	3
	4	3
	5	2
Profesores de diferentes áreas de investigación.	6	5
	7	4
	8	3
	9	3
	10	3
	11	2
Profesores de diferentes áreas de investigación y alumnos de nivel licenciatura en modalidad presencial.	12	2
	13	3
Profesores de diferentes áreas de investigación y alumnos de nivel posgrado en modalidad presencial.	14	2
	15	4
Profesores de una misma área de investigación y alumnos de nivel licenciatura en modalidad presencial.	16	3
	17	3
	18	3
	19	2
	20	4
Profesores de una misma área de investigación y alumnos de nivel posgrado en modalidad presencial.	21	2
	22	3
	23	2
	24	2
	25	4
Profesores de una misma área de investigación y alumnos de nivel posgrado en modalidad en línea.	26	2
	27	3
Alumnos de nivel licenciatura en modalidad presencial.	28	2
	29	4
	30	2
	31	3
	32	4
	33	2
	34	2
Alumnos de nivel posgrado en modalidad presencial.	35	4
	36	2
	37	3
	38	2
	39	5
Alumnos de nivel posgrado en modalidad en línea.	40	5
	41	2
	42	4
	43	3
	44	5
	45	2
	46	2
	47	2
	48	2
	49	4
	50	3

Finalmente, se asignó el rol del mediador a una persona de cada uno de los grupos. El mediador de cada grupo estaba encargado de establecer y anunciar la temática concreta de la búsqueda del grupo. Por ejemplo, el mediador del grupo 1 estableció una búsqueda de OA relacionados con *e-learning*. En cuanto a los pesos de los criterios de filtrado todos los grupos utilizaron los valores por defecto proporcionados por DELPHOS.

Una vez definida la configuración de la búsqueda, los mediadores de los grupos también establecieron unos períodos concretos de tiempo máximo para poder completar las tareas correspondientes de búsqueda y valoración, las cuales se describen:

- **Búsqueda y valoración individual.** Se estableció un tiempo máximo de 2 semanas para que los miembros de los grupos realizaran las búsquedas de los OA relacionados con cada grupo y lo añadieran al grupo correspondiente. Además de añadir los OA a los grupos correspondientes, se tenía que realizar la asignación de las valoraciones o puntuaciones por parte de los miembros. Dichas valoraciones se han representado con una escala de 1 a 5 estrellas (1: No recomendable, 2: Poco recomendable, 3: Recomendable, 4: Muy recomendable y 5: Muy altamente recomendable). Las valoraciones emitidas podían visualizarse por todos los miembros del mismo grupo.
- **Valoración final consensuada.** Se estableció un período máximo de una semana para establecer la valoración final para cada uno de los OA agregados al grupo mediante consenso entre todos los miembros del grupo. Esta actividad es coordinada por el mediador, el cual se encargó de asignar la valoración final tras previo acuerdo. Los miembros de cada grupo contaron con dos herramientas de comunicación como el chat y el correo electrónico para ayudarlos en la toma de decisiones a distancia de forma tanto síncrona como asíncrona, respectivamente.

Finalmente, tras una semana más de aplazamiento, se pudieron obtener todas las valoraciones finales de los OA mediante consenso o método democrático entre los miembros de todos los grupos en un período de un mes.

5.3.2 SELECCIÓN DEL MÉTODO DE AGREGACIÓN

La metodología que se propuso para seleccionar de forma automática el método de agregación de valoraciones idóneo para un nuevo grupo de usuarios está basada en meta-aprendizaje, la cual se describió en el epígrafe 3.2.5. Para ello, inicialmente se creó un conjunto de datos que contenía toda la información recogida de las características de los 50 grupos (ver Tabla 5.12): Actividad (nivel de actividad del grupo), Experiencia (nivel de experiencia del grupo), Social (nivel de contacto social) y Opinión (nivel de disimilaridad del grupo). Para cada una de ellas se ha utilizado tanto el valor de la Media (promedio entre los valores de los miembros) como de la Varianza (desviación de los valores con respecto a la media).

Tabla 5.12 Características de los 50 grupos conformados

ID grupo	Actividad (Media)	Actividad (Varianza)	Experiencia (Media)	Experiencia (Varianza)	Social (Media)	Social (Varianza)	Opinión (Media)	Opinión (Varianza)
1	3	1,5	3,4	0,77	4,3	0,43	2,98	1,61
2	2,25	1,58	3,5	0,4	4,5	0,27	2,80	1,17
3	2,33	0,33	3,5	0,27	5	0	3	0,84
4	2,33	2,33	3,5	0,45	4,16	0,16	2,83	1,17
5	2,5	0,5	3,25	0,78	5	0	2,95	1,43
6	3	2,5	3,5	0,68	3,18	2,29	2,84	1,40
7	2,5	1,66	3,56	0,26	2,58	3,53	2,80	1,39
8	2	1	3,41	0,44	1,66	0,66	2,96	1,46
9	2,66	4,33	3,25	0,93	1,5	0,7	2,66	1,05
10	2,33	2,33	3,25	0,57	1,83	0,16	2,54	1,88
11	3	2	3,62	0,55	3,5	0,5	2,54	1,04
12	3	2	3,37	0,55	2,5	0,5	2,91	1,38
13	2,33	2,33	3	1,09	1,83	0,96	2,86	0,80
14	1,66	0,33	2,83	1,24	1,33	0,26	2,58	1,56
15	2	0,66	3,12	1,18	1,66	1,15	3,08	1,39
16	2,66	1,33	3,25	0,93	1,66	0,66	2,90	0,86
17	1,66	0,33	2,33	0,96	1,33	0,66	3,19	1,13
18	1,33	0,33	2,33	1,15	1,75	2,25	2,88	1,58
19	2,5	0,5	2,75	1,35	1,5	0,5	3,04	1,34
20	2,5	4,5	2,75	1,35	1,5	0,5	2,87	1,15
21	3	2	2,62	1,12	1,5	0,5	2,73	1,40
22	2	1	3,75	0,38	2,16	2,16	2,78	1,48
23	2,66	0,33	3,13	0,98	2,66	1,86	2,80	1,47
24	2,5	0,5	2,87	1,26	2,5	0,5	3,25	1,15
25	3	2	3,5	1,14	1	0	2,87	0,72
26	3	2	2,87	0,98	1	0	3	1,04
27	3,33	2,33	2,66	0,96	1,83	1,36	2,79	1,43
28	2,5	0,5	3	1,14	1	0	3	1,65
29	1,5	0,5	1,87	0,41	1	0	2,96	1,47
30	1,5	0,5	1,75	0,21	2,5	0,5	2,88	1,62
31	1,5	0,5	1,75	0,21	4	0	2,79	1,47
32	1,5	0,5	1,87	0,41	1,5	0,5	2,89	1,80
33	1	0	2	0,57	2,5	0,5	2,57	1,45
34	1,5	0,5	2	0,57	1,5	0,5	2,80	1,44
35	1,5	0,5	1,87	0,41	2	0	2,85	1,68
36	1,5	0,5	2,25	0,5	3,5	0,5	2,87	1,24
37	1,5	0,5	2,62	1,12	2,5	0,5	3	1,52
38	1	0	2,62	0,55	2,5	0,5	2,84	1,49
39	1	0	2,25	0,5	3,5	0,5	3,10	1,43
40	1,5	0,5	2,25	1,07	1,5	0,5	2,96	1,07
41	1,5	0,33	2,43	0,92	1,83	0,33	3,08	1,46
42	1,25	0,25	2,87	1,58	1,91	0,26	3,05	1,14
43	2	1	2,33	0,96	2,66	1,06	3,15	1,34
44	1,5	0,5	2	0,57	2,5	0,5	2,75	1,30
45	2	0	2,75	1,35	2,5	0,5	2,78	1,06
46	1	0	2,37	0,83	4,5	0,5	3,35	1,35
47	2,5	0,5	2,12	0,98	3	2	3,04	1,34
48	1,5	0,5	3,12	1,83	3,5	0,5	3	1,40
49	2	0	2,62	0,55	2,5	0,5	3,03	1,14
50	1,5	0,5	2,62	0,83	2,5	0,5	3,14	1,16

Para cada conjunto de datos se dispone de la decisión final o puntuación consensuada del grupo (valor entre 1 y 5) para cada uno de los OA seleccionados. Este valor representa la valoración real que un grupo proporciona a cada OA. También se calculó de forma simulada los valores que se obtendría al utilizar cada uno de los 12 métodos de agregación de

puntuaciones propuestos en la bibliografía. Este valor representa la valoración predicha por un método para cada OA. Finalmente, se obtuvo el valor del error cometido por cada método de agregación con respecto a la valoración real a través del valor de RMSE (ver fórmula 3.12).

A partir de este valor de RMSE, para cada grupo se ordenaron los 12 métodos de agregación de menor a mayor valor. Por ejemplo, para el grupo con el número de identificación 1 de la Tabla 5.12, el ranking y valor del RMSE obtenido por los métodos fue el siguiente (ver Tabla 5.13):

Tabla 5.13 Ranking obtenido por los métodos de agregación en el grupo 1

Ranking	Método de agregación	RMSE
1	Promedio ponderado (usuario más activo)	0,23299295
2	Mediana	0,26726124
3	Promedio	0,27548658
4	Promedio ponderado (usuario más social)	0,29760952
5	Promedio ponderado (usuario con más experiencia)	0,30705979
6	Voto pluralista	0,53452248
7	Valor mínimo	0,70710678
8	El más votado	0,75592895
9	Equitativo	0,75592895
10	Votación por aprobación_(umbral 2)	0,96362411
11	Promedio descartando el valor mínimo (umbral 2)	1,65156852
12	Votación por aprobación (umbral_3)	2,05287255

A partir de este ranking, se seleccionó sólo el primer método (el de menor porcentaje RMSE) como el mejor método de valoración para el grupo y también para recomendárselo a un nuevo grupo que tuviera características similares. Para poder realizar esta recomendación se utilizó un nuevo conjunto de datos que contenía todos los metadatos, es decir, tanto las características de los grupos (Tabla 5.12) como el rendimiento de los métodos de agregación (en nuestro caso el primer método del ranking que sería el que se recomienda al usuario). A partir de este conjunto de datos, se puede predecir cuál es el mejor método de agregación que debe de utilizar un nuevo grupo de usuarios mediante clasificación.

En este problema de predicción o clasificación la clase o atributo a predecir es precisamente el método de agregación que obtiene el mejor ranking. Para ello se van a utilizar diferentes algoritmos de clasificación proporcionados por el software WEKA (Witten & Frank, 2005), el cual fue seleccionado para este experimento por tres motivos: primero, contiene un código de distribución libre, segundo, dispone de una gran cantidad de algoritmos de clasificación y finalmente es una de las herramientas más populares y utilizadas de minería de datos.

Concretamente, se ha elegido un algoritmo representativo de cada uno de los distintos tipos de algoritmos de clasificación disponibles en WEKA:

- **JRip del tipo reglas.** Es una versión optimizada del algoritmo de reglas de asociación *IREP*, el cual implementa reglas por medio de listas de decisión (Cohen, 1995).
- **J48 del tipo árboles.** Es una implementación del algoritmo C4.5, el cual basa la inducción de su clasificador en un conjunto de árboles de decisión (Quinlan, 1993).
- **NaiveBayesSimple del tipo bayes.** Es un algoritmo clasificador de Bayes con características independientes en el que los atributos numéricos se modelan mediante una distribución normal (Duda et al., 2000).
- **SMO del tipo funciones.** Es un algoritmo secuencial específico de optimización mínima para entrenar un clasificador de vectores de soporte (Platt, 1999).
- **IBK del tipo relajado.** Es una implementación de un algoritmo KNN, el cual propone que un nuevo caso se va a clasificar en la clase más frecuente a la que pertenecen sus K vecinos más cercanos (Aha et al., 1991).

Se han ejecutado todos los anteriores algoritmos de clasificación utilizando los valores de sus parámetros por defecto y una validación cruzada de 10 iteraciones. Esto significa, que el conjunto de datos se divide aleatoriamente en 10 subconjuntos disjuntos de igual tamaño de una manera estratificada (manteniendo la distribución original de la clase). El algoritmo es ejecutado 10 veces y en cada repetición, uno de los 10 subconjuntos se utiliza como el conjunto de prueba y los otros 9 subconjuntos se combinan para formar el conjunto de entrenamiento. Los valores de rendimiento de clasificación se calculan promediando los resultados obtenidos en los conjuntos de entrenamiento.

Finalmente para evaluar los resultados obtenidos con los 5 algoritmos y determinar el algoritmo de clasificación que mejor se comporta en este problema con nuestros datos, se han utilizado dos medidas que han sido recientemente propuestas precisamente para evaluar la recomendación de algoritmos de clasificación (Song et al., 2012). La primera medida, se denomina *ARE* (*Average Recommendation Error*) que mide el error promedio de la recomendación actual (método de agregación predicho) con respecto a la mejor y peor recomendación (mejor y peor método de agregación de la lista de métodos ordenada de menor a mayor RMSE), tal como se expresa en la Fórmula 5.3:

$$ARE = \frac{1}{|G|} \sum_{i=1}^{|G|} \frac{ERROR_{actual_i} - ERROR_{peor_i}}{ERROR_{mejor_i} - ERROR_{peor_i}} \quad (5.3)$$

Donde:

- G = El número de instancias o grupos.
- $ERROR_{actual_i}$ = Es el error del método de agregación predicho por el algoritmo para el grupo.
- $ERROR_{peor_i}$ = Es el error del método de agregación posicionado en el lugar más bajo del ranking del grupo.
- $ERROR_{mejor_i}$ = Es el error del método de agregación posicionado en primer lugar del ranking del grupo.

Por tanto, a menor ARE mejor será el algoritmo de clasificación/predicción utilizado.

La segunda medida utilizada es el *Average Reciprocal Hit Rate* también conocida como *Mean Reciprocal Rank (MRR)*, que mide la posición media que ocupa el método actualmente predicho para cada uno de los grupos en la lista completa de métodos ordenada por RMSE, y se calcula a través de la Fórmula 5.4:

$$MRR = \frac{1}{|G|} \sum_{i=1}^{|G|} \frac{1}{rank_i} \quad (5.4)$$

Donde:

- G = El número de instancias o grupos
- $Rank_i$ = El rango o posición que ocupa el método actualmente recomendado en la lista del grupo i .

De forma que a posiciones más altas del ranking o menor MRR, mejor será el algoritmo de clasificación/predicción utilizado.

A continuación se muestran los resultados obtenidos para el ARE (ver Figura 5.4), como para el MRR (ver Figura 5.5) obtenidos por los 5 algoritmos de clasificación utilizados.

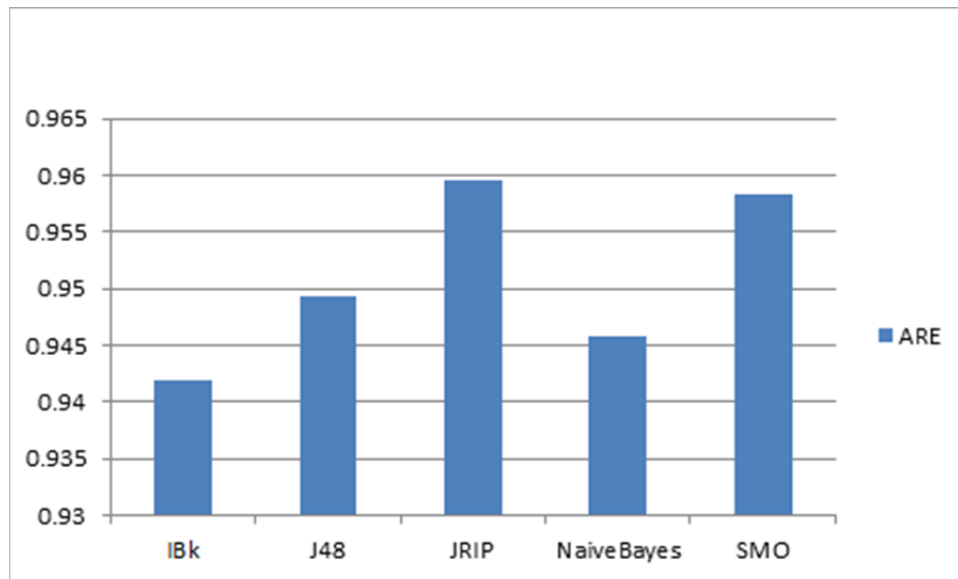


Figura 5.4 Comparativa del Average Recommendation Error obtenido por los algoritmos de clasificación.

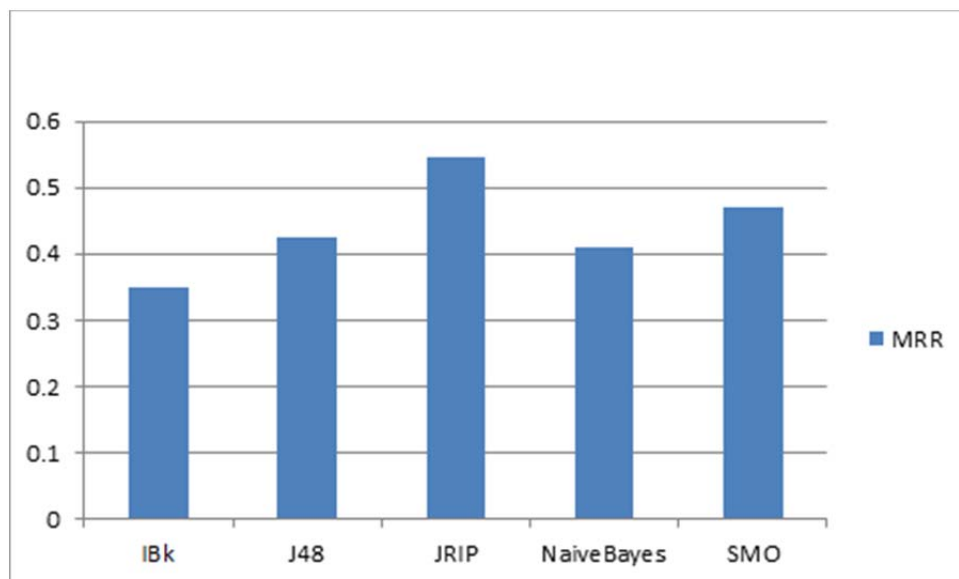


Figura 5.5 Comparativa del Mean Reciprocal Rank obtenida por los algoritmos de clasificación.

Como se observa en las Figuras 5.4 y 5.5, el algoritmo que obtuvo los mejores resultados (menor valores de ARE y MRR) fue el IBK, seguido del NaiveBayes y después el J48. Sin embargo, debido a que se ha utilizado la versión más simple del IBK (con $K=1$) que sólo tiene en cuenta el vecino más cercano, se pretende realizar más pruebas con distintos tamaños de K , para determinar si mejores resultados se pueden obtener con un mayor número de vecinos considerados.

Finalmente, como el algoritmo IBK es el que ha obtenido los mejores resultados, es el que se debe de utilizar como método de clasificación/predicción (basado en el vecino o grupo más cercano) para recomendar a cada nuevo grupo el método de agregación de las valoraciones finales de los OA de forma automática. Prescindiendo de esta manera si el moderador lo desea de realizar la decisión/votación final consensuada y utilizando en su lugar la predicha por el método de agregación seleccionado por el algoritmo IBK.

5.4 CONCLUSIONES

En este capítulo se han descrito los experimentos y resultados obtenidos al evaluar el MRH-OA-IC implementado en el sistema de recomendación híbrido DELPHOS. Para ello, se ha contado con la participación de profesores y alumnos registrados en el sistema.

Los resultados obtenidos en los experimentos muestran en general un impacto positivo en el enfoque presentado en esta tesis. Por una parte, se ha comprobado que los valores por defecto de los criterios de recomendación pueden adaptarse bien a cualquier tipo de búsqueda. Por otra parte, los participantes de las pruebas han retroalimentado con sus comentarios sobre el entorno del sistema y contribuyeron al desarrollo de nuevas funcionalidades. Finalmente, las pruebas de la recomendación colaborativa permitieron obtener resultados preliminares sobre el algoritmo más adecuado para efectuar las recomendaciones a los grupos de búsqueda colaborativa de OA.

6 CONCLUSIONES Y TRABAJO FUTURO

Este capítulo tiene como finalidad presentar una valoración global de los resultados alcanzados con el desarrollo de este trabajo. Se inicia analizando la consecución de los objetivos planteados, para proseguir con una revisión de las principales aportaciones generadas como resultado de la investigación. También se presentan las distintas publicaciones realizadas en estos últimos años que corroboran los resultados obtenidos. Finalmente, se establecen las líneas futuras de investigación.

6.1 ANÁLISIS DE LA CONSECUCIÓN DE OBJETIVOS

En el primer capítulo de esta tesis (epígrafe 1.2), se han definido varios objetivos parciales con la finalidad de lograr el objetivo central del marco de la tesis que es:

Objetivo general

Establecer un modelo híbrido de recomendación de objetos de aprendizaje tanto para un sólo usuario individual como para un grupo de usuarios. Se considera un modelo híbrido, en el sentido de utilizar un modelo de conocimiento que permita describir tanto las características de los objetos como los perfiles de los usuarios utilizando simultáneamente diversas técnicas de filtrado y ordenación o ranking.

A continuación se presentan, para cada uno de los objetivos parciales, los elementos que permiten verificar su cumplimiento.

1. Explorar las diversas fuentes de información que almacenan los OA.

Las fuentes de información tradicionales que almacenan los OA son los metadatos. En esta tesis se propusieron otras fuentes provenientes de la interacción que se produce entre los usuarios y los propios objetos. Como ejemplo, pueden señalarse: las actividades de gestión (publicación y descarga de OA), las evaluaciones de calidad pedagógica, las valoraciones asignadas de acuerdo a una temática específica, las etiquetas y los comentarios. La utilización de toda esta información ha permitido mejorar y personalizar las búsquedas, de forma que se puedan ofrecer recomendaciones basadas no sólo en objetos de contenidos similar, como normalmente haría un sistema de búsqueda tradicional, sino también basada en los objetos que han sido más utilizados, o bien los mejor valorados o incluso en los objetos creados por usuarios afines a nosotros mismos (*epígrafe 3.1.2*).

2. Implementar un mecanismo de recomendación de OA para el establecimiento de una ordenación o rankings entre los objetos seleccionados.

Se implementó un mecanismo de recomendación de OA, el cual contempla las modalidades de búsqueda individual y colaborativa. Para ello, se utilizaron distintas fuentes de información provenientes de los OA, los perfiles de registro de los usuarios y las interacciones entre ambos. Este mecanismo está conformado por cuatro criterios de recomendación basados en: similitud de contenido, uso, evaluación y similitud de perfil de usuario. Para cada uno de ellos, se desarrolló un conjunto de métricas que interactúan de manera efectiva para priorizar la recuperación de OA de acuerdo a los intereses del usuario. Adicionalmente, a cada uno de los criterios de recomendación se incorporaron pesos adaptivos y dinámicos que representan la densidad de los valores asociados a ese filtrado (*epígrafes 3.1 y 3.2*).

3. Proponer un método para ayudar a los profesores y alumnos en la búsqueda individual de los OA más apropiados para cada uno.

Se ha propuesto una metodología para la búsqueda individual de objetos de aprendizaje. Para ello, se han considerado las modalidades simple y compuesta, de

modo que cada una se realiza de forma independiente y a través de métodos que involucran métricas y filtrados de información que se ajustan a las necesidades del solicitante (*epígrafes 3.1.1 y 3.1.2*).

- 4. Proponer un método de búsqueda colaborativa de OA que permita ayudar a la creación de grupos de usuarios con intereses comunes así como la recomendación de OA para dichos grupos.**

Se implementó una metodología para la búsqueda colaborativa de OA, la cual está conformada de los siguientes elementos: un módulo para la gestión de grupos, un componente social (etiquetas y comentarios) y métodos de agregación para determinar la valoración final de los objetos seleccionados (*epígrafe 3.2*).

- 5. Implementar una metodología capaz de adaptarse a diferentes entornos de e-learning, como repositorios o sistemas de gestión del aprendizaje.**

Se ha desarrollado una metodología de recomendación híbrida para la búsqueda individual y colaborativa de OA, la cual se ha implementado en el sistema de recomendación híbrido DELPHOS, una aplicación software orientada a la asistencia de los usuarios en las tareas de búsqueda, selección y valoración de objetos de aprendizaje (*epígrafe 4*).

- 6. Definir un conjunto de instrumentos e indicadores para medir la eficiencia y efectividad de la propuesta.**

Se definió un conjunto de experimentos para validar la propuesta. Para ello, se utilizaron medidas para comprobar la eficiencia y el rendimiento de las recomendaciones. También, se implementaron cuestionarios para medir la usabilidad de la interfaz del sistema DELPHOS (*epígrafe 5*).

- 7. Realizar un conjunto de experimentos basados en cantidades suficientes de datos capturados de las actividades de gestión realizada por profesores y alumnos en condiciones reales.**

Se realizaron pruebas experimentales con profesores y alumnos de la Universidad Autónoma de Yucatán, México, en las que fueron validadas las distintas modalidades

de búsqueda. Los resultados confirmaron el rendimiento efectivo y la usabilidad del sistema desde el punto de vista de los usuarios. También se consiguieron resultados preliminares para la recomendación colaborativa y se visualizan nuevas versiones de los experimentos (*epígrafe 5*).

Considerando que se han logrado los objetivos parciales, puede decirse que el objetivo general de este trabajo se ha cumplido, puesto que se ha definido, desarrollado y evaluado un modelo híbrido de recomendación de OA.

La veracidad de la hipótesis presentada en el *epígrafe 1* se basaba en el cumplimiento del objetivo general. Así pues, a partir de los resultados obtenidos, puede concluirse que la hipótesis ha quedado demostrada y por tanto puede afirmarse que

Hipótesis inicial

Los sistemas de recomendación, basados en fuentes de información asociadas a los objetos de Aprendizaje y perfiles de usuario, optimizan la búsqueda, selección y recuperación de dichos objetos de aprendizaje.

6.2 ANÁLISIS DE LAS APORTACIONES

En esta tesis se han presentado los distintos elementos que conforman la propuesta y cuya realización han permitido lograr el objetivo planteado. A través de la investigación se han obtenido varias aportaciones conceptuales y prácticas, las cuales se resumen a continuación:

- Se diseñó un modelo híbrido de recomendación de objetos de aprendizaje, el cual está orientado en mejorar los mecanismos de recuperación y clasificación de objetos de aprendizaje en repositorios.
- Se realizó un estudio sistemático del estado del arte en temas afines a la línea de investigación, como es el caso de los objetos de aprendizaje, los métodos de filtrado de información y sus aplicaciones en entornos educativos, entre otros. Se destaca la comparativa realizada entre 14 sistemas de recomendación de OA, la cual permitió conocer a fondo las principales características que consideran estos sistemas y poder

detectar los puntos de mejora.

- Se planteó una metodología de extracción del conocimiento almacenada en las distintas fuentes de información asociadas a los OA. Para ello, se han implementado cuatro criterios de filtrado de información basados en: similitud por contenido, uso, evaluaciones pedagógicas y similitud de perfil de usuario.
- Se formuló un enfoque de la búsqueda individual simple y compuesta de objetos de aprendizaje. Para ello, se desarrollaron un conjunto de métricas para establecer el ranking de los OA de acuerdo a los criterios de filtrado de información, los cuales son multiplicados de forma correspondiente por un peso que representa la densidad de los valores asociados a ese filtrado.
- Se propuso la obtención automática de estos pesos de manera adaptativa y dinámica. Es decir, el valor de estos pesos se ajusta o adapta a la densidad de información disponible sobre cada criterio de filtrado.
- Se planteó un enfoque de búsqueda colaborativa, cuyo objetivo es la conformación de grupos con intereses comunes y recomendaciones de objetos de aprendizaje para los miembros de dichos grupos.
- Se diseñó un modelo de clasificación/recomendación para predecir el mejor método de agregación para un grupo de reciente creación, en función de las características de los miembros que lo componen y sus valoraciones individuales.

Las aportaciones prácticas se constituyen de la siguiente forma:

- Se diseñó una arquitectura abierta y distribuida orientada a la recomendación de objetos de aprendizaje, la cual contiene tres capas correspondientes a las aplicaciones, los servicios y la gestión de la información.
- Se implementó el sistema de recomendación híbrido DELPHOS, un sistema que facilita la búsqueda de objetos de aprendizaje en repositorios. Una primera versión se encuentra disponible para su acceso a través de la dirección <http://smile.esi.uclm.es/delphospruebas/>.

- En el entorno del sistema DELPHOS, se implementaron funcionalidades para la búsqueda colaborativa, tales como: un módulo para la gestión de grupos y un panel que facilita las actividades colaborativas de los miembros del grupo.
- Se diseñaron e implementaron experimentos que validaron la efectividad y el rendimiento de la búsqueda individual y colaborativa. También, se evaluó la usabilidad de la interfaz de usuario. Para ello, se contó con la participación de profesores y alumnos registrados en el sistema.

6.3 DIVULGACIÓN DE RESULTADOS

Con el propósito de divulgar los distintos resultados obtenidos de la investigación, se han generado varias publicaciones, las cuales han sido presentadas como artículos de revista (6), capítulos de libro (5) y ponencias (11) en diferentes foros científicos. A continuación se lista cada uno de ellos agrupados por categorías y ordenados cronológicamente en forma descendente.

6.3.1 ARTÍCULOS EN REVISTAS INDIZADAS EN EL JOURNAL CITATION REPORTS

1. Zapata, A. Menéndez, V.H., Prieto, M.E., Romero, C. (2013). A Framework for Recommendation in Learning Object Repositories: An Example of Application in Civil Engineering. *Advances in Engineering Software*, 56, 1–14. doi:<http://www.sciencedirect.com/science/article/pii/S0965997812001433> (*factor de impacto 1.092*).
2. Ferreira-Satler, M., Romero, F.P., Menéndez, V.H., Zapata, A. Prieto, M.E. (2012). Fuzzy Ontologies and User Profiles applied to enhance e-Learning Activities. *Soft Computing*. DOI 10.1007/s00500-011-0788-y (*factor de impacto 1.88*).

6.3.2 ARTÍCULOS EN OTRAS REVISTAS

1. Zapata, A., Menéndez, V.H., Prieto, M.E., Romero, C. (2011). A Hybrid Recommender Method for Learning Objects. *IJCA Proceedings on Design and Evaluation of Digital Content for Education* (Vol. 2011, pp. 1-7): Foundation of Computer Science.

2. Menéndez, V.H., Castellanos, M.E., Zapata, A., Prieto, M.E. (2011). Generación de Objetos de aprendizaje empleando un enfoque asistido. *Píxel-Bit. Revista de Medios y Educación*, 38, 141-153.
3. Menéndez, V.H., Castellanos, M.E., Zapata, A., Prieto, M.E., Romero, F.P. (2011). Characterising the Reuse of Learning Objects in a Specialised Repository. *IJCA Proceedings on Design and Evaluation of Digital Content for Education* (Vol. 2011, pp. 20-24): Foundation of Computer Science.
4. Menéndez, V.H, Prieto, M.E, Zapata, A. (2010). Sistemas de Gestión Integral de Objetos de Aprendizaje. *IEEE-RITA. Revista Iberoamericana de Tecnologías del Aprendizaje*, 5(2), 56-62.

6.3.3 CAPÍTULOS EN LIBROS INTERNACIONALES

1. Prieto, M.E., Zapata, A., Menéndez, V.H. (2010). Data Mining Learning Objects. C. Romero, S. Ventura, M. Pechenizkiy y R. Baker (Eds.), *Handbook of Educational Data Mining* (pp. 481-492). Boca Raton, FL: CRC Press.
2. Menéndez, V.H, Zapata, A., Vidal, C., Segura, A., Prieto M.E (2010). An Approach to Metadata Generation for Learning Objects. M. D. Lytras, P. Ordonez De Pablos, A. Ziderman, A. Roulstone, H. Maurer y J.B. Imber (Eds.), *Knowledge Management, Information Systems, E-Learning, and Sustainability Research* (Vol. 111, pp. 190-195). Berlin, Germany: Springer Berlin Heidelberg.
3. Vidal, C., Segura, A., Menéndez, V.H, Zapata, A., Prieto M.E. (2010). Metadata and Ontologies in Learning Resources Design. M. D. Lytras, P. Ordonez De Pablos, A. Ziderman, A. Roulstone, H. Maurer y J.B. Imber (Eds.), *Knowledge Management, Information Systems, E-Learning, and Sustainability Research* (Vol. 111, pp. 105-114). Berlin, Germany: Springer Berlin Heidelberg.
4. Segura, A., Vidal, C., Menéndez, V.H, Zapata, A., Prieto, M.E (2009). Exploring Characterizations of Learning Object Repositories Using Data Mining Techniques. F. Sartori, M. Á. Sicilia y N. Manouselis (Eds.), *Metadata and Semantic Research* (Vol. 46, pp. 215-225). Berlin, Germany: Springer Berlin Heidelberg.

6.3.4 CAPÍTULOS EN LIBROS NACIONALES

1. Prieto, M.E., Zapata, A., Menéndez, V.H. (2012). Modelos de conocimiento para la recomendación de valores de metadatos en Objetos de Aprendizaje. En Álvarez, FJ., Muñoz, A. Ed(s) Avances en Objetos de Aprendizaje, experiencias en redes de colaboración en México. Universidad Autónoma de Aguascalientes. ISBN:9786078227006.

6.3.5 PONENCIAS EN CONGRESOS INTERNACIONALES

1. Menéndez, V.H., Zapata, A., Prieto, M.E., Romero, C., Serrano, J. (2011). A Similarity-based Approach to Enhance Learning Objects Management Systems, 11th International Conference on Intelligent Systems Design and Applications. Córdoba, España (*CORE C*).
2. Zapata, A., Menéndez, V.H., Prieto, M.E., Romero, C. (2011). Using data mining in a recommender system to search for learning objects in repositories, 4th International Conference on Educational Data Mining, Eindhoven, Holanda (*CORE B*).
3. Ferreira-Satler, M., Romero, F.P., Menéndez, V.H., Zapata, A., Prieto, M.E. (2010). A fuzzy ontology approach to represent user profiles in e-Learning Environments. Fuzzy Systems, 2010 IEEE International Conference, Barcelona, España (*CORE A*).
4. Zapata, A., Menéndez, V.H., Prieto, M.E. (2009). Discovering Learning Objects Usability Characteristics. Intelligent Systems Design and Applications ISDA2009. Ninth International Conference on, Pisa, Italia (*CORE C*).

6.3.6 PONENCIAS EN CONGRESOS NACIONALES

1. Zapata, A., Menéndez, V.H., Prieto, M.E., Romero, C. (2012). Búsqueda colaborativa para el apoyo a la recomendación de Objetos de Aprendizaje. En: Multidisciplinary symposium on the design and evaluation of digital content for education (SPDECE-2012), Alicante, España.
2. Menéndez, V.H., Castellanos, M.E., Zapata, A., Prieto, M.E. (2012). La plataforma AGORA. Una propuesta para la gestión e intercambio de Objetos de Aprendizaje.

En: XVII Jornadas de Enseñanza Universitaria de la Informática (JENUI 2012), Cuidad Real, España.

3. Menéndez, V.H, Castellanos, M.E., Zapata, A., Prieto, M.E (2010). Modelo de asistencia para la generación de objetos de aprendizaje a partir de recursos digitales. M.E. Prieto, J.M. Doderó y D.O. Villegas (Eds.), Recursos digitales para la educación y la cultura (Vol. SPDECE, pp. 151-157). Cádiz, España: Universidad de Cádiz.
4. Zapata, A., Ferreira-Satler, M., Menéndez, V.H., Prieto, M.E., Olivas, J.A., Vidal, C. (2010). Un Sistema de recomendaciones basado en técnicas de soft computing para el filtrado de objetos de aprendizaje. III Simposio sobre Lógica Fuzzy y Soft Computing LFSC2010, Valencia, España.
5. Ferreira-Satler, M., Menéndez, V.H., Romero, F.P., Zapata, A., Prieto, M.E. (2010). Ontologías Borrosas para representar Perfiles de Usuario en una Herramienta de Gestión de Objetos de Aprendizaje. Congreso Español sobre Tecnologías y Lógica Fuzzy ESTYLF2010, Huelva, España.
6. Zapata, A., Menéndez, V.H, Prieto, M.E. (2009). Aplicación de Minería de Datos para la extracción de reglas en Objetos de Aprendizaje. M. E. Prieto, S. Sanchez-Alonso, S. Pech y X. S. Ochoa (Eds.), Recursos Digitales para el Aprendizaje (pp. 109-117). Mérida, México: Editorial UADY.
7. Zapata, A., Menéndez, V.H, Eguigure, Y., Prieto, M.E (2009). Quality Evaluation Model for Learning Objects from Pedagogical Perspective. A Case of Study. International Conference of Education, Research and Innovation ICERI2009, Madrid, España.

6.4 TRABAJO A FUTURO

A lo largo del desarrollo de esta tesis se han identificado diferentes líneas de trabajo que permiten la ampliación y mejora de los resultados obtenidos a este momento. Concretamente, pueden identificarse tres líneas principales: la primera se enfoca a las

mejoras en las técnicas utilizadas. Por otro lado, se proyectan mejoras en el sistema de recomendación híbrido DELPHOS y una académica, orientada a la consolidación y divulgación del conocimiento adquirido. A continuación se describen:

6.4.1 MEJORAS EN LAS TÉCNICAS UTILIZADAS

Existen diferentes elementos del marco metodológico presentado en este trabajo que pueden mejorarse. Algunos de ellos son descritos a continuación:

- **Personalización en la búsqueda de objetos de aprendizaje.** Se propone afinar el grado de personalización en el proceso de la búsqueda a través del uso de técnicas de clustering o agrupamiento. En esta implementación, se podría realizar utilizando por ejemplo el algoritmo de clasificación K-Nearest Neighbor (Fix, 1951) a las fuentes de información provenientes de la gestión de los OA y los perfiles de usuario (Zapata et al., 2011).
- **Gestión de grupos de búsqueda colaborativa.** La metodología propuesta contempla un módulo de gestión de grupos, donde se define que el mediador del grupo envíe las peticiones de incorporación a otros usuarios para formar parte del mismo. Estas peticiones podrían realizarse a través de sugerencias del sistema, las cuales estarían basadas de información implícita proporcionada por los usuarios, los OA y las interacciones entre ambos. Para ello, se plantean diferentes aproximaciones:
 - o *Minería de datos.* Se propone implementar algoritmos de agrupamiento como el K-Means (Macqueen, 1967), para crear grupos de forma automática y asignar a un nuevo posible miembro a grupos existentes.
 - o *Similitud de perfil.* Basándose en el perfil de registro del mediador del grupo, pueden generarse sugerencias para la incorporación de nuevos miembros a partir de la similitud de sus atributos. También se podría utilizar la información disponible de los usuarios en las redes sociales, tales como: Facebook, Twitter ó redes de tipo profesionales/académica como: LinkedIn, ResearchGate, entre otras. A través de la interfaz de programación de

aplicaciones de cada una de estas redes sociales, se puede extraer conocimiento sobre las relaciones existentes entre los usuarios.

- **Método de agregación para las valoraciones finales de OA.** El experimento desarrollado en el *epígrafe 5.3.2*, permitió obtener resultados preliminares sobre el algoritmo más adecuado para efectuar las recomendaciones a los grupos de búsqueda colaborativa de OA. Sin embargo, se plantea una nueva fase de pruebas basadas en dos vertientes:
 - o *El algoritmo IBK.* Los mejores resultados del experimento se obtuvieron con este algoritmo utilizándolo en su versión más simple (con $K=1$). Se prevee ampliar las pruebas utilizando distintos tamaños de K en un rango de [2 a 10] vecinos, con el objetivo de optimizar sus resultados y descubrir el mejor valor de K para nuestros datos.
 - o *Ampliar la muestra del experimento.* Se desea llevar cabo más pruebas con una mayor cantidad y diversidad de grupos para realizar para búsqueda colaborativa de OA. También, se considerarán nuevos algoritmos de clasificación para la etapa de meta-aprendizaje.

6.4.2 MEJORAS EN EL SISTEMA DE RECOMENDACIÓN HÍBRIDO DELPHOS

El sistema DELPHOS ha resultado ser una herramienta muy útil, debido a que facilita a los usuarios realizar diversas actividades individuales y colaborativas para la búsqueda, recuperación y valoración de objetos de aprendizaje. Su arquitectura permite extender los servicios para implementar mejoras, algunas de las cuales se listan a continuación:

- **Interfaz adaptativa a dispositivos móviles.** Ante la diversidad de dispositivos móviles que hoy en día se utilizan para acceder a Internet, se hace necesario desarrollar interfaces adaptativas que puedan ajustarse a sus limitaciones de tamaño para visualizar información. Un rediseño en la interfaz de usuario de DELPHOS, permitiría incorporar estas características, de manera que la presentación de información sea flexible, tanto en relación a las necesidades del usuario, como a las características del dispositivo o navegador utilizado para acceder a la aplicación.

- **Fortalecer la interoperabilidad con otros sistemas de gestión.** El sistema DELPHOS está basado en tecnologías y protocolos estándares de libre distribución y fue diseñado para implementarse en entornos reales de repositorios de objetos de aprendizaje. Se plantea desarrollar un componente que permita integrarse al entorno de sistemas de gestión del aprendizaje como Moodle o Dokeos.
- **Integrar el modelo de clasificación seleccionado mediante meta-aprendizaje dentro de DELPHOS.** Se implementará un clasificador basado en los vecinos más cercanos, el cual permita determinar que técnica de agregación de valoraciones (promedio, mediana, equitativo, valor mínimo, promedio descartando el valor mínimo, el más votado, voto pluralista, votación por aprobación y promedio ponderado) funciona mejor para cada tipo de grupo.

6.4.3 DIVULGACIÓN DEL CONOCIMIENTO ADQUIRIDO

La línea de investigación desarrollada contribuirá a la identificación de problemas, la generación de nuevas líneas de investigación y la implementación de métodos desde una perspectiva científica. Por otra parte, permitirá apoyar desde la docencia a los diversos programas académicos que se imparten en la Facultad de Educación de la Universidad Autónoma de Yucatán, México. A continuación se describen las principales acciones concretas que se desarrollarán:

- Mantener y fortalecer los vínculos creados con los grupos de investigación SMILe y KDIS, para fomentar el desarrollo de la línea de investigación que sustenta esta tesis, así como el desarrollo de nuevos proyectos.
- Desarrollar proyectos de investigación con la participación de grupos multidisciplinarios en México.
- Consolidar un grupo de desarrollo de tecnologías e-Learning que motiven la participación de estudiantes y profesores en actividades formales de investigación.
- Realizar actividades de tutoría para el desarrollo de tesis de posgrado.
- Analizar el impacto de las nuevas tecnologías en la mejora de los procesos

académicos y administrativos, haciendo un uso eficiente de la tecnología existente.

- Continuar publicando en conferencias y revistas internacionales. De hecho, actualmente se está terminado de escribir un artículo titulado: "An approach of Collaborative search to support learning object reusability" que se va a enviar a la revista: International Journal of Human Computer Studies que tiene un factor de impacto de 1.171.

BIBLIOGRAFÍA

ADL. (2004). Sharable Content Object Reference Model 2004.

Agrawal, R., Imieliński, T., Swami, A. (1993). Mining association rules between sets of items in large databases. *Proceedings of the 1993 ACM SIGMOD international conference on Management of data - SIGMOD '93* (Vol. 22, pp. 207–216). New York, New York, USA: ACM Press. doi:10.1145/170035.170072

Aha, D., Kibler, D., Albert, M. (1991). Instance-Based Learning Algorithms. *Machine Learning*, 6(1), 37–66.

Al-Khalifa, H.S. (2008). Building an Arabic learning object repository with an ad hoc recommendation engine. *Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services - iiWAS '08* (pp. 390–394). New York, New York, USA: ACM Press. doi:10.1145/1497308.1497378

Al-Khalifa, H.S., Davis, H.C. (2006). The evolution of metadata from standards to semantics in E-learning applications. *Proceedings of the seventeenth conference on Hypertext and hypermedia - HYPERTEXT '06* (p. 69). New York, New York, USA: ACM Press. doi:10.1145/1149941.1149956

Alavi, M., Leidne, D. (2001). Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *Management Information Systems Research Center, University of Minnesota*, 25(1), 107–136.

Alcalá-Fdez, J., Sánchez, L., García, S., Jesus, M.J., Ventura, S., Garrell, J.M., Otero, J., et al. (2008). KEEL: a software tool to assess evolutionary algorithms for data mining problems. *Soft Computing*, 13(3), 307–318. doi:10.1007/s00500-008-0323-y

Amer-yahia, S., Roy, S.B., Chawla, A., Das, G., Yu, C. (2009). Group Recommendation : Semantics and Efficiency *. *Proc. VLDB Endow.*, 2(1), 754–765.

- Amershi, S., Morris, M.R. (2008). CoSearch : A System for Co-located Collaborative Web Search. *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems (CHI '08)* (pp. 1647–1656).
- Anderson, M., Ball, M., Boley, H., Greene, S., Howse, N., McGrath, S., Lemire, D. (2003). RACOFI: A Rule-Aplying Collaborative Filtering System. *Proceedings of COLA'03*.
- Ankur, G., Sanil, P., Jones, Q., Borcea, C. (2004). Automatic identification of informal social groups and places for geo-social recommendations. *International Journal Mobile Network and Innovation*, 2(3/4), 159–171.
- Apache Friends Project. (2002). Xampp. Retrieved November 24, 2012, from <http://www.apachefriends.org/es/xampp.html>
- ARIADNE Foundation. (2001). Retrieved from <http://www.ariadne-eu.org/>.
- Avancini, H., Straccia, U. (2005). User recommendation for collaborative and personalised digital archives. *International Journal of Web Based Communities*, 1(2), 163–175.
- Ba-omar, H., Petrounias, I., Anwar, F. (2007). A Framework for Using Web Usage Mining to Personalise E-learning. *Proc. Int. Conf. Adv. Learn. Technol.*, (Vol. 1, pp. 937–938).
- Basu, C., Ave, P., Room, A., Park, F. (1998). Using Social Recommendation as Classification : and Content-Based Information in Haym Hirsh Recommendation William Cohen. *AAAI Technical Report WS-98-08*.
- Bian, J., Liu, Y., Gichtein, E., Hongyuan, Z. (2008). Finding the Right Facts in the Crowd : Factoid Question Answering over Social Media. *Proceedings of the 17th international conference on World Wide Web* (pp. 467–476).
- Biolchini, J., Mian, P.G., Natali, A.C., Travassos, G.H. (2005). *Systematic Review in Software Engineering. Technical Report ES 679/05*.
- Blanco, J.J., Galisteo del Valle, A., et al. (2008). Perfil de aplicación LOM-ES V.1.0.

- Bohl, O., Schellhase, J., Sengler, R., Winand, P.U. (2002). The Sharable Content Object Reference Model (SCORM) – A Critical Review. *Proceedings International Conference on Computers in Education*. (pp. 17–18).
- Boratto, L., Carta, S. (2010). State-of-the-Art in Group Recommendation and New Approaches for Automatic Identification of Groups. In S. et Al. (Ed.), *Inform. Retrieval and Mining in Distrib. Environments* (pp. 1–20).
- Boratto, L., Carta, S., Chessa, A., Agelli, M., Clemente, M.L. (2009). Group Recommendation with Automatic Identification of Users Communities. *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, 3, 547–550. doi:10.1109/WI-IAT.2009.346
- Bozo, J., Alarc, R., Iribarra, S. (2010). Recommending Learning Objects According to a Teachers ' Contex Model. *Learning*, 470–475.
- Brazdil, P., Giraud-Carrier, C., Soares, C., & Vilalta, R. (2009). *Metalearning: Applications to Data Mining*.
- Brogan, P. (1999). Using the Web for Interactive Teaching and Learning. *Macromedia white paper*, 1–38.
- Brooke, J. (1996). SUS - A quick and dirty usability scale. In P. Jordan, B. Thomas, B. Weerdmeester, & A. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). London: Taylor y Francis.
- Burke, R. (1999). Knowledge-based recommender systems. *Encyclopedia of Library and Information Science*, 1–23.
- Burke, R. (2002). Hybrid Recommender Systems : Survey and Experiments. *User Modeling and User-Adapted Interaction*, 12(4), 331–370.

- Campos, L.M., Fernández-Luna, J.M., Huete, J.F., Rueda-Morales, M.A. (2008). Managing uncertainty in group recommending processes. *User Modeling and User-Adapted Interaction*, 19(3), 207–242. doi:10.1007/s11257-008-9061-1
- CanCore Learning Resource Metadata Initiative. (2000). Retrieved from <http://cancore.athabasca.ca/en/>
- Casali, A., Gerling V., Deco, C., Bender, C. (2011). Chapter 8: Recommender System for Personalized Retrieval of Learning Objects. *Book of Educational Recommender Systems and Technologies: Practices and Challenges* (pp. 182–210).
- CETIS. (2008). UK LOM Core. Retrieved from <http://zope.cetis.ac.uk/profiles/uklomcore/>
- Chen, H. (2001). *Knowledge Management Systems: A text Mining perspective*. University of Arizona, Tucson, Arizona. Knowledge Computing Corporation.
- Chen, Y.L., Cheng, L.C., Chuang, C.N. (2008). A group recommendation system with consideration of interactions among group members. *Expert Systems with Applications*, 34(3), 2082–2090. doi:10.1016/j.eswa.2007.02.008
- Claypool, M., Le, P., Waseda, M., Brown, D. (2001). Implicit Interest Indicators. *Proceedings of ACM Intelligent User Interfaces Conference (IUI)*.
- Cohen, W. (1995). Fast Effective Rule Induction. In *Proceedings of the Twelfth International Conference on Machine Learning* (pp. 115–123).
- Connor, M.O., Cosley, D., Konstan, J.A., Riedl, J. (2001). PolyLens: A Recommender System for Groups of Users. *Proceedings of the seventh conference on European Conference on Computer Supported Cooperative Work*, 199–218.
- Coyle, M., Smyth, B. (2008). (Web Search)shared: Social Aspects of a Collaborative, Community-Based Search Network. *Proceedings of the 5th international conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH '08)* (pp. 103–112).

- Cremonesi, P., Milano, P., Koren, Y. (2010). Performance of Recommender Algorithms on Top-N Recommendation Tasks. *Proceedings of the fourth ACM conference on Recommender systems* (pp. 39–46).
- DBMiner Technology Inc. (2003). DBMiner. Retrieved from <http://www.dbminer.com/>
- De Praetere, T., Lobelle, C. (2010). *The Dokeos e-Learning Project Management Guide* (pp. 2007–2010).
- Decker, S., Harmelen, F. Van, Broekstra, J. (2000). The Semantic Web - on the respective Roles of XML and RDF. *Internet Computing, IEEE*, 4(5), 63–73.
- Dempster, A.P., Laird, N.M, Rubin, D. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1), 1–38.
- Dougiamas, M., Taylor, P.C. (2003). Moodle : Using Learning Communities to Create an Open Source Course Management System. *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 171–178.).
- Downes, S. (2005). Feature: E-learning 2.0. *eLearn Magazine, ACM Press*, 10.
- Dublin Core Metadata Initiative. (1995). Retrieved from <http://dublincore.org/>
- Duch, W. (2007). Towards Comprehensive Foundations of Computational Intelligence. In Duch, W., Mandziuk, J. (Eds.), *Challenges for Computational Intelligence* (Vol. 63, pp. 261–316).
- Duch, W. (2011). Meta-learning. In Dubitzky (Ed.), *Encyclopedia of Systems Biology*. W. Dubitzky, O. Wolkenhauer, H. Yokota, K-H Cho.
- Duda, R., Hart, P., Stork, D. (2000). *Patter Classification*. Wiley-Interscience.
- Duval, E., Forte, E., Cardinaels, K., Verhoeven, B. (2001). The Ariadne knowledge Pool System. *Communications of the ACM*, 44(5), 73–78.

- Education Services Australia. (2008). ANZ-LOM Metadata Application Profile. Retrieved from http://www.ndlrm.edu.au/verve/_resources/ANZ-LOM_1_02_file.pdf
- Edwards A.L. (1976). *Introduction to Linear Regression and Correlation*. (A series of books in psychology.) W. H. Freeman and Comp., San Francisco.
- Eiben, A. E., Smith, J. . (2003). What is an Evolutionary Algorithm? *Introduction to Evolutionary Computing* (pp. 15–35).
- Ellis, C.A., Gibbs, S.J., Rein, G. (1991). Groupware - Some Issues and Experiences. *Communications of the ACM*, 34(1), 38–58.
- EUN. (2007). The EUN Learning Resource Exchange Metadata Application Profile. Retrieved from <http://fire.eun.org/node/6>
- Fayyad, U., Piatetsky-shapiro, G., Smyth, P. (1996). From Data Mining to Knowledge Discovery in, 17(3), 37–54.
- Fiaidhi, J. (2004). A Model for Collaboratively Filtering Java Learning Objects. *International Journal of Instructional Technology and Distance Learning*, 1(7), 35–50.
- Fieller, E.C., Hartley, H.O., Pearson, E.S. (2008). Tests for rank correlation coefficients. *Biometrika*, 44(3), 470–481.
- Fischer, S. (2001). Course and exercise sequencing using metadata in adaptive hypermedia learning systems. *Journal of Educational Resources in Computing*, 1(1). doi:10.1145/376697.376700
- Fisher, D.H. (1987). Knowledge Acquisition Via Incremental Conceptual Clustering. *Machine learning*, 2(1980), 139–172.
- Fix, E. Hodges, J. (1951). *Discriminatory analysis. nonparametric discrimination: Consistency properties*.
- Friesen, N. (2004). *International LOM Survey: Report (Draft)* (pp. 1–31).

- Fuks, H., Raposo, A.B., Gerosa, M.A., Lucena, C.J. (2005). Applying the 3C Model to Groupware Development. *International Journal of Cooperative Information Systems (IJCIS)*, 14(2-3), 299–328.
- García, E., Romero, C., Ventura, S., De Castro, C. (2007). An architecture for making recommendations to courseware authors using association rule mining and collaborative filtering, (Dm).
- Gartrell, M., Xing, X., Lv, Q., Beach, A., Han, R., Mishra, S., & Seada, K. (2010). Enhancing group recommendation by incorporating social relationship interactions. *Proceedings of the 16th ACM international conference on Supporting group work - GROUP '10* (pp. 97–106). New York, New York, USA: ACM Press. doi:10.1145/1880071.1880087
- Giraud-Carrier, C., Vilalta, R., Brazdil, P. (2004). Introduction to the Special Issue on Meta-Learning. *Machine Learning*, 54(3), 187–193. doi:10.1023/B:MACH.0000015878.60765.42
- Goldberg, D., Nichols, D., Oki, B.M., Terry, D. (1992). Using collaborative filtering to weave an information Tapestry . *Communication of the ACM*, 61(10), 61–70.
- Golovchinsky, G., Adcock, J., Pickens, J., Qvarfordt, P., Back, M. (2008). Cerchiamo : a collaborative exploratory search tool. *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '08)*, 315–322.
- Guo, Q., Zhang, M. (2009). Implement web learning environment based on data mining. *Knowledge-Based Systems*, 22(6), 439–442. doi:10.1016/j.knosys.2009.06.001
- Hamers, L., Hemeryck, Y., et al. (1989). Similarity measures in scientometric research: the Jaccard index versus Salton's cosine formula. *Information Processing & Management*, 25(3), 315–318.

- Hanani, U.R.I., Shapira, B., Shoval, P. (2001). Information Filtering : Overview of Issues , Research and Systems. *User Modeling and User-Adapted Interaction*, 11(3), 203–259.
- Harper, B., Bennett, S., Lockyer, L., Harper, B., Agostinho, S., Lukasiak, J. (2005). Constructing high quality learning environments using learning designs and learning objects ., (July), 266–270.
- Heiner, C., Baker, R., Yacef, K. (2006). Proceedings of the Workshop on Educational Data Mining at the 8th International Conference on Intelligent Tutoring Systems.
- Hernández, J., Ramírez, M.J., Ferri, C. (2004). *Introducción a la Minería de Datos*. Pearson Prentice Hall.
- Hosmer, D.W., Stanley, L. (2000). *Applied Logistic Regression*. (John Wiley & sons, Ed.).
- Hyndman, R.J., Koehler, A.B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688.
- IEEE-LTSC. (2002). Draft Standard for Learning Object Metadata, (July), 1–44. Retrieved from <http://www.ieeeeltsc.org:8080/Plone>
- IMS Global Learning Consortium. (1997). Retrieved from <http://www.imsglobal.org/>
- IMS Global Learning Consortium. (2001). IMS-Content Packaging Specification. Retrieved from <http://www.imsglobal.org/content/packaging/>
- IMS Global Learning Consortium. (2002). IMS Reusable Definition of Competency or Educational Objective Specification. Retrieved from <http://www.imsglobal.org/competencies/>
- IMS Global Learning Consortium. (2003a). IMS Learning Design. Information Model, Best Practice and Implementation Guide, Binding document, Schemas, (June), 1–37.
- IMS Global Learning Consortium. (2003b). IMS Question & Test Interoperability Specification. Retrieved from <http://www.imsglobal.org/question/>

- Itmazi, J. (2005). *Sistemas flexible de gestión del E-Learning para soportar el aprendizaje en las universidades tradicionales y abiertas*.
- Jacko, J., Sears, A. (2003). *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*. L. Erlbaum Associates Inc.
- Jankowski, N., Duch, W., Grąbczewski, K. (2011). *Meta-learning in Computational Intelligence* (p. 127). Springer Berlin / Heidelberg.
- Jumper Networks Inc. (2011). Jumper 2.0. Retrieved from http://en.wikipedia.org/wiki/Jumper_2.0
- Jung, S., Herlocker, J.L., Webster, J. (2007). Click data as implicit relevance feedback in web search. *Information Processing & Management*, 43(3), 791–807. doi:10.1016/j.ipm.2006.07.021
- Kapp, K.M. (2003). Five Technological Considerations When Choosing an E-Learning Solution. *eLearn Magazine, ACM Press*, 7.
- KERIS. (2004). Korea Educational Metadata (KEM) Profile for K–12. Retrieved from <http://www.keris.or.kr/datafiles/data/RM2004-22.pdf>
- Khan, B. (1997). Web-based Instruction (WBI): What is it and Why is it? In Khan, B. (ed), *Web-based Instruction, Educational Technology Publications*.
- Kitchenham, B., Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering, Version 2.3. Keele University and Durham University, EBSE Technical Report, EBSE-2007-01: UK*.
- Knolmayer, G.F. (2003). Decision Support Models for Composing and Navigating through e-Learning Objects, 00(C), 1–10.
- Kohavi, R., Karl, P. (1994). : A Machine Learning Library in C ++. *International Journal on Artificial Intelligence Tools*, 6(4), 537–566.

- Kohonen, T. (1988). Self-organized formation of topologically correct feature maps. In E. Anderson, J.A. Rosenfeld (Ed.), *Neurocomputing: foundations of research* (pp. 509–521).
- Kowalski, G. (1997). *Information Retrieval Systems: Theory and Implementation*. Kluwer Academic Publishers.
- Krulwich, B. (1997). Using Large-Scale Demographic Data. *AI Magazine*, 18(2), 37–46.
- Kumar, V., Nesbit, J., Winnie, P., Hadwin, A., Jamieson-Noel, D. Han, K. (2007). Quality Rating and Recommendation of Learning Objects. In S. Pierre (Ed.), *E-Learning Networked Environments and Architectures, Advanced Information and Knowledge Processing* (pp. 337–373).
- Lemire, D., Maclachlan, A. (2005). Slope one predictors for online rating-based collaborative filtering. *Proceedings of SIAM Data Mining (SDM'05)*.
- Lewis, J.R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1), 57–78. doi:10.1080/10447319509526110
- Lynch, T.D., Lynch, C.E. (2002). Web-Based Education. *The Innovation Journal: The Public Sector Innovation Journal*, 8(4), 1–28.
- Macqueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 281–297.
- Manouselis, N., Vuorikari, R., Van Assche, F. (2010). Collaborative recommendation of e-learning resources: an experimental investigation. *Journal of Computer Assisted Learning*, 26(4), 227–242. doi:10.1111/j.1365-2729.2010.00362.x

- Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H. (2011). Recommender Systems in Technology Enhanced Learning. In K. P. Ricci F, Rokach L, Shapira B (Ed.), *Recommender Systems Handbook* (pp. 387–415).
- Masthoff, J. (2011). Group Recommender Systems : Combining individual models. In P. B. Ricci, Francesco and Rokach, Lior and Shapira, Bracha and Kantor (Ed.), *Recommender Systems Handbook* (pp. 677–702).
- McGreal, R. (2004). Learning Objects : A Practical Definition. *International Journal of Instructional Technology and Distance Learning*, 1(9), 21–32.
- Mclean, N., Lynch, C. (2003). Interoperability between Information and Learning Environments: Bridging the Gaps. A Joint White Paper on behalf of the IMS Global Learning Consortium and the Coalition for Networked Information. DRAFT – Version of June 28, 2003. Retrieved from www.imsglobal.org/DLims_white_paper_publicdraft_1.pdf
- Menendez-Dominguez, V.H., Zapata, A., Prieto-Mendez, M.E., Romero, C., Serrano-Guerrero, J. (2011). A similarity-based approach to enhance learning objects management systems. *2011 11th International Conference on Intelligent Systems Design and Applications* (pp. 996–1001). Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/ISDA.2011.6121788
- Menéndez, V.H. (2011). *Marco conceptual para la gestión integral de Objetos de Aprendizaje*. Tesis doctoral, Universidad de Castilla-La Mancha.
- Merrill, M.D. (1996). Instructional Transaction Theory : An Instructional Design Model based on Knowledge Objects What is Instructional Design Theory? *Educational Technology*, 36(3), 30–37.
- Michie, E.D., Spiegelhalter, D.J., Taylor, C.C. (1994). *Machine Learning , Neural and Statistical Classification*. Ellis Horwood.

- Mohan, P. (2004). Reusable online learning resources: problems, solutions and opportunities. *IEEE International Conference on Advanced Learning Technologies, 2004. Proceedings.* (pp. 904–905). Ieee. doi:10.1109/ICALT.2004.1357713
- Morris, M.R. (2008). A survey of collaborative web search practices. *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08* (pp. 1657–1660). New York, New York, USA: ACM Press. doi:10.1145/1357054.1357312
- Morris, M.R., Horvitz, E. (2007). SearchTogether: An Interface for Collaborative Web Search. *Proceedings of the 20th annual ACM symposium on User interface software and technology (UIST '07)* (pp. 3–12).
- Olivas, J.A. (2011). Sistemas de recuperación de información. *Búsqueda eficaz de información en la Web.*
- Pazzani, M.J., Billsus, D. (2007). Content-Based Recommendation Systems. In and W. N. P. Brusilovsky, A. Kobsa (Ed.), *The Adaptive Web: Methods and Strategies of Web Personalization* (pp. 325–341).
- Platt, J.C. (1999). Fast Training of Support Vector Machines using Sequential Minimal Optimization. *Advances in Kernel Methods: Support Vector Learning* (pp. 185–208). MIT Press.
- Porcel, C., López-Herrera, a. G., Herrera-Viedma, E. (2009). A recommender system for research resources based on fuzzy linguistic modeling. *Expert Systems with Applications*, 36(3), 5173–5183. doi:10.1016/j.eswa.2008.06.038
- Preyne, J., Smyth, B. (2004). Collaborative Search; Deployment Experiences. *The 24th SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence* (pp. 121–134).

- Prieto, M.E., Menéndez, V.H. (2009). Gestión Integral de Recursos para el Aprendizaje. In Prieto M.E., Sanchez-Alonso, S., Pech, S., Ochoa, X. (Eds.), *Recursos Digitales Para el Aprendizaje* (pp. 27–33).
- Project Seeks. (2011). Seeks. Retrieved from <http://www.seeks-project.info/site/>
- Quinlan, J.R. (2007). Induction of Decision Trees. *Machine learning*, 1(1), 81–106.
- Quinlan, J.R. (1993). *C4.5: Programs for Machine Learning* (Vol. 240, pp. 235–240). Morgan Kaufmann Publishers.
- Rafaeli, S., Dan-Gur, Y., Barak, M. (2004). QSIA- a Web-based environment for learning, assessing and knowledge sharing in communities. *Computers & Education*, 43(3), 273–289. doi:10.1016/j.compedu.2003.10.008
- Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (2011). *Recommender Systems Handbook*.
- Rištofič, A.K. (2005). Recommender System for Adaptive Hypermedia Applications. *Proc. Stud. Res. Conf. Informat. Inf. Technol.* (pp. 229–234).
- Rocchio, J.J. (1971). Relevance feedback in information retrieval. *The SMART retrieval system*, 313–323.
- Romero, C., Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33(1), 135–146. doi:10.1016/j.eswa.2006.04.005
- Romero, C., Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(February), 12–27. doi:10.1002/widm.1075
- Romero, C., Ventura, S. (2010). Educational Data Mining : A Review of the State of the Art. *Trans. Sys. Man Cyber Part C*, 40(6), 601–618.
- Romero, C., Ventura, S., Zafra, A., De Bra, P. (2009). Applying Web usage mining for personalizing hyperlinks in Web-based adaptive educational systems. *Computers & Education*, 53(3), 828–840. doi:10.1016/j.compedu.2009.05.003

- Rosenfeld, L., Morville, P. (2002). *Information architecture for the world wide web*. O'Reilly \& Associates, Inc.
- Ruiz, F., Polo, M., Piattini, M. (2002). Utilización de Investigación-Acción en la Definición de un Entorno para la Gestión del Proceso de Mantenimiento del Software. *1er. Workshop en: Métodos de Investigación y Fundamentos Filosóficos en Ingeniería del Software y Sistemas de Información*.
- Ruiz-Iniesta, A., Jimenez-Diaz, G., Gomez-Albarran, M. (2009). Recommendation in Repositories of Learning Objects: A Proactive Approach that Exploits Diversity and Navigation-by-Proposing. *2009 Ninth IEEE International Conference on Advanced Learning Technologies*, 543–545. doi:10.1109/ICALT.2009.23
- Rumelhart, D.E., Hinton, G.E., Williams, R.J. (1986). Learning internal representations by error propagation. In J. L. Rumelhart, David E. and McClelland (Ed.), *Parallel distributed processing: explorations in the microstructure of cognition* (pp. 318–362).
- Salton, G., McGill, M.J. (1986). *Introduction to Modern Information Retrieval*. McGraw-Hill, Inc.
- Sarwar, B., Karypis, G., Konstan, J., Riedl, J. (2001). Item-Based Collaborative Filtering Recommendation. *Proceedings of the 10th international conference on World Wide Web* (pp. 285–295).
- Schafer, J. (2001). E-Commerce Recommendation Applications, 115–153.
- Scheffer, T. (2004). Finding Association Rules that Trade Support Optimally Against Confidence. *Intell. Data Anal.*, 9(4), 381–395.
- Shell, G., Max, B. (2002). Merlot: A Repository of e-Learning Objects for Higher Education. *e-Service Journal*, 1(2), 53–64.

- Smeaton, A.F., Lee, H., Foley, C. (2005). Fischlelar-DiamondTouch : Collaborative Video Searching on a Table. *SPIE Electronic Imaging - Multimedia Content Analysis, Management, and Retrieval*, 6073, 15–19.
- Smyth, B., Briggs, P., Coyle, M., Mahony, M.P (2009). A Case-Based Perspective on Social Web Search. *Proceedings of the 8th International Conference on Case-Based Reasoning: Case-Based Reasoning Research and Development (ICCBR '09)* (pp. 494–508).
- Smyth, B., Coyle, M., Briggs, P. (2011). Communities, Collaboration, and Recommender Systems in Personalized Web Search. In B. . K. P. B. Ricci, F.; Rokach L.; Shapira (Ed.), *Recommender Systems Handbook* (pp. 579–644).
- Song, Q., Wang, G., Wang, C. (2012). Automatic recommendation of classification algorithms based on data set characteristics. *Pattern Recognition*, 45(7), 2672–2689. doi:10.1016/j.patcog.2011.12.025
- Stefaner, M., Vecchia, E.D., Condotta, M., Wolpers, M., Specht, M., Apelt, S., Duval, E. (2007). MACE – Enriching Architectural Learning Objects for Experience Multiplication. In R. K. and M. W. E. Duval (Ed.), *Creating New Learning Experiences on a Global Scale*. (pp. 322–336). Springer Berlin / Heidelberg.
- Studio Smart Agent Technologies. (2010). *Easyrec*. Retrieved from http://sat.researchstudio.at/home_en.html
- Sánchez, J.G., Rosas, A.C., Ricardo, J., Terrats, A. (2006). en México: un primer acercamiento (1). *Virtual educa* (pp. 1–17).
- Sánchez-Alonso, Sicilia, M.A. (2005). Normative Specifications of Learning Objects and Learning Processes: Towards Higher Levels of Automation in Standardized e-Learning. *International Journal of Instructional Technology and Distance Learning*, 2(3), 3–11.

- Tang, T.Y., Mccalla, G. (2005). Smart Recommendation for an Evolving E-Learning System. *International Journal on E-Learning*, 4, 105–129.
- Toronto University. (2002). ATutor: Learning Management System. Retrieved September 25, 2012, from <http://atutor.ca/atutor/>
- Traxler, J. (2005). Defining mobile learning. *IADIS International Conference Mobile Learning 2005* (pp. 261–266).
- Tsai, K., Chiu, T., Lee, M., Wang, T. (2006). A Learning Objects Recommendation Model based on the Preference and Ontological Approaches. *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)*, 36–40. doi:10.1109/ICALT.2006.1652359
- Tsai, S., Machado, P. (2002). E-Learning Basics: Essay: E-learning, online learning, web-based learning, or distance learning: unveiling the ambiguity in current terminology. *eLearn Magazine, ACM Press*, 3.
- UNESCO. (2010). OER useful resources/Repositories. Retrieved September 4, 2012, from http://oerwiki.iiep.unesco.org/index.php/OER_useful_resources/Repositories
- Universidad de Castilla-La Mancha. (2012). SMILe (Soft Management of Internet and Learning). Retrieved September 3, 2012, from <http://smile.esi.uclm.es/smile/>
- Universidad de Córdoba. (2012). KDIS (Knowledge Discovery and Intelligent Systems). Retrieved October 17, 2012, from <http://www.uco.es/grupos/kdis/>
- Universidad de Dortmund. (2001). RapidMiner. Retrieved November 22, 2012, from <http://rapid-i.com/content/view/181/190/>
- University of cologne. (2004). Didactical concept of ILIAS. Retrieved November 2, 2012, from http://www.ilias.de/docu/goto_docu_lm_392.html

- Valipour, M.H., Amirzafari, B., Maleki, K.N., Daneshpour, N. (2009). A brief survey of software architecture concepts and service oriented architecture. *2009 2nd IEEE International Conference on Computer Science and Information Technology* (pp. 34–38). Ieee. doi:10.1109/ICCSIT.2009.5235004
- Vialardi, C., Bravo, J., Shafti, L., Ortigosa, A. (2009). Recommendation in Higher Education Using Data Mining Techniques. *Educational Data Mining 2009* (pp. 190–199).
- Walker, A., Recker, M., Lawless, K., Wiley, D. (2004). Collaborative Information Filtering : a review and an educational application. *International Journal of Artificial Intelligence in Education*, 14, 1–26.
- Wan, X., Jamaliding, Q., Okamoto, T. (2011). Analyzing Learners' Relationship to Improve the Quality of Recommender System for Group Learning Support. *Journal of Computers*, 6(2), 254–262. doi:10.4304/jcp.6.2.254-262
- Wang, F.H., Shao, H.M. (2004). Effective personalized recommendation based on time-framed navigation clustering and association mining. *Expert Systems with Applications*, 27(3), 365–377. doi:10.1016/j.eswa.2004.05.005
- Wang, P. (1998). Why Recommendation is Special ? *Workshop on Recommender Systems part of the 15th National Conference on Artificial Intelligence AAAI*, 15, 111–113.
- Wiley, D.A. (2002). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. *The Instructional Use of Learning Objects*. D. A. Wiley., 2830(435), 1–35.
- Witten I.H., Frank E. (2005). *Data mining: Practical machine learning tools and techniques (second edition)*. Morgan Kaufmann Publishers.
- Witten I.H., Frank E., Hall, M.A. (2011). *Data Mining: Practical Machine Learning Tools and Techniques (Third Edition)*.

- Ye, P., Baldwin, T. (2006). Semantic Role Labelling of Prepositional Phrases, 5(3), 228–244.
- Yu, Z., Zhou, X., Zhang, D. (2005). An Adaptive In-Vehicle Multimedia Recommender for Group Users. *Vehicular Technology Conference* (Vol. 5, pp. 2800–2804).
- Zafra, A., Romero, C., Ventura, S., Herrera-Viedma, E. (2009). Multi-instance genetic programming for web index recommendation. *Expert Systems with Applications*, 36(9), 11470–11479. doi:10.1016/j.eswa.2009.03.059
- Zapata, A., Menendez, V.H, Eguigure, Y., Prieto, M.E (2009). Quality Evaluation Model for Learning Objects from Pedagogical Perspective . A case of study. *International Conference of Education, Research and Innovation* (pp. 2228–2238).
- Zapata, A., Menendez, V.H, Prieto, M.E, Romero, C. (2011). A Hybrid Recommender Method for Learning Objects. *IJCA Proceedings on Design and Evaluation of Digital Content for Education (DEDCE)* (pp. 1–7).
- Zapata, A., Menéndez, V.H., Prieto, M.E., Romero, C. (2011). Using Data mining in a recommender system to search for learning objects in repositories. *Proceedings of the 4th International Conference on Educational Data Mining* (pp. 321–322). Pechenizkiy, M., Calders, T., Conati, C., Ventura, S., Romero, C., and Stamper, J. (Eds.).
- Zapata, A. (2009). *Extracción de conocimiento a partir de Recursos Digitales para el Aprendizaje*. Tesis de Máster, Universidad de Castilla-La Mancha.
- Zapata, A. Menéndez, V.H., Prieto, M.E., Romero, C. (2012). Búsqueda colaborativa para el apoyo a la recomendación de Objetos de Aprendizaje. In P. Such, M., Pernías (Ed.), *SPDECE-2012. Ninth multidisciplinary symposium on the design and evaluation of digital content for education* (pp. 57–67).
- Zapata, A., Menéndez, V.H., Prieto, M.E., Romero, C. (2013). A Framework for Recommendation in Learning Object Repositories: An Example of Application in Civil

Engineering. *Advances in Engineering Software*, 56, 1–14.
doi:<http://www.sciencedirect.com/science/article/pii/S0965997812001433>

Zhang, H. (2004). The Optimality of Naive Bayes. In Z. Barr, V., Markov (Ed.), *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference*.

Zhang, L., Liu, X., Liu, X. (2008). Personalized Instructing Recommendation System Based on Web Mining. *Proc. Int. Conf. Young Comput. Sci.* (pp. 2517–2521).