

Obteniendo Casos de Uso centrados en la Calidad de los Datos desde Procesos de Negocio descritos con BPMN

Alfonso Rodríguez ¹, Angélica Caro ¹

alfonso@ubiobio.cl, mcaro@ubiobio.cl

¹ Departamento de Ciencias de la Computación y Tecnologías de Información, Universidad del Bío-Bío, Andrés Bello s/n, Chillán, Chile.

DOI: 10.4304/risti.10.65-80

Resumen: La calidad de datos es considerada un aspecto importante en relación con el éxito o fracaso de las tareas cotidianas en una organización. Hoy en día, la mayoría de estas tareas se encuentran soportadas por aplicaciones de software. La especificación temprana de los requisitos que deben cumplir esas aplicaciones es un desafío para la ingeniería del software. En este trabajo se aborda la especificación temprana de requisitos de software, poniendo especial atención en la calidad de datos. Dichos requisitos serán capturados desde modelos de procesos de negocio descritos con BPMN (Business Process Model and Notation), y expresados mediante casos de uso de UML (Unified Model Language). Con este propósito se propone un método mediante el cual, de manera ordenada y sistemática, los expertos del negocio pueden modelar procesos de negocio consciente de la calidad de datos y obtener desde dichos modelos artefactos útiles para la creación de software.

Palabras-claves: Calidad de Datos; Procesos de Negocio; Casos de Uso; BPMN; UML; Requisitos de Software.

Abstract: Data quality is considered an important aspect for the success or failure of routine tasks in an organization. Nowadays, most of these tasks are supported by software applications. The early specification of the requirements for these applications is a challenge for software engineering. This paper focuses on early software requirements specification, paying special attention on data quality. These requirements will be captured from business process models described with BPMN (Business Process Model and Notation), and expressed through use cases of UML (Unified Model Language). For this purpose we propose a method whereby business experts can model business processes aware of data quality and software engineers can get artifacts useful for creating software.

Key-words: Data Quality; Business Process; Use Cases; BPMN; UML; Software Requirements.

1. Introducción

Las Tecnologías de Información (TI), en general, y los Sistemas de Información (SI), en particular, juegan un rol fundamental en la gestión de los Procesos de Negocio (BP, Business Process) debido al importante número de actividades dentro de una organización que son apoyadas por los SI. A nivel organizacional, los BP son esenciales para entender la forma en que operan las organizaciones y también tienen un rol importante en el diseño y creación de SI flexibles (Weske, 2007).

Por otro lado, la gestión de la Calidad de Datos (DQ, Data Quality) es un aspecto relevante que debe ser considerado cuando se quiere mejorar el desempeño general de las organizaciones (Redman, 2008). Sólo aquellas organizaciones que logran incorporar una estrategia de gestión de DQ efectiva en su estrategia de negocios serán capaces de convertir sus datos en una ventaja competitiva real, entregando valor a corto y largo plazo para asegurar el éxito y la sustentabilidad de su negocio (el Abed, 2009).

La especificación temprana de los requisitos de un SI, constituye un desafío permanente para la ingeniería de software. Con el propósito de incluir requisitos de DQ en conjunto con la descripción del BP, en trabajos previos (Rodríguez, Caro, Cappiello, & Caballero, 2012) se ha extendido la capacidad expresiva de BPMN (Object Management Group, 2011). Complementariamente, se ha propuesto un método, BPiDQ (Caro, Rodríguez, Cappiello, & Caballero, 2012), que permite llevar a cabo la especificación de requisitos de DQ en BP y a partir de allí obtener diagramas de casos de uso UML (Object Management Group, 2007) relacionados con la calidad de datos. En base a estos trabajos y a una propuesta para la transformación de las descripciones de BP en casos de uso de UML (Rodríguez, Fernández-Medina, & Piattini, 2007b), en este artículo se presenta una adaptación del método, llamada BPiDQ* (Business Process including Data Quality), para orientar los resultados obtenidos hacia la ingeniería de software. En la primera versión del método se ha privilegiado la mejora del modelo del proceso de negocio propiamente dicho, teniendo en cuenta la calidad de datos. Por ejemplo, se incorporan nuevas actividades y/o se cambia el flujo del BP para garantizar que los aspectos de calidad de datos especificados en el BP puedan ser incluidos. La adaptación del método propuesta en este artículo, tiene como objetivo obtener casos de uso (generales y relacionados con la DQ) desde la especificación del BP, lo que permitirá avanzar en el proceso de desarrollo del software.

El resto del artículo se encuentra organizado de la siguiente forma. En la sección 2 se presentan trabajos relacionados. En la sección 3 se describe la metodología BPiDQ*. Un ejemplo ilustrativo del uso de la metodología es entregado en la sección 4. Finalmente, la sección 5 contiene las conclusiones y el trabajo futuro.

2. Trabajos relacionados

Un Proceso de Negocio es un conjunto de actividades que se realizan de manera coordinada para cumplir un objetivo de negocio en un contexto tecnológico y organizacional (Weske, 2007). Los BP, desde el punto de vista de la ingeniería de software, pueden ser vistos como una descripción del dominio del software y como una fuente de requisitos para el desarrollo de éste. Así, el modelado de procesos de negocio

podría considerarse como una primera etapa en la especificación de requisitos de software (Liew, Kontogiannis, & Tong, 2004).

Los lenguajes más usados para el modelado de BP son UML y BPMN, siendo este último el más utilizado en la industria por lo que es ampliamente reconocido como el estándar *de facto* para el modelado de BP (Harmon & Wolf, 2011; Recker, 2010). BPMN ha demostrado una gran expresividad, que además puede ser extendida para dar cabida a nuevos aspectos relacionados con los BP. Entre las extensiones propuestas en la literatura se pueden mencionar algunas orientadas a: expresar algunas necesidades de los clientes, tales como tiempo, costo y fiabilidad (Saeedi, Zhao, & Falcone Sampaio, 2010), especificar propiedades no funcionales como desempeño y fiabilidad (Bocciarelli & D'Ambrogio, 2011), modelar requisitos de seguridad en BP (Rodríguez, Fernández-Medina, & Piattini, 2007a), representar explícitamente restricciones legales (Goldner & Papproth, 2011), analizar el desempeño de los procesos de negocio (Lodhi, Veit, & Saake, 2011) y expresar requisitos de calidad de datos (Rodríguez et al., 2012), entre otras.

En particular, la representación de requisitos de DQ en modelos procesos de negocio no había sido abordada hasta (Rodríguez et al., 2012). La calidad de datos ha sido definida como un concepto multidimensional, dependiente del contexto y que representa el hecho que los datos sean "*apropiados para el uso*" (Strong, Lee, & Wang, 1997; Wang & Strong, 1996). Esto último enfatiza la idea de que son los usuarios quienes deciden si los datos son útiles o no. En la literatura pocos trabajos han estudiado la especificación temprana de aspectos de DQ en BP, en especial en las etapas de modelado y diseño de BP. En la Tabla 1 se muestra un resumen de algunos trabajos en los que se ha abordado el tema y sus respectivas propuestas.

En el contexto de los SI, la representación temprana de requisitos de DQ en BP puede tener una importante repercusión para la definición de los requisitos de software. Es así como en (Liew et al., 2004) se propone un marco de trabajo que genera artefactos UML (casos de uso, diagramas de colaboración y despliegue) desde modelos de BP usando BPMN. Los autores introducen una notación adicional en BPMN a partir de la cual es posible extraer datos y patrones que permiten guiar el diseño del software. En (Coskuncay, Aysolmaz, Demirors, Bilen, & Dogani, 2010) se presenta una propuesta (que incluye dos notaciones, un proceso y una herramienta) que permite la ejecución simultánea del modelado de BP y el análisis de requisitos para el desarrollo de software. A partir del modelo de BP se generan requisitos de software en lenguaje natural. Otro trabajo, (Rodríguez et al., 2007b), propone una transformación desde modelos de BP con requisitos de seguridad hacia casos de uso. Esta transformación se realiza usando un conjunto de reglas QVT (Object Management Group, 2004), listas de control y reglas de refinamiento. Por último, en (Caro et al., 2012), se propone la generación de casos de uso que representan requisitos de DQ desde modelos de BP extendidos para representar requisitos de DQ. Estos dos últimos trabajos constituyen la motivación y fundamento para la propuesta presentada en este artículo, tal como se describe en la siguiente sección.

Tabla 1 – Trabajos que abordan la especificación temprana de requisitos de DQ en BP

Autores	Propuesta
<i>(Bringel, Caetano, & Tribolet, 2004)</i>	Patrón de BP para asegurar la DQ en una organización, reusable en distintos escenarios. Se definen dimensiones de DQ asociadas a las entidades de información, con diversos significados dependiendo de la vista del negocio y de diferentes dimensiones organizacionales.
<i>(Bagchi, Bai, & Kalaganam, 2006)</i>	Marco de trabajo para el modelado de BP considerando la estimación y gestión cuantitativa de DQ en los SI. Basado en el flujo del BP se estiman errores que surgen de la transacción de datos y el impacto de su propagación.
<i>(el Abed, 2009)</i>	Marco de trabajo que describe una metodología, procesos y roles, para lograr el máximo valor de negocio. El objetivo es la mejora continua de los BP mediante el uso de reglas de negocio que representan requisitos de DQ.
<i>(Soffer, 2010)</i>	Explora la inexactitud de los datos y sus potenciales consecuencias. Se proveen las bases para generar BP que eviten problemas de inexactitud.
<i>(Heravizadeh, Mending, & Rosemann, 2009)</i>	Marco de trabajo que identifica atributos de calidad en BP basado en cuatro dimensiones. Entre ellas la dimensión de “ <i>calidad de objetos de entrada y salida</i> ” para la cual se identifican once atributos de DQ.

3. BPiDQ*: Un método para la obtención de requisitos de software centrados en DQ desde especificaciones de BP.

El objetivo de BPiDQ* es soportar la especificación temprana de requisitos de DQ en BP y, a partir de dicha especificación, obtener requisitos de software centrados en la DQ, expresados como casos de uso. En la Figura 1, en color gris, se resume el conjunto de elementos que forman parte esta propuesta. Concretamente, en la parte central se muestra el nuevo método BPiDQ*, con sus cuatro etapas, la extensión dqBP que permite agregar requisitos de DQ en modelos de BP descritos con BPMN, repositorios tanto para las actividades de calidad de datos como para los casos de uso dedicados a representar los requisitos de calidad de datos y, finalmente, los casos de uso que se derivan de la aplicación del método BPiDQ*.

Por otro lado, esta propuesta se enmarca en el enfoque de la Arquitectura Dirigida por Modelos, promovido por OMG (Object Management Group, 2003), al lado izquierdo de la Figura 1, y en el proceso de desarrollo de software Proceso Unificado (Rational Software, 2001), al lado derecho de la misma figura. Consecuentemente, un modelo de procesos de negocio especificado con BPMN extendido con requisitos de DQ está en el nivel CIM (Computation Independent Model) y los casos de uso generados a partir de modelos de BP en el nivel PIM (Platform Independent Model). Asimismo, en el Proceso Unificado, el modelo de BP se sitúa en la etapa de descripción del “*Modelo del Negocio*” y los casos de uso en la etapa de “*Requisitos y Análisis & Diseño*”. En las subsecciones siguientes se describen los componentes que soportan el método BPiDQ* y el detalle de las etapas que lo conforman.

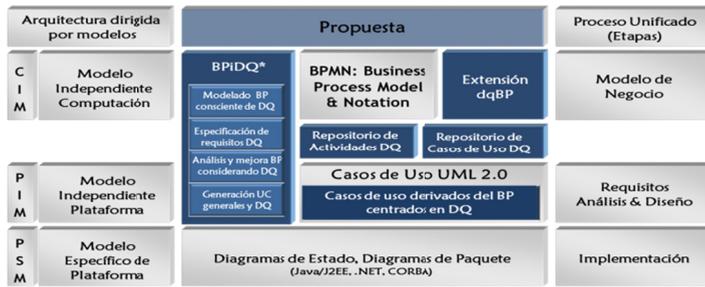


Figura 1 – Vista general de la propuesta

3.1. Componentes del método BPIDQ*

Para que la aplicación del método BPIDQ* sea posible es necesario contar con tres componentes que apoyan las etapas definidas. Estos componentes son la extensión de la notación BPMN, un conjunto de actividades relacionadas con el tratamiento de los requisitos de calidad de datos a nivel de proceso de negocio y un catálogo de casos de uso estándar que permiten abordar las especificaciones de DQ.

3.1.1. La extensión dqBP

La extensión dqBP (Rodríguez et al., 2012) tiene por objetivo agregar capacidad expresiva a la notación BPMN 2.0, permitiendo la representación de requisitos de DQ en un modelo de BP. En la Figura 2 se muestra el metamodelo en que aparece la nueva clase dqFlag y el vínculo que ésta tiene con los elementos de BPMN.

Dado que BPMN es una notación en que se privilegia la representación simbólica de los distintos aspectos del negocio, se ha asociado un símbolo a la clase dqFlag que consiste en la fusión de las letras DQ (DQ). Este símbolo deberá ser usado para marcar los elementos de BPMN en los cuales es posible asociar requisitos de calidad de datos. La forma en que se representa este nuevo símbolo en conjunto con los elementos de BPMN y el significado de dicha representación se muestran en la Tabla 2.

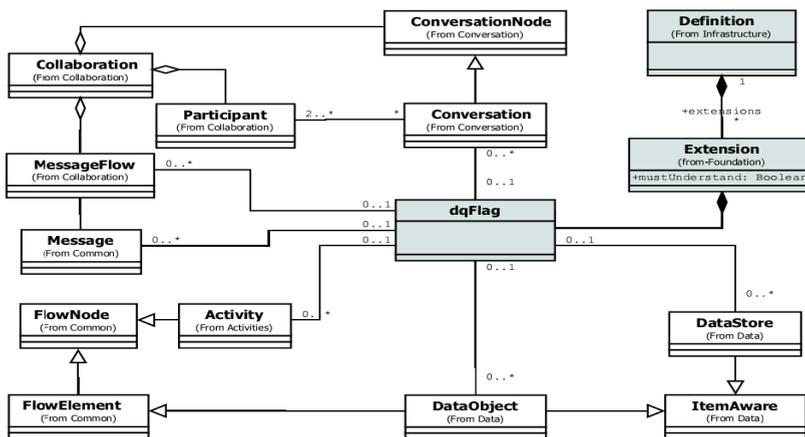
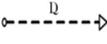


Figura 2 – Extensión de BPMN 2.0 para incluir aspectos de calidad de datos

Tabla 2 – Elementos de BPMN y la especificación de requisitos de DQ

Vista gráfica	Significado
 <i>Message</i>	Representa los datos contenidos en un mensaje, los cuales deberían satisfacer ciertos requisitos de DQ necesarios para el éxito del proceso de negocio. Por ejemplo, completitud y consistencia en una receta médica enviada por el médico al paciente
 <i>Message flow</i>	Representa los datos implícitos en un flujo de mensaje, los cuales deberían satisfacer ciertos requisitos de DQ. Por ejemplo, la vigencia de una autorización para una tarjeta de crédito
 <i>Conversation</i>	Representa los datos que están contenidos en una conversación compuesta por un conjunto de mensajes, los cuales deberían satisfacer ciertos requisitos de calidad de datos. Por ejemplo, seguridad y exactitud en los datos intercambiados entre los clientes y una línea aérea durante la compra de los billetes del vuelo.
 <i>Data Object</i>	Representa los datos que se encuentran contenidos en un objeto de datos. Estos datos deberían satisfacer ciertos requisitos de calidad de datos, como por ejemplo, completitud, consistencia y exactitud de los datos (nombre, dirección) necesarios para entregar un paquete a un cliente.
 <i>Data Store</i>	Representa los datos contenidos en una base de datos, los cuales deberían satisfacer ciertos requisitos de calidad de datos, como por ejemplo, completitud de los datos actualizados acerca de una venta de productos.
 <i>Activity</i>	Representa los datos que son usados y/o producidos en una actividad. Estos datos deberían satisfacer ciertos requisitos de calidad de datos como por ejemplo, precisión y exactitud acerca de los presupuestos de gastos generados como salida de la actividad.

3.1.2. El repositorio de actividades de DQ

El segundo componente es un repositorio que contiene actividades en el nivel de BP orientadas a satisfacer requisitos de DQ. Un requisito de calidad de datos expresado en el modelo de BP con el símbolo \mathcal{DQ} (DQ-Flag) puede estar compuesto por una o más dimensiones de DQ. Cada una de las dimensiones de DQ se asocia a un conjunto de actividades de DQ contenidas en el repositorio. En la Tabla 3 se muestran, a modo de ejemplo, las dimensiones de DQ exactitud, oportunidad y completitud. Para cada una de ellas se entrega una definición de acuerdo con diferentes autores, un conjunto (no completo) de actividades que se podrían incluir en el modelo de BP para la mejora del mismo, teniendo en cuenta la DQ, y algunos ejemplos de la aplicación de estas actividades en el contexto de un BP.

Tabla 3 – Dimensiones de DQ y actividades de mejora asociadas

Dimensiones de DQ	Actividades de Mejora	Ejemplos
<p>Exactitud: Grado en que los datos reflejan una vista del mundo real en un contexto y un proceso de negocio específico (el Abed, 2009; Heravizadeh et al., 2009; Soffer, 2010).</p>	<ul style="list-style-type: none"> - Determinar el conjunto de datos que requieren exactitud. - Definir el dominio válido para los datos - Verificar los datos con el dominio correcto. - Verificar los datos en distintas fuentes. - Limpiar las bases de datos para alcanzar los niveles de exactitud requeridos. - Mejorar los datos hasta alcanzar la exactitud requerida 	<ul style="list-style-type: none"> - El precio recibido por el cliente para una reserva de hotel debe ser exacto. - En una prescripción médica, el nombre del medicamento puede ser confrontado con el Vademecum. - El peso del paquete a entregar debe estar en el rango predeterminado.
<p>Oportunidad: Grado en que los datos están lo suficientemente actualizados y disponibles como para ser útiles en un contexto y en un proceso de negocio específico (Bringel et al., 2004; el Abed, 2009; Heravizadeh et al., 2009).</p>	<ul style="list-style-type: none"> - Verificar si el dato tiene requisitos de vigencia para una determinada tarea. - Para diferentes fuentes de datos, seleccionar una que provea los datos con la actualización requerida por el proceso. - Verificar que el dato sea entregado en el tiempo requerido. 	<ul style="list-style-type: none"> - Comprobar si los mismos datos están en diferentes fuentes dentro de la empresa, y si es así elegir aquel dato con la actualización más reciente. - La verificación de validez de una tarjeta de crédito, por parte de una entidad financiera, debe obtenerse antes de confirmar una venta.
<p>Completitud: Grado en que los datos tienen todos los valores necesarios para la ejecución exitosa de un proceso de negocio en un contexto y dominio específico (Bringel et al., 2004; el Abed, 2009; Heravizadeh et al., 2009).</p>	<ul style="list-style-type: none"> - Especificar los datos que son obligatorios - Verificar que todos los datos obligatorios tengan valores - Completar los datos obligatorios con otras fuentes de datos - Usar un procedimiento para forzar la entrega de todos los datos obligatorios. 	<ul style="list-style-type: none"> - Los resultados de un examen de laboratorio debe contener valores para todas las pruebas solicitadas por el médico tratante. - Para entregar un paquete, los datos relacionados con la dirección e identificación del cliente deben estar completos.

3.1.3. Repositorio de casos de uso de DQ

El tercer componente del método es un repositorio que contiene los casos de uso estándar para cada dimensión de DQ que puede ser especificada como requisito de DQ en un BP. Estos casos de uso estándar han sido definidos en base a (i) la definición de cada dimensión de DQ, (ii) el conjunto de actividades que serán realizadas en función de los requisitos especificados (repositorio de actividades de DQ) y (iii) el conocimiento extraído de la literatura y de la experiencia de desarrolladores. En la Figura 3 se muestran algunos ejemplos de casos de uso estándar para las dimensiones de DQ exactitud y completitud.

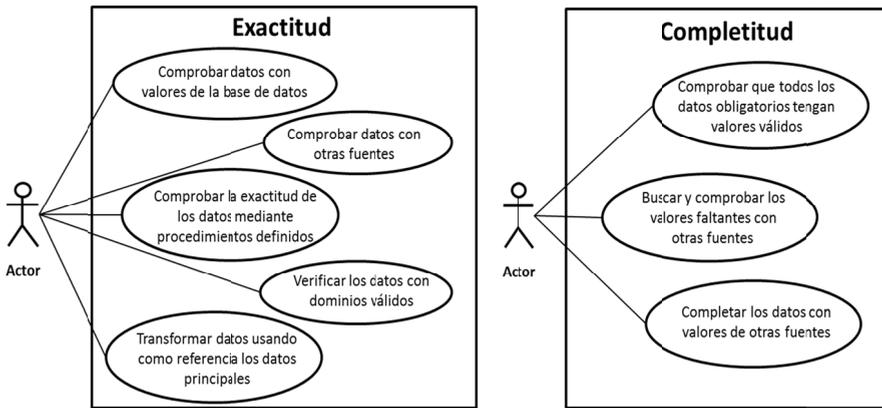


Figura 3 – Ejemplos de casos de uso estándar

Basados en estos casos de uso estándar de DQ, los trabajadores deberán hacer los ajustes necesarios de acuerdo a las características propias del BP y relacionarlos con los otros casos de uso obtenidos desde el BP propiamente dicho.

3.2 Etapas del método BPiDQ*

En las subsecciones siguientes se describen en detalle cada una de las etapas que componen el método BPiDQ*. Como se dijo anteriormente, el método sólo varía de la propuesta original (Caro et al., 2012) en las dos últimas, que es cuando se pone énfasis en la obtención de artefactos útiles para el desarrollo de software. En la Figura 4 se muestra una vista completa del método BPiDQ*.



Figura 4 – Las cuatro etapas del método BPiDQ*

En la Tabla 4 se muestra, en forma resumida, cada una de las etapas del método indicando las entradas, los trabajadores que participan, las herramientas que se utilizan y las salidas que se generan.

Tabla 4 – Resumen de las etapas de BPiDQ*

	BPiDQ-S1	BPiDQ-S2	BPiDQ*-S3	BPiDQ*-S4
Entradas	- Modelo BP	- Modelo de BP con DQ-Flags	- Modelo BP con DQ-Flags - Detalle de requisitos DQ	- Modelo BP con actividades DQ agregadas
Trabajadores	- Experto del Negocio/ Analista BP	- Analista BP - Experto DQ	- Diseñador BP - Experto DQ	- Analista de sistemas - Experto DQ
Herramientas	- BPMN 2.0 extendido con dqBP	- BPMN 2.0 extendido con dqBP	- BPMN 2.0 extendido con dqBP - Repositorio Actividades DQ	- Casos de uso UML - Repositorio de Casos de uso DQ
Salidas	- Modelo de BP con DQ-Flags	- Modelo de BP con DQ-Flags - Detalle de requisitos DQ	- Modelo BP con actividades DQ agregadas	- Diagramas de casos de uso del sistema

3.2.1 Etapa 1: Modelado de procesos de negocio consciente de la calidad de datos

Esta etapa está dedicada a la captura temprana de requisitos de DQ, los que son representados en un modelo de BP a nivel descriptivo de BPMN (Silver, 2009). Durante el modelado se incorporan marcas (DQ-Flags) donde se estime que la calidad de los datos involucrados en el BP es relevante para el éxito del negocio. Los elementos de entrada de esta etapa son el estándar BPMN y la extensión que permite incluir requisitos de DQ. Los trabajadores de esta etapa son el experto del negocio y/o el analista de procesos de negocio, quienes tienen la responsabilidad de definir las necesidades del negocio y, desde esa perspectiva, la importancia que tiene la DQ para el desempeño del mismo. El resultado de esta etapa es una descripción del proceso de negocio en la cual se han incluido marcas (DQ-Flags) que denotan el interés de los expertos del negocio por profundizar en la definición de los requisitos de DQ que son importantes para el buen desempeño del proceso de negocio. Junto con ello, también se deben identificar los elementos de datos involucrados en las marcas y una estimación del nivel de influencia (baja, media o alta) que tienen los datos asociados a esas marcas en el desempeño total del BP.

3.2.2 Etapa 2: Especificación de requisitos de calidad de datos

El principal objetivo de esta etapa es obtener una especificación detallada de los requisitos de DQ definidos en el proceso de negocio. El único elemento de entrada en esta etapa es el modelo de BP con requisitos de DQ (DQ-Flags). Los trabajadores involucrados en esta etapa son el analista de procesos de negocio y el experto en calidad de datos. Estos trabajadores determinan el conjunto final de DQ-Flags y especifican en forma detallada los requisitos de DQ asociados a cada uno de ellos. Las salidas de esta etapa son (i) el modelo del BP con requisitos de DQ (DQ-Flags) y (ii) para cada DQ-Flag, una especificación detallada que contiene: el elemento del BP en que se ha puesto el DQ-Flag, la importancia del requisito de DQ en el BP (alta, media o baja), la probabilidad de ejecución de la actividad asociada a la especificación del requisito de DQ, las dimensiones de DQ asociadas, la sobrecarga para el BP debido a la

incorporación de nuevas actividades asociadas a las dimensiones de DQ, el nombre del elemento de dato involucrado en el requisito de DQ, su descripción, medio de soporte y origen.

3.2.3 Etapa 3: Análisis y mejora de Procesos de Negocio relacionado con la Calidad de Datos

En esta etapa se analiza y deciden las mejoras que se pueden hacer al modelo del BP teniendo en cuenta los requisitos de DQ especificados. Esta etapa ha variado respecto de la primera versión del método ya que originalmente estaba centrada en mejorar el modelo del BP en sí mismo (reorganización de actividades, inclusión de nuevas actividades, ajuste de los flujos de ejecución, etc.). No obstante, en la versión BPiDQ* sólo se ha considerado la introducción de las nuevas actividades relacionadas con las dimensiones de DQ derivadas de la especificación de los requisitos de DQ. Las entradas en esta etapa son: la descripción del BP con especificaciones de calidad de datos (DQ-Flags), un detalle de las especificaciones de DQ del BP y un repositorio con las actividades que se relacionan con las dimensiones de DQ. Los trabajadores involucrados en esta etapa son el diseñador de procesos de negocio y el experto en DQ. Las dimensiones de DQ son utilizadas para seleccionar el conjunto de actividades de DQ que se deberán agregar a la descripción del proceso de negocio. El resultado de esta etapa es una descripción del proceso de negocio en la cual se han incluido nuevas actividades que consideran los requisitos de DQ.

3.2.4 Etapa 4: Generación de Diagramas de Casos de Uso

Esta etapa también ha sido modificada respecto del método original, ya que ahora no sólo se generan los casos de uso relacionados con DQ (obtenidos desde el repositorio de casos de uso estándar de DQ) sino que también se obtendrán casos de uso generales que se corresponden con el resto de los requisitos representados en el BP y que posteriormente serán implementados como parte del sistema de información. Para ello se ha tenido en cuenta una propuesta que permite obtener casos de uso desde la descripción de un BP (Rodríguez et al., 2007b). Las entradas de esta etapa son: la descripción del BP con las actividades de calidad de datos agregadas en la etapa anterior y un repositorio con los casos de uso estándar que se relacionan la DQ. Los trabajadores involucrados en esta etapa son el analista de sistemas y el experto en calidad de datos. Las actividades relacionadas con calidad de datos se usan para seleccionar el conjunto de casos de uso de DQ estándar y el resto de los casos de uso se obtienen en forma directa desde la descripción del BP. El resultado de esta etapa es un conjunto de casos de uso que pueden ser usados en un proceso de desarrollo de software. Los casos de uso estándar no tienen asociados actores específicos debiendo ser integrados con los casos de uso que representan las actividades del BP (que representan todos los requisitos de la aplicación que soportará el BP). De manera que los casos de uso relacionados con las dimensiones de DQ serán considerados como casos de uso «*include*».

4. Ejemplo ilustrativo

Para ejemplificar esta propuesta se ha considerado un proceso de negocio que describe el pago y la entrega de un pedido de productos. El proceso comienza con el pago de los productos. El pago se puede realizar de dos formas diferentes: con tarjeta de crédito o con efectivo (o cheque). Si el pago es realizado mediante tarjeta de crédito, es necesario pedir una autorización para la tarjeta de crédito a una «Institución Financiera». Si la autorización es rechazada, el pago con la tarjeta de crédito no será posible y el proceso finaliza. Por el contrario, si el pago con tarjeta de crédito es autorizado o si el pago es realizado en efectivo (o cheque), el «Departamento de Distribución» prepara el paquete y lo envía al cliente, después de lo cual el proceso termina. En la Figura 5 se muestra el BP descrito usando BPMN y la extensión para representar DQ. A continuación se describe, etapa por etapa, la forma en que se ha aplicado el método BPiDQ*.

En la primera etapa, (*BPiDQ-S1: Modelado de Procesos de Negocio consciente de la Calidad de Datos*), los expertos del negocio y/o analista del negocio identifican los elementos de BPMN en el modelo del BP que necesitan mayor atención en cuanto a calidad de datos para lograr el éxito del proceso. Ellos marcarán cada uno de esos elementos mediante el símbolo gráfico Ω . En el ejemplo ilustrativo, se incluyeron dos marcas (DQ-Flags). La primera, denominada DQFlag1 (el número de secuencia asignado a la marca obedece a la lectura del modelo de arriba a abajo y de izquierda a derecha), fue asociada con el Data Object de entrada a la actividad “*Entregar el paquete al Cliente*” (ver Figura 5). Este Data Object contiene el elemento de datos denominado “*Orden de Entrega*” que contiene la información del cliente necesaria para hacer la entrega del paquete (identificación, dirección). La segunda marca, denominada DQFlag2, fue asociada al elemento de BPMN Message Flow que va desde el pool «Institución Financiera» hasta el lane «Ventas». Este Message Flow contiene un mensaje con la respuesta de la «Institución Financiera» a la solicitud de aprobación o rechazo del pago con tarjeta de crédito. La salida de esta etapa es el modelo del BP enriquecido con las marcas asociadas a los requisitos de DQ (DQ-Flags).

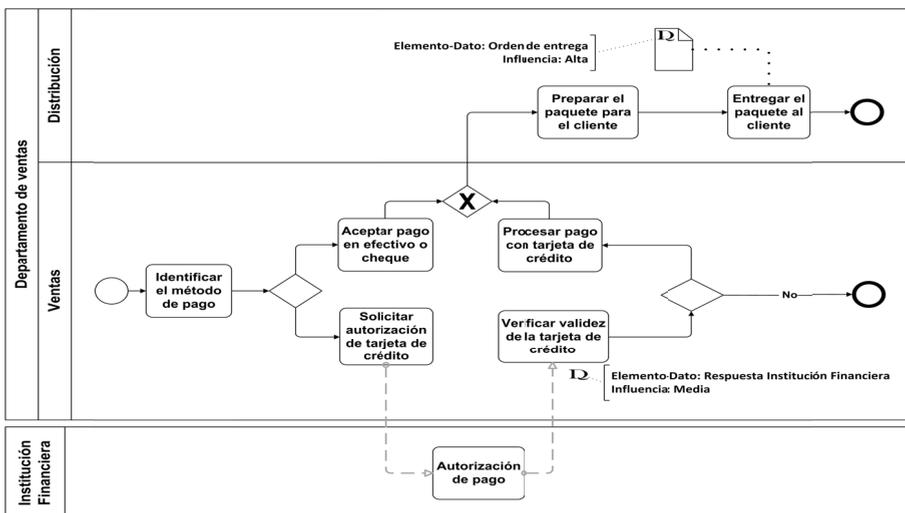


Figura 5 – Ejemplo Ilustrativo: Modelo BPMN con DQ-Flags

En la segunda etapa, (*BPiDQ-S2: Especificación de requisitos de Calidad de Datos*), los trabajadores (analista de BP y experto de DQ) registran información acerca del BP y de los DQ-Flags. Por cada uno de los DQ-Flags los trabajadores deben especificar los requisitos de DQ en forma más detallada (dimensiones de DQ y su importancia). El requisito de DQ que se ha marcado en la “*Orden de Entrega*” involucra dos dimensiones de DQ: exactitud y completitud. Por otro lado, para el DQFlag2 se define sólo la dimensión de DQ actualidad. Adicionalmente, para los dos DQ-Flags se obtienen o calculan la probabilidad de ejecución y la sobrecarga que implica la agregación de actividades al BP. En la Tabla 5 se muestra el detalle de las especificaciones realizadas por cada DQ-Flag. Tomando en cuenta la información disponible, los trabajadores de esta etapa, decidirán el conjunto definitivo de dimensiones de DQ para los elementos de datos en cada DQ-Flag. En el ejemplo ilustrativo el DQFlag1 (asociado a la “*Orden de Entrega*” en el Data Object) tiene un alto impacto en el éxito del BP. La probabilidad de ejecución de la actividad con que se relaciona el DQFlag1 es de un 75% (tomando en cuenta las bifurcaciones previas y considerando que algunas veces la actividad puede no ejecutarse). La sobrecarga calculada es de 25% porque para poder satisfacer los requisitos de DQ se deben incluir dos actividades nuevas (ver en la Figura 6, en el lane «Distribución», las actividades nuevas en color gris). Por su parte el DQFlag2 tiene un impacto medio en el éxito del BP. La probabilidad de solicitar la autorización de pago es del 50% porque cuando no se paga con tarjeta de crédito la actividad relacionada con el DQ-Flag no es ejecutada. La sobrecarga para este DQ-Flag es de 12.5% porque para satisfacer las dimensiones de DQ se debe incluir sólo una nueva actividad en el proceso (ver en la Figura 6, en el lane «Ventas», la nueva actividad en color gris).

Tabla 5 – Detalle de las especificadores para cada requisito de DQ

Especificación de DQFlag-1		Especificación de DQFlag-2			
Elemento de BPMN	Data Object	Elemento de BPMN	MessageFlow		
Influencia sobre el BP	Alta	Influencia sobre el BP	Media		
Probabilidad de Ejecución	75%	Probabilidad de Ejecución	50%		
Sobrecarga del BP	$2/8 * 100 = 25\%$	Sobrecarga del BP	$1/8 * 100 = 12.5\%$		
Requisitos DQ	Elemento de Dato	Requisitos DQ	Elemento de Dato		
	Nombre		Orden de Entrega	Nombre	Respuesta de la institución financiera
	Descripción		Nombre y dirección del cliente	Descripción	Aprobación/rechazo del pago
	Medio soporte		Electrónico	Medio soporte	Electrónico
	Origen		Interno	Origen	Interno
	Dimensiones de Calidad de Datos		Dimensiones de Calidad de Datos		
	Exactitud		Alta	Actualidad	Alta
Completitud	Media				

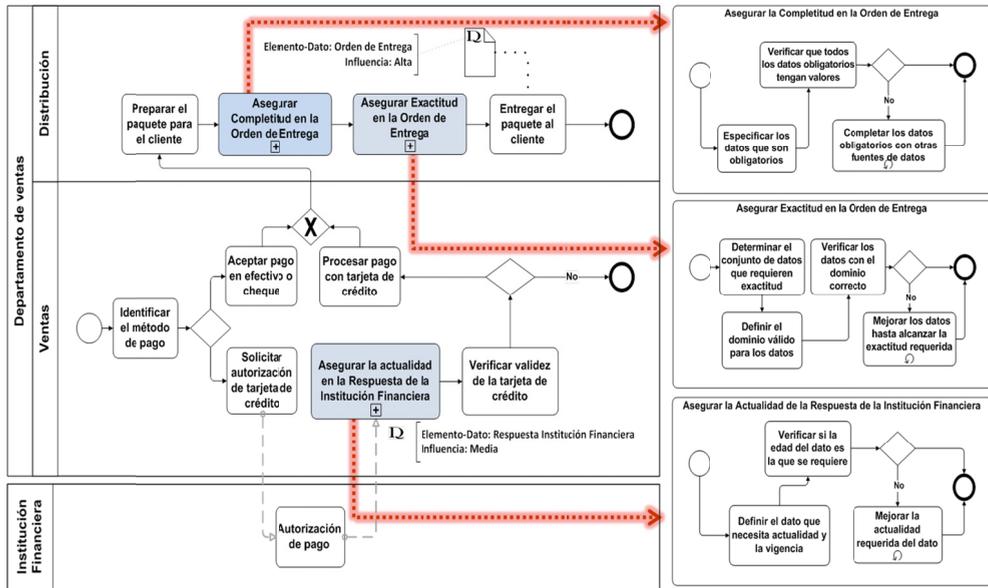


Figura 6. Modelo de BP mejorado con nuevas actividades de DQ

En la tercera etapa, (BPiDQ*-S3: *Análisis y mejora del Proceso de Negocio relacionado con la Calidad de Datos*), el diseñador de BP y el experto de DQ deben decidir cuál es el conjunto final de dimensiones de DQ que será considerado para cada requisito de DQ especificado. Luego, para cada dimensión de DQ se seleccionarán las actividades más adecuadas al BP las que serán extraídas desde el repositorio de actividades de DQ. En el ejemplo, se han agregado tres actividades de mejora (lado izquierdo de la Figura 6 en oscuro) que se presentan en el BP en forma colapsada y cuya representación detallada se puede observar a la derecha de la Figura 6.

Finalmente, en la cuarta y última etapa, (BPiDQ*-S4: *Generación de Diagramas de Casos de Uso*), el analista de sistemas y el experto de DQ deberán analizar los casos de uso generados en forma automática desde la descripción del proceso de negocio. Estos casos de uso deberán servir de base para la elaboración de los casos de uso definitivos que serán utilizados en la construcción del software. En la Figura 7. Se muestra el diagrama de casos de uso derivado del BP. En gris se han marcado aquellos casos de uso que se relacionan directamente con las especificaciones de requisitos de DQ y que han sido derivados desde el modelo de BP.

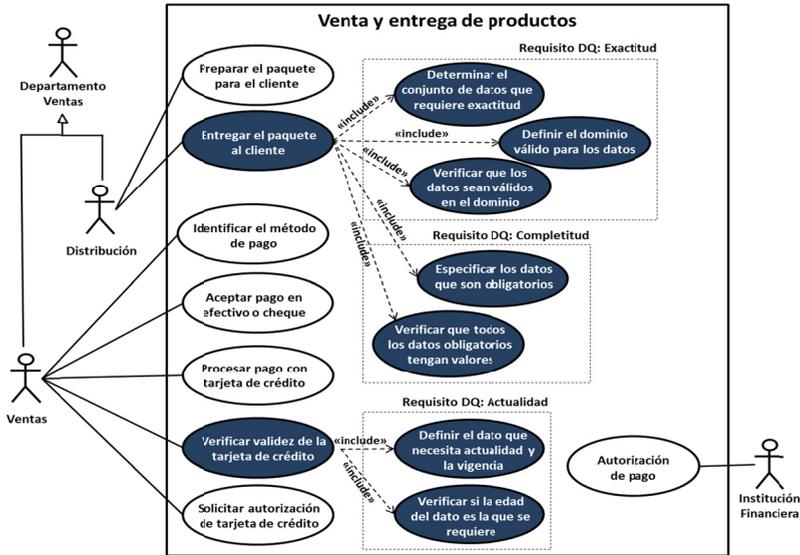


Figura 7. Casos de uso derivados del BP mejorado

5. Conclusiones

En este artículo se ha presentado el método BPiDQ* cuyo objetivo es generar casos de uso centrados en la DQ a partir de modelos de BP descritos en BPMN. Esta propuesta, tiene como propósito cubrir dos necesidades del campo de los sistemas de información. En primer lugar, la obtención de artefactos útiles para el desarrollo de software a partir de modelos de BP. Y en segundo lugar, la necesidad de especificar tempranamente requisitos de DQ para un sistema de información.

Como trabajo futuro se planea a corto plazo la implementación de una herramienta que soporte la aplicación del método y que, por tanto, facilite su uso por parte de los distintos trabajadores involucrados. Asimismo, se realizarán casos de estudio que permitan ajustar y mejorar el método.

6. Agradecimientos

Proyecto (MECESUP-UBB0704) orientado al Fortalecimiento de Núcleos Académicos en Programas de Postgrado en la Universidad del Bío-Bío.

7. Referencias

Bagchi, S., Bai, X., & Kalagnanam, J. (2006). Data quality management using business process modeling.

- Bocciarelli, Paolo , & D'Ambrogio, Andrea. (2011). A BPMN extension for modeling non functional properties of business processes. Paper presented at the Proceedings of the 2011 Symposium on Theory of Modeling & Simulation: DEVS Integrative M&S Symposium, Boston, Massachusetts.
- Bringel, H., Caetano, A., & Tribolet, J. (2004, April 14-17, 2004). Business Process Modeling Towards Data Quality Assurance. Paper presented at the 6th International Conference on Enterprise Information Systems, Porto, Portugal.
- Caro, Angelica, Rodriguez, Alfonso, Cappiello, Cinzia, & Caballero, Ismael. (2012). Designing Business Processes able to satisfy Data Quality Requirements. Paper presented at the 17th International Conference on Information Quality (ICIQ), Paris, France.
- Coskuncay, Ahmet, Aysolmaz, Banu, Demirors, Onur, Bilen, Omer, & Dogani, Idris. (2010). Bridging the gap between business process modeling and software requirements analysis: A case study Paper presented at the Proceedings of MCIS 2010. Paper 20.
- el Abed, W. (2009). Data Governance: A Business Value-Driven Approach.
- Goldner, Sascha, & Papproth, Alf. (2011). Extending the BPMN Syntax for Requirements Management. Paper presented at the Business Process Model and Notation.
- Harmon, Paul, & Wolf, Celia. (2011). Business Process Modeling Survey. Business Process Trends (<http://www.bptrends.com/>). <http://www.bptrends.com/>
- Heravizadeh, M., Mendling, J., & Rosemann, M. (2009). Dimensions of business processes quality (QoBP).
- Liew, P., Kontogiannis, K., & Tong, T. (2004, 17-19 Sept. 2005). A framework for business model driven development. Paper presented at the Software Technology and Engineering Practice, 2004. STEP 2004. The 12th International Workshop on.
- Lodhi, Azeem, Veit, Köppen, & Saake, Gunter. (2011). An Extension of BPMN Meta-model for Evaluation of Business Processes. J. Riga Technical University, 43, 27-34.
- Object Management Group. (2004). Meta Object Facility (MOF) 2.0 Query/View/Transformation Specification.
- Object Management Group. (2007). Unified Modeling Language: Superstructure Version 2.1.1 (formal/2007-02-05), .
- Object Management Group. (2011). Business Process Model and Notation (BPMN) Version 2.0.
- Recker, J. (2010). Opportunities and constraints: the current struggle with BPMN. Business Process Management Journal, 16(1), 181-201.
- Redman, Thomas. (2008). Data Driven: Harvard Business School Press.

- Rodríguez, Alfonso, Caro, Angelica, Cappiello, Cinzia, & Caballero, Ismael. (2012). A BPMN extension for including data quality requirements in business process modeling. Paper presented at the 4th International Workshop on the Business Process Model and Notation, Vienna, Austria.
- Rodríguez, Alfonso, Fernández-Medina, Eduardo, & Piattini, Mario. (2007a). A BPMN extension for the modeling of Security Requirements in Business Processes. *IEICE Transactions on Information and Systems*, 90(4), 745-752.
- Rodríguez, Alfonso, Fernández-Medina, Eduardo, & Piattini, Mario. (2007b). Towards CIM to PIM transformation: from Secure Business Processes defined by BPMN to Use Cases. Paper presented at the 5^o International Conference on Business Process Management (BPM), Brisbane, Australia.
- Saeedi, Kawther , Zhao, Liping, & Falcone Sampaio, Pedro R. . (2010). Extending BPMN for Supporting Customer-Facing Service Quality Requirements. Paper presented at the Proceedings of the 2010 IEEE International Conference on Web Services
- Silver, Bruce. (2009). *BPMN Method & Style: A levels-based methodology for BPM process modeling and improvement using BPMN 2.0*: Cody-Cassidy Press.
- Soffer, P. (2010). Mirror, mirror on the wall, can i count on you at all? exploring data inaccuracy in business processes. *Enterprise, Business-Process and Information Systems Modeling*, 14-25.
- Strong, Diane, Lee, Yang, & Wang, Richard. (1997). Data Quality in Context. *Communications of the ACM*, Vol. 40, N^o 5, 103 -110.
- Wang, R., & Strong, D. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems; Armonk; Spring*, 12(4), 5-33.
- Weske, Mathias. (2007). *Business Process Management: Concepts, Languages, Architectures* (ed.): Springer-Verlag Berlin Heidelberg.