

Modelo de búsqueda de procesos de negocios basado en unidades mínimas estándar *Business Process Modeling Notation**

Daniela Pianda*
David Moncayo**✉

Fecha de recepción: 19/08/2013
Fecha de aceptación: 05/12/2013

Cómo citar este artículo / To reference this article / Para citar este artigo: Pianda, D. & Mocayo, D. (2013). Modelo de búsqueda de procesos de negocios basado en unidades mínimas estándar *Business Process Modeling Notation*. *Revista UNIMAR*, 31 (2), pp. 31-40.

RESUMEN

En este artículo se presenta un modelo de búsqueda de procesos de negocio, basado en las unidades mínimas estándar en la notación de modelado de procesos de negocio, el cual realiza una recuperación efectiva y rápida de los procesos de negocio consultados por parte del usuario. Está compuesto por tres unidades: Capa de datos, capa de indexación y capa de consulta, cada una de ellas con unos componentes y funciones diferentes. Se describe además la variedad de repositorios y lenguajes para ejecutar los procesos de negocio modelados, se presenta la herramienta implementada, y finalmente, los resultados de la evaluación a la que fue sometido.

Palabras clave: modelo, lenguaje, proceso de negocio, recuperación de información, repositorio.

Search model of business processes based in minimum standard units *Business Process Modeling Notation*

ABSTRACT

This article presents a search model of business processes based on minimum standard units in the notation for modeling of business processes, which does a quick and effective recovery of business processes consulted by the user. It consists of three components: 1) layer of data, 2) indexing layer and 3) layer of consultation, each with its different units and functions. In addition, it describes the variety of repositories and languages to execute modeled business processes. Finally it presents the tool implemented and the results of the assessment to which it was subjected.

Key words: model, language, business process, information retrieval, repository.

Modelo de pesquisa dos Processos de negócios baseado em unidades mínimas padrão *Business Process Modeling Notation*

RESUMO

Este artigo apresenta um modelo de busca dos processos de negócios com base em unidades mínimas padrão na cotação de modelagem nos processos de negócio, que executa uma recuperação rápida e eficaz dos processos de negócio consultados pelo usuário. O modelo é composto por três componentes: 1) camada de dados, 2) camada de indexação e 3) camada de consulta, cada uma com elementos e funções diferentes. Além disso, é descrita a variedade de repositórios e idiomas para executar processos de modelagem de negócio e apresenta a ferramenta que foi aplicada no modelo e, finalmente, os resultados da avaliação a que foi submetido.

Palavras-chave: modelo, linguagem, processos de negócios, recuperação de informação, repositório.

♦ Artículo de investigación e innovación.

* Ingeniera de Sistemas, Universidad Mariana, San Juan de Pasto, Nariño, Colombia. Correo electrónico: daniela90@hotmail.com

**✉ Ingeniero de Sistemas, Universidad Mariana, San Juan de Pasto, Nariño, Colombia. Correo electrónico: d.alonsomt@gmail.com

1. Introducción

En la actualidad existen empresas que soportan los objetivos de la organización en su infraestructura de las tecnologías de información (TI), para garantizar la agilidad ante el mercado y la rápida recuperación de la inversión en las diferentes áreas de negocio; basan sus fortalezas en arquitecturas que resuelven los problemas de heterogeneidad e interoperabilidad, que facilitan el aprovisionamiento de los servicios típicos de cualquier infraestructura de la tecnología de información. En este sentido la *Service Oriented Architecture* (Arquitectura Orientada a Servicios, SOA), es una alternativa que garantiza la rápida salida de los servicios al mercado, como respuesta a la competitiva presión global. SOA es una evolución de la computación distribuida y diseñada para permitir la interacción de componentes *software*, llamados servicios Web, a través de la red (Aalst, 2010).

Un servicio web es definido como una aplicación *software* que utiliza interfaces estándares que permiten a cualquier servicio inter-operar con otras aplicaciones en la Web. Una de las principales ventajas de estos servicios es la capacidad de convertirse en servicios más complejos, a partir de funciones suministradas por diferentes componentes, que están dinámicamente integrados para garantizar una tarea de negocio más robusta, estos servicios web complejos reciben el nombre de procesos de negocio (Chaves, 2010). Teniendo en cuenta lo dicho anteriormente, las aplicaciones SOA son creadas como una composición de varios servicios web o procesos de negocio compartidos entre múltiples aplicaciones (Bernstein & Dayal, 1994).

En este marco, diversos autores indican que el descubrimiento de procesos de negocios se puede realizar de forma automática o manual. Algunos afirman que la técnica más utilizada actualmente es la localización manual, que consiste en buscar las descripciones en repositorios y registros para posteriormente invocarlos, pero la recuperación manual se torna algo compleja cuando el número de éstos es considerable, lo que implica que utilizar esta técnica es menos eficiente que descubrir procesos de manera automática. La recuperación automática

reduce considerablemente el tiempo invertido en la selección de unos procesos con el fin de ser reutilizados; esto implica una significativa reducción en los costos de desarrollo de nuevas aplicaciones o servicios de valor agregado.

Es así como, en este artículo se propone un modelo de repositorio, basado en unidades mínimas estándar del lenguaje *Business Process Modeling Notation* (BPMN) (en español Notación para el Modelado de Procesos de Negocio), que permita el almacenamiento y análisis de distintos tipos de *Business Process* (BP) existentes en una organización, para obtener una visión completa de los procesos que se ejecutan actualmente. En la siguiente sección se presenta algunos trabajos relacionados, posteriormente la metodología relacionada con la temática y el modelo propuesto, para luego, mostrar una evaluación parcial del modelo basado en precisión, recuerdo y medida F, medidas ampliamente utilizadas en recuperación de información, y finalmente, las conclusiones y el trabajo a futuro.

A nivel nacional se está innovando en la tercerización de procesos, lo que conlleva al uso de un modelo para la notación de procesos de negocios; uno de los ejemplos claros sobre esto es *Suite Business Process* (BP-Suite), un conjunto de herramientas basadas en la ejecución de procesos especificadas en *Business Process Execution Language* (BPEL). Consta de *Business Process Query Language* (BP-QL) un lenguaje para las definiciones de procesos. *Monitoring Process Business* (BP-Mon) es una herramienta para monitorear las instancias de procesos de negocio en ejecución y *Execution Process Business* (BP-Ex), que es una herramienta para analizar las anotaciones de las ejecuciones (Mazanek & Hanus, 2011).

BP-Suite soporta los aspectos de actividades, flujo de control, datos, recursos, estructura organizacional, monitoreo y autorización, además del almacenamiento de modelos de procesos, instancias de procesos e información histórica acerca de las instancias. Los modelos de procesos son almacenados en formato BPEL (Mazanek & Minas, 2009). Además de las funcionalidades estándar de una base

de datos para el almacenamiento y recuperación de modelos de procesos, *BP-Suite* soporta consultas sobre los 3 tipos de procesos, y puede almacenar definiciones, instancias en ejecución, anotaciones, entre otras (Elhadad, 2008).

Con respecto a esta investigación, se tiene en cuenta el BPMN como lenguaje propietario para consultas, instancias y manejo de historial a través del IDE Eclipse; en este caso se necesita del IDE Eclipse para ser ejecutado, pero no permite realizar consultas semánticas.

Dentro de un contexto internacional se encontró ORYX, herramienta de modelado de procesos basada en Web que soporta la búsqueda de usuarios, creación, almacenamiento y actualización de modelos de procesos en línea. La herramienta utiliza un repositorio para almacenar los modelos de procesos de negocio que son creados en él (Decker & Weske, 2008). ORYX se enfoca principalmente en el aspecto del flujo de control y almacena modelos de procesos específicos de compañías, además soporta muchas notaciones de modelado de procesos, incluyendo: BPMN, EPC, redes de Petri, redes de Workflows, FMC Block Diagrams y XForms. Internamente, los procesos son almacenados en una base de datos; externamente representa modelos de procesos en un formato RDF (Smirnov, Weidlich, Mendling & Weske, 2012).

Además se tiene en cuenta algunas empresas que trabajan con BPMN, como por ejemplo IBM, donde “los recursos que *Business Process Management* provee a los usuarios especializados en el negocio son aquellos referidos a la entrega de la información necesaria para buscar, evaluar, armar, desarrollar y manejar soluciones de integración de negocio multi-productos, basado en *Software de BPM WebSphere*” (IBM, s.f.). Otra compañía que trabaja con BPMN es LANSÁ Composer, la cual brinda el poder de la tecnología en la integración de procesos de negocio en un ambiente altamente visual para que diseñadores y analistas puedan automatizar e integrar rápidamente los procesos de negocio. LANSÁ Composer es una plataforma de diseño y transformación de datos y procesamiento de negocio personalizado (LANSÁ, s.f.).

Por otra parte La Rosa y Cols (2011, citado por Figueroa, Corrales & Ramirez-Gonzalez, 2012), refieren que:

Apromore es un repositorio avanzado que mantiene, analiza y reutiliza grandes colecciones de modelos de BP. Además, es una plataforma de código abierto, implementada de acuerdo a la arquitectura SOA y por lo tanto accesible a través de servicios web. La representación de los procesos está basada en un formato canónico que utiliza EPC y BPMN como lenguajes de modelado. Actualmente existe una versión de un prototipo disponible en la red que contiene funcionalidades básicas como: importar, exportar, buscar, clasificar y comparar modelos de BP (p. 37).

Estos trabajos relacionados fueron tomados como referencia para el desarrollo de la investigación, debido a que son una base de optimización en la búsqueda de los procesos de negocios, ampliando la perspectiva de lo que se puede obtener de dichos procesos, y ensamblando un repositorio que ayude a la mayoría de empresas a formular sus propios procesos, teniendo en cuenta que esto ayuda a disminuir los gastos de una empresa, y a cambio se obtiene una serie de modelos optimizados y reutilizables según el ámbito en que se encuentren.

2. Metodología

Para el proceso de esta investigación se aplicó el modelo PHVA (Planear, Hacer, Verificar y Actuar) (García et al., 2001), debido a su gran aporte en el orden de ejecución de la comprobación del modelo planteado; además, permite un control permanente de las diferentes etapas de evaluación. En la planeación se define los métodos con los que podemos alcanzar los objetivos propuestos, recopilando los datos y teorías necesarias; en la etapa de hacer se plantea un modelo a realizar para el desarrollo de una herramienta prototipo, continuando con la fase de verificación o pruebas, y así, en la última etapa de actuar se concluyó la investigación. La primera etapa de planeación permite el siguiente análisis:

La Tabla 1 muestra la comparación entre los repositorios investigados y el propuesto, permitiendo así un análisis del lenguaje que utilizan, el tipo de almacenamiento y el tipo de consulta.

Tabla 1. *Comparación entre repositorios.*

Repositorios	Lenguajes				Tipo Almacenamiento		Tipo de Consulta	
	BPMN	BPEL	XPDL	Redes de PETRI	Base de datos	INDEX	GRAFOS	Semántica
APROMORE	X				X		X	
PROSERO	X	X			X		X	
BP-SUITE		X			X		X	
ORYX	X				X		X	
REPOX					X		X	
BPMN REPOSITORY	X				X		X	
IPM		X			X		X	
PVR		X					X	
MIT PROCESS HANDBOOK	X				X		X	
IBM BPEL REPOSITORY		X			X		X	
MODELO DE BUSQUEDA BASADO EN BPMN	X		X		X		X	

En la Tabla 2 se observa las ventajas y desventajas de los repositorios instigados y el propuesto en esta investigación.

Tabla 2. *Ventajas y desventajas de los repositorios.*

Repositorios	Ventajas	Desventajas
APROMORE	Mantiene, analiza y reutiliza grandes colecciones de BP. Utiliza BPMN y EPC como lenguajes de modelado. Implementado según la arquitectura SOA. Plataforma de código abierto.	No proporciona búsquedas semánticas. Su prototipo disponible en la red y contiene unas funcionalidades básicas.
PROSERO	Soporta aspectos dentro de un BP, los cuales son: actividades, flujos de control, datos, recursos, autorización y estructura organizacional.	No permite realizar búsquedas semánticas. No posee un lenguaje definido para la recuperación de BP.

Repositorios	Ventajas	Desventajas
BP-SUITE	<p>Soporta aspectos como actividades, flujo de control, datos, recursos, estructura organizacional, monitoreo y autorización.</p> <p>Los BP son almacenados en formato BPEL.</p> <p>Soporta consultas sobre 3 tipos de procesos que se puede almacenar (definiciones, instancias en ejecución, anotaciones).</p>	<p>Necesita del IDE Eclipse para ser ejecutado.</p> <p>No permite realizar búsquedas semánticas.</p>
ORYX	<p>Soporta distintos tipos de notaciones de modelado como BPMN, redes de Petri, XFroms, etc.</p> <p>Permite almacenar los BP creados en ORYX.</p> <p>Puede importar y exportar distintos formatos como ERDF, XPDL, XHTML.</p>	<p>Para añadir semántica al contenido se debe combinar RDFSschema y OWL.</p> <p>No permite realizar búsquedas semánticas.</p>
REPOX	<p>Modelo de repositorio basado en XML.</p> <p>Se puede intercambiar los BP utilizando documentos XML.</p> <p>Los BP son almacenados en una BD orientada a objetos.</p> <p>Almacena los aspectos del flujo de control de los procesos junto con los datos usados en los procesos.</p>	<p>No cuenta con una notación específica.</p> <p>No permite realizar búsquedas semánticas.</p> <p>Todas las consultas están basadas en XQuery.</p>
BPMN R EPOSITORY	<p>Está dirigido a otros repositorios de BP que soporten procesos inter organizacionales.</p> <p>Soporta aspectos de BP como flujo de control, datos estructuras organizacionales, monitoreo y controles.</p>	<p>No tiene un lenguaje de consulta definido.</p> <p>No permite realizar consultas semánticas.</p>
IPM	<p>Permite la gestión de versiones y configuración de un BP.</p> <p>Gestiona los BP a través del ciclo de vida del mismo.</p> <p>Permite intercambiar los BP con el repositorio a través del lenguaje IPM EPDL.</p> <p>Utiliza como lenguaje base XML o una BD relacional.</p>	<p>Es propietario tanto el lenguaje de modelado como de consulta.</p> <p>Está diseñado para modelar los procesos específicos de las compañías.</p>
PVR	<p>Soporta actividades, flujo de control, datos, recursos, aspectos de monitoreo y autorización.</p> <p>Almacena en la BD procesos específicos, instancias e historiales de BP.</p>	<p>El lenguaje de modelado es propietario.</p> <p>No utiliza un lenguaje específico para el almacenamiento y recuperación.</p>
MIT PROCESS HANDBOOK	<p>Se basa en un esquema para la organización de modelos de procesos y sus actividades.</p> <p>Organiza el conocimiento relacionado con los procesos.</p> <p>El almacenamiento de los BP ingresa en forma de texto en la BD.</p>	<p>Se enfoca en almacenar las referencias a los procesos para tenerlos disponibles como una base de conocimiento generado.</p> <p>Sólo hace una descripción del BP.</p> <p>No tiene un orden o tipo de almacenamiento predefinido.</p>
IBM BPEL REPOSITORY	<p>Permite crear, leer, escribir y borrar datos almacenados en archivos XML como objetos JAVA.</p> <p>Utiliza lenguajes como BPEL, WSDL, XSD, entre otros.</p> <p>Está creado para ser adaptable a nuevas funcionalidades o extensiones como OCL o EMF.</p>	<p>Sólo permite almacenar archivos en formatos BPEL y XML.</p> <p>No tiene características de búsqueda semántica.</p> <p>Para realizar las consultas utiliza objetos OCL, obligando a tener un motor OCL ligado al IDE Eclipse.</p>

En la segunda etapa se desarrolló el siguiente modelo:

2.1 Modelo propuesto

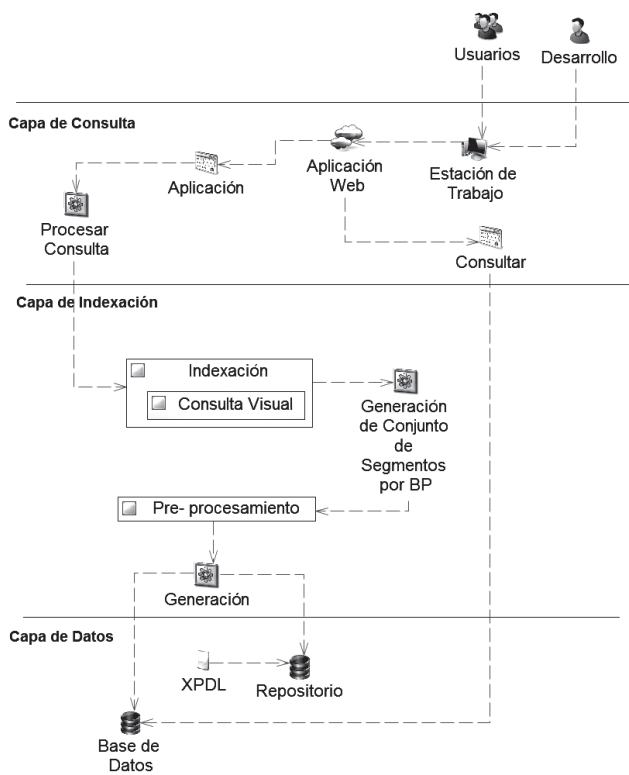


Figura 1. Arquitectura del modelo propuesto.

Los procesos de negocios almacenados en la base de datos son reutilizables en cualquier momento, por ende, la información contenida en estos BP sirve como base para replantear o modelar nuevos BP que se ajusten a los requerimientos actuales.

Este modelo se centra en aplicar una estrategia de búsqueda que utiliza información de la estructura secuencial de los BP contenidos en el repositorio, con el fin de aumentar el nivel de relevancia en los resultados de las consultas realizadas al repositorio.

El modelo está compuesto por tres capas (ver Figura 1): capa de datos, capa de indexación y capa de consulta, las cuales se describe a continuación:

2.2 Capa de datos

Esta capa contiene el repositorio que almacena los BP, que permite tener acceso a los atributos de in-

formación presente en cada BP -roles, descripciones de actividades, temporizadores, mensajes, llamadas a servicios- (Narayanan, Jayaraman, Luo & Swaminathan, 2011). BP modelos están modelados en BPMN y descritos en sintaxis XPDL (XML Process Definition Language), también, incorpora un algoritmo parser que construye un conjunto de componentes secuenciales del flujo de cada uno de los BP denominado MCS (Matriz de Componentes Secuenciales). Los principales componentes y procesos son descritos a continuación.

Repositorio: es usado como unidad central de almacenamiento y gestión, con similitud a una base de datos que comparte información acerca de los artefactos de ingeniería producida o utilizada por una empresa. Contiene 120 patrones de BP (ver Figura 2). Además, permite la visibilidad de tareas, subprocesos y flujos de control que hacen parte de cada uno de los BP almacenados en el MCS.

Archivos XPDL: es un formato de archivo basado en lenguaje XML que se usa para el intercambio de BPM entre distintas herramientas.

2.3 Capa de indexación

En esta capa se crea un índice para la organización de archivos; esta estructura se realiza indexando uno por uno los BP (ver Figura 2) para realizar su respectivo pre-procesamiento, donde se traduce todas las palabras en minúsculas, se elimina palabras vacías y caracteres especiales y, finalmente, se aplica derivada (Algoritmo de Porter) para convertir cada uno de los componentes textuales BP a su raíz léxica.

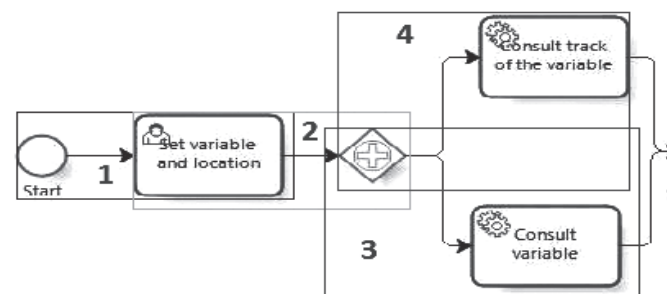


Figura 2. Creación de Codebook o Patrón.

La Figura 3 en un ejemplo de un proceso de negocio con su respectiva secuencia de *codebook* o patrones.

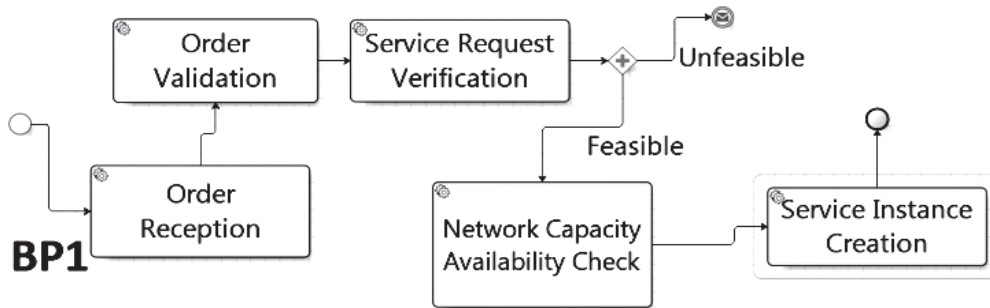


Figura 3. Modelo de BP1.

El proceso de indexación inicia tomando cada uno de los BP almacenados en el repositorio ($BP_1, BP_2, BP_n, BP_1, BP_2, BP_n, \dots$), y luego se crea los componentes y las características textuales. A continuación se puede observar los componentes de la Matriz de Codebooks (MCs) de la Figura 3, representados e indexados en la matriz **MCd**:

$MCd = \{[StartEvent_TaskService, TaskService_TaskService, TaskService_TaskService, TaskService_RouteParallel, RouteParallel_TaskService, RouteParallel_TriggerResultMessage, TaskService_TaskService, TaskService_EndEvent]\}$

En concordancia con lo anterior, se muestra la manera en la que se forma la matriz donde cada celda W_{ij} muestra la importancia del componente textual en su raíz léxica o la matriz de componentes secuenciales MCS versus BP y se basa en la **Ecuación 1**, propuesta por Salton (Manning, Raghavan & Schütze, 2008; Baeza-Yates & Ribeiro-Neto, 1999), donde F_{ij} es el componente frecuencia observada o codebookj en el BP_i , $\text{Max } F_i$ es la frecuencia más observada en el BP_i , donde n es el número de BP en la colección y n_i es el número de BP en la que aparece el componente textual o codebookj. Por último, podemos ver la matriz de indexación $MI = \{MCd_{ij}\}$ en la Figura 4 con cada uno de los componentes de MI.

Ecuación 1

$$W_{ij} = \frac{F_{ij}}{\max(F_i)} \times \log \left[\frac{N}{n_j + 1} \right]$$

	1	2	3	4	5	6
	1	Cd1	Cd2	Cd3	Cd4	Cdm
BP1	2					
BP2	3	W_{ij}				
BP3	4		W_{ij}			
BP4	5			W_{ij}		
BPn	n				W_{ij}	

Figura 4. Matriz índice.

BP1	0.201	0.301	0.301	0.602	0.602
BP2	0.201	0.398	0.398	0.602	0.398
BPn	0.201	0.301	0.398	0.398	0.301

Figura 5. Matriz de indexación.

Teniendo en cuenta la Figura 5, se puede observar los mejores resultados que se consigue al momento de realizar la consulta. Los resultados que se encuentran resaltados (en gris) tienen un mayor nivel de relevancia por columnas o por codebook, porque se recuperan BP que relacionan los datos compuestos por el índice de la matriz.

2.4 Capa de consulta

Esta capa es responsable de proporcionar una interfaz de usuario donde se realiza las consultas. Estas búsquedas se representan como un vector de términos $q = \{t_1, t_2, t_3, t \dots n\}$, de manera que

los términos del vector q se convirtió en minúsculas, se eliminaron los acentos, caracteres especiales y, finalmente, es aplicada la derivada (Algoritmo Porter) para convertir cada uno de los términos de q a su raíz léxica. Con la consulta procesada, una cadena de búsqueda se ejecuta, y devuelve los resultados filtrados a través de la **Ecuación 2**, que se describe a continuación.

Ecuación 2

$$score(q, d) = coord(q, d) * Qnorma(q) \sum_{t \text{ in } q} (tf(t \text{ in } d) * idf(t)^2 * t.getBoost * norm(t, d))$$

En la ecuación anterior t es un término de la consulta q y d es el documento visualizado, $tf(t \text{ en } d)$ frecuencia del término, que se define como el número de veces que el término t aparece en el BP d .

En esta medida, los documentos clasificados más altos son a menudo para contener el término con más frecuencia; $idf(t)^2$ la frecuencia inversa del término t en un documento (número de documentos en los que aparece término t); $coord(q, d)$ es el factor de puntuación en función de los términos de consulta se encuentran en el documento, los documentos contienen los términos de consulta con mayor puntuación; $Qnorma(q)$ es un factor de normalización utilizado para hacer las puntuaciones (para este modelo se toma el valor de 1, lo que hace es no afectar a la puntuación de cada documento evaluado); $t.getBoost()$ es una ponderación en el término t de consulta y en este caso es igual a 1, porque todos los términos de la consulta tienen el mismo peso; $norma(t, d)$ es una ponderación en el índice, tomado de W_{ij} en el índice.

Una vez que los resultados se ordenan y se filtran, se enumeran en orden de acuerdo a la similitud que aparecen en relación a la consulta realizada por el usuario, que puede seleccionar y visualizar cada uno de los modelos BP recuperados.

2.5 Herramienta que soporta el modelo propuesto

La herramienta cuenta con una interfaz centrada en el usuario final, procurando incorporar los principales atributos que componen la usabilidad. Atributos objetivos como facilidad de aprendizaje, faci-

lidad de memorización, eficacia, eficiencia o tiempo empleado para completar una tarea, operabilidad, y facilidad de comprensión; así como también, atributos subjetivos orientados a la satisfacción del usuario como: accesibilidad, funcionalidad, utilidad, estética y credibilidad. Además, cuenta con una interfaz sencilla y usable que está conformada por paneles, que forman las diferentes funcionalidades del modelo.

Por otra parte, el usuario está en la capacidad de elegir un modelo de BP de la lista de resultados para visualizar, y así comprobar la validez de la consulta.

En la Figura 8 se muestra una ejecución de proceso de consulta realizada por proceso de negocio, en esta los resultados son desplegados (recuadro rojo) una vez el usuario ha ejecutado la consulta, los resultados contienen los modelos de BP más relevantes según el nivel de similitud entre la consulta y los BP contenidos en el repositorio.

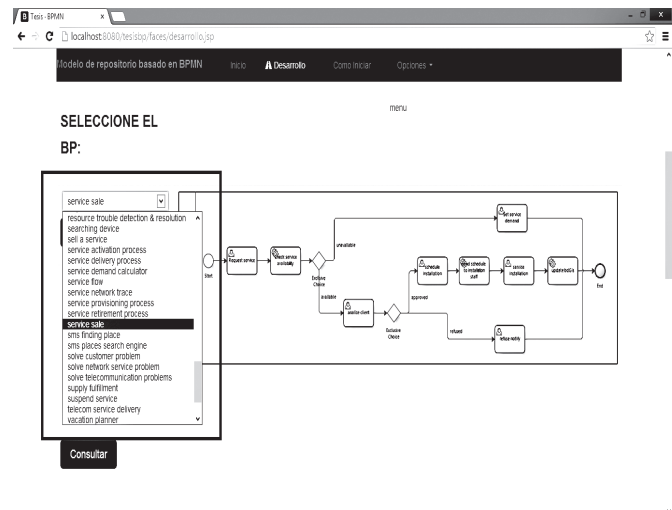


Figura 6. Opción de consulta por proceso de negocio.

Continuando con la fase de verificación o de pruebas, se concluyó la de la siguiente manera la investigación:

Para determinar la calidad del entorno fue necesario someterlo a un proceso de evaluación experimental, con el objetivo de verificar la eficiencia en el proceso de descubrimiento de BP con base al modelo de similitud definido para las opciones de consulta que éste permite. La evaluación se aplicó a un grupo de diez estudiantes a los cuales se les

hizo una capacitación previa para la realización de la prueba. Se muestra los resultados obtenidos en la siguiente figura.

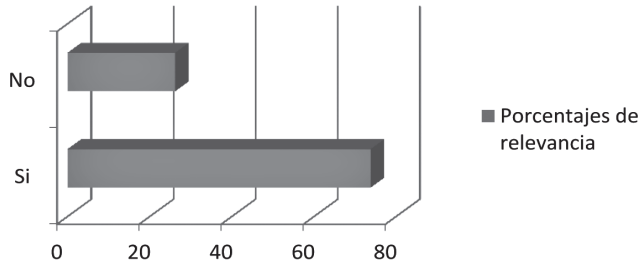


Figura 7. Resultados experimentales generales.

La Figura 7 muestra los valores obtenidos en la evaluación que realizaron los usuarios bajo su criterio subjetivo, en cuanto a que si los resultados arrojados por la herramienta fueron relevantes o no. Los resultados que se obtuvo fueron de un 74% para SI, afirmando que son resultados óptimos y relevantes según su criterio, y un 26% para NO, donde los resultados no son relevantes según el criterio del usuario.

La finalidad de la evaluación consiste en generar un ordenamiento (Ranking) de los 10 primeros modelos BP (dispuestos por orden de similitud), retornados para satisfacer una petición definida por medio de una de las opciones de consulta. En este sentido, es posible evaluar la calidad de los resultados obtenidos en la ejecución de esta operación del sistema, a partir de la aplicación de medidas estadísticas ampliamente empleadas en la evaluación de sistemas de recuperación de información (Baeza-Yates & Ribeiro-Neto, 1999). Éstas son las medidas de efectividad: Precisión P, Recall R y Medida F (Manning, Raghavan & Schütze, 2008). Para esto se tiene:

- D, conjunto de documentos.
- R, conjunto de documentos relevantes.
- $R^- = D - R$, conjunto de documentos no relevantes.
- A, conjunto de documentos recuperados.
- $A \cap R$, conjunto de documentos relevantes recuperados.

Entonces:

La **precisión** mide la porción de documentos recuperados que son relevantes.

$$precision = \frac{A \cap R}{A}$$

Recall mide la porción de documentos relevantes que son recuperados.

$$recall = \frac{A \cap R}{R}$$

Media F mide la armonía entre la precisión y el recall.

$$F = \frac{2 \times recall \times precision}{recall + precision}$$

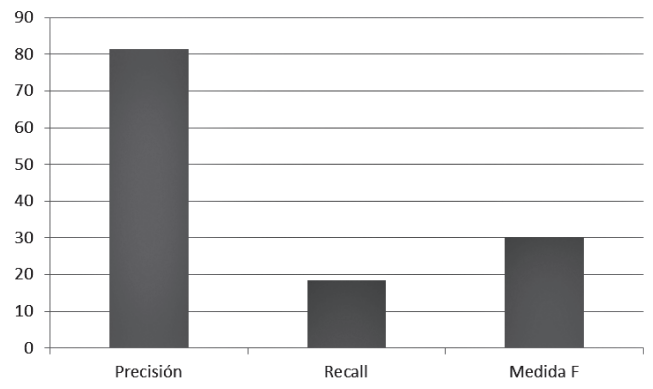


Figura 8. Resultados de evaluación.

La Figura 8 muestra los valores obtenidos en la evaluación, donde se puede visualizar que el modelo alcanzó 81% de precisión, demostrando válidas las tareas de consulta. Por otra parte, el 19% alcanzado en el Recall demuestra que el modelo deja de recuperar un número elevado de falsos positivos (BP que se recupera erróneamente); además, la Media F muestra que se tiene un 30% de armonía en los resultados retornados por el modelo, debido a que alcanzó un tiempo mínimo y resultados óptimos en la muestra de resultados en las búsquedas realizadas.

3. Conclusiones

Se definió un modelo de búsqueda el cual se aplicó a la herramienta prototipo que permite la consulta de procesos de negocio y *codebook* almacenados, que son solicitados por el usuario. El modelo propuesto demuestra que se puede obtener procesos de negocio que sirven como base para el replanteamiento de uno nuevo, y así reduce el tiempo del modelado.

El planteamiento del modelo de repositorio, representado sobre un modelo de recuperación de información permitió disminuir significativamente el tiempo de ejecución de la búsqueda, debido a que las consultas realizadas son representadas en forma de matrices de términos con los *codebook* y en forma de listas en la consulta de BP.

Los resultados de la evaluación permiten evidenciar el alto grado de validez del modelo alcanzado con el 80% de nivel de precisión en las consultas realizadas. Además, se comprueba que entre mayor sea el número de BP recuperados, aumenta el número de falsos positivos entre el conjunto de BP y *codebook*.

El desarrollo de la herramienta que implementó el modelo de descubrimiento de procesos de negocio planteado permitió realizar la evaluación de manera coherente tanto con los usuarios encuestados como la comparación del modelo de evaluación, realizado por jueces humanos expertos en la temática de recuperación de procesos de negocio planteado en Figueroa y Sandino (2011).

Referencias Bibliográficas

- Aalst, W.M.P. (2010). Configurable Services in the Cloud: Supporting Variability While Enabling Cross-Organizational Process Mining. *On the Move to Meaningful Internet Systems: OTM*, 6426, 8-25.
- Baeza-Yates, R. & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. New York: Addison-Wesley Longman Limited.
- Bernstein, P. & Dayal, U. (1994). An overview of repository technology. *Proceedings of VLDB*, 707-713.
- Chaves, M. (2010). Arquitectura orientada a servicios. *Revista: Ciencias Económicas*, 28 (1), 297-312.
- Decker, G. & Weske, M. (2008). Oryx An Open Modeling Platform for the BPM Community. *Proceedings of BPM*, 5240, 382-385.
- Elhadad, M. (2008). Effective business process outsourcing: The prosero approach. *International Journal of Interoperability in Business Information Systems*, 3 (1).
- Figueroa, L. & Sandino, C. (2011). Plataforma para evaluar sistemas de recuperación de procesos de negocio. *Revista de Investigaciones UCM*, 64-76.
- Figueroa, C., Corrales, J. & Ramirez-Gonzalez, G. (2012). *Recuperación Multinivel de Procesos de Negocio Basada en Semántica del Comportamiento*. Publisher Researchandinnovationbook.
- García, M. et al. (2001). Serie de Normas NTP ISO 9000:2001. *Rev. Industrial Data - Instituto de Investigación FII - UNMSM*, (8).
- IBM. (s.f.) BPM: Gestión de Procesos de Negocio. Recuperado el 12 de mayo de 2012, en <http://www-01.ibm.com/software/co/websphere/info/bpm/business/>
- LANZA. (s.f.) LANSА Composer. Recuperado el 12 de mayo de 2012, en <http://www.lansa.com/es/products/composeroverview.htm>
- Manning, H., Raghavan, P. & Schütze, H. (2008). *An Introduction to Retrieval Information*. Cambridge, England: Cambridge University Press.
- Mazanek, S. & Hanus, M. (2011). Constructing a Bidirectional Transformation between BPMN and BPEL with a Functional Logic Programming Language. *Journal of Visual Language & Computing*, 22 (1), 66-89.
- Mazanek, S. & Minas M. (2009). Constructing a Bidirectional Transformation between BPMN and BPEL with Functional-logic Graph Parser Combinators. Recuperado el 25 de febrero de 2011 de http://is.tm.tue.nl/staff/pvgorp/events/grabats2009/submissions/grabats2009_submission_9.pdf
- Narayanan, S., Jayaraman, V., Luo, Y. & Swaminathan, J. (2011). The antecedents of process integration in business process outsourcing and its effect on firm performance. *Journal of Operations Management*, 29 (1-2), 3-16.
- Smirnov, S., Weidlich, M., Mendling, J. & Weske, M. (2012). Action patterns in business process model repositories. *Computers in Industry*, 63 (2), 98-111.