

**COMPARACIÓN SEMÁNTICA DE TAREAS ENTRE DOS PROCESOS DE NEGOCIO
DE TELECOMUNICACIONES**



Universidad
del Cauca

**ADRIANA XIMENA BASTIDAS NARVÁEZ
LEANDRO ORDÓÑEZ ANTE**

Monografía presentada para optar al título de Ingeniero en
Electrónica y Telecomunicaciones

Director: PhD. Ing. Juan Carlos Corrales Muñoz

**UNIVERSIDAD DEL CAUCA
FACULTAD DE INGENIERÍA ELECTRÓNICA Y TELECOMUNICACIONES
DEPARTAMENTO DE TELEMÁTICA
POPAYÁN
2010**

CONTENIDO

	pág.
1. INTRODUCCIÓN	1
1.1. CONTEXTO	1
1.2. DECLARACIÓN DEL PROBLEMA	2
1.3. ESCENARIO DE MOTIVACIÓN	3
1.4. ALCANCE	4
1.5. CONTRIBUCIONES	4
1.6. CONTENIDO DE LA MONOGRAFÍA	6
2. ESTADO DEL ARTE	9
2.1. INTRODUCCIÓN	9
2.2. CONCEPTOS GENERALES	9
2.2.1. Ontologías	9
2.2.2. Procesos de Negocio	11
2.3. TRABAJOS RELACIONADOS	13
2.3.1. Descubrimiento de Servicios Web Semánticos	14
2.4. RESUMEN	22
3. DEFINICIÓN DE LENGUAJES SEMANTICOS Y ONTOLOGÍAS DE DOMINIO	23
3.1. INTRODUCCIÓN	23
3.2. LENGUAJES DE DESCRIPCIÓN DE ONTOLOGÍAS	23
3.2.1. RDF (Resource Description Framework)	24
3.2.2. RDFS (RDF Schema)	24
3.2.3. OWL (Web Ontology Language)	25
3.2.4. WSML (<i>Web Service Modeling Language</i>)	25
3.2.5. Selección del lenguaje semántico de descripción de ontologías.	26
3.3. LENGUAJE DE NOTACIÓN SEMÁNTICA PARA PROCESOS DE NEGOCIO	27
3.3.1. BPMO (Business Process Modelling Ontology)	27
3.4. ONTOLOGÍAS DE DOMINIO DE TELECOMUNICACIONES	28
3.4.1. eTOM - enhanced Telecom Operations Map.	29
3.4.2. SID - Shared Information/Data Model	31
3.5. ADAPTACIÓN DE ONTOLOGÍAS DE DOMINIO	33
3.5.1. Mecanismo de enriquecimiento semántico de Tareas	34
3.6. RESUMEN	42

4. ALGORITMOS DE COMPARACIÓN SEMÁNTICA DE TAREAS	43
4.1. INTRODUCCIÓN	43
4.2. DISTANCIA Y EMPAREJAMIENTO SEMÁNTICOS	43
4.2.1. Distancia y Emparejamiento entre Entradas/Salidas de las Tareas	44
4.2.2. Distancia y Emparejamiento entre Identificadores de las Tareas	49
4.3. SIMILITUD SEMÁNTICA	52
4.3.1. Similitud Semántica entre Conceptos	52
4.3.2. Similitud Semántica entre Tareas	55
4.4. RESUMEN	60
5. PROTOTIPO Y EVALUACIÓN	61
5.1. INTRODUCCIÓN	61
5.2. DESCRIPCIÓN DEL PROTOTIPO	61
5.2.1. Modelo de Casos de Uso del Sistema	62
5.2.2. Descripción de la Arquitectura de Referencia	64
5.3. Pruebas del Prototipo	74
5.3.1. Medidas de Desempeño	75
5.3.2. Plan de Pruebas y Resultados Obtenidos	77
5.4. RESUMEN	85
6. CONCLUSIONES Y TRABAJO FUTURO	87
6.1. CONTRIBUCIONES	87
6.2. CONCLUSIONES	89
6.3. TRABAJOS FUTUROS	92

INDICE DE FIGURAS

	pág.
Figura 1. Representación de Tareas y Subprocesos en BPMN	13
Figura 2. Categorías de Correspondencia.	15
Figura 3. Mapa de Operaciones de Telecomunicaciones Mejorado – eTOM.	30
Figura 4. Modelo de Información/Datos Compartidos – SID.....	32
Figura 5. Proceso “Solve Network Service Problem” en WSMO Studio.	35
Figura 6. Adición de Entradas a las Tareas BPMO	36
Figura 7. Selección de Data Type - Entrada ServiceName.....	37
Figura 8. Enriquecimiento Entrada NetworkServiceLocalization	37
Figura 9. Entrada “NetworkServiceLocalization” enriquecida	38
Figura 10. Ontología del Proceso de Negocio “Solve Network Service Problem” en WSMT	39
Figura 11. Definición de la variante de WSML	39
Figura 12. Adición de ontologías de dominio al namespace.....	40
Figura 13. Adición manual de anotaciones semánticas a la tarea Assess Network Service	41
Figura 14. Resultado de la consulta sobre la ontología del proceso	41
Figura 15. Ontología de dominio	45
Figura 16. Distancia Semántica entre Identificadores.....	50
Figura 17. Diagrama de Casos de Uso del Sistema	64
Figura 18. Arquitectura de Referencia del Sistema	65
Figura 19. Arquitectura para la Comparación Semántica de Tareas.....	67
Figura 20. Configuración de Parámetros de Comparación	71
Figura 21. Interfaz de Selección de las Tareas de Consulta a Comparar	72
Figura 22. Interfaz de Resultados de Comparación y Ranking de Tareas.	73
Figura 23. Interfaz de Resultados Detallados de Comparación.	73
Figura 24. Gráfica de rendimiento del sistema en la Transformación de Procesos de Negocio BPMO a Grafos y discriminación de sus tareas.	79
Figura 25. Gráfica del rendimiento del sistema en la operación de Comparación de Tareas.	80
Figura 26. Medidas de Desempeño: Precision, Recall y Overall para el Sistema.....	81
Figura 27. Medidas de Desempeño: Precision, Recall y Overall para un sistema de recuperación de servicios basado en el emparejamiento sintáctico interfaces.....	82
Figura 28. Gráfica de Precision vs. Recall	82
Figura 29. Gráfica de <i>top-k precision</i> vs. k.....	83
Figura 30. Gráfica de <i>top-p precision</i> vs. k	84

INDICE DE TABLAS

	pág.
Tabla 1. Valores de p según la ontología de dominio.....	54
Tabla 2. Plan de Pruebas	77
Tabla 3. Especificaciones Técnicas del Equipo empleado para Pruebas del Prototipo.....	78
Tabla 4. Ejemplo Comparación de Tareas	85
Tabla 5. Configuración de parámetros del prototipo.	85

ANEXOS

	pág.
ANEXO A. Modelo de Despliegue y Modelo de Diseño del Sistema.....	103
ANEXO B. Descripción de la Herramienta para la Comparación Manual de Tareas de Procesos de Negocio.....	119
ANEXO C. Banco de Procesos Publicados.....	123

1. INTRODUCCIÓN

1.1. CONTEXTO

Gracias a la generalización de las redes de conmutación de paquetes, la aplicación masiva del protocolo de internet (IP) en las comunicaciones, el auge de la arquitectura IMS, así como la tendencia creciente hacia el desarrollo de software soportado sobre arquitecturas en capas, las telecomunicaciones están atravesando por una etapa de transición dirigida hacia la convergencia. Desde el enfoque de los servicios, esta convergencia constituye un progreso significativo respecto al modelo de los sistemas de telecomunicaciones del pasado, el cual presentaba serios inconvenientes relacionados con la marcada dependencia que presentaba el despliegue de servicios respecto a la tecnología subyacente, de ahí la inherente falta de flexibilidad de estos sistemas, en los que la introducción de nuevas características requería cambios notables en la red del operador. En la actualidad, el modelo de las telecomunicaciones está orientado hacia la oferta de servicios que satisfagan eficazmente las necesidades de los usuarios, un enfoque que sustenta la evolución hacia la convergencia, a partir de la cual se pretende facilitar la introducción de nuevos e innovadores servicios de valor agregado, fomentando la reusabilidad de componentes y disminuyendo el tiempo necesario para el despliegue de los servicios en el mercado (*time to market*), lo cual además propiciaría la interacción e integración *Business to Business* y la generación de cadenas de valor entre los operadores de telecomunicaciones y empresas de otros sectores (Kwiatkowski, et al., 2007).

El fenómeno de convergencia en los servicios de telecomunicaciones actualmente se sustenta en la implementación de arquitecturas como SOA (*Service Oriented Architecture*), las cuales promueven y habilitan características de reutilización, interoperabilidad, escalabilidad y flexibilidad en el despliegue de servicios. Gracias a esto es posible articular servicios complejos de telecomunicaciones a partir de la composición de servicios web¹, definiendo de esta manera, procesos de negocio² en este sector.

Por otra parte, además del progreso que ha supuesto la implementación del modelado orientado a servicios sobre el nivel de "Tecnología de la Información" (*Information Technology*, IT) del ciclo de desarrollo de los procesos empresariales, en los últimos años se ha generado un interés creciente hacia el desarrollo de la metodología de Gestión de Procesos de Negocio (*Business Process Management*, BPM) y su aplicación dentro de las organizaciones, tendencia que no ha sido ajena a los proveedores de servicios de telecomunicaciones. El desarrollo de esta metodología impulsó en 2005 la unión del *Object Management Group* (OMG) y la *Business Process Management Initiative*

¹ Un servicio web es una aplicación software modular e independiente, que puede ser descrita, publicada, localizada e invocada a través de una red (W3C, 2010).

² Un *proceso de negocio* es un conjunto de procedimientos o actividades conectados, los cuales cumplen colectivamente un objetivo de negocio, definiendo roles y relaciones funcionales (WfMC, 1999). En el contexto del presente proyecto es entendido también como un servicio, en la medida en que, si bien está compuesto por actividades (algunas de las cuales invocan servicios), su función final es la de proveer información o una funcionalidad relevante para el cliente que lo ejecuta.

(BPMI)(BPMI, 2008), organizaciones que en búsqueda de eliminar la brecha existente entre los niveles de diseño e implementación de los procesos de negocio, han promovido el desarrollo de la *Business Process Management Notation* (BPMN), la cual surge como una notación estándar para modelar los procesos de negocio de las compañías, buscando facilitar un entendimiento común por parte de todos los involucrados en el desarrollo y operación de los procesos, desde los analistas del negocio hasta el personal técnico encargado de llevar a cabo su implementación (OMG, 2008).

Así mismo, terminológicamente alineada con BPMN, surge la *Business Process Modelling Ontology* (BPMO), cuyo propósito es proveer una plataforma estable para la definición de procesos de negocio enriquecida semánticamente, mediante la creación de una ontología capaz de representar artefactos de las múltiples metodologías empleadas para el modelado de procesos de negocio, promoviendo de esta manera, una interpretación única y una representación unificada de los mismos (Domingue, 2008). Actualmente, aprovechando iniciativas de estandarización como la que promueve BPMO en la descripción de procesos empresariales, es posible pensar en mecanismos que permitan optimizar el despliegue de dichos procesos a partir de la identificación y recuperación de componentes reutilizables (tales como actividades atómicas) presentes en la estructura de procesos de negocio existentes dentro de las organizaciones. En este contexto, el desarrollo de herramientas que viabilicen este procedimiento de identificación y recuperación es de suma importancia.

Tanto en BPMN como en BPMO, una actividad atómica (también conocida como tarea) es la representación más concreta y específica del trabajo llevado a cabo dentro del proceso de negocio, la cual hace referencia únicamente al servicio que provee y no al flujo de control del proceso. Las tareas están definidas a partir de sus atributos, entre ellos: identificador o nombre, conjunto de entradas y salidas, descripción en lenguaje natural, entre otras.

Teniendo en cuenta que la implementación de las actividades de los procesos de negocio mediante servicios web es una práctica común, el presente trabajo parte del estudio de las técnicas utilizadas para llevar a cabo los procedimientos de búsqueda y descubrimiento de servicios web, para determinar en qué medida pueden ser adaptadas, y posteriormente aplicadas sobre la comparación de tareas de un proceso de negocio descrito en BPMO.

1.2. DECLARACIÓN DEL PROBLEMA

En los últimos años el desarrollo del sector de las telecomunicaciones ha reflejado una tendencia hacia la globalización, propiciando la alianza y fusión de empresas e instituciones, con el fin de satisfacer las cambiantes necesidades de los usuarios, mediante la generación de nuevos servicios, la adición de valor agregado y la transición hacia el despliegue y prestación de servicios convergentes.

En estos procesos de alianza y fusión se han presentado algunas dificultades en cuanto a la interoperabilidad entre los procesos de negocio, dada la heterogeneidad de los sistemas de información de las compañías. Lo anterior ha propiciado el desarrollo e

implementación de SOA sobre los procesos de negocio de las empresas, a partir de la utilización de servicios web, que proporcionan datos o información relevante tanto para el usuario como para otros servicios, mediante interfaces bien definidas que facilitan la interoperabilidad.

Con el fin de sobresalir en el entorno competitivo propio de este sector, se debe procurar que el tiempo necesario para el despliegue de procesos empresariales, que soporten la demanda del mercado, sea mínimo. Sin embargo, hoy en día las actividades relacionadas con la implementación, desarrollo, mantenimiento e integración de nuevas soluciones a partir de componentes existentes, presenta limitaciones en cuanto a la agilidad en el proceso de descubrimiento, puesto que es necesaria la intervención de personal técnico, que lleve a cabo dicha tarea, a partir de su interpretación subjetiva del proceso de negocio (Pérez, et al., 2007).

Por otro lado, los mecanismos existentes que permiten el descubrimiento de servicios (refiriéndose en este caso no a servicios web, sino a procesos de negocio que describen servicios complejos) se centran en la comparación de palabras clave (las cuales hacen referencia a las etiquetas que nombran a las actividades del proceso de negocio), teniendo en cuenta criterios sintácticos. El proceso de descubrimiento soportado en este tipo de comparación, no es eficiente en tanto que no tiene la capacidad de reducir la ambigüedad de sus resultados (recuperando incluso, servicios que no corresponden con el requerimiento y en ocasiones dejando de lado servicios relevantes) debido a que, al realizar la comparación únicamente con criterios sintácticos, no se tiene en cuenta el significado de aquello que realmente se está solicitando realice el servicio.

El problema central abordado en este proyecto, está relacionado con la adaptación de mecanismos empleados en el descubrimiento y recuperación de servicios web semánticos, que pueden ser ajustados y aplicados para llevar a cabo la comparación semántica de tareas entre procesos de negocio del dominio de las telecomunicaciones. Durante el desarrollo de la presente investigación se ha dado respuesta al siguiente interrogante: ¿Cómo enriquecer semánticamente los procesos de negocio de un operador de telecomunicaciones, para facilitar el descubrimiento o búsqueda de servicios de telecomunicaciones?

1.3. ESCENARIO DE MOTIVACIÓN

Considere un operador de telecomunicaciones, que modele sus operaciones mediante procesos de negocio compuestos por un conjunto definido de tareas. Suponga que el operador desea evaluar los procesos de negocio existentes en su repositorio, para determinar si es posible mejorar su desempeño en cuanto a costos, calidad, servicio y tiempo. Se considera entonces, que es necesario identificar las tareas que pueden adaptarse y reutilizarse para satisfacer nuevas necesidades, o aquellas que deben ser reemplazadas o suprimidas de los procesos de negocio, porque constituyen una funcionalidad redundante dentro de los mismos. Un mecanismo automático de comparación semántica de tareas, permitiría al operador realizar estas actividades con mayor rapidez y con precisión aceptable, lo cual facilitaría la optimización del consumo de recursos en el desarrollo de actividades de análisis, gestión y reingeniería de procesos.

1.4. ALCANCE

En el presente proyecto, se analizaron algunos mecanismos existentes, para la determinación del grado de correspondencia y similitud semántica empleados comúnmente en procesos de descubrimiento y recuperación de servicios web, los cuales fueron adaptados y aplicados al ámbito de la comparación de tareas de procesos de negocio de telecomunicaciones, constituyendo la base para la determinación de la similitud semántica general entre las mismas, a partir de la comparación de sus entradas, salidas e identificadores. Este proyecto solo considera la comparación semántica entre tareas, enriquecidas con las mismas ontologías de dominio, es decir, el caso en el cual el grado de correspondencia y la similitud semántica pueden calcularse entre conceptos de la misma ontología. El proceso de evaluación y pruebas, al cual fue sometido el prototipo, permitió estimar el nivel de desempeño (en cuanto a eficiencia y calidad) de los resultados de similitud y de recuperación de tareas arrojados por el mecanismo de comparación semántico formulado.

Un requerimiento significativo, para garantizar la precisión en los resultados de la aplicación del mecanismo propuesto, consiste en asegurar la adecuada caracterización semántica de las tareas de consulta y publicada, la cual debe estar soportada en la utilización de ontologías del dominio de telecomunicaciones.

1.5. CONTRIBUCIONES

- **Selección y adaptación de dos ontologías de dominio.** Considerando el dominio de conocimiento, sobre el cual se centra el proyecto que se documenta en esta monografía, se llevó a cabo una exploración de las ontologías existentes relacionadas con el sector de las telecomunicaciones. A partir de esta exploración, se determinó emplear dos modelos formales, que reúnen el conocimiento relativo a los procesos y operaciones, de cualquier empresa proveedora de servicios de este sector: *SeTOM*³ y *SSID*⁴, ontologías basadas en los modelos *eTOM* (*enhanced Telecom Operations Map*) y *SID* (*Shared Information/Data Model*), de la iniciativa *NGOSS* (*New Generation Operations Systems and Software*)⁵ del *TMF* (*TeleManagement Forum*) (TMF - *eTOM*, 2005)(TMF - *SID*, 2005). Estas ontologías se adaptaron al mecanismo de enriquecimiento semántico mencionado, mediante la definición de la relación existente entre el conocimiento dispuesto en estos modelos y la naturaleza de la información contenida en cada atributo de las tareas, considerado en el procedimiento de comparación propuesto.
- **Definición y aplicación de un mecanismo de enriquecimiento semántico de tareas.** A partir de la identificación de un lenguaje formal, para la descripción

³ Semantic eTOM: Ontología basada en el modelo eTOM.

⁴ Semantic SID: Ontología basada en el modelo SID.

⁵ NGOSS es un conjunto de modelos y recomendaciones definidos por el TMF, convertidos en estándares de facto para el sector de las telecomunicaciones, en www.tmforum.org/SolutionFrameworks/1911/home.html

semántica de procesos de negocio, se estableció un mecanismo que permitiera asociar conceptos de una ontología de dominio, a cada uno de los atributos propios de las tareas que los constituyen (los atributos que determinan la funcionalidad de la tarea: identificador, entradas y salidas). La aplicación de este mecanismo de enriquecimiento, es un requisito que debe satisfacerse, para habilitar el procedimiento de comparación semántica de tareas propuesta en este proyecto.

- **Definición de un mecanismo de comparación semántica de tareas.** El estudio y apropiación de los métodos de emparejamiento semántico de servicios web existentes, soportaron la formulación de un conjunto de algoritmos, que constituyen un mecanismo para la comparación de tareas, basado en la similitud semántica de los parámetros que definen su funcionalidad (identificador, entradas y salidas).
- **Prototipo de comparación semántica de tareas.** Durante el trabajo de grado se implementó un prototipo, para la comparación semántica de tareas, entre procesos de negocio del dominio de los servicios de telecomunicaciones, el cual está fundamentado en el mecanismo propuesto.
- **Herramienta de comparación manual de tareas entre procesos de negocio.** Para abordar el proceso de evaluación del prototipo implementado, fue necesario el desarrollo previo de una herramienta, que permitió recoger el criterio de comparación de personas, con amplio conocimiento en el dominio de los procesos de negocio de telecomunicaciones. La herramienta permitió de esta manera, la construcción de una base de comparaciones de referencia (o *benchmark de referencia*), a partir de la cual, fue posible llevar a cabo la evaluación del desempeño del sistema propuesto.
- Se destaca la producción de un artículo expuesto en el desarrollo del 5to Seminario Nacional “Tecnologías Emergentes en Telecomunicaciones y Telemática – TET2010”, el cual aborda la descripción del mecanismo de comparación propuesto (temática documentada en los Capítulos 3 y 4 de esta monografía), así como definición de la Arquitectura del Comparador Semántico de Tareas (a la cual se dedica parte del Capítulo 5 del presente documento).
- El trabajo desarrollado dentro del presente proyecto, soporta el componente de análisis semántico del sistema que se propone en la tesis de maestría: **Descubrimiento Automático de Procesos de Negocio Basado en Semántica del Comportamiento**, desarrollada dentro del Grupo de Ingeniería Telemática de la Universidad del Cauca.

1.6. CONTENIDO DE LA MONOGRAFÍA

Capítulo 2. Estado del Arte

En este capítulo, se abordan las definiciones formales de conceptos claves para el entendimiento del proyecto realizado, tales como ontologías, procesos de negocio, estimación de la distancia y similitud semántica entre conceptos, emparejamiento semántico de servicios web, entre otros; orientando la definición de estos términos, hacia el dominio de los servicios de telecomunicaciones y abordando los trabajos relacionados correspondientes.

Capítulo 3. Definición de Lenguajes Semánticos y Ontologías de Dominio

Este capítulo define los lenguajes semánticos existentes, para la especificación de procesos de negocio y ontologías, y expone cuáles de ellos se ajustan adecuadamente al enfoque de comparación que se formula en el presente proyecto. Así mismo, presenta una descripción de las ontologías de dominio de telecomunicaciones empleadas y resalta además, la importancia del mecanismo de enriquecimiento semántico de los procesos de negocio, como requisito fundamental para la comparación semántica de sus tareas. Por último, define la forma en que se adaptaron las ontologías de dominio al prototipo desarrollado.

Capítulo 4. Algoritmos de Comparación Semántica de Tareas

Corresponde a una descripción de los algoritmos de emparejamiento y similitud semántica de servicios web existentes, que pueden adaptarse a la comparación de tareas. Presenta una descripción y justificación de los algoritmos seleccionados, y define detalladamente, el mecanismo de similitud propuesto para la comparación semántica de las tareas, entre dos procesos de negocio del dominio de los servicios de telecomunicaciones.

Capítulo 5. Prototipo y Evaluación

Este capítulo está dedicado, a la descripción detallada del prototipo desarrollado, a partir del mecanismo de comparación semántica presentado en el Capítulo 4. Se especifica cada uno de los módulos que componen la arquitectura y la forma como se llevó a cabo su implementación en el prototipo.

Además, se aborda la documentación del proceso de evaluación de desempeño aplicado sobre el prototipo implementado, partiendo de la descripción de la base de comparaciones de referencia generada, para confrontar los resultados arrojados por el mecanismo y determinar la calidad de los mismos. Posteriormente, se identifican las medidas de desempeño empleadas y finalmente se expone el análisis de los resultados obtenidos en la aplicación de dichas medidas.

Capítulo 6. Conclusiones y Trabajo Futuro

Por último, se analizan los resultados del trabajo realizado, se detallan las principales contribuciones obtenidas en la ejecución del proyecto y se expone un conjunto de recomendaciones importantes para el desarrollo de trabajos futuros.

2. ESTADO DEL ARTE

2.1. INTRODUCCIÓN

En este capítulo se presentan las definiciones formales y bases teóricas, necesarias para comprender la temática esencial del proyecto realizado. A continuación, se exponen los principales conceptos, en los que se fundamenta la propuesta de comparación de tareas entre procesos de negocio del dominio de las telecomunicaciones, y posteriormente se presenta el estado actual de los trabajos e iniciativas relacionadas con esta temática. Finalmente, se realiza un resumen que expone los principales aportes de este capítulo.

2.2. CONCEPTOS GENERALES

2.2.1. Ontologías

Aunque el término ontología tiene su origen en la filosofía, ha sido adoptado en el contexto de la informática y la ingeniería de software. Actualmente no existe una definición estándar de lo que es una ontología, sin embargo una aproximación muy aceptada es la propuesta de (Gruber, 1993), según la cual “una ontología es la especificación formal y explícita de una conceptualización compartida”. En esta definición, *conceptualización* hace referencia a un modelo abstracto de algún fenómeno del mundo que representa sus conceptos más relevantes; *explícito* se refiere a la necesidad de especificar de forma consciente los distintos conceptos que conforman una ontología; *formal* indica que la especificación debe representarse por medio de un lenguaje de representación formalizado, y *compartida* refleja que una ontología debe, en el mejor de los casos, dar cuenta de un conocimiento aceptado. Otra definición destacada es la que propone la W3C⁶ (World Wide Web Consortium): “una ontología define los términos utilizados para describir y representar un área de conocimiento”(W3C, 2004). Así también, para el SUO WG⁷ (*Standard Upper Ontology Working Group*) “una ontología consiste en un conjunto de conceptos, axiomas y relaciones que describen un dominio⁸ de interés” (IEEE).

Tomando como referencia estas definiciones, se puede concluir que una ontología es una representación de un vocabulario compartido, que permite describir el conocimiento sobre el mundo o una parte del mismo, y que define los elementos relevantes (conceptos, relaciones) para el dominio en estudio, proveyendo significado al vocabulario y formalizando restricciones en su uso.

De esta manera, las ontologías proporcionan una definición formal de conceptos que asegura la interpretación correcta del conocimiento compartido y brindan una forma de

⁶Disponible en la dirección www.w3.org/Consortium/

⁷ Disponible en la dirección www.suo.ieee.org/

⁸Área específica de interés o área de conocimiento.

representar y compartir información haciendo uso de un vocabulario común (Samper, 2005). Por lo tanto, se usan para favorecer la comunicación entre personas, organizaciones y aplicaciones, ya que proveen una comprensión común de un dominio, reduciendo las ambigüedades y confusiones conceptuales y terminológicas.

Las ontologías son elementos computacionalmente comprensibles, cuyo propósito es servir de soporte a diversas aplicaciones, que requieren de conocimiento específico sobre una materia, y por tanto permiten la automatización de muchos procesos por ejemplo de integración, búsqueda y recuperación inteligente de información, gestión de conocimiento, interoperabilidad entre sistemas informáticos, validación de datos, entre otros (Villa, et al., 2009) (Delgado, et al., 2009). En este sentido, las ontologías juegan un papel fundamental en el contexto del presente proyecto, puesto que permiten realizar el enriquecimiento y la inferencia sobre las tareas que conforman los procesos de negocio de telecomunicaciones, proporcionando los elementos necesarios para dotar de semántica a dichas tareas, y reduciendo la ambigüedad en la definición de sus funcionalidades, para facilitar la correcta y eficiente interpretación en las comparaciones.

2.1.1.1. Componentes de las ontologías

Las ontologías tienen los siguientes componentes, que sirven para representar el conocimiento de algún dominio (Gruber, 1993):

- **Conceptos:** son las ideas básicas que se intentan formalizar, es decir los elementos básicos de la terminología de dominio que representan clases de objetos del mundo, tanto abstracto como real, y comparten propiedades específicas.
- **Relaciones:** representa la interacción y enlace entre los conceptos del dominio. Suelen formar la taxonomía del dominio.
- **Funciones:** son un tipo concreto de relación, donde se identifica un elemento mediante el cálculo de una función, que considera varios elementos de la ontología.
- **Instancias:** se utilizan para representar objetos determinados de un concepto.
- **Axiomas:** son reglas de restricción o teoremas, que se declaran sobre relaciones que deben cumplir los elementos de la ontología y permiten inferir conocimiento, que no está indicado explícitamente en ella.

2.1.1.2. Clasificación de Ontologías

En la comunidad científica aún no hay consenso acerca de la clasificación de las ontologías, por lo cual, cada autor define su propia clasificación de acuerdo a diferentes criterios. Por ejemplo, Van Heijst propone dos dimensiones según las cuales es posible clasificar las ontologías: la cantidad y tipo de la conceptualización (que hace referencia a la motivación de la ontología), y el objeto de la conceptualización (es decir el tipo de conocimiento contenido en ella) (Van Heijst, et al., 1997). De acuerdo al primer criterio, se distinguen tres tipos fundamentales de ontologías:

- **Terminológicas:** especifican los términos que son usados para representar conocimiento en el universo de discurso. Suelen ser usadas para unificar vocabulario en un dominio determinado.

- **De Información:** especifican la estructura de almacenamiento de bases de datos. Ofrecen un marco para el almacenamiento estandarizado de información.
- **De modelado del conocimiento:** especifican conceptualizaciones del conocimiento. Contienen una rica estructura interna, y suelen estar ajustadas al uso particular del conocimiento que describen.

Conforme a la segunda dimensión, existen cuatro categorías en las que se puede clasificar las ontologías:

- **De Representación:** son ontologías que estructuran los conceptos relacionados a un paradigma o representación del conocimiento. Definen, los elementos básicos que conforman una ontología más específica.
- **Genéricas:** son meta-ontologías, que definen, términos generales que pueden ser usados en diferentes dominios.
- **De Dominio:** describen conceptos relevantes de un dominio particular. Normalmente estas ontologías reutilizan ontologías genéricas, para especializar los conceptos y las relaciones.
- **De Aplicación:** contienen los conceptos específicos, necesarios en una aplicación. Son construidas de la reutilización de ontologías de dominio.

Considerando esta clasificación, nuestro mayor interés se centra en las ontologías de dominio, puesto que el proyecto está enmarcado en el área particular de las telecomunicaciones, siendo necesario, obtener una visión general de la terminología y las operaciones que llevan a cabo los operadores del sector en la prestación de sus servicios, mediante el conocimiento contenido en dichas ontologías. La descripción detallada de las ontologías utilizadas dentro del proyecto se encuentra en el capítulo 3.

2.2.2. Procesos de Negocio

Un proceso de negocio, puede definirse, como un conjunto de actividades que se realizan coordinadamente en un ambiente organizacional y técnico, las cuales en forma conjunta realizan un objetivo del negocio (Weske, 2007). Los procesos de negocio son parte fundamental de las empresas, ya que son el medio a través del cual se pueden proporcionar los productos y/o servicios a los clientes, generando cadenas de valor, que permiten cumplir las metas estratégicas de las compañías. En el contexto del presente proyecto, estos pueden ser entendidos también como servicios, en la medida en que, si bien están compuestos por tareas (algunas de las cuales invocan servicios), su función final es proveer información o ejecutar las operaciones necesarias para satisfacer las necesidades del cliente.

Los procesos de negocio, son la base para comprender, la forma en que opera un negocio o empresa en sus diferentes áreas, por lo cual su estructura suele ser amplia y compleja, requiriendo ser modelada para facilitar su entendimiento y ejecución. En otras palabras, se hace necesario modelar los procesos de negocio, para organizar y documentar la información relativa a las actividades que estos ejecutan, así como los departamentos y el personal encargado que involucran. De esta forma, se pretende mejorar la eficacia organizativa y la calidad al interior de las empresas, facilitando el desarrollo de metodologías para la gestión y optimización de los procesos.

En cuanto al modelado de procesos de negocio, uno de los estándares más reconocidos es BPMN, la cual fue inicialmente desarrollada por la organización BPMP, y es actualmente mantenida por el OMG.

2.2.2.1. BPMN

BPMN es una notación gráfica, para modelar procesos de negocio en formato de flujo de trabajo (*workflow*). Ha sido especialmente diseñada, para coordinar la secuencia de los procesos y los mensajes que fluyen entre los participantes de las diferentes actividades.

El principal objetivo de BPMN, es proveer una notación estándar, que sea fácilmente legible y entendible, por parte de todos los involucrados e interesados del negocio (OMG, 2008). Entre estos interesados están los analistas de negocio (quienes definen y redefinen los procesos), los desarrolladores técnicos (responsables de implementar los procesos) y los gerentes y administradores del negocio (quienes monitorean y gestionan los procesos). BPMN tiene la finalidad de servir como lenguaje común, para disminuir la brecha de comunicación, que frecuentemente se presenta entre el diseño de los procesos de negocio y su implementación.

De esta manera BPMN permite modelar los procesos a través de diagramas sencillos denominados BPD (Business Process Diagrams), los cuales permiten representar gráficamente la secuencia de todas las actividades que ocurren durante un proceso, e incluyen además toda la información que se considera necesaria para su análisis. Dentro de un diagrama BPD se utilizan un conjunto de elementos gráficos, que se encuentran agrupados en cuatro categorías (OMG, 2008):

- **Objetos de flujo:** Eventos, Actividades, Rombos de control de flujo (*Gateways*).
- **Objetos de conexión:** Líneas de Secuencia, Líneas de Mensaje, Asociaciones.
- **Diagramas de calles (Swimlanes):** Piscina (*pool*), Carril (*lane*).
- **Artefactos:** Objetos de Datos, Grupo, Anotaciones.

De este conjunto de elementos, en el presente proyecto se analizaron únicamente las *actividades*, las cuales representan el trabajo que se lleva a cabo en los procesos de negocio. En BPMN, pueden modelarse dos tipos de actividades: las *tareas* (actividades atómicas) y los *subprocesos* (actividades compuestas). Las *tareas* son usadas cuando el trabajo en el proceso no puede descomponerse o subdividirse, y se representan mediante un rectángulo con puntas ovaladas. Los *subprocesos* son agrupaciones de tareas, que ejecutan una operación un poco compleja y pueden comprimirse en un solo rectángulo para simplificar el modelo (Figura 1). Teniendo en cuenta que los subprocesos están compuestos a su vez por tareas, siendo estas la unidad elemental de trabajo de un proceso, se decidió profundizar el análisis únicamente en estas últimas.

En BPMN las tareas están descritas a partir de atributos como: identificador, tipo, especificación de entradas y salidas, recurso que las ejecuta, asociaciones de datos de entrada y de salida, en algunos casos la tecnología, las referencias a otras operaciones o servicios que se ejecuten dentro de ellas, entre otros; los cuales permiten describir o representar la información acerca de las operaciones que ejecutan y la función que cumplen dentro del proceso de negocio.

Figura 1. Representación de Tareas y Subprocesos en BPMN



Sin embargo, el estudio de todos estos atributos en la comparación de tareas entre procesos de negocio, puede resultar un procedimiento extenso y complejo, por lo cual para simplificar el análisis, se escogieron los atributos más relevantes para definir su funcionalidad: el identificador o nombre de la tarea que refleja la operación que esta desempeña, las entradas que son la “materia prima” necesaria para llevar a cabo la tarea, y las salidas que son el resultado final que producen. A partir de esta definición, se pudo establecer una relación de correspondencia entre los atributos que describen las tareas y los atributos que definen los servicios web, ya que ambos comparten este conjunto de parámetros (identificador, entradas y salidas).

De esta manera, teniendo en cuenta esta relación y considerando que en la actualidad no hay suficiente documentación acerca de la comparación semántica de tareas entre procesos de negocio, se realizó un estudio de los principales trabajos de investigación relacionados con el descubrimiento de servicios web semánticos para identificar los aportes que pudieran ser aplicados en el ámbito de las tareas, con el fin de proponer un mecanismo de similitud para facilitar la comparación semántica de las tareas que constituyen procesos de negocio de telecomunicaciones.

2.3. TRABAJOS RELACIONADOS

La inclusión de SOA, en la definición de servicios y procesos de negocio en entornos empresariales, ha facilitado procesos de alianza e integración entre las compañías, y en especial, gracias al auge de esta arquitectura, el desarrollo de servicios web, dentro de la implementación de los procesos de negocio, se ha constituido en una práctica común para promover la competitividad, al facilitar la publicación y el consumo de servicios, tanto para proveedores como para clientes. Actualmente, se desarrollan algunos trabajos de investigación, que pretenden automatizar las tareas de descubrimiento de servicios web, los cuales están dirigidos hacia la adición de semántica sobre sus documentos descriptores, buscando asegurar la fiabilidad y la precisión en el cumplimiento de una solicitud específica de servicio. Esta sección presenta el análisis de los principales trabajos de investigación relacionados con esta temática, exponiendo su aporte y aplicación en el desarrollo del presente proyecto.

2.3.1. Descubrimiento de Servicios Web Semánticos

El descubrimiento de servicios puede definirse como la capacidad de encontrar y utilizar posteriormente un servicio, basado en alguna descripción publicada de su funcionalidad y parámetros operacionales (Ayomi, et al., 2007).

Algunos de los mecanismos existentes, para el descubrimiento automático de servicios, se basan en interfaces, y se orientan hacia la utilización de criterios sintácticos para la comparación de palabras clave, que describen las actividades del servicio (Bianchini, et al., 2006) (Budanitsky, et al., 2006). El proceso de descubrimiento, soportado en este tipo de comparación, no tiene en cuenta la semántica de la solicitud del servicio, respecto a las funcionalidades esperadas, en consecuencia, dicha técnica tiene alta probabilidad de recuperar servicios ambiguos, que no corresponden directamente con el criterio de búsqueda.

Por esta razón, la mayoría de trabajos de investigación, actualmente están orientados hacia el uso de inferencia semántica sobre las descripciones de los servicios web, para lo cual, es necesario que dichas descripciones estén realizadas con lenguajes que incorporen componentes semánticos y sean computacionalmente comprensibles, es decir, que permitan que los sistemas de búsqueda puedan interpretar, las relaciones existentes entre los conceptos, que definen la funcionalidad de los servicios web, minimizando en lo posible la intervención humana y mejorando la eficacia en la recuperación de servicios que satisfagan adecuadamente los requerimientos del cliente.

2.3.1.1. Emparejamiento Semántico

Uno de los mecanismos más importantes, para automatizar el descubrimiento de servicios web semánticos, es el emparejamiento semántico, el cual se basa esencialmente en el uso de inferencia lógica, para identificar relaciones de equivalencia o subsunción entre los parámetros, que definen las descripciones de una solicitud de servicio y aquellos que describen un servicio suministrado por un proveedor (servicio publicado). En este sentido, se han identificado los tres enfoques más representativos, que permiten abordar el proceso de emparejamiento de servicios web. En consecuencia, y dado que las tareas poseen atributos similares a los de los servicios, estos enfoques pueden adecuarse, para su aplicación sobre la comparación semántica de tareas de procesos de negocio.

Enfoque de Emparejamiento basado en Relaciones Terminológicas: consiste en explotar relaciones terminológicas, para comparar las descripciones de los servicios, utilizando sistemas de léxico como Wordnet o Thesaurus⁹, para contrastar los conceptos de la solicitud con los del servicio publicado y de esta forma encontrar relaciones terminológicas por ejemplo de sinonimia, especialización/generalización y composición (Bianchini, et al., 2006).

Enfoque de Emparejamiento basado en Categorías: este enfoque propone la implementación de un mecanismo de categorización cualitativa, basado en la forma en

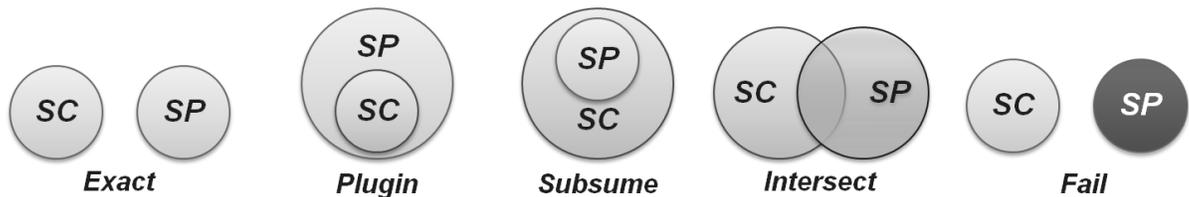
⁹Thesaurus es una obra que lista palabras agrupadas de acuerdo a la similitud de su significado (contiene sinonimos y algunos antonimos), en contraste con un diccionario, que contiene definiciones.

que un servicio publicado en un repositorio, satisface una solicitud de servicio (Li, et al., 2003) (Paolucci, et al., 2002). Este mecanismo evalúa, la relación existente entre los parámetros definidos en los documentos descriptores del servicio publicado y solicitado, tales como, los conceptos que enriquecen sus entradas y salidas (los cuales pertenecen a ontologías de dominio), con el fin de determinar la correspondencia entre ambos. La evaluación realizada, sobre la relación que existe entre los conceptos que hacen parte del enriquecimiento de los servicios, se lleva a cabo mediante la utilización de motores de razonamiento de subsunción, que permiten inferir sobre la jerarquía de conceptos de la ontología.

Una vez identificada la relación existente entre los conceptos del enriquecimiento, esta es clasificada dentro de cinco categorías denominadas grados o niveles de correspondencia (Figura 2):

- **Exact:** Si el servicio publicado, satisface completamente los requerimientos del servicio solicitado.
- **Plug-in:** Si el servicio publicado, abarca la funcionalidad del servicio solicitado, sin embargo, no satisface algunos de sus requerimientos específicos.
- **Subsume:** Si el servicio publicado, satisface parcialmente los requerimientos del servicio solicitado, por lo tanto, no puede reemplazarlo plenamente en su funcionalidad.
- **Intersection:** Cuando las funcionalidades de los servicios publicado y solicitado, no son del todo incompatibles.
- **Fail (Disjoint):** en cualquier otro caso, la correspondencia es establecida como fail.

Figura 2. Categorías de Correspondencia.



Enfoque de Emparejamiento basado en la Clasificación (Ranking) de los servicios: este enfoque sugiere, que la clasificación de la correspondencia basada en categorías, no es suficiente para determinar el grado de relación entre los servicios comparados. Por lo cual, propone cuantificar esta relación, a partir del cálculo de la similitud entre las entradas/salidas del servicio solicitado y las entradas/salidas de los servicios publicados, de manera que pueda realizarse una clasificación o ranking de servicios, de acuerdo a dicho valor de similitud (Xu, et al., 2006).

A continuación, se presentan los trabajos de investigación más relevantes en cuanto a emparejamiento de servicios:

- El trabajo presentado en (Paolucci, et al., 2002), es uno de los primeros proyectos en introducir el enfoque de categorías o niveles de correspondencia y ha servido como

punto de partida para muchas otras investigaciones. En este trabajo, se aborda una propuesta de descubrimiento de servicios web, a partir del análisis de sus capacidades, proponiendo un algoritmo para determinar la correspondencia entre los servicios solicitado y publicado (matchmaking), mediante la comparación de sus entradas y salidas, la cual es clasificada en 4 categorías: exact, plugin, subsume y fail. Este artículo define además, que si existe una correspondencia entre el servicio solicitado y un servicio publicado, esta debe almacenarse y se le debe asignar un valor de acuerdo al grado de similitud existente, de manera que las correspondencias se ordenan al final, de acuerdo con dicho valor, para facilitar la localización de los servicios que mejor corresponden con el servicio requerido. La contribución de este trabajo para el proyecto realizado, consiste en la definición de las categorías de correspondencia, sin embargo no tiene en cuenta información relevante respecto a la funcionalidad del servicio, que puede estar contenida en su identificador.

- El trabajo descrito en (Çelik, et al., 2006), presenta un Buscador Inteligente de Servicios Web Semánticos (IntelligentSemantic Web ServiceFinder ISWSF), que utiliza un sistema de Agente de Búsqueda Semántica (SemanticSearchAgent SSA), para recuperar servicios web, y usa el lenguaje OWL-S para añadir la descripción semántica a los servicios. Este sistema se desarrolla en dos partes principales: el mejoramiento de la semántica de los términos de entrada, a partir de la adición de sinónimos de las palabras claves de la petición del cliente, y un algoritmo de correspondencia (matchmaking) aplicado sobre las entradas y salidas (I/O) del servicio web semántico recuperado, y las I/O de la solicitud del cliente. Este algoritmo clasifica las relaciones de correspondencia en cuatro categorías, al igual que en (Paolucci, et al., 2002), y si el sistema concreta cualquier similitud entre los términos, da una puntuación de correspondencia de acuerdo a dichas categorías. Uno de los aspectos más representativos de este trabajo, consiste en enriquecer las palabras clave de la solicitud del cliente, adjuntando sus sinónimos y términos relacionados antes de enviarlos al ISWSF (soportado en el uso de ontologías de dominio), permitiendo que el proceso de búsqueda de servicios se realice con mayor precisión. Sin embargo, esta característica hace que el número de operaciones de comparación se incremente significativamente, lo cual puede ser crítico en cuanto al tiempo. Otra desventaja que presenta este sistema es que no permite recuperar aquellos servicios que son compatibles con la solicitud pero no presentan correspondencias tipo exact, plug-in o subsume.
- En (Li, et al., 2003) se presenta un caso de estudio para una aplicación de comercio electrónico (e-commerce), utilizando la ontología DAML-S, para lo cual proponen un mecanismo de emparejamiento, donde agentes implementados en JADE (Java Agent Development Framework) pueden buscar y publicar servicios que correspondan con las descripciones semánticas de la solicitud del cliente (generadas mediante DAML-S), las cuales se encuentran almacenadas en un repositorio. Asimismo, utiliza un razonador basado en lógica de descripción (Racer DL), para calcular correspondencias semánticas entre solicitudes de servicios y servicios publicados. La característica más atractiva de este trabajo, consiste en la definición de una nueva categoría de correspondencia conocida como **Intersection**, que tiene en cuenta aquellos servicios publicados, que no son del todo incompatibles con el servicio requerido, pero que no se clasifican como plugin o subsume. Esta nueva categoría proporciona mayor flexibilidad al mecanismo de emparejamiento, sin embargo, en este

trabajo al igual que en (Paolucci, et al., 2002), el proceso de emparejamiento solo tiene en cuenta el conjunto de entradas y salidas de los servicios solicitado y publicado, por lo que se omite la información que puede suministrar el identificador del servicio

- En (Choi, et al., 2005) se propone, una serie de algoritmos para encontrar la correspondencia entre los servicios publicado y solicitado, a través de la utilización de un motor de correspondencia, que discrimina la relación entre servicios en cinco niveles: exact, plug-in, subsume, intersection y fail. Según esta relación se define una puntuación, que permite realizar posteriormente un ranking de los resultados de búsqueda. El principal aporte de este trabajo, consiste en la proposición de una función extra, a parte del mecanismo de correspondencia, que permite clasificar servicios que están dentro del mismo nivel, utilizando la relación entre niveles consecutivos de la ontología, es decir, la cercanía vertical y horizontal entre los términos de la descripción del servicio web.
- En (Xu, et al., 2006), se propone la utilización de distancia semántica, para determinar el grado de correspondencia entre las peticiones de servicio y los servicios publicados en un repositorio de servicios web. En este trabajo, se alejan del enfoque de clasificar el grado de similitud de dos servicios en las categorías tradicionales, debido a que se considera que esta clasificación es muy general y no permite distinguir, por ejemplo, cuál de los servicios pertenecientes a una misma categoría, es el más óptimo de acuerdo a los requerimientos del solicitante. De esta manera, se desarrolla un matchmaker semántico basado en 4 algoritmos de distancia semántica (URD, WRD, DFD, SRD), cada uno de los cuales, califican cuantitativamente la correspondencia entre los conceptos de una ontología. El aporte de este trabajo para el presente proyecto, consiste en la aproximación al cálculo de la distancia semántica, a partir de la comparación de los conceptos en la ontología, teniendo en cuenta su ubicación en la jerarquía de la ontología y el camino más corto que los conecta.
- En (Bianchini, et al., 2006), se propone un enfoque combinado, para flexibilizar el proceso de emparejamiento de servicios, que presenta diferentes modelos de correspondencia, para llevar a cabo una comparación flexible de servicios solicitado y publicado. En primer lugar, se considera un enfoque basado en relaciones terminológicas, para medir el grado de correspondencia entre los servicios, y a continuación, se describe un enfoque basado en categorías, capaz de encontrar la información faltante entre la solicitud y cada una de las ofertas, mediante la aplicación de un operador de diferencia basado en lógica. Por último, se propone una posible combinación de estos dos enfoques, con el fin de definir la similitud de los servicios publicado y solicitado. El matchmaker, expuesto en este trabajo, se divide en cuatro componentes: modelo de correspondencia, métricas, clasificación y optimización. De estos componentes, el más relevante para el proyecto desarrollado, es el modelo de correspondencia, el cual explota tanto las relaciones semánticas (de acuerdo a la ontología de dominio) como la afinidad terminológica (de acuerdo con el diccionario thesaurus), para clasificar el tipo de correspondencia entre la solicitud y el servicio publicado en los cinco niveles, y establecer algunas reglas para calcular la similitud entre dichos servicios. El principal aporte de este trabajo, consistió en la idea de retornar la información faltante entre el servicio solicitado y publicado, además de

identificar el tipo de correspondencia, con el fin de ofrecer mejores resultados y mayor información en el proceso de descubrimiento.

- En (Okkyung, et al., 2008), se documenta un trabajo de investigación, que tiene como propósito, permitir la búsqueda eficiente y la composición de servicios web, mediante el diseño e implementación de un sistema, que emplea una técnica de búsqueda basada en reglas y un algoritmo de correspondencia, que permite ofrecer calidad de servicio. Sin embargo, la complejidad debida a la gran variedad de algoritmos de activación utilizados, puede incrementar mucho el tiempo de búsqueda, lo cual resulta inconveniente para su aplicación en el presente proyecto.

Análisis:

La mayoría de los trabajos aquí presentados, no abordan el proceso de emparejamiento desde un enfoque particular, sino que son el producto de una combinación de dos o más enfoques, dependiendo de la complejidad del desarrollo y la calidad de los resultados que se desea obtener en el proceso de descubrimiento.

Los enfoques de emparejamiento expuestos en este capítulo, no son los únicos existentes (Bellur, et al., 2008), sin embargo, para el propósito del presente proyecto, los enfoques aquí descritos, se consideran los más adecuados desde el punto de vista de la complejidad, implementación y adaptabilidad a las necesidades de la solución propuesta.

A partir de la exploración realizada a través de los tres enfoques de emparejamiento identificados, y considerando la orientación del proyecto desarrollado, el cual se basa, en la utilización de ontologías de dominio de telecomunicaciones; se optó por aprovechar, las ventajas propias del segundo y tercer enfoque (enfoque basado en categorías y enfoque basado en el ranking de los servicios), para determinar la correspondencia entre el enriquecimiento de las tareas de procesos de negocio de telecomunicaciones. Teniendo en cuenta que el enfoque basado en relaciones terminológicas, si bien se caracteriza por una alta flexibilidad, en algunos casos puede provocar que los resultados obtenidos, del proceso de emparejamiento, incluyan servicios web que no satisfagan eficientemente el requerimiento del cliente, lo cual se traduce en medidas de precisión y recall limitadas (Budanitsky, et al., 2006). Respecto al enfoque basado en categorías, se puede decir que, gracias a que se soporta en la inferencia sobre ontologías de dominio, la precisión del proceso de emparejamiento, es mejor en comparación con el enfoque anterior. Sin embargo, tal como se argumenta en (Choi, et al., 2005), la sola clasificación de la correspondencia en categorías, no permite establecer cuál de los servicios clasificados, dentro de una misma categoría, satisface mejor la solicitud. Por último, el enfoque basado en la clasificación de los servicios, permite determinar cuantitativamente la capacidad de un servicio publicado, para satisfacer un requerimiento de servicio determinado, a través de algoritmos de distancia semántica, aplicados sobre los conceptos que componen la descripción de los servicios, por lo tanto, puede aplicarse para solucionar el inconveniente del segundo enfoque.

De esta manera, se definió que es adecuado, en primera instancia, aplicar un algoritmo de emparejamiento basado en categorías, que permita caracterizar cualitativamente, la correspondencia existente entre dos tareas (tarea publicada y tarea solicitada). Posteriormente, es necesario cuantificar estas relaciones y emplear criterios de

clasificación (ranking), que permitan reflejar la capacidad de las tareas, para satisfacer un requerimiento de otra, mediante la definición de similitudes semánticas entre estas. Para esto, se hizo necesario estudiar los mecanismos existentes, para determinar la distancia y similitud semántica aplicadas sobre servicios web, los cuales se describen a continuación.

2.3.1.2. Distancia Semántica y Similitud Semántica

En el proceso de descubrimiento de servicios, es fundamental filtrar adecuadamente la información contenida en las descripciones de los recursos publicados, para poder retornar los más adecuados a la solicitud del usuario. Para lograrlo, se hace necesario emplear mecanismos eficientes, para medir la similitud¹⁰ entre los conceptos que hacen parte de dichas descripciones, los cuales permitan identificar asociaciones entre conceptos, que pueden no ser inmediatamente evidentes para el usuario, o encontrar correspondencias o equivalencias parciales que otorguen cierta flexibilidad en los resultados de las consultas realizadas (d'Amato, 2007).

Ahora bien, dada la utilización cada vez más frecuente de ontologías como soporte en determinadas aplicaciones, es necesario emplear una nueva variable para representar esta relación de semejanza, pero desde el punto de vista semántico. En el contexto del presente proyecto, esta nueva variable se conoce como Similitud Semántica, y concretamente se refiere la proximidad que existe entre dos conceptos, dentro de una ontología (d'Amato, 2007) (Inkpen, et al., 2005). Así mismo, otra forma de cuantificar la relación entre conceptos, es a través de la distancia semántica, la cual es una medida del grado de diferencia entre el significado (sentido) de dos conceptos en una ontología, y puede considerarse como el inverso de la similitud semántica (Al-Mubaid, et al., 2006).

Existen muchas formas de definir matemáticamente estas relaciones, las cuales pueden agruparse en dos enfoques: uno basado en el contenido de la información y otro basado en el conocimiento codificado en ontologías (Al-Mubaid, et al., 2006). El primero, expresa que la similitud está determinada por la frecuencia en que los términos comparados aparecen dentro de un documento descriptor. Este enfoque tiene una visión estadística de la información para establecer la proximidad de los dos términos, y la distancia y similitud están determinadas, por la probabilidad de encontrar los conceptos dentro del documento dado. Algunos trabajos relacionados que hacen parte de este enfoque se encuentran en (Lord, et al., 2003) (Lin, 1998) (Resnik, 1995) y (Resnik, 1993). Sin embargo, este enfoque no es considerado relevante para el presente proyecto, puesto que su aplicación actualmente es muy limitada, ya que no tiene en cuenta la semántica de las descripciones aportada por las ontologías, por consiguiente nuestro estudio se centra únicamente en el segundo enfoque. Este último consiste, en determinar la similitud semántica entre los dos términos dentro de una ontología, empleando estrategias ya sea para medir la longitud del trayecto entre dos conceptos o para otorgar pesos a las aristas del camino (una vez se haya identificado), los cuales, dependen del número de nodos secundarios en la jerarquía de la ontología (profundidad). Los trabajos relacionados más representativos dentro de este enfoque se presentan a continuación:

¹⁰ la similitud hace referencia a la relación de semejanza existente entre los dos términos comparados.

- En (Blázquez, et al., 2006), se presenta una medida de similitud semántica entre elementos de una taxonomía, que puede ser empleada en aplicaciones que involucran consultas semánticas sobre bases de datos, las cuales, dada su característica de heterogeneidad en el proceso de integración de datos, hacen uso de ontologías. En este trabajo, se aborda el problema de la similitud entre conceptos, en el contexto de la expansión de consultas, en primer lugar, en el caso en el que la comparación se hace entre un concepto y uno de sus subconceptos, para posteriormente aplicarlo sobre dos conceptos que no presentan este tipo de subsunción. El principal inconveniente de este trabajo, consiste en que la estimación de la similitud se reduce a un simple cálculo del número de saltos, entre los conceptos comparados, y no tiene en cuenta la contribución de la distancia semántica de los trayectos, que hay entre ellos en el resultado final.
- En (Ge, et al., 2008), se propone un método de emparejamiento de servicios semánticos, basado en la comparación de la similitud entre los conceptos que definen sus descripciones, soportándose en el cálculo de la distancia semántica existente entre los mismos. De esta manera, se presenta un algoritmo que permite determinar el grado de correspondencia entre dos conceptos pertenecientes a una ontología, teniendo en cuenta, tanto relaciones de herencia, como la distancia semántica que los separa. El algoritmo, calcula la similitud semántica partir de la ejecución de cuatro fases: en la primera, se asignan pesos a los caminos existentes entre cada nodo (concepto) de la ontología y el nodo raíz. En la asignación de pesos, se tienen en cuenta varios criterios deseables, en cuanto a la ubicación en profundidad de los conceptos, dentro de la jerarquía definida en la ontología. En la segunda fase, se crea una “tabla de enrutamiento de nodos” en la cual, se consigna para cada concepto en la ontología, todos los caminos existentes hacia el nodo raíz, junto con los pesos asignados entre los nodos (conceptos) que hacen parte de cada camino. La tercera fase, lleva a cabo el cálculo de la distancia semántica, para lo cual emplea la información recopilada en la tabla de enrutamiento, determinando la distancia entre los dos conceptos, a partir de la suma de los pesos del camino más corto que los conecta. Finalmente, en la cuarta fase, se realiza el cálculo de la similitud semántica entre conceptos, la cual consideran como una medida que normaliza los resultados obtenidos de la fase anterior y permite identificar mejor, la relación semántica existente entre los conceptos comparados. Los planteamientos que se exponen en este trabajo, constituyeron un aporte considerable para el proyecto realizado, puesto que se ajustan a las características de las ontologías utilizadas, haciendo de este un enfoque ampliamente adaptable a la comparación de tareas entre procesos de negocio.
- En (Hau, et al., 2005), se propone una métrica para la medición de similitud entre servicios semánticos anotados con OWL, con el fin de establecer la compatibilidad entre servicios y evaluar su posible composición. En el enfoque propuesto en este trabajo, la similitud entre dos servicios se determina, a partir de la comparación de sus descripciones semánticas, conformadas por un conjunto de tripletas RDF. De esta manera, la medida de similitud entre dos servicios, se fundamenta en el cálculo de la razón del número de sentencias que son comunes para ambos, así, entre más información común haya en las descripciones de los dos servicios, más similares son. En este caso, la similitud entre dos servicios no depende de una medida de distancia

semántica entre conceptos, sino de una suma ponderada de la similitud entre las descripciones del *ServiceProfile*¹¹, el *ServiceModel*¹² y el *ServiceGrounding*¹³. De esta manera, este trabajo presenta un mecanismo interesante para abordar la comparación de servicios semánticos, sin embargo, la orientación hacia el estándar de descripción de servicios OWL-S, constituyó una limitación para llevar a cabo su aplicación, dentro del proyecto realizado, puesto que esa propuesta depende completamente de dicho lenguaje.

- En (Lee, et al., 2008), se presenta un análisis comparativo entre tres estrategias, para medir la similitud semántica entre conceptos del dominio de la medicina, específicamente, para la comparación de 20 enfermedades pertenecientes a la ontología de dominio SNOMED-CT. Este trabajo, aborda los dos enfoques explicados anteriormente, para compararlos con los resultados obtenidos de la opinión de médicos expertos. El primer mecanismo analizado, corresponde al enfoque basado solo en la ontología, en el cual se explotan la característica de localización de los conceptos, dentro de un grupo de nodos, para encontrar la distancia del camino más corto entre dos conceptos. El segundo mecanismo, consiste en medir la distancia entre dos pacientes, basado en las enfermedades que cada uno presenta, la cual es análoga a la distancia entre dos enfermedades en la ontología. Por consiguiente, para encontrar la Distancia Descendiente entre dos conceptos de la ontología, se asigna un peso a cada enlace del trayecto, de acuerdo con el número de descendientes para cada nodo (en el trayecto) y posteriormente se suman estos pesos. La tercera estrategia analizada, propone una medida de distancia, a partir del enfoque basado en el contenido de la información, a través de la frecuencia del término en un documento, es decir, dependiendo del número de veces que se emplea un concepto dentro del mismo. Por último, este trabajo concluye que de todos los mecanismos estudiados, el que más correspondencia tiene con la evaluación de los expertos, es el que está basado solamente en la localización de los conceptos, en la ontología para encontrar la distancia semántica, el cual no tiene correlación con las otras estrategias analizadas.

Análisis:

En esta sección, se ha presentado un conjunto de mecanismos, para cuantificar la relación existente entre dos conceptos y determinar su distancia o similitud. Se ha encontrado que existen dos enfoques para realizar estas medidas, uno está basado en la información contenida en un documento descriptor (cuerpo de conocimiento), y otro basado en el conocimiento codificado en ontologías. Sin embargo, dadas las características del presente proyecto, en el cual, el enriquecimiento de los procesos es realizado con conceptos de una ontología del dominio de las telecomunicaciones, se consideró que el enfoque más adecuado para determinar la similitud semántica entre tareas, es el que está basado solo en la ontología. Esto se definió teniendo en cuenta el trabajo realizado en (Lee, et al., 2008), a partir del cual se pudo concluir que de los dos enfoques estudiados, el que más se aproxima a los resultados de la opinión de los

¹¹El *ServiceProfile* de OWL-S permite describir las capacidades de un servicio en términos de sus funciones, limitaciones, aplicabilidad, calidad y requerimientos de uso.

¹²El *ServiceModel* de OWL-S permite describir como un servicio puede ser utilizado.

¹³El *ServiceGrounding* de OWL-S permite describir detalles de cómo puede accederse a un servicio.

expertos de un dominio, en cuanto a la distancia semántica, y consecuentemente a la similitud entre dos conceptos, es aquel que está basado enteramente en la ontología, dada su simplicidad, escalabilidad y a que es independiente del tamaño, la calidad y la disponibilidad del documento de texto (lo cual es indispensable para el enfoque basado en el contenido de la información).

Dentro de este enfoque, se han analizado un conjunto de estrategias, para medir la similitud y distancias semánticas, y se ha encontrado que el mejor mecanismo para llevar a cabo esta tarea, es el que considera la contribución de cada segmento del trayecto que hay entre los dos conceptos comparados, el cual además de tener en cuenta las relaciones dadas por la jerarquía, en la que están dispuestos los conceptos en la ontología, considera también la influencia de la profundidad de la localización de los conceptos, en el cálculo de la similitud. La utilización de este enfoque se justifica en que es independiente del estándar de descripción de servicios que se esté utilizando, al mismo tiempo que no requiere tener a disposición información sobre la probabilidad de ocurrencia (o de utilización) de conceptos de una ontología de dominio determinada, por lo tanto, se constituye en el enfoque más adaptable, a las características de la solución construida en torno a la comparación de tareas, entre procesos de negocio de telecomunicaciones.

2.4. RESUMEN

Este capítulo, corresponde a una contextualización en torno a la temática de comparación semántica de tareas, entre procesos de negocio de telecomunicaciones, comenzando por definir, los conceptos fundamentales necesarios para comprender el campo de aplicación del presente proyecto, para posteriormente, hacer un recorrido sobre los principales trabajos de investigación relacionados con esta propuesta.

Dada la escasez de documentación respecto a mecanismos de comparación semántica de tareas, fue necesario recurrir a otras líneas de investigación afines, con el propósito de adaptar, tales propuestas al campo de aplicación de interés. Una alternativa adecuada se identificó, al tener en cuenta la correspondencia existente entre los principales atributos que describen la funcionalidad de las tareas y los parámetros definidos para los servicios web (identificador o nombre, entradas y salidas). De esta manera, resultó de gran interés, el estudio del estado actual de los trabajos de investigación relacionados con el descubrimiento de servicios web semánticos, haciendo especial énfasis, en los algoritmos de emparejamiento, distancia y similitud semánticos aplicados sobre los conceptos que enriquecen sus descripciones, los cuales se soportan en la utilización de ontologías.

En este capítulo se realizó una selección de los enfoques que enmarcaron la propuesta presentada, y se identificaron los algoritmos de emparejamiento y similitud semánticos que constituyeron el punto de partida del presente proyecto, los cuales posteriormente, fueron adaptados y aplicados, al ámbito de la comparación de tareas entre procesos de negocio del dominio de telecomunicaciones, tal como se describe en los siguientes capítulos.

3. DEFINICIÓN DE LENGUAJES SEMANTICOS Y ONTOLOGÍAS DE DOMINIO

3.1. INTRODUCCIÓN

En este capítulo, se destaca la importancia de la semántica para el enriquecimiento de los procesos de negocio de un operador de telecomunicaciones. En primer lugar, se definen los lenguajes semánticos para la descripción de procesos de negocio y ontologías existentes. Posteriormente, se presenta el estudio y selección de las ontologías de dominio de telecomunicaciones más apropiadas, para el enriquecimiento de las tareas que conforman los procesos de negocio. Por último, se realiza una descripción detallada de la adaptación de dichas ontologías dentro del presente proyecto, mediante la definición de un mecanismo de enriquecimiento semántico de tareas.

3.2. LENGUAJES DE DESCRIPCIÓN DE ONTOLOGÍAS

El crecimiento descontrolado en la generación y compartición del conocimiento en internet, ha propiciado la especificación formal de la información, a través de modelos computacionalmente comprensibles como las ontologías. En consecuencia, los lenguajes primitivos para el intercambio de información tales como KIF, Ontolingua, Frame Logic, entre otros (Gomez-Perez, et al., 2004), han sido reemplazados por lenguajes de descripción más avanzados como RDF (Resource Description Framework), DAML (DARPA Agent Markup Language), OIL (Ontology Inference Layer) y OWL (Web Ontology Language), los cuales han permitido explotar las características de la web de una forma más eficiente, y han facilitado los procesos de razonamiento, inferencia y evaluación automática sobre las descripciones de los recursos.

La gestión de servicios y procesos de negocio no ha sido ajena a esta tendencia, respecto al manejo de la información, en los últimos años, las organizaciones han promovido la utilización de ontologías, para el enriquecimiento semántico de los documentos descriptores de sus recursos (servicios y procesos de negocio), como una medida para reducir la ambigüedad y facilitar el entendimiento común de los objetivos del negocio, en todos los niveles de la empresa.

Es por esto, que en la actualidad existen numerosos lenguajes estándares de descripción de ontologías, empleados tanto para el manejo de la información disponible en la web, como para la gestión semántica de procesos de negocio. A continuación, se presenta una breve descripción de los lenguajes más conocidos, y ampliamente utilizados para la descripción de ontologías.

3.2.1. RDF (Resource Description Framework)

RDF es un lenguaje de propósito general desarrollado por el W3C¹⁴ (World Wide Web Consortium), para representar información de los recursos que pueden ser identificados en la web. RDF proporciona un marco común para expresar esta información, de manera que pueda ser intercambiada entre aplicaciones sin perder su significado o sentido (Manola, et al., 2004), mediante la especificación de enunciados simples acerca de los recursos empleando propiedades y valores (tripletras RDF).

Sin embargo, este lenguaje no permite precisar cuando las descripciones de los recursos se refieren a un tipo o clase específica de un dominio (Manola, et al., 2004), por lo cual se hizo necesario crear un mecanismo que permitiera definir vocabularios especializados, para ser utilizados en los enunciados, esto es lo que se conoce como Lenguaje de Descripción de Vocabularios RDF: RDFS (*RDF Vocabulary Description Language* o RDF Schema) (Brickley, et al., 2004).

3.2.2. RDFS (RDF Schema)

RDFS es una extensión semántica de RDF, que define un mecanismo para describir las propiedades y las clases de los recursos RDF. RDFS provee terminologías definidas sobre RDF, que permiten el modelado de objetos con una semántica definida, gracias a la cual, es posible especificar las propiedades aplicables a las clases de objetos (Brickley, et al., 2004).

RDFS proporciona un conjunto de términos que permiten: construir las expresiones o sentencias acerca de los recursos, representar sus propiedades y establecer, mediante relaciones taxonómicas, jerarquías de generalización entre las clases. Mediante este lenguaje es posible instanciar objetos, a partir de las clases usando propiedades, y establecer algunas restricciones sobre dichas propiedades (Dinos, 2004).

Este lenguaje permite construir ontologías simples sobre las cuales se puede realizar consultas y razonamiento automático, sin embargo, no es lo suficientemente expresivo para ontologías complejas, ya que solo permite realizar inferencia sobre la herencia de propiedades y no posee la capacidad para generar axiomas (Corcho, et al., 2000) (Silva, 2002) (Hardman, et al., 2002) (Hernández, et al., 2009).

Para solucionar los inconvenientes presentados con RDF y RDFS, se desarrollaron otros lenguajes como OIL, DAML-ONT (DARPA Agent Markup Language - Lenguaje de Marcas de Agentes de DARPA) y DAML+OIL. OIL permite mejor interoperabilidad semántica entre recursos web que RDF, y emplea un modelo para la representación del conocimiento, basado en lógica descriptiva (axiomas y reglas) y en sistemas basados en marcos o frames (taxonomía de clases y atributos). Las principales carencias de este lenguaje son: la falta de expresividad para declarar axiomas y para soportar dominios concretos. Por su parte, DAML-ONT fue desarrollado como una extensión de XML y RDF para extender el nivel de expresividad de RDFS. Estos lenguajes se fundieron en DAML+OIL, el cual hereda muchas de las características de OIL, alejándose un poco del modelo basado en marcos y potenciando la lógica descriptiva (Bechhofer, et al., 2001). A partir de esta unión,

¹⁴ Disponible en la dirección www.w3.org/Consortium/

se definió el lenguaje OWL, con el propósito de reunir todas las ventajas de DAML+OIL y estandarizar el conocimiento extendido de RDF.

3.2.3. OWL (Web Ontology Language)

OWL es un lenguaje de marcado semántico, creado por la W3C para publicar y compartir ontologías en la web. En relación a sus predecesores, OWL añade más vocabulario para describir propiedades y clases, permite establecer relaciones adicionales entre las clases (por ejemplo de disjunción), así como atribuir ciertas propiedades a las relaciones, por ejemplo de cardinalidad, simetría, transitividad o relaciones inversas. OWL incluye toda la capacidad expresiva de RDFS, y la extiende con la posibilidad de utilizar expresiones lógicas (Castells, 2004).

Si bien, estas características hacen de OWL un lenguaje poderoso en expresión, también generan dificultades, en cuanto a la complejidad y decidibilidad computacional (García, 2007) (McGuinness, et al., 2004), lo cual limita la capacidad de los razonadores para inferir nuevo conocimiento a partir de sentencias dadas. Es por esto, que se crearon 3 variantes o sublenguajes de OWL, que proveen expresividad incremental con propiedades computacionales diferentes: OWL Lite, OWL DL y OWL Full (García, 2007).

OWL es el lenguaje de representación de ontologías, más utilizado en el contexto de la web semántica, principalmente porque se constituyó en un estándar avalado por el W3C. En la actualidad, existe una nueva versión denominada OWL 2 (Motik, et al., 2009), la cual adiciona algunas características respecto a OWL 1, enriqueciendo la expresividad con mas tipos y rangos de datos, incluyendo cadenas de propiedad, restricciones de cardinalidad calificadas y capacidades de anotación mejoradas (Golbreich, et al., 2009) (Horridge, et al., 2009) (Cuenca Grau, et al., 2009).

Paralelamente a OWL, surgió otro lenguaje de gran importancia para la descripción de ontologías, el cual, si bien no es considerado como tecnología de la web semántica, presenta grandes capacidades, en cuanto expresividad y es comúnmente utilizado en sistemas de gestión de procesos de negocio, este lenguaje se denomina WSML.

3.2.4. WSML (Web Service Modeling Language)

WSML es un lenguaje de modelado, que define un mecanismo formal para describir los elementos definidos en WSMO (Ontología de Modelado de Servicios Web - *Web Service Modelling Ontology*).

WSMO es una ontología formal, desarrollada por un conjunto de empresas coordinadas por DERI¹⁵, que proporciona un modelo conceptual, para la descripción de diversos aspectos relacionados con servicios web semánticos. WSMO identifica cuatro elementos de alto nivel, que agrupan los conceptos principales, que deben ser descritos para poder definir un servicio web semántico: las *Ontologías*, las cuales proporcionan la terminología usada por otros elementos de WSMO; los *Servicios Web*, que representan entidades computacionales capaces de proporcionar acceso a los servicios; las *Metas*, que describen aspectos relacionados con los deseos de los usuarios, respecto a la

¹⁵ Digital Enterprise Research Institute – www.deri.org

funcionalidad solicitada; y los *Mediadores*, cuyo objetivo es el manejo automático de problemas de interoperabilidad, entre diferentes elementos de WSMO (Keller, et al., 2004).

Si bien, WSMO proporciona el modelo conceptual para describir los anteriores elementos (entre los que se encuentran las ontologías), dicho lenguaje carece de una formalización real de los conceptos, para que puedan ser computacionalmente procesables. Es por esto, que la W3C recomienda utilizar WSML, como el lenguaje formal para la escritura, almacenamiento y comunicación de los conceptos definidos en WSMO (de Bruijn, et al., 2005).

WSML está basado en diferentes lógicas formales denominadas: Lógica de Descripción (Description Logics), Lógica de Primer Orden (First-Order Logic) y Programación Lógica (Logic Programming), las cuales se utilizan para el modelado de servicios web semánticos. Con base en estas lógicas formales, se han definido cinco variantes llamadas WSML-Core, WSML-DL, WSML-Flight, WSML-Rule y WSML-Full, cada una con distinto nivel de expresividad lógica. De estas cinco variantes, la más adecuada para describir ontologías es WSML-Flight, ya que es la variante más expresiva y combina tanto la lógica de descripción, como la de primer orden, con una complejidad computacional manejable (de Bruijn, et al., 2005).

Una de las ventajas más importantes que presenta WSML, es que se especifica en una sintaxis legible, y adicionalmente tiene una sintaxis XML y RDF, para el intercambio de datos entre servicios, lo que le brinda interoperabilidad con otras aplicaciones basadas en RDF. Además, WSML discrimina claramente, entre sintaxis conceptual y sintaxis de expresión lógica, permitiendo modelar distintivamente los elementos conceptuales de WSMO (servicio web, ontologías, metas y mediadores), de manera independiente de la descripción de restricciones adicionales y axiomas (Lausen, et al., 2005)

3.2.5. Selección del lenguaje semántico de descripción de ontologías.

Tanto las ontologías como los servicios, necesitan ser especificados en lenguajes formales con el fin de permitir su procesamiento automatizado. El lenguaje que recomienda la W3C para la descripción de ontologías es OWL, sin embargo, a pesar de que este lenguaje posee un gran potencial, en cuanto a expresividad lógica, presenta limitaciones tanto a nivel conceptual, como en la definición de algunas de sus propiedades formales (Fensel, et al., 2005) (de Bruijn, et al., 2005). Por esta razón, se seleccionó un lenguaje conceptual que permitió resolver el problema de la integración de aplicaciones, mediante la definición de una tecnología coherente y con buenas propiedades computacionales, para describir servicios web semánticos, WSMO.

Uno de los elementos principales de WSMO son las ontologías, sin embargo este modelo conceptual, por si solo, es insuficiente para automatizar su procesamiento, y por ende, requiere de un lenguaje formal de descripción como WSML. Este lenguaje proporciona una expresividad lógica razonable y una semántica formal bien definida, que facilitan el razonamiento lógico sobre las ontologías y demás elementos de WSMO (Feier, et al., 2005).

Teniendo en cuenta lo descrito anteriormente, se definió WSML como el lenguaje semántico para describir las ontologías utilizadas dentro del proyecto, ya que es más fácilmente comprensible en relación con OWL, y es uno de los lenguajes más avanzados, dentro de la industria de los servicios web semánticos. Además existen numerosas herramientas¹⁶ disponibles en la web que facilitan su utilización, tales como WSMO Studio, WSML Rule Reasoner, WSMO4J, WSMT (Web Service Modeling Toolkit), entre otras.

3.3. LENGUAJE DE NOTACIÓN SEMÁNTICA PARA PROCESOS DE NEGOCIO

En los modelos de negocio actuales son cada vez más comunes los procesos de fusión, alianza y colaboración entre empresas, especialmente cuando se trata de mejorar las ofertas comerciales y brindar servicios más complejos al cliente. Para sobrellevar estas tendencias, se requiere que las organizaciones realicen la gestión del negocio de manera eficiente y responsable, orientando su esfuerzo hacia la evolución en el manejo de sus procesos, con la ayuda de herramientas especializadas, que faciliten el modelado y análisis de los mismos. Sin embargo, estas herramientas generalmente requieren un perfil muy técnico para su uso, y sus representaciones son demasiado complejas, para ser comprendidas por los responsables del negocio.

Se hace necesario entonces, integrar las herramientas de gestión de procesos de negocio con componentes semánticos, que permitan expresar los procesos en un lenguaje que facilite su comprensión al personal no técnico. Uno de esos componentes puede estar dado por un lenguaje que soporte la definición de procesos de negocio, enriquecidos semánticamente como BPMO.

3.3.1. BPMO (Business Process Modelling Ontology)

BPMO es un lenguaje para la definición de procesos de negocio, que busca disminuir la brecha de entendimiento existente, entre los profesionales del negocio y los expertos en TI (Jędrzejczak, et al., 2008). Provee una plataforma, que abarca la descripción de los procesos a un nivel alto de abstracción, facilitando tanto la gestión de los procesos a nivel de negocios, como la conexión a nivel de implementación técnica, con el fin de dar soporte a la ejecución automatizada de los procesos.

BPMO está constituido por varios elementos, necesarios para la descripción de procesos de negocio semánticos (Cimpian, 2007), entre los que se encuentra una ontología BPMO que define las entidades u objetos involucrados en dichos procesos (actores, organizaciones, patrones, subprocessos, tareas, mensajes, entre otros) (Cabral, et al., 2009). Al estar fundamentado en BPMN, este lenguaje permite modelar los procesos de negocio con dicha notación y enriquecerlos con conceptos de la ontología BPMO, la cual proporciona la semántica necesaria para facilitar y reducir la ambigüedad en su interpretación.

¹⁶ Disponibles en <http://www.wsmo.org/wsml/>

En el presente proyecto se utiliza BPMO, como lenguaje de notación semántico para describir procesos de negocio, sin embargo, aun cuando dichos procesos estén enriquecidos con la ontología BPMO (siendo comprensibles en distintos niveles), es necesario añadir un componente semántico propio del dominio de las telecomunicaciones, para enmarcarlos en el área específica de interés. Este componente corresponde a las ontologías de dominio, cuyo estudio se presenta en las siguientes secciones.

3.4. ONTOLOGÍAS DE DOMINIO DE TELECOMUNICACIONES

Una de las claves, para lograr la interoperabilidad de los sistemas o aplicaciones, es contar con un lenguaje estandarizado, que permita el entendimiento común de los conceptos. En esta medida, surge la necesidad de encontrar un modelo, que permita obtener una visión general de los conceptos más utilizados por los operadores de telecomunicaciones, y que además sea computacionalmente comprensible (es decir que pueda ser interpretado y procesado por una aplicación o sistema software); estas características son proporcionadas por las ontologías de dominio de telecomunicaciones. Dichas ontologías, son necesarias para llevar a cabo el propósito del presente proyecto, puesto que constituyen la herramienta para enriquecer semánticamente, los parámetros que definen las funcionalidades de las tareas de los procesos de negocio (descritas en términos de sus entradas, salidas e identificadores).

Dentro de las investigaciones revisadas se encontraron varios proyectos relacionados con ontologías de dominio de telecomunicaciones (Frankowski, et al., 2007) (Dawidziuk, et al., 2007) (Rój, 2007) (Watkins, et al., 2005) (Wahle, 2008) (Devitt, et al., 2006) (TeLQAS) (Qiao, et al., 2010); sin embargo gran parte de las ontologías utilizadas en estos se limitan a modelar el conocimiento de un sector muy específico o de una empresa particular, por lo tanto su aplicación dentro del presente proyecto no es viable, en la medida en que no favorece el carácter genérico y la escalabilidad de la solución propuesta.

No obstante, uno de los proyectos importantes que se encontraron, es un marco de referencia semántico para el sector de las telecomunicaciones denominado YATOSP (Yet Another Telecommunication Ontologies, Process And Services Framework), el cual busca proporcionar un puente entre la arquitectura SUPER¹⁷ (Semantics Utilized for Process Management within and between Enterprises) y un conjunto de ontologías basadas en la iniciativa NGOSS del TMF denominadas: SeTOM, SSID y STAM¹⁸, que definen procesos de negocio semánticos en el dominio de telecomunicaciones, mediante la extracción de información común entre las compañías de este sector (Martinez, et al., 2008) (Heymans, 2007). A pesar de que existe una buena documentación, en cuanto a la descripción de la arquitectura de YATOSP, este trabajo carece de ejemplos de ejecución que permitan verificar la validez de sus planteamientos y restringe el acceso a algunos documentos de gran relevancia para su posible implementación.

¹⁷ Disponible en <http://ip-super.org/>

¹⁸ Semantic TAM: Ontología basada en el modelo TAM.

Sin embargo, el marco de referencia YATOSP sirvió para centrar la atención en la iniciativa NGOSS, especialmente en el conjunto de modelos propuestos para la gestión de los procesos de negocio, la reingeniería de procesos y la integración de aplicaciones empresariales, propias del sector de las telecomunicaciones: el Mapa de Operaciones de Telecomunicaciones Mejorado (eTOM, enhanced Telecom Operations Map), el Modelo de Información/Datos Compartidos (SID, Shared Information/Data Model), el Mapa de Aplicaciones de Telecomunicaciones (TAM, Telecom Applications Map) y la Arquitectura Neutral de Tecnologías (TNA, Technology Neutral Architecture) (Fleck, 2003).

Teniendo como referencia estas propuestas, fue posible establecer una relación de correspondencia, entre la forma en que se describen las tareas de un proceso de negocio y dos de los modelos desarrollados dentro de esta iniciativa, así: el identificador de una tarea es un parámetro que en el dominio específico de las telecomunicaciones puede, modelarse a partir del conocimiento dispuesto en el estándar eTOM, mientras que las entradas/salidas, pueden asimilarse a las entidades que conforman el modelo SID. Para entender mejor estas relaciones, a continuación se presenta una descripción de las principales características de estos dos modelos (eTOM y SID) y de las ontologías desarrolladas a partir de ellos (SeTOM y SSID).

3.4.1. eTOM - enhanced Telecom Operations Map.

eTOM es un marco de referencia de procesos de negocio, que proporciona un modelo para la categorización de todas las actividades, involucradas en la operación de un proveedor de servicios de telecomunicaciones. Provee una visión común (sustentada en una terminología común) de la industria de las telecomunicaciones, en cuanto a procesos e información, la cual facilita y soporta las relaciones operador-a-operador, operador-a-cliente y operador-a-socio/proveedor respondiendo con esto, a la necesidad que representa el rápido despliegue de servicios y el manejo de problemas que exige el entorno competitivo actual (TMF - eTOM, 2005).

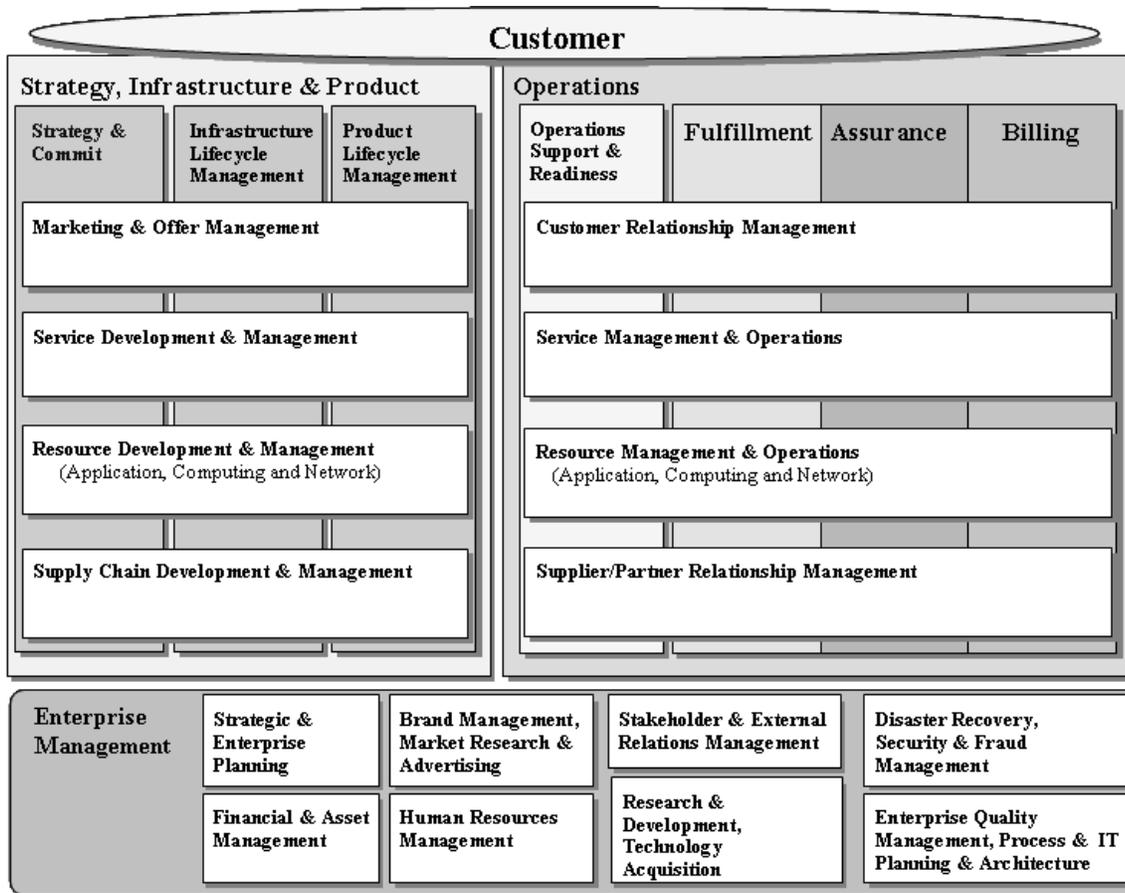
De esta manera, eTOM ordena de una forma jerárquica coherente, los conceptos que constituyen el conocimiento relativo a los procesos, subprocesos y actividades que se llevan a cabo dentro de un proveedor de servicios de telecomunicaciones. Estos procesos pueden clasificarse en cuatro niveles, cada uno de los cuales detalla a su vez procesos más específicos.

El nivel 0 corresponde a las tres grandes áreas que abarca el modelo eTOM: Estrategia-Infraestructura-Producto, Operaciones y Gestión Empresarial (Figura 3).

El área de Estrategia-Infraestructura-Producto abarca los procesos involucrados en la elaboración de estrategias, a partir de los cuales, la empresa planifica el desarrollo y la gestión tanto de la entrega, como de la cadena de suministro y el mejoramiento de infraestructura y productos, entendiendo infraestructura, no solo como los recursos de tecnologías de la información y de red, que soportan el despliegue de productos y servicios, sino también incluyendo la infraestructura operativa y organizacional, necesarias para sustentar la comercialización, ventas y procesos de servicios y cadena de suministro.

Por otra parte, el área de Operaciones comprende aquellos procesos que proveen soporte directo al cliente, en ella pueden identificarse cuatro agrupaciones de procesos de extremo a extremo: Aseguramiento, Cumplimiento, Facturación, y Soporte y Disposición de Operaciones. Ésta última se encarga de asegurar que las tres agrupaciones previas puedan responder a los requerimientos del cliente, preparando la información, productos, servicios y recursos, así como los proveedores y socios, para entregar y sustentar las instancias de servicio del cliente.

Figura 3. Mapa de Operaciones de Telecomunicaciones Mejorado – eTOM.



Fuente: Enhanced Telecom Operations Map (eTOM) – The business process framework (TMF - eTOM, 2005).

Por último, en el área de Gestión Empresarial, se incluyen los procesos de negocio que soportan directamente la administración y el funcionamiento de cualquier empresa, comprendiendo aquellos procesos que están dirigidos hacia el establecimiento y alcance de metas y objetivos corporativos y el soporte de servicios requeridos en la operación de toda la organización. Entre estos procesos se encuentran, la Gestión Financiera y la Gestión de Recursos Humanos.

Los procesos pertenecientes a las áreas de Estrategia-Infraestructura-Producto y Operaciones, están organizados a su vez, dentro de una matriz (nivel 1): Las divisiones verticales, representan flujos de proceso extremo a extremo, que atraviesan capas

horizontales que representan cada una, interfaces con el cliente, servicios, recursos y socios/proveedores (Rodríguez, et al., 2008). En cada capa horizontal se ubican agrupamientos de procesos u operaciones más específicos (nivel 2), y dentro de estos bloques, se distribuyen las actividades, en las cuales se soportan dichos procesos (nivel 3).

Actualmente eTOM, es el estándar más ampliamente empleado y aceptado para procesos de negocio en la industria de las telecomunicaciones (Burrows, et al., 2008), por lo cual, ha sido utilizado dentro del marco de referencia semántico YATOSP, para formular una ontología basada en esta iniciativa de estandarización:

Ontología SeTOM: Recoge y modela en WSMML, los más de trescientos conceptos contenidos en el estándar eTOM de NGOSS, los cuales definen con gran detalle las áreas funcionales y actividades que se desarrollan en un operador de Telecomunicaciones. A modo de ejemplo aparecen conceptos como: Autorizar Crédito, Crear Factura, Analizar calidad del servicio, etc (Martínez, et al., 2008).

El carácter de estándar de eTOM, sumado a la amplia aceptación dentro del sector de las telecomunicaciones, hacen que la utilización de la ontología construida con base en el conocimiento contenido dentro de este marco de referencia (*SeTOM*), se constituya en una alternativa adecuada para abordar el enriquecimiento semántico de los identificadores de las tareas, de los procesos de negocio de un operador de telecomunicaciones, lo cual está establecido como uno de los principales objetivos del proyecto ejecutado. La utilización de la ontología *SeTOM* está enmarcada, en el enriquecimiento de procesos, que están fundamentalmente relacionados con el desarrollo, soporte y despliegue de servicios de telecomunicaciones.

A continuación, se aborda una descripción general del marco de referencia SID, cuyo conocimiento, al igual que eTOM, ha sido modelado en forma de una ontología de dominio (*SSID*).

3.4.2. SID - Shared Information/Data Model

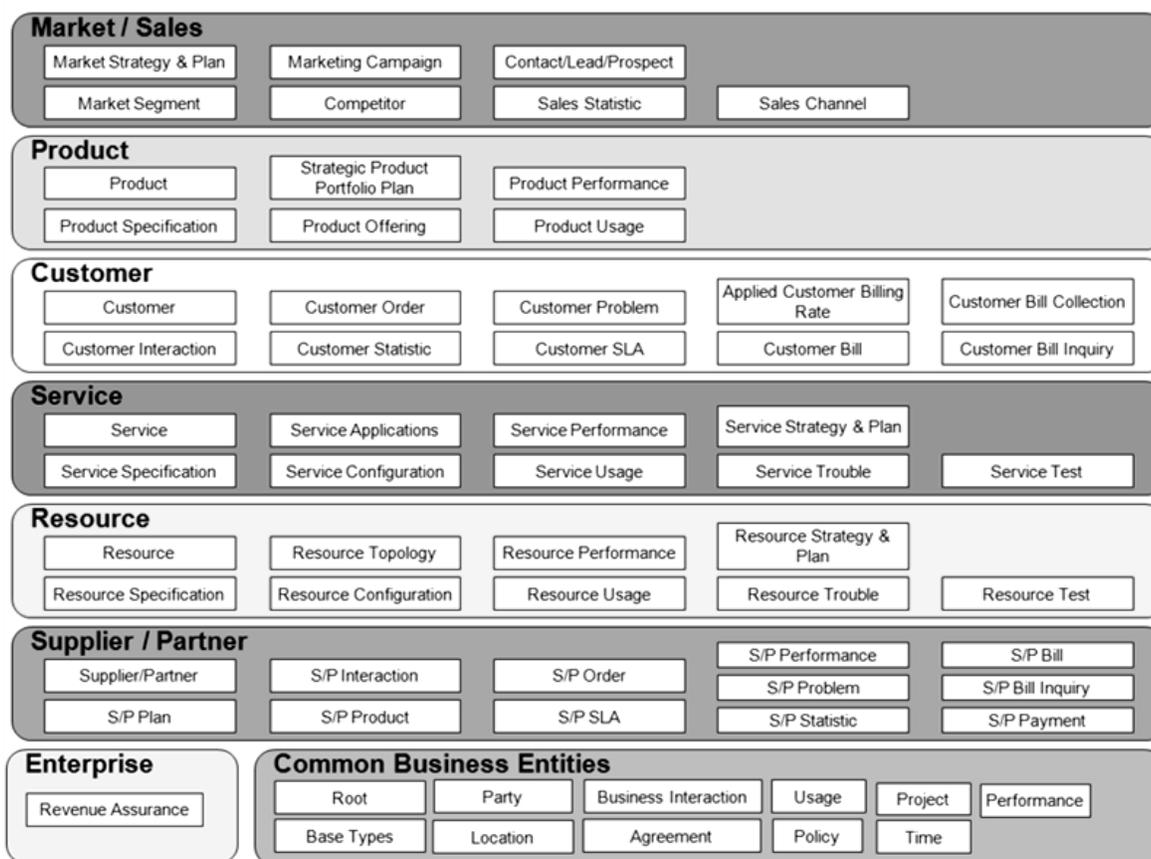
El modelo de Información/Datos compartidos (SID) es un modelo unificado de referencia de datos, que proporciona un conjunto único de términos, que son comúnmente utilizados en procesos de negocio básicos del dominio de las telecomunicaciones, así como las conexiones y relaciones entre ellos (TMF - SID, 2005). El modelo SID pretende facilitar la comunicación entre los encargados del negocio, los desarrolladores de software, e incluso los mismos empleados de las empresas, promoviendo la utilización de estos términos para referirse o describir la misma información en los procesos de negocio.

SID captura los conceptos y principios necesarios en la definición de un modelo de información compartida, y define en detalle, muchos elementos del negocio (conocidos en SID como "Entidades") y sus atributos asociados. Una entidad de negocio, es un objeto de interés en las empresas tales como cliente, producto, servicios o red, mientras que sus atributos son características que describen la entidad. Juntos, estos términos proporcionan una perspectiva orientada a la información y los datos del negocio que se necesitan para el funcionamiento de la empresa (Lambea, 2009).

Para clasificar los datos en una forma usable, la arquitectura de SID está diseñada como un modelo de capas o niveles, las cuales dividen los datos/información en ocho dominios: Mercado/Ventas, Producto, Cliente, Servicio, Recurso, Suministrador/Socio, Empresa y Entidades de Negocio Comunes (Figura 4).

En la capa más superficial, cada uno de los ocho dominios de información está alineado con el marco de referencia eTOM. Dentro de cada dominio, existe un alto grado de cohesión entre las entidades del negocio, y entre dominios existe un acoplamiento flexible. Lo anterior, facilita la segmentación del problema del negocio en piezas manejables y permite que los recursos se concentren en un área particular de interés. En otras palabras, para un proceso de negocio particular, que se esté automatizando, gracias a SID, se puede identificar qué información es necesaria, dentro de la arquitectura para soportar ese proceso.

Figura 4. Modelo de Información/Datos Compartidos – SID.



Fuente: Information Framework (SID) In Depth (TMF - SID, 2005)

Además de contener, las definiciones de las entidades empresariales, que participan en la gestión de las operaciones de telecomunicaciones, el modelo SID desempeña un papel importante al permitir la integración y la interoperabilidad de los modelos de la iniciativa NGOSS, dado que es el medio, por el cual la información útil (es decir la que es relevante

para un determinado conjunto de procesos de negocio), puede ser compartida y empleada, en múltiples procesos de negocio dentro de todo el sistema (NGOSS, 2002).

Una de las ventajas que tiene la utilización de SID, es que proporciona una fusión de la terminología manejada en las empresas de telecomunicaciones, condensada en un modelo simple de información del negocio, la cual hace posible el desarrollo de aplicaciones OSS/BSS con interfaces verdaderamente interoperables. SID constituye un marco de referencia para la representación de información/datos, que pueden ser compartidos y/o reutilizados por aplicaciones OSS/BSS brindados por múltiples proveedores (Rodríguez, et al., 2008).

La ontología SSID propuesta en YATOSP, presenta más de mil conceptos organizados en diez niveles de jerarquía, que conforman un vocabulario común, ampliamente utilizado dentro del sector de las telecomunicaciones, los cuales cubren términos como cliente, producto, servicio, recurso, proveedor, etc (Martínez, et al., 2008). Por esta razón, es recomendable la utilización de esta ontología dentro del presente proyecto, puesto que sus conceptos están estrechamente relacionados con la información que se comparte tanto entre aplicaciones informáticas, como entre el personal al interior de un operador de telecomunicaciones. De esta manera es posible realizar el enriquecimiento semántico de las entradas y salidas, de las tareas de procesos de negocio, con conceptos pertenecientes a esta ontología, dado que estos corresponden a la información que se intercambia para realizar las distintas actividades del negocio.

3.5. ADAPTACIÓN DE ONTOLOGÍAS DE DOMINIO

Las ontologías de dominio, utilizadas en el presente proyecto, fueron definidas dentro del marco de referencia para las telecomunicaciones YATOSP del proyecto SUPER, y están basadas en el conocimiento provisto en los modelos eTOM y SID de la NGOSS. Es importante resaltar, que estos modelos han tenido gran aceptación en la industria de las telecomunicaciones, a tal punto, que uno de los organismos regulatorios más importantes del sector como es la Unión Internacional de Telecomunicaciones (ITU), los han acogido para la formulación de varias recomendaciones, en el caso de eTOM correspondientes a la serie M.3050 y en el caso de SID, está siendo estudiada la propuesta para la recomendación M.3190. Esto refleja, que el potencial de estas ontologías, es bastante prometedor e involucra a muchos actores de la industria, razón por la cual actualmente las ontologías de eTOM y SID son confidenciales y están siendo estudiadas a fondo, para ser modificadas y mejoradas.

De esta manera, la adaptación de estas ontologías dentro del presente proyecto no está enfocada en la alteración de los conceptos, instancias y axiomas que las conforman, dado que la construcción y evolución de este tipo de ontologías, son procesos políticos que involucran empresas del sector de telecomunicaciones, entes regulatorios, grupos de investigación, entre otros; principalmente debido al deseo de estandarizar y masificar el uso de estos recursos en distintas aplicaciones. Por esta razón las ontologías de dominio seleccionadas, fueron adaptadas al proyecto específicamente, para enriquecer las tareas de los procesos de negocio de telecomunicaciones, permitiendo ampliar el conocimiento, respecto a las funcionalidades que desempeñan las tareas, y definiendo puntualmente la

ontología más adecuada, para enriquecer cada parámetro de las mismas (identificador, entradas y salidas).

En este sentido, de acuerdo a las propuestas de la NGOSS descritas en la sección anterior, se estableció una clara relación entre los parámetros que describen las tareas de un proceso de negocio y dos de los modelos propuestos dentro de esta iniciativa. Como se mencionó anteriormente, el identificador de una tarea, el cual hace referencia a la operación que esta ejecuta sobre sus entradas para obtener sus salidas, es un parámetro, que en el dominio de las telecomunicaciones, puede modelarse a partir del conocimiento dispuesto en el estándar eTOM, y por ende la ontología más adecuada para enriquecer semánticamente los identificadores de las tareas es SeTOM, cuyos conceptos e instancias corresponden a verbos o acciones que se llevan a cabo dentro de los procesos de negocio.

Por otro lado, las entradas y salidas de las tareas pueden asimilarse a las entidades que conforman el modelo SID, en tanto corresponden a la información que se comparte entre las tareas de un proceso de negocio, es decir, a los objetos del negocio que se intercambian entre dichas tareas. Como consecuencia, se definió que la ontología más apropiada para enriquecer semánticamente las entradas y salidas de las tareas es SSID, cuyos conceptos corresponden a sustantivos u objetos propios del dominio de telecomunicaciones.

A continuación, se presenta una descripción detallada del procedimiento llevado a cabo para adaptar las ontologías de dominio al proyecto, realizada mediante el enriquecimiento de las tareas de los procesos de negocio de telecomunicaciones.

3.5.1. Mecanismo de enriquecimiento semántico de Tareas

El propósito del presente proyecto, es desarrollar un mecanismo de comparación entre tareas de dos procesos de negocio, del dominio de los servicios de telecomunicaciones a partir del cálculo de su similitud semántica. Un requisito previo fundamental para conseguir este propósito, consiste en enriquecer semánticamente los parámetros que definen las capacidades o funcionalidades de las tareas, descritas en términos de sus entradas, salidas e identificadores, para posteriormente evaluar la similitud existente entre estos parámetros.

En este sentido, se hace necesario enriquecer semánticamente los procesos de negocio para llevar a cabo la inferencia sobre los componentes que los constituyen. La automatización de operaciones, como la comparación y el descubrimiento de servicios (descritos a partir de un proceso de negocio), dependen en gran medida de la correcta realización de dicha inferencia, y por lo tanto, de la pertinencia del enriquecimiento semántico realizado sobre los elementos que integran el proceso, es decir, de las tareas.

En el presente proyecto, la labor de enriquecer semánticamente las tareas de un proceso de negocio, se abordó a partir de la utilización de la herramienta WSMO Studio, la cual permite modelar gráficamente el proceso en BPMO, al tiempo que genera la ontología que define el proceso en particular. Posteriormente, empleando las ontologías de dominio de telecomunicaciones, es posible especificar los conceptos que representan los

identificadores, entradas y salidas de las tareas del proceso, para enmarcarlo en el área específica de interés.

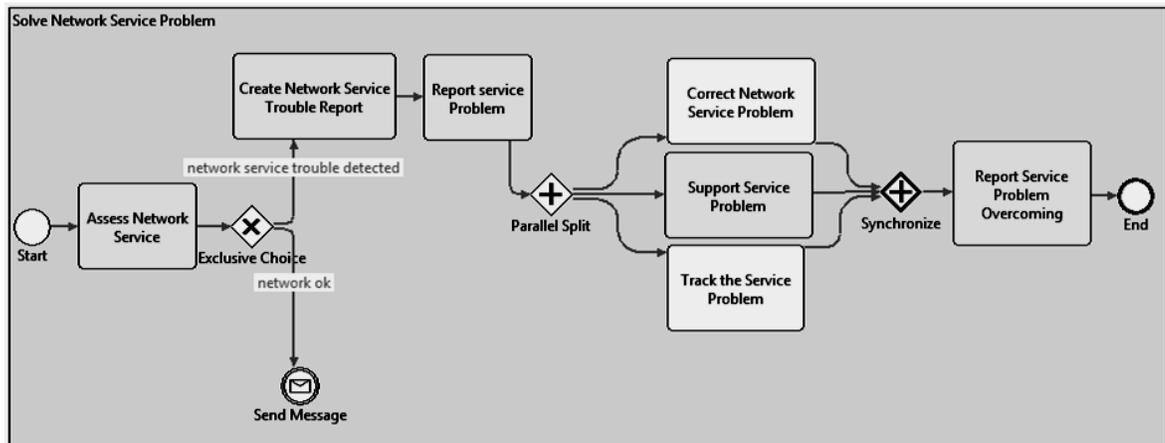
A continuación, se da a conocer la forma como fue realizado dicho enriquecimiento sobre las tareas de procesos de negocio, a partir de la implementación de un caso representativo del dominio de las telecomunicaciones. Para esto, se dispuso de las ontologías de dominio definidas en el marco de YATOSP: SeTOM y SSID (versiones disponibles en la página oficial de SUPER a principios del año 2009), y se utilizó el entorno de desarrollo WSMT¹⁹, para el enriquecimiento de los identificadores y para realizar operaciones de razonamiento, sobre la ontología del proceso de negocio de estudio.

3.5.1.1. Definición del Caso de Estudio y Modelado del Proceso de Negocio.

El caso estudio, que se presenta en esta sección, hace parte del conjunto de procesos de negocio diseñados e implementados dentro del presente proyecto, y corresponde al procedimiento ejecutado, para solucionar un problema de red, que influye en la prestación de un servicio de telecomunicaciones.

De esta manera, utilizando la capacidad de modelado de procesos de negocio sobre BPMO que provee WSMO Studio, se implementó el proceso “Solve Network Service Problem”, tal como se muestra en la Figura 5. Este modelo describe, las operaciones necesarias para detectar y corregir un problema en la red del operador de telecomunicaciones, incluyendo actividades de evaluación, reporte y seguimiento del problema hasta asegurar su solución.

Figura 5. Proceso “Solve Network Service Problem” en WSMO Studio.



Una vez modelado el proceso de negocio, fue posible empezar a enriquecer semánticamente las tareas que lo constituyen. En primer lugar, se enriquecieron sus

¹⁹ *Web Service Modelling Toolkit*, una descripción básica de las funcionalidades de este entorno se presenta en el Anexo “Pruebas de Razonamiento sobre WSML2Reasoner”.

entradas y salidas, y posteriormente se enriquecieron sus identificadores. Ambos procedimientos son explicados en las siguientes secciones.

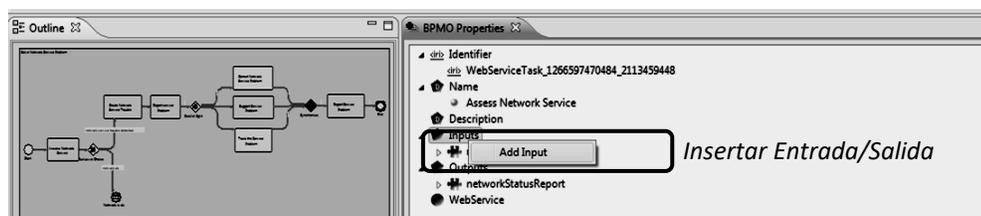
3.5.1.2. **Enriquecimiento Semántico de Entradas/Salidas de las Tareas**

Existen dos formas de enriquecer semánticamente las entradas/salidas de las tareas de un proceso de negocio, esto es: definiendo el tipo de dato que las representa, o definiendo el concepto que las identifica. El primer caso, se aplica cuando las entradas/salidas corresponden a datos básicos, tales como enteros, flotantes, cadenas (strings), booleanos, etc, mientras que el último caso, es aplicable cuando las entradas/salidas representan entidades de un dominio específico, es decir que pertenecen al conjunto de conceptos de una ontología.

Para ilustrar esto, considere la tarea “Assess Network Service”, del proceso anteriormente modelado, la cual se encarga de verificar, si los servicios proporcionados en la red del operador funcionan de manera adecuada, y de reportar las fallas que puedan presentarse. Esta tarea recibe como entrada el nombre del servicio (“ServiceName” dentro del modelo), que corresponde en este caso, a un dato básico de tipo string. También recibe como entrada, la localización del servicio en la red (denominada “NetworkServiceLocalization”), la cual corresponde a una entidad específica del dominio de telecomunicaciones. Con esta información, la tarea evalúa el estado de los servicios y verifica si todos los componentes están operando apropiadamente en la red. Una vez realizado dicho análisis, la tarea entrega a la salida un reporte del estado de la red, llamado “NetworkStatusReport”.

Para reflejar estas características dentro del modelo BPMO del proceso, es necesario adicionar las entradas y salidas correspondientes a la tarea (Figura 6).

Figura 6. Adición de Entradas a las Tareas BPMO



Una vez creadas las entradas y salidas se debe especificar el tipo de dato correspondiente a cada una de ellas. En el caso de la entrada “ServiceName”, se debe seleccionar el *Tipo de Dato* string entre un conjunto de tipos básicos predefinidos, de acuerdo con la especificación de WSMO (Figura 7).

Figura 7. Selección de Data Type - Entrada ServiceName

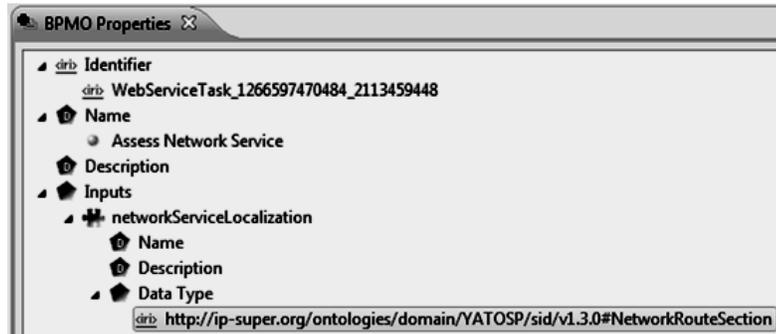


En caso de que las entradas/salidas de la tarea, representen entidades que hacen parte del conjunto de conceptos una ontología, por ejemplo, la entrada "NetworkServiceLocalization" (la cual puede ser asociada a una instancia del concepto *NetworkRouteSection* definido dentro de la ontología de SID), se debe especificar el *Tipo de Concepto*, en lugar del *Tipo de Dato*. Para el caso particular que se está abordando, se selecciona el concepto *NetworkRouteSection* presente en la ontología de SID, tal como se muestra en las Figuras 8 y 9.

Figura 8. Enriquecimiento Entrada NetworkServiceLocalization



Figura 9. Entrada “NetworkServiceLocalization” enriquecida



3.5.1.3. Enriquecimiento semántico de los identificadores de las Tareas

Es posible realizar el enriquecimiento semántico sobre los identificadores de las tareas que componen el proceso de negocio de tal manera que estas se identifiquen con conceptos presentes en una ontología de dominio. Sin embargo, realizar este tipo de anotación en WSMO Studio no es un procedimiento tan intuitivo como el que se observó para las entradas/salidas de las tareas, ya que es necesario trabajar directamente sobre el código de la ontología correspondiente al proceso de negocio modelado, de manera que se requiere tener un conocimiento básico de la estructura de una ontología, especificada en WSML (de Bruijn, et al., 2005).

En el caso que se está considerando como ejemplo, es posible identificar que la funcionalidad de la tarea “Assess Network Service”, corresponde a la descripción del concepto *testServiceEnd_to_End* de la ontología de eTOM, el cual se refiere a “poner a prueba los servicios específicos, para asegurar que todos los componentes estén operando dentro de parámetros normales, y que el servicio esté trabajando dentro de los niveles de rendimiento acordados”. De esta manera es posible determinar que la tarea “Assess Network Service” representa una instancia del concepto *testServiceEnd_to_End* de SeTOM, dentro de la ontología del proceso de negocio, y por lo tanto su identificador puede ser enriquecido con dicho concepto.

Para realizar, efectivamente este enriquecimiento, puede utilizarse cualquier editor de texto, WSMO Studio también cuenta con esta herramienta, sin embargo se recomienda utilizar el editor provisto en el entorno WSMT (Figura 10), ya que es más amigable y permite realizar pequeñas pruebas de razonamiento, sobre la ontología del proceso, para verificar que el procedimiento de enriquecimiento se ha realizado satisfactoriamente.

Para realizar el enriquecimiento de los identificadores, en primer lugar, se debe definir la variante de WSML sobre la cual, se va a inferir en la ontología del proceso de negocio. Teniendo como referencia la sección 3.1.4 de este capítulo, se determinó utilizar la variante WSML-Flight, debido a que es la más adecuada para la descripción de ontologías (Figura 11).

Figura 10. Ontología del Proceso de Negocio “Solve Network Service Problem” en WSMT

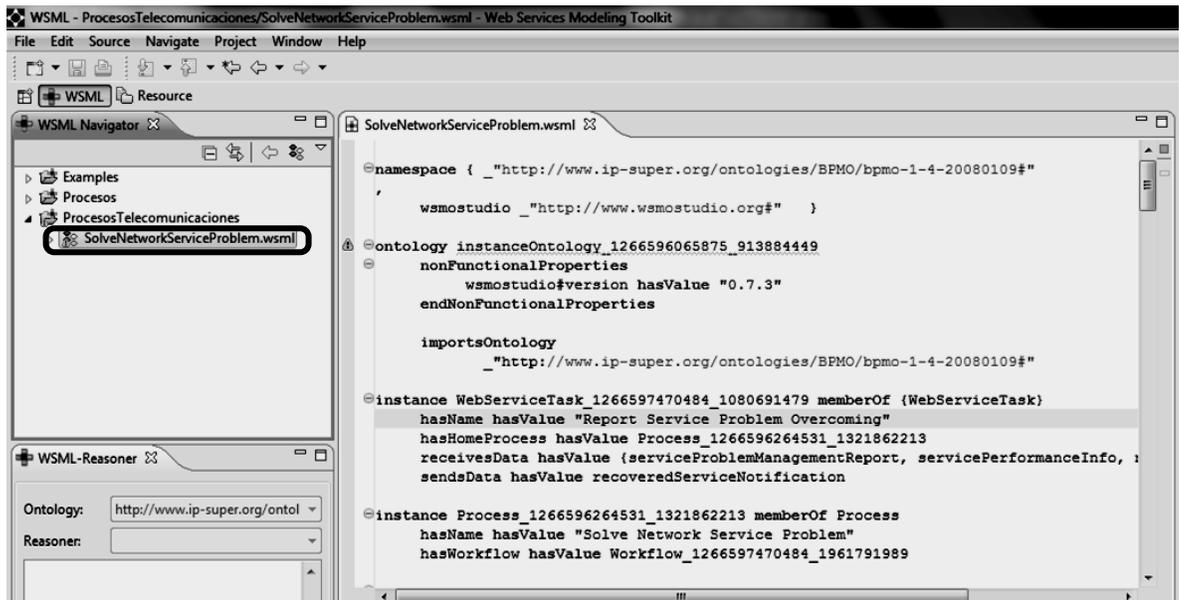
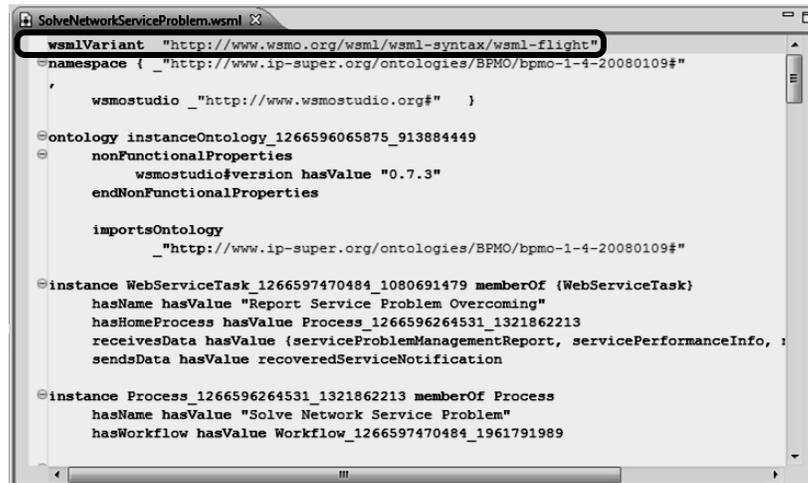


Figura 11. Definición de la variante de WSML



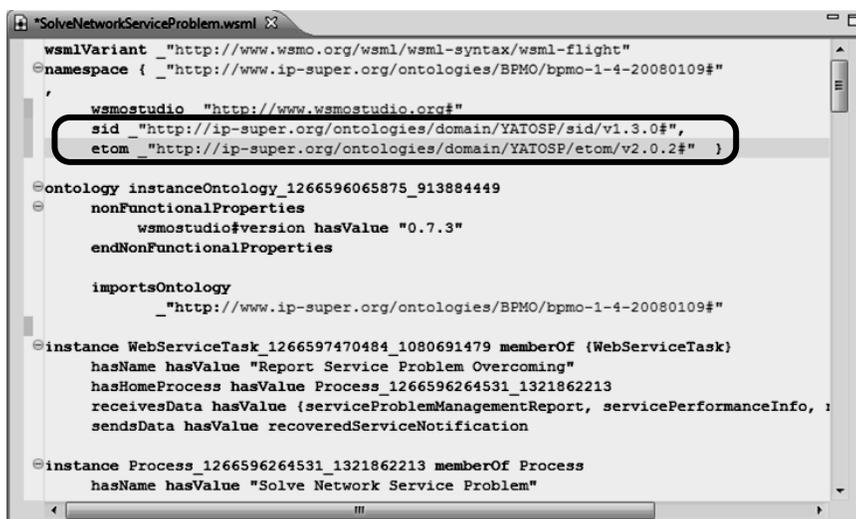
Posteriormente, se debe añadir en el namespace²⁰ las ontologías de dominio que se van a utilizar, para realizar el enriquecimiento, para el presente proyecto son las ontologías basadas en los estándares SID y eTOM (Figura 12). Una vez son añadidas las ontologías de dominio a la ontología de proceso, se pueden utilizar sus conceptos para realizar el enriquecimiento semántico de las tareas.

²⁰ Concepto heredado de XML y otras recomendaciones del W3C para denotar un ámbito que agrupa un conjunto de identificadores únicos.

Ahora bien, la ontología que describe el proceso de negocio, está constituida por instancias de los conceptos definidos en la ontología BPMO (proceso, tareas, eventos, conectores, etc.), es decir que el proceso de negocio modelado, así como los elementos que lo conforman, son instancias de los conceptos contenidos en dicha ontología, y por lo tanto, todas las tareas contenidas en el proceso son instancias de conceptos de BPMO. Adaptando este principio al ámbito de las ontologías de dominio (de Francisco, et al., 2009), las tareas de los procesos de negocio pueden definirse como instancias de conceptos de la ontología de eTOM, es decir, que se puede asociar el identificador de una tarea a un concepto de SeTOM, de la misma manera, como se asocia con el concepto de BPMO.

En el ejemplo que se ha venido trabajando, la tarea “Assess Network Service”, puede ser identificada como una instancia del concepto *testServiceEnd_to_End* de la ontología de eTOM, por lo tanto, se debe especificar dentro de la ontología del proceso de negocio que la instancia correspondiente a la tarea “Assess Network Service” es miembro del concepto *testServiceEnd_to_End*.

Figura 12. Adición de ontologías de dominio al namespace



Esto se puede realizar al modificar la definición de la instancia, adicionando el prefijo correspondiente de la ontología de dominio (como aparece en el namespace) y el concepto de dicha ontología al que se desea asociar la instancia (tarea) (Figura 13).

Por último, en la Figura 14 se muestra el resultado de la consulta en WSMML (*?x memberOf ?var0*) que se realizó en el razonador proporcionado por la herramienta WSMT, para comprobar la validez del procedimiento anterior. En esta figura se puede observar que la tarea “Assess Network Service”, cuyo identificador es “WebServiceTask_1266597470484_2113459448” es miembro del concepto *testServiceEnd_to_End* de eTOM.

Figura 13. Adición manual de anotaciones semánticas a la tarea Assess Network Service

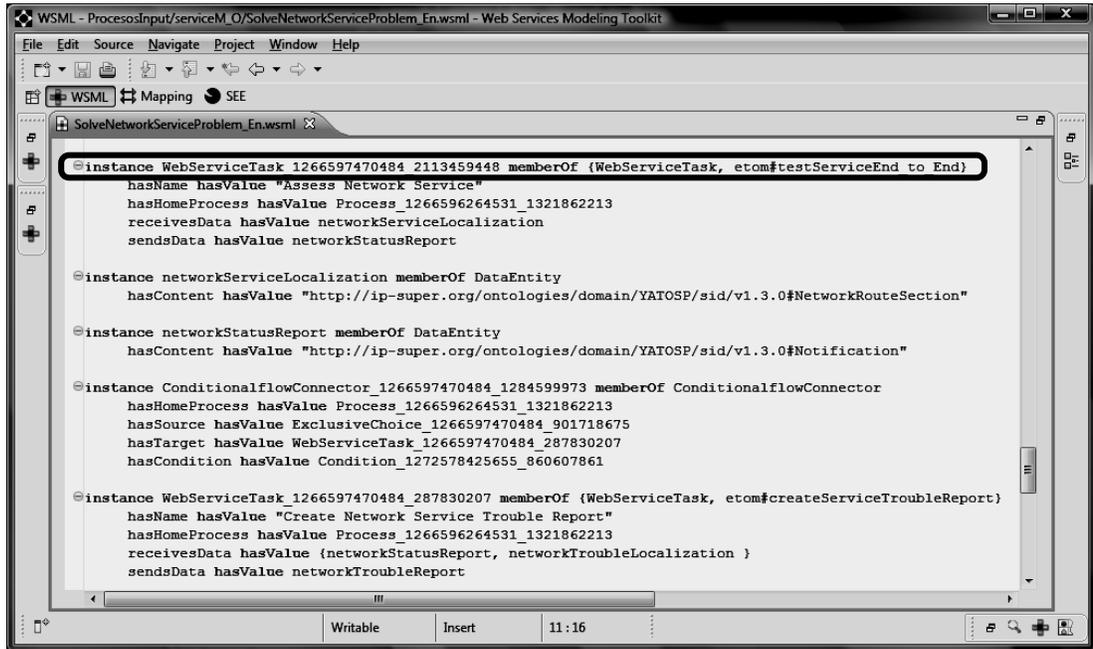
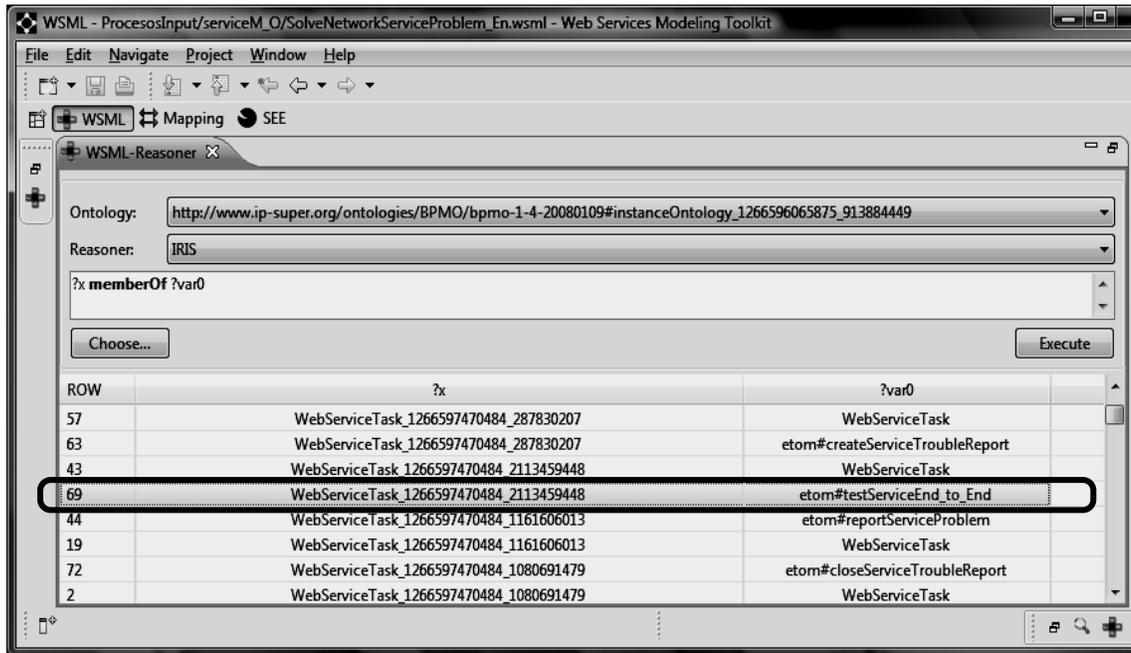


Figura 14. Resultado de la consulta sobre la ontología del proceso



3.6. RESUMEN

Para facilitar el entendimiento y gestión eficiente de los procesos de negocio, dentro de las empresas de telecomunicaciones, es importante contar con un lenguaje de notación semántico como BPMO, el cual permite enriquecer las descripciones de dichos procesos, con elementos que los hacen comprensibles, tanto para el personal técnico experto en TI como para los profesionales del negocio. Adicionalmente, es necesario enriquecer los procesos con conceptos propios del sector de telecomunicaciones, utilizando recursos como las ontologías de dominio, que proporcionan un vocabulario común dentro de esta área específica de conocimiento. Para favorecer la interoperabilidad entre las aplicaciones y sistemas de gestión, estas ontologías en lo posible deben estar fundamentadas en modelos ampliamente reconocidos. En el caso particular del sector de las telecomunicaciones, una buena alternativa son las propuestas de la NGOSS: eTOM y SID, que constituyen la base fundamental para las ontologías SeTOM y SSID de YATOSP. Junto con BPMO, estas ontologías juegan un papel fundamental dentro del presente proyecto, puesto que son las herramientas que permiten modelar los procesos de negocio de telecomunicaciones, y enriquecer semánticamente las tareas que los conforman, lo cual se constituye en un requisito indispensable para realizar la comparación semántica entre tareas de procesos de negocio de telecomunicaciones. El estudio de los modelos propuestos por la NGOSS (eTOM y SID) permitió definir la relación existente entre el conocimiento dispuesto en las ontologías de dominio seleccionadas y los atributos que definen la funcionalidad de las tareas (identificador, entradas y salidas). De esta manera, se hizo posible la adaptación de las ontologías SeTOM y SSID al mecanismo de enriquecimiento descrito en este capítulo.

4. ALGORITMOS DE COMPARACIÓN SEMÁNTICA DE TAREAS

4.1. INTRODUCCIÓN

En este capítulo se describe el conjunto de algoritmos propuestos, para realizar la comparación semántica de tareas entre procesos de negocio de telecomunicaciones, basados en el cálculo de la similitud semántica, existente entre los parámetros que definen su funcionalidad.

La primera parte, aborda los procedimientos para encontrar la distancia semántica y el grado de correspondencia entre conceptos, que conforman el enriquecimiento de las tareas, los cuales se derivan de algunos mecanismos comúnmente empleados en procesos de descubrimiento y recuperación de servicios web semánticos. Estos mecanismos, han sido ajustados y aplicados al ámbito de la comparación de tareas, constituyéndose en la base fundamental, para la formulación de los algoritmos, que determinan la similitud semántica general entre tareas a partir de la comparación de sus identificadores, entradas y salidas.

4.2. DISTANCIA Y EMPAREJAMIENTO SEMÁNTICOS

Como requisito fundamental, para asegurar la fiabilidad en los resultados de comparación de tareas de procesos de negocio, es indispensable garantizar, una correcta definición de la correspondencia entre el enriquecimiento de las tareas, y por ende de la similitud entre las mismas. En este sentido, se hace necesario contar con que dicho enriquecimiento haya sido asignado previamente de manera apropiada, es decir, que los conceptos ontológicos se hayan asociado fielmente a los atributos que definen la funcionalidad de las tareas (identificadores, entradas y salidas). Una vez verificado este requisito, es esencial definir, en primer lugar, la distancia semántica existente entre los conceptos que conforman dicho enriquecimiento, puesto que de ella depende la clasificación de las relaciones en categorías, o niveles de correspondencia y la posterior determinación de la similitud semántica, entre los parámetros que definen la funcionalidad de las tareas.

A continuación, se presenta una descripción detallada de los algoritmos empleados para definir la distancia y la correspondencia semánticas, entre los conceptos del enriquecimiento de las tareas comparadas, tanto a nivel de identificadores, como de entradas y salidas, basados en los trabajos de investigación mencionados en el capítulo 2.

4.2.1. Distancia y Emparejamiento entre Entradas/Salidas de las Tareas

4.2.1.1. Distancia Semántica

La distancia semántica, en el marco del presente proyecto, puede ser entendida como una medida que refleja el grado de diferencia en el significado o sentido de dos conceptos pertenecientes a la misma ontología de dominio.

Para calcular la distancia semántica, en el caso de las entradas y salidas, se debe tener en cuenta que su enriquecimiento es realizado con conceptos de la ontología SSID, los cuales están distribuidos en una jerarquía de 10 niveles. En este sentido, es necesario definir unos criterios para garantizar que el valor de la distancia, entre los conceptos comparados, refleje efectivamente la información en cuanto a la diferencia del significado implícita en la jerarquía de la ontología (Al-Mubaid, et al., 2006). Estos criterios se enumeran a continuación:

1. La diferencia semántica entre conceptos de niveles superiores, es mayor que la de conceptos de niveles inferiores en la ontología de dominio, en otras palabras, dos conceptos generales son menos similares que dos conceptos especializados.
2. La distancia entre conceptos hermanos (horizontalmente adyacentes) es mayor que entre conceptos padre e hijo (verticalmente adyacentes).
3. La profundidad del concepto raíz es cero, y la profundidad de los otros conceptos es igual a la longitud de su trayectoria (camino) hasta el nodo raíz.

Teniendo como referencia estos criterios, se determinó el siguiente procedimiento para calcular la distancia semántica entre dos conceptos del enriquecimiento de dos entradas/salidas, de las tareas de consulta²¹ y publicada²²:

En primer lugar, se debe calcular la *Distancia por Salto* (D_{salto}), en cada uno de los niveles de la ontología que separan a los dos conceptos comparados, lo cual se realiza tomando un concepto de cada nivel y calculando su profundidad (*depth*), para posteriormente aplicar la siguiente fórmula (Ge, et al., 2008):

$$D_{\text{salto}} = 1 + \frac{1}{2^{\text{depth}}} \quad \text{Ec. 1.}$$

Donde *depth*, es el número de saltos en la ontología que hay desde el *concepto raíz*²³ hasta el *concepto objetivo*²⁴. De esta manera, D_{salto} se calcula para cada concepto, presente en el trayecto entre el concepto del enriquecimiento de la tarea de consulta y el concepto del enriquecimiento de la tarea publicada.

²¹ Se refiere a una tarea perteneciente a un proceso de negocio que el operador desea implementar.

²² Hace referencia a una tarea perteneciente a un proceso de negocio disponible e implementado dentro del operador.

²³ Es el primer concepto en la jerarquía de la ontología a partir del cual se ramifican los demás conceptos.

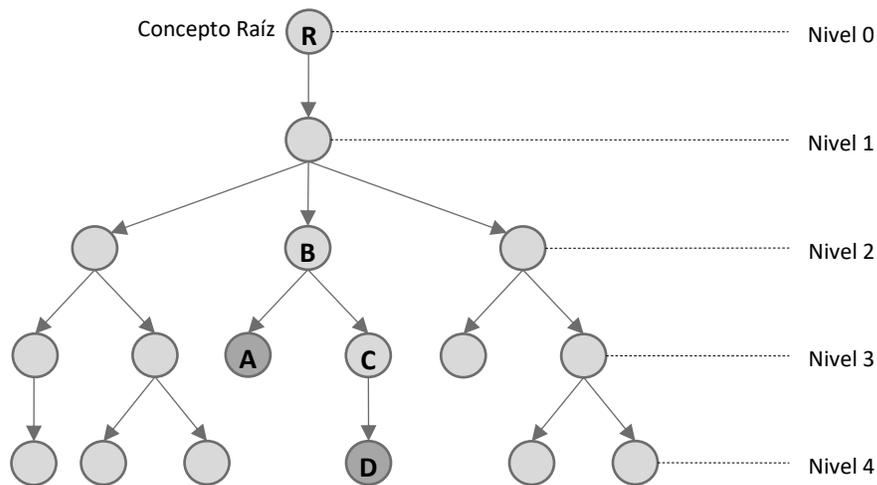
²⁴ Entre los dos conceptos de cada trayecto, el concepto objetivo es el que está más alejado del nodo raíz.

Posteriormente, se calcula la Distancia Semántica (SD) entre los conceptos comparados a través de la sumatoria de los valores de D_{salto} , asignados a cada uno de los trayectos que los separa (Ge, et al., 2008):

$$SD = \sum D_{\text{salto}} \quad \text{Ec. 2.}$$

Para ilustrar mejor este procedimiento, considere un ejemplo donde se desea calcular la distancia semántica entre una entrada de la tarea de consulta, enriquecida mediante el concepto A, y una entrada de la tarea publicada enriquecida con el concepto D, de la ontología mostrada en la Figura 15.

Figura 15. Ontología de dominio



De acuerdo a lo descrito anteriormente, es necesario en primera instancia calcular las distancias (D_{salto}) que hay entre los segmentos A-B, B-C y C-D, para posteriormente, sumarlos y obtener la Distancia Semántica total entre A y D. Estas distancias por salto se obtienen de la siguiente manera:

- Para el segmento A-B: el concepto objetivo es A, dado que es el más alejado del nodo raíz, por lo cual la profundidad es 3, ya que corresponde al número de saltos que hay desde A a R, es decir $\text{depth}_A = 3$. Aplicando la ecuación 1 se tiene:

$$D_{\text{salto (A-B)}} = 1 + \frac{1}{2^3} = 1.125$$

- Para el segmento B-C: el concepto objetivo es C, por lo cual la profundidad es el número de saltos que hay desde C a R, es decir $\text{depth}_C = 3$. Aplicando la ecuación 1:

$$D_{\text{salto (B-C)}} = 1 + \frac{1}{2^3} = 1.125$$

- Para el segmento C-D: el concepto objetivo es D, por lo cual $\text{depth}_D = 4$. Aplicando la ecuación 1 se tiene:

$$D_{\text{salto (C-D)}} = 1 + \frac{1}{2^4} = 1.0625$$

Finalmente, la Distancia Semantica entre A y D, se obtiene sumando los valores de estos segmentos (ecuación 2), así:

$$SD = \sum D_{\text{salto}} = D_{\text{salto (A-B)}} + D_{\text{salto (B-C)}} + D_{\text{salto (C-D)}} = 3.3125$$

Mediante este ejemplo, es posible comprobar el cumplimiento de cada uno de los criterios anteriormente definidos, teniendo en cuenta que: la distancia entre el segmento A-B es mayor que la del segmento C-D (criterio 1), la distancia entre A-C es mayor que entre A-B o B-C (criterio 2), y finalmente, los valores de profundidad, para el concepto raíz y los nodos involucrados en la comparación son: $\text{depth}_R = 0$, $\text{depth}_A = 3$, $\text{depth}_C = 3$ y $\text{depth}_D = 4$ (criterio 3).

La forma como se implementó este procedimiento dentro del prototipo, fue a través de la definición de un mecanismo, para calcular la distancia semántica entre dos conceptos pertenecientes a la misma ontología de dominio (SSID), dicho mecanismo, consiste en la ejecución de consultas sobre la ontología, para encontrar los superconceptos²⁵ de los dos términos comparados. Para llevar a cabo estas consultas (las cuales están definidas en sintaxis wsml) se emplea el razonador WSML2Reasoner (Grimm, et al., 2006) soportado sobre un motor de inferencia de Datalog²⁶, denominado IRIS (Bishop, et al., 2008).

Como resultado de las consultas realizadas, el razonador arroja un conjunto de superconceptos, por cada concepto comparado, el cual se almacena en un arreglo de cadenas de caracteres (strings). Posteriormente, los arreglos obtenidos son analizados, para determinar la distancia semántica entre los dos conceptos comparados. Sin embargo, para proceder con el análisis, es necesario tener en cuenta que el razonador no entrega los resultados de la consulta de superconceptos en un orden coherente, por lo que es necesario ordenar los conceptos contenidos en estos arreglos en forma ascendente (es decir desde el término más específico al más general), de acuerdo con la disposición establecida en la ontología SSID. De esta forma, el concepto comparado siempre queda en el primer campo del arreglo, seguido por el concepto inmediatamente superior y así sucesivamente, hasta llegar al concepto más general, perteneciente a la misma "línea de ascendencia" del concepto comparado.

Este mecanismo se realiza con el fin de facilitar la comparación de los dos conceptos, correspondientes al enriquecimiento semántico de las tareas de consulta y publicada, además, como posteriormente se verá, este proceso se constituye en el punto de partida para determinar también, el grado de correspondencia y la similitud semántica entre los dos conceptos.

²⁵ Se refiere a los conceptos que se encuentran por encima del concepto a comparar y en la misma línea en la ontología de dominio.

²⁶ Datalog es un lenguaje de consulta para bases de datos deductivas.

Tomando como referencia este mecanismo y el procedimiento explicado anteriormente, se diseñó un algoritmo, para calcular la distancia semántica entre dos conceptos del enriquecimiento, asociado a parejas de entradas/salidas de las tareas comparadas (Algoritmo 1). El algoritmo empieza por evaluar, si el concepto del enriquecimiento de la tarea de consulta (C_Q), es diferente al concepto de la tarea publicada (C_T), ya que si esta condición no se cumple, significa que los dos enriquecimientos corresponden exactamente al mismo concepto, por lo cual la distancia entre ellos toma el valor de 0.

En el caso en que C_Q y C_T sean distintos, se procede a encontrar sus superconceptos (denotados por SC_Q y SC_T respectivamente) y el concepto que es común en estos dos arreglos (CC), esto se realiza con el fin de definir el trayecto más corto entre los dos conceptos comparados. Posteriormente, se calcula la sumatoria de todas las distancias por salto que hay entre CC y C_Q (en el arreglo SC_Q) y entre CC y C_T (en el arreglo SC_T). Por último, se suma estos dos valores para obtener la Distancia Semántica total (SD) entre los dos conceptos.

Algoritmo 1. Calcular Distancia Semántica entre parejas de entradas/salidas

```

procedure findSemanticDistance( $C_Q$ ,  $C_T$ )
   $SD := 0$ 
  if  $C_Q \neq C_T$  then
    begin
       $hops := 0$ 
      { $hops$  is the number of hops in the domain ontology that there are between  $C_Q$  and  $C_T$ }
       $SC_Q [ ] := getSuperConcepts(C_Q)$ 
       $SC_T [ ] := getSuperConcepts(C_T)$ 
       $CC := findCommonConcept(SC_Q, SC_T)$ 
      { $CC$  is the closest common concept between  $SC_Q$  and  $SC_T$  in the domain ontology}
       $dist1 := 0$ 
      { $dist1$  is the number of hops between  $C_Q$  and  $CC$  in the domain ontology}
       $dist2 := 0$ 
      { $dist2$  is the number of hops between  $C_T$  and  $CC$  in the domain ontology}
      for  $j := 1$  to  $indexOf(CC, SC_Q)$ 
        { $indexOf(CC, SC_Q)$  returns the index value for  $CC$  in  $SC_Q$ }
         $dist1 := dist1 + 1 + 1/(2^{(size(SC_Q)-j)})$ 
      for  $j := 1$  to  $indexOf(CC, SC_T)$ 
         $dist2 := dist2 + 1 + 1/(2^{(size(SC_T)-j)})$ 
       $SD := dist1 + dist2$ 
    end

```

4.2.1.2. Emparejamiento Semántico

A partir de los trabajos presentados en el capítulo 2, y teniendo en cuenta el mecanismo para calcular la distancia semántica, explicado anteriormente, se definieron 5 categorías para clasificar, el grado de correspondencia entre los conceptos del enriquecimiento de las entradas/salidas de las tareas comparadas (segundo enfoque de emparejamiento):

- **Exact:** Cuando los dos términos del enriquecimiento comparados, pertenecen al mismo concepto en la ontología SSID.
- **Plugin:** Cuando el concepto del enriquecimiento, de la tarea publicada, pertenece al conjunto de superconceptos del enriquecimiento, de la tarea de consulta, y la separación entre conceptos es menor a dos saltos en la ontología de dominio.
- **Subsume:** Cuando el concepto del enriquecimiento, de la tarea de consulta, pertenece al conjunto de superconceptos del enriquecimiento, de la tarea publicada, y la separación entre conceptos es menor a cuatro saltos en la ontología de dominio.
- **Intersect:** Cuando los conceptos del enriquecimiento semántico de las tareas de consulta y publicada, tienen superconceptos comunes, y la separación entre conceptos es menor a cuatro saltos en la ontología de dominio.
- **Fail:** Cuando la separación entre conceptos es mayor a cuatro saltos en la ontología de dominio (una distancia semántica de cuatro implica una disimilitud semántica considerable, teniendo en cuenta que la ontología de SID tiene diez niveles de jerarquía).

Tomando como referencia estas categorías, se diseñó un algoritmo para determinar o clasificar, la relación entre parejas de entradas/salidas de tareas comparadas (Algoritmo 2). De acuerdo con este algoritmo, cuando el concepto de la tarea de consulta (C_Q) sea igual al concepto de la tarea publicada (C_T), el grado de correspondencia es establecido como *exact*. De lo contrario, se procede a calcular la distancia semántica entre los dos conceptos C_Q y C_T , si este valor resulta mayor que cuatro, el grado de correspondencia se define como *fail*.

Si ninguna de las anteriores condiciones se cumple, es necesario encontrar los superconceptos de C_Q y C_T (los cuales se denominan SC_Q y SC_T respectivamente). Luego se evalúa, si el concepto de la tarea publicada pertenece al conjunto de superconceptos del concepto de la tarea de consulta (es decir si C_T está contenido en SC_Q) y se verifica si la distancia semántica entre C_Q y C_T es menor que 2, si ambas condiciones se cumplen se define el grado de correspondencia como *plugin*. En caso de que estas condiciones tampoco se cumplan, se procede a examinar, si el concepto de la tarea de consulta es superconcepto del concepto de la tarea publicada (C_Q está contenido en SC_T), si el resultado es positivo el grado de correspondencia se determina como *subsume*.

De no cumplirse las condiciones para las cuales el grado de correspondencia sea *exact*, *plugin*, *subsume* o *fail*, el grado de correspondencia se define como *intersect*.

Algoritmo 2. Determinar el Grado de Correspondencia entre parejas de entradas/salidas

```

procedure SemanticMatchmaking( $C_Q, C_T$ )
if  $C_Q = C_T$  then matchingDegree := exact
else
begin

```

```

SD := findSemanticDistance(CQ,CT)
if SD < 4 then
begin
  SCQ := getSuperConcepts(CQ)
  {SCQ is a set of superconcepts of CQ}
  SCT := getSuperConcepts(CT)
  {SCT is a set of superconcepts of CT}
  if (CT ⊆ SCQ and SD < 2) then matchingDegree := plugin
  else if CQ ⊆ SCT then matchingDegree := subsume
  else matchingDegree := intersect
  end
else matchingDegree := fail
end
{the relation between CQ and CT is defined by matchingDegree}

```

4.2.2. Distancia y Emparejamiento entre Identificadores de las Tareas

Determinar la distancia semántica y el grado de correspondencia, entre conceptos del enriquecimiento de identificadores de las tareas, requiere de procesos diferentes a los utilizados para las entradas y salidas, dado que la ontología de dominio, utilizada para enriquecer dichos identificadores (SeTOM) no está organizada en niveles de jerarquía. En lugar de esto, SeTOM proporciona un conjunto de axiomas, que permiten inferir sobre otras relaciones existentes entre los conceptos de la ontología, por ejemplo respecto al agrupamiento según las áreas funcionales (niveles) del modelo eTOM (sección 3.4.1.). De esta manera, mediante la utilización de algunos de estos axiomas, fue posible establecer los mecanismos para determinar la distancia semántica y la correspondencia entre conceptos.

4.2.2.1. Distancia Semántica

La distancia semántica, entre conceptos pertenecientes a la ontología SeTOM, se calculó mediante el análisis de los elementos que conforman los conceptos del nivel 2, del modelo eTOM. Se escogió este nivel como referencia, puesto que los conceptos que hacen parte de él, proporcionan una visión intermedia de las operaciones que se llevan a cabo dentro de la empresa, por debajo de la perspectiva gerencial de los procesos otorgada por el nivel 1, y por encima del nivel de detalle de las actividades que los componen dada por el nivel 3 (razón por la cual, la mayoría de identificadores de las tareas fueron enriquecidos con conceptos de nivel 2 y 3).

De esta manera, se pudo determinar que la cantidad de componentes de nivel 3 que integran uno de nivel 2, da cuenta de la distribución del trabajo que se lleva a cabo por parte del componente superior. En este sentido, asumiendo que la contribución de los conceptos componentes (pertenecientes al nivel 3) de un concepto de nivel 2 (concepto contenedor) es equivalente, es decir, la funcionalidad total de una operación determinada (identificada por un concepto de nivel 2), está distribuida equitativamente entre cada una de las suboperaciones que la componen (identificadas por conceptos de nivel 3), puede entonces relacionarse, la distancia que separa un concepto componente de un concepto

contenedor, como el número de suboperaciones (conceptos componentes) necesarias para llevar a cabo la operación identificada por el concepto contenedor. De esta manera, entre menor sea el número de suboperaciones, que conforma un concepto contenedor, menor será la distancia semántica entre este último y sus conceptos componentes. Así mismo, la distancia entre conceptos componentes depende de su cantidad, en otras palabras entre mayor sea el número de elementos, que componen un concepto de nivel superior, menor será su dependencia funcional y por tanto mayor su distancia semántica. Para ilustrar lo anterior, considere la Figura 16, en la cual se tiene un concepto contenedor compuesto por cinco conceptos componentes A, B, C, D y E; en este caso, la distancia semántica (entre contenedor y componentes, así como entre componentes) es equivalente a cinco (5).

Figura 16. Distancia Semántica entre Identificadores.



Teniendo en cuenta lo anterior, el algoritmo propuesto que permite encontrar la distancia semántica, entre conceptos del enriquecimiento de identificadores (Algoritmo 3), comienza por evaluar, si el concepto del enriquecimiento del identificador de la tarea de consulta (I_Q), es diferente al concepto de la tarea publicada (I_T), ya que si esta condición no se cumple significa que los dos enriquecimientos corresponden al mismo concepto, en la ontología SeTOM, por lo cual la distancia entre ellos toma el valor de 0.

En el caso en que I_Q e I_T , sean distintos se procede a encontrar el nivel de eTOM al que pertenecen, utilizando los axiomas de descomposición y agrupamiento de la ontología (estos niveles se denotan como L_{I_Q} y L_{I_T}). Posteriormente, se verifica que tanto I_Q como I_T pertenezcan a los niveles 2 o 3 de la misma área funcional, con lo cual se define que la Distancia Semántica (SD) entre ellos es igual al número de componentes de nivel 3 que integran al de nivel 2. De no cumplirse las anteriores condiciones, la distancia semántica es fijada en un valor lo suficientemente grande, para reflejar que no hay semejanza entre los dos conceptos comparados.

Algoritmo 3. Calcular Distancia Semántica entre parejas de identificadores

```

procedure findIDSemanticDistance( $I_Q$ ,  $I_T$ )
 $SD := 0$ 
if  $I_Q \neq I_T$  then
begin
     $L_{I_Q} := getLevel(I_Q)$ 
     $L_{I_T} := getLevel(I_T)$ 
    if ( $(L_{I_Q} \neq 2$  and  $L_{I_Q} \neq 3)$  or ( $L_{I_T} \neq 2$  and  $L_{I_T} \neq 3$ )) then  $SD := \infty$ 

```

```

else
begin
  SLCQ := getSecondLevelConcept(IQ)
  {SLCQ is the second level concept of the IQ}
  SLCT := getSecondLevelConcept(IT)
  {SLCT is the second level concept of the IT}
  if (SLCQ = SLCT) then SD := getNumComponents(SLCQ)
  else SD := ∞
end
end

```

4.2.2.2. *Emparejamiento Semántico*

Teniendo en cuenta el segundo enfoque de emparejamiento y el análisis realizado a los niveles 2 y 3 del modelo eTOM, se definieron 5 categorías, en las que se puede clasificar la correspondencia entre identificadores:

- **Exact:** Cuando los dos términos del enriquecimiento comparados, pertenecen al mismo concepto en la ontología SeTOM.
- **Plugin:** Para clasificar el grado de correspondencia, en esta categoría, es necesario cumplir 3 condiciones:
 1. El concepto del enriquecimiento de la tarea de consulta debe pertenecer al nivel 3 del modelo eTOM.
 2. El concepto del enriquecimiento de la tarea publicada debe pertenecer al nivel 2 del modelo eTOM.
 3. El concepto del enriquecimiento de la tarea de consulta debe hacer parte del conjunto de componentes que conforman el concepto del enriquecimiento de la tarea publicada.
- **Subsume:** Para clasificar el grado de correspondencia en esta categoría es necesario cumplir 3 condiciones:
 1. El enriquecimiento de la tarea publicada, debe pertenecer al nivel 3 de eTOM.
 2. El enriquecimiento de la tarea de consulta, debe pertenecer al nivel 2 de eTOM.
 3. El enriquecimiento de la tarea de consulta, debe hacer parte del conjunto de componentes que conforman el enriquecimiento de la tarea publicada.
- **Intersect:** Cuando los conceptos del enriquecimiento semántico de las tareas de consulta y publicada son de nivel 3, y hacen parte del mismo componente de nivel 2.
- **Fail:** Cuando ninguna de las condiciones anteriores se cumple.

Tomando como referencia estas categorías, se diseñó un algoritmo para clasificar el grado de correspondencia entre parejas de identificadores de tareas comparadas (Algoritmo 4). Según este algoritmo, cuando el concepto de la tarea de consulta (I_Q), es igual al concepto de la tarea publicada (I_T), el grado de correspondencia es establecido

como *exact*. De lo contrario se procede a encontrar el nivel al que pertenecen dichos conceptos (L_{I_Q} y L_{I_T}), y se verifica que tanto I_Q como I_T pertenezcan a los niveles 2 o 3, de la misma área funcional. Si esto se cumple se evalúa si tanto el concepto de la tarea de consulta (I_Q), como el de la tarea publicada (I_T), pertenecen al nivel 3, con lo cual la correspondencia se establece como *intersect*. De no ser así, se evalúa si el concepto de la tarea publicada (I_T) pertenece a nivel 2 determinando la correspondencia como *plugin*, o si el concepto de la tarea de consulta (I_Q) pertenece a nivel 2, definiendo la correspondencia como *subsume*.

De no cumplirse las condiciones para las cuales el grado de correspondencia sea *exact*, *plugin*, *subsume* o *intersect*, el grado de correspondencia se define como *fail*.

Algoritmo 4. Determinar el Grado de Correspondencia entre parejas de identificadores

```

procedure SemanticMatchmaking ( $I_Q, I_T$ )
if  $I_Q = I_T$  then matchingDegree := exact
else
begin
     $L_{I_Q}$  := getLevel( $I_Q$ )
     $L_{I_T}$  := getLevel( $I_T$ )
    if (( $L_{I_Q} \neq 2$  and  $L_{I_Q} \neq 3$ ) or ( $L_{I_T} \neq 2$  and  $L_{I_T} \neq 3$ )) then matchingDegree := fail
    else
    begin
         $SLC_Q$  := getSecondLevelConcept( $I_Q$ )
        { $SLC_Q$  is the second level concept of the  $I_Q$ }
         $SLC_T$  := getSecondLevelConcept( $I_T$ )
        { $SLC_T$  is the second level concept of the  $I_T$ }
        if ( $SLC_Q = SLC_T$ ) then
        if ( $L_{I_Q} = 3$  and  $L_{I_T} = 3$ ) then matchingDegree := intersect
        else if ( $L_{I_T} = 2$ ) then matchingDegree := plugin
        else matchingDegree := subsume
        else
        matchingDegree := fail
    end
    end
end
{the relation between  $C_Q$  and  $C_T$  is defined by matchingDegree}

```

4.3. SIMILITUD SEMÁNTICA

4.3.1. Similitud Semántica entre Conceptos

Una vez establecidas la Distancia Semántica y la categoría en que se clasifica la correspondencia, entre el enriquecimiento de las tareas comparadas (sección 4.2.), es

necesario ponderar estas relaciones, para determinar la semejanza existente entre dichos conceptos. Esto se realiza mediante la definición de la Similitud Semántica, la cual se refiere, al grado de proximidad que hay entre dos conceptos dentro de una ontología (d'Amato, 2007) (Inkpen, et al., 2005), que para el caso del presente proyecto, corresponde a una ontología del dominio específico de las telecomunicaciones.

De esta manera, tomando como referencia, los trabajos de investigación descritos en la sección 2.3.1.2. de esta monografía, se definieron algunos criterios a tener en cuenta a la hora de calcular la similitud semántica entre dos conceptos:

- La similitud semántica es inversamente proporcional a la distancia semántica.
- El valor de similitud semántica está en el rango de cero a uno.
- Cuando la distancia semántica es máxima, la similitud semántica debe ser igual a CERO.
- Cuando la distancia semántica es mínima (cero), la similitud semántica debe ser igual a UNO.

Según lo expuesto en dicha sección, existen muchos métodos que permiten garantizar estos criterios, al medir la similitud semántica entre dos conceptos, los cuales pueden ser agrupados dentro de dos enfoques, de los cuales, el que más se aproxima a los resultados de la opinión de los expertos de un dominio, es aquel que está basado enteramente en la ontología, en tanto estas representaciones de la información o el conocimiento, permiten establecer relaciones conceptuales más complejas y reducen la ambigüedad de los resultados obtenidos.

Por lo tanto, el mecanismo establecido para calcular la similitud semántica entre dos conceptos, se definió a partir de sus relaciones, en cuanto a distancia semántica y grado de correspondencia dados según la ontología. Matemáticamente, este mecanismo está representado por la siguiente fórmula (Ge, et al., 2008):

$$SS = \frac{1}{p \cdot SD + 1} \quad \text{Ec. 3.}$$

Donde SD, es la medida de la Distancia Semántica entre los dos conceptos en la ontología de dominio, y p es un factor que permite condicionar el impacto de la distancia semántica en el valor de la similitud, dependiendo del grado de correspondencia que hay entre los dos conceptos comparados. De esta forma se asignaron los valores de p de la siguiente manera (teniendo en cuenta, que a mayor correspondencia, el valor de similitud debe ser mayor): si el grado de correspondencia es *exact* el valor de p es 0, si es *plugin* p = 0.2, si es *subsume* p = 0.6, si es *intersect* p = 0.8 y si es *fail* p = 1.

El factor p surgió de la necesidad de diferenciar la similitud entre dos conceptos, cuando el valor de la distancia semántica entre ellos es igual, pero su ubicación en la ontología es diferente, por ejemplo en el caso de dos comparaciones entre dos conceptos C_Q y C_T , donde C_Q sea superconcepto de C_T (el grado de correspondencia es *subsume*) y viceversa (el grado de correspondencia es *plugin*). En este ejemplo el valor de SD es el mismo sin importar el orden jerárquico que hay entre C_Q y C_T , por lo tanto si no se tiene en cuenta el factor p, la similitud semántica sería igual para ambas comparaciones, lo que

se constituiría en un error puesto que SS debería ser mayor, en el caso en que C_T sea superconcepto de C_Q , es decir en el caso en el que C_T proporcione las funcionalidades de C_Q y añada unas nuevas.

Por el contrario, si se tiene en cuenta el factor p , el valor de la similitud cambia drásticamente, ya que si C_Q es superconcepto de C_T (el grado de correspondencia *subsume*), entonces p sería igual a 0.6, mientras que si C_T es superconcepto de C_Q (el grado de correspondencia *plugin*) el factor p se establecería como 0.2, por lo cual al hacer el cálculo de SS, a pesar de que SD sea igual para ambos, el resultado será mayor para este último caso, garantizando que la medida de similitud concuerde con la relación de correspondencia funcional que hay entre los dos conceptos.

Los valores de p fueron asignados, en un principio, de acuerdo a lo que en teoría representa el grado de correspondencia, sin embargo, después de realizar las pruebas del prototipo (descritas en el capítulo 5), se determinó que para optimizar el desempeño del sistema, estos valores debían ser ajustados según la ontología de dominio utilizada en el enriquecimiento de las tareas (Tabla 1).

A partir de este mecanismo para calcular la similitud semántica entre conceptos del enriquecimiento se definió el Algoritmo 5, el cual recibe como entradas, los dos conceptos a comparar (C_Q y C_T) y la Ontología de Dominio a la que pertenecen dichos conceptos (DO).

Tabla 1. Valores de p según la ontología de dominio.

	exact	plugin	subsume	intersect	fail
SeTOM	0	0.2	0.5	0.9	1
SSID	0	0.7	0.9	1	1

En primer lugar, el algoritmo evalúa si el grado de correspondencia es *exact*, caso en el cual hace $p = 0$, con lo que la similitud semántica (SS) se establece en 1. De lo contrario, calcula la distancia semántica (denotada por SD) y el grado de correspondencia entre C_Q y C_T (MD). Luego, de acuerdo con el resultado obtenido en MD y dependiendo de la ontología registrada en DO, ajusta el parámetro p que interviene en el cálculo de la similitud semántica junto con SD.

Algoritmo 5. Determinar la Similitud Semántica entre Conceptos

procedure *findSemanticSimilarity* (C_Q, C_T, DO)

{ C_Q is the enrichment concept of the query task, C_T is the enrichment concept of the target task and DO is de domain ontology}

$p := 0$

{ p is a factor that depends on the matching degree}

$SD := 0$

{ SD is the semantic distance value between C_Q and C_T }

```

MD := findMatchingDegree (CQ, CT)
{MD is the matching degree between CQ and CT}
if MD ≠ exact then
begin
  SD:= findSemanticDistance (CQ, CT)
  if MD = plugin then
    if DO = SSID then p := 0.7
    else p := 0.2
  else if MD = subsume then
    if DO = SSID then p := 0.9
    else p := 0.5
  else if MD = intersect then
    if DO = SSID then p := 0.1
    else p := 0.9
  else p := 1
end
SS := 1/(p*SD + 1)
{SS is the semantic similarity value between CQ and CT}

```

4.3.2. Similitud Semántica entre Tareas

Los algoritmos propuestos anteriormente, son la base fundamental para determinar la Similitud Semántica entre tareas de procesos de negocio de telecomunicaciones, dado que los parámetros, que definen la funcionalidad de dichas tareas, están enriquecidos con conceptos dispuestos en ontologías de dominio. En las siguientes secciones, se explican los procedimientos utilizados para encontrar los valores de similitud para cada parámetro o atributo de las tareas comparadas (identificador, entradas y salidas), y finalmente se explica, cómo a partir de estos valores se determina la similitud total entre las mismas.

4.3.2.1. Similitud Semántica de Identificadores

Dado que las tareas de procesos de negocio de telecomunicaciones poseen un sólo identificador, el cual está enriquecido con un único concepto de la ontología SeTOM, el mecanismo para determinar la Similitud Semántica de los identificadores de dos tareas comparadas (tarea de consulta y tarea publicada), corresponde al mismo método desarrollado para calcular la similitud semántica entre conceptos (Algoritmo 5).

4.3.2.2. Similitud Semántica de Entradas y Salidas

Debido a que las tareas comparadas pueden contener varias entradas y varias salidas, el mecanismo para calcular la Similitud Semántica respecto a estos parámetros, es un poco más complejo que en el caso de los identificadores. Este problema de comparación de las entradas/salidas de las tareas, se aborda mediante una adaptación del mecanismo propuesto en (Benatallah, et al., 2003), en el cual a partir de una petición de servicio, se pretende encontrar la combinación de servicios que “mejor cubra” las capacidades requeridas. Dicha combinación de servicios debe compartir el mayor número posible de

entradas y salidas con el servicio solicitado, sin exceder, en lo posible, las entradas del mismo.

En cuanto a la comparación de tareas de dos procesos de negocio, la adaptación de este enfoque consiste en determinar la tarea publicada (T_T) (NO la combinación de tareas), más similar a una tarea de consulta (T_Q), teniendo en cuenta las siguientes condiciones:

1. T_T debe compartir el mayor número posible de salidas con T_Q .
2. T_T debe compartir y, en lo posible, no exceder las entradas de T_Q .

Decir que T_T comparte una entrada/salida con T_Q , no implica necesariamente que dicha entrada/salida sea la misma para ambas tareas, es decir, estas entradas/salidas pueden ser sintácticamente diferentes, pero semánticamente similares.

En este sentido, el mecanismo propuesto para calcular la similitud semántica de las entradas/salidas de dos tareas, consiste en evaluar en primera instancia las similitudes individuales entre parejas de entradas/salidas, y posteriormente analizar esos resultados, para cumplir con las dos condiciones establecidas. Este mecanismo se explica a continuación, para el caso de las entradas, considerando que para las salidas se aplica el mismo procedimiento, pero con algunas variaciones que serán definidas posteriormente.

La similitud entre parejas de entradas, de la tarea de consulta y la tarea publicada, se calcula, a partir del mecanismo para la comparación de conceptos explicado en el Algoritmo 5. Posteriormente, estos resultados se organizan en una matriz de similitud, cuyas filas corresponden a las entradas de la tarea de consulta y las columnas a las entradas de la tarea publicada.

Esta matriz, es analizada para determinar las *similitudes máximas* entre parejas de entradas, de las dos tareas comparadas, las cuales hacen referencia a los mayores valores de similitud, que pueden obtenerse entre todas las posibles combinaciones de parejas de entradas (una entrada de la tarea de consulta y otra de la publicada), garantizando que cada entrada, de la tarea de consulta, está relacionada con una única entrada de la tarea publicada y viceversa.

De esta manera, si se define S como la matriz de similitud de entradas de dimensión $m \times n$ (donde m hace referencia a las entradas de la tarea de consulta y n a las de la tarea publicada), se determina el conjunto de *similitudes máximas* como la agrupación de similitudes $S_{i,j}$ que cumplen con la siguiente propiedad:

$$S_{i,j} : \forall S_{k,j} (k = 1, 2, \dots, m) S_{i,j} > S_{k,j} \wedge \forall S_{i,l} (l = 1, 2, \dots, n) S_{i,j} > S_{i,l}$$

Una vez definido este conjunto de *similitudes máximas*, se procede a encontrar las entradas que quedaron aisladas de las parejas establecidas en estas relaciones (entradas perdidas), las cuales se presentan cuando el número de entradas totales de la tarea de consulta (m) excede a las de la tarea publicada (n) o viceversa.

A partir de la especificación de estos parámetros, y teniendo en cuenta las condiciones del análisis de cobertura descritas al inicio de esta sección, se define la Similitud Semántica entre entradas o salidas (SIOS) como:

$$SIOS = \frac{\sum S_{i,j}}{\min(m,n) + \beta * |m-n|} \quad \text{Ec. 4.}$$

Donde:

- $S_{i,j}$ son las *similitudes máximas*
- $\min(m,n)$ es el máximo número de parejas de entradas, que puede establecerse entre las dos tareas comparadas según la propiedad de las $S_{i,j}$.
- $\beta * |m-n|$ es un parámetro que permite condicionar el valor de la similitud cuando existen entradas perdidas.

En esta ecuación, la relación $\sum S_{i,j} / \min(m,n)$, refleja la máxima proporción de similitudes de entrada, que puede encontrarse al relacionar cada una de las entradas de la tarea de consulta, con una única entrada de la tarea publicada. Sin embargo esta relación no permite cumplir la condición 2, del análisis de cobertura, por lo tanto, se agregó el parámetro $\beta * |m-n|$, para controlar la influencia del número de entradas de las tareas comparadas, en el valor de la similitud promedio de entradas. De esta manera, cuando la tarea de consulta tiene más entradas que la tarea publicada se cumple la condición 2 por lo tanto β toma el valor de 0.8, de lo contrario significa que la tarea publicada excede el número de entradas de la tarea a de consulta, por lo cual se hace $\beta = 1$.

A partir de este mecanismo, se definió un algoritmo para calcular la similitud semántica de entradas (Algoritmo 6), el cual recibe como parámetros las dos tareas a comparar (T_Q , T_T). El algoritmo opera de la siguiente forma: en primer lugar extrae las entradas de la tarea de consulta en el arreglo $I_Q []$ y asigna a m el número de entradas de dicha tarea. De la misma manera extrae las entradas de la tarea publicada para almacenarlas en $I_T []$ y guarda el tamaño de este arreglo en n . Posteriormente, obtiene el enriquecimiento de las entradas de la tarea de consulta en el arreglo $IE_Q[m]$, y de la tarea publicada en el arreglo $IE_T[n]$.

Con el fin de verificar si existen entradas a comparar, se evalúa si los arreglos $I_Q []$ e $I_T []$ están vacíos, es decir, si ninguna de las tareas comparadas tiene entradas, con lo cual se establece que la similitud de entradas (S/S) es igual a 1. De lo contrario, se examina si alguno de ellos está vacío (si cualquiera de las dos tareas no tiene entradas), con lo cual S/S se hace igual a 0, puesto que si alguna de las tareas tuviera entradas y la otra no, no sería posible realizar la comparación en cuanto a este parámetro.

Si ninguna de las anteriores condiciones se presenta, se procede a obtener la matriz de similitudes de entradas ($ISM [m][n]$), a partir de la comparación de los arreglos de enriquecimiento ($IE_Q [m]$ e $IE_T [n]$). Luego se analiza esta matriz, para obtener las máximas similitudes ($MSA []$).

Una vez reunida toda esta información, se evalúa si la tarea de consulta tiene más entradas que la tarea publicada (con el fin de cumplir con la condición 2 del análisis de cobertura), en cuyo caso se calcula la similitud semántica de entradas a partir de la ecuación 4, teniendo en cuenta que β toma el valor de 0.8. Si la tarea de consulta no excede las entradas de la tarea publicada, se calcula la similitud tomando β igual a 1.

Algoritmo 6. Determinar la Similitud Semántica de Entradas

```
procedure findSemanticInputsSimilarity (TQ, TT)
{ TQ is the query task, TT is the target task }
begin
IQ[ ] := getInputs(TQ)
m := sizeof(IQ[ ])
IT[ ] := getInputs(TT)
n := sizeof(IT[ ])
{ IQ[ ] is the query task inputs array, and IT[ ] is the target task inputs array }
IEQ[m] := getInputsEnrichment(TQ)
IET[n] := getInputsEnrichment(TT)
{ IEQ[ ] is the query task inputs enrichment array, and IET[ ] is the target task inputs
enrichment array }
if (IQ[ ] = ∅ and IT[ ] = ∅) then S/S := 1
else if (IQ[ ] = ∅ or IT[ ] = ∅) then S/S := 0
  else
  begin
  ISM [m][n] := getInputSimMatrix(IEQ[m], IET[n])
  { ISM [m][n] is the inputs similarities matrix }
  MSA [ ] := getMaxSimsArray(ISM [ ][ ])
  { MSA [ ] is the maximum similarities array }
  MSS := addMaxSims(MSA [ ])
  { MSS is the maximum similarities sum }
  if (m > n) then S/S := MSS / ( n + 0.8 * |m - n| )
  else S/S = MSS / ( m + |m - n| )
  end
end
{ S/S is the semantic inputs similarity value between TQ and TT }
```

Con base en este mecanismo, se definió el algoritmo para calcular la similitud semántica de salidas entre las dos tareas comparadas (Algoritmo 7), el cual funciona de la misma manera que el algoritmo de las entradas. La única variación de este procedimiento respecto al anterior, consiste en modificar el valor de β , para cumplir la condición 1 del análisis de cobertura, de esta manera se definió que cuando la tarea publicada excede el número de salidas de la tarea de consulta, el valor de β es 0.8, en caso contrario β se hace igual a 1, esto permite condicionar el impacto de las salidas perdidas en el valor de similitud.

Algoritmo 7. Determinar la Similitud Semántica de Salidas

```
procedure findSemanticOutputsSimilarity (TQ, TT)
{ TQ is the query task, TT is the target task }
begin
OQ[ ] := getOutputs(TQ)
```

```

m := sizeof(OQ[ ])
OT[ ] := getOutputs(TT)
n := sizeof(OT[ ])
{OQ[ ] is the query task outputs array, and OT[ ] is the target task outputs array}
OEQ[m] := getOutputsEnrichment(TQ)
OET[n] := getOutputsEnrichment(TT)
{OEQ[ ] is the query task outputs enrichment array, and OET[ ] is the target task outputs enrichment array}
if (OQ[ ] = ∅ and OT[ ] = ∅) then SOS := 1
else if (OQ[ ] = ∅ or OT[ ] = ∅) then SOS := 0
    else
    begin
    OSM [m][n] := getOutputSimMatrix(OEQ[m], OET[n])
    {OSM [m][n] is the outputs similarities matrix}
    MSA [ ] := getMaxSimsArray(OSM [ ][ ])
    {MSA [ ] is the maximum similarities array}
    MSS := addMaxSims(MSA [ ])
    {MSS is the maximum similarities sum}
    if (m>n) then SOS := MSS / ( n +|m-n|)
    else SOS = MSS / ( m + 0.8*|m-n|)
    end
end
{SOS is the semantic outputs similarity value between TQ and TT}

```

4.3.2.3. Similitud Semántica General de Tareas

Por último, una vez obtenidos los valores de similitud de identificadores, entradas y salidas, mediante los algoritmos explicados anteriormente, se procede a calcular la similitud semántica general (OSS), entre las dos tareas comparadas, a través de una suma ponderada de dichos resultados:

$$OSS = SS * wId + SIS * wIn + SOS * wO \quad \text{Ec. 5.}$$

Donde:

- **SS** es la similitud semántica de identificadores.
- **wId** es el porcentaje que determina la contribución de la similitud de identificadores en la similitud semántica general de las tareas.
- **SIS** es la similitud semántica de entradas.
- **wIn** es el porcentaje que determina la contribución de la similitud de entradas en la similitud semántica general de las tareas.
- **SOS** es la similitud semántica de salidas.
- **wO** es el porcentaje que determina la contribución de la similitud de salidas en la similitud semántica general de las tareas.

Los valores de **wId**, **wIn** y **wO** son proporcionados por los usuarios de acuerdo a su criterio. De acuerdo con los resultados obtenidos en el proceso de evaluación al cuál se

sometió el mecanismo propuesto, los pesos sugeridos (para los cuales se alcanza el mejor rendimiento del sistema) se presentan en el siguiente capítulo.

El mecanismo para calcular la similitud semántica entre dos tareas de procesos de negocio de telecomunicaciones, descrito en este capítulo, puede aplicarse para la comparación entre las tareas del proceso de negocio de consulta, y las tareas de los procesos disponibles en el operador. De esta manera, mediante los valores de similitud, es posible establecer un ranking de tareas publicadas, para cada una de las tareas de consulta, dependiendo de la capacidad que tienen las primeras para cumplir los requerimientos de la solicitud.

4.4. RESUMEN

En este capítulo se describieron los principales algoritmos empleados para determinar la similitud semántica, entre tareas de procesos de negocio de telecomunicaciones, los cuales, se soportan en tres mecanismos fundamentales aplicados sobre la comparación del enriquecimiento de las tareas: la distancia, el emparejamiento y la similitud entre conceptos. Estos mecanismos se basan en algunos procedimientos comúnmente empleados en descubrimiento y recuperación de servicios web semánticos, y han sido adaptados al ámbito de la comparación de tareas, para constituir la base sobre la que se sustentan los algoritmos para calcular la similitud semántica general entre tareas, a partir de la comparación de los principales atributos, que definen su funcionalidad (identificadores, entradas y salidas). La aplicación de este mecanismo, puede simplificar la identificación de tareas disponibles dentro del operador, que pueden adaptarse y reutilizarse en la satisfacción de nuevas necesidades, facilitando así, la optimización del consumo de recursos en el desarrollo de actividades de análisis, gestión y reingeniería de procesos.

5. PROTOTIPO Y EVALUACIÓN

5.1. INTRODUCCIÓN

Tal como se definió en el *Escenario de Motivación*, presentado en el Capítulo 1 de este documento, es posible para un operador de telecomunicaciones, agilizar el despliegue de nuevos servicios, mediante la reutilización de componentes pertenecientes a procesos de negocio, ya implementados dentro de la organización. Sin embargo, la identificación de estos componentes reutilizables puede constituirse en una tarea considerablemente dispendiosa, dependiendo de la cantidad de procesos de negocio disponibles y de que tan ordenada sea la gestión realizada sobre los mismos.

Para superar esta limitación y hacer más eficiente este procedimiento de identificación, en el capítulo anterior, se propuso un mecanismo para llevar a cabo la comparación semántica entre tareas de procesos de negocio, del dominio de las telecomunicaciones descritos en BPMO, teniendo en cuenta para esto, un conjunto representativo de los atributos que definen la funcionalidad de las mismas, esto es: sus entradas, salidas e identificadores.

En este capítulo, se abordará la descripción detallada del trabajo realizado alrededor del desarrollo de un sistema software, que implementa el mecanismo de comparación propuesto. Esta descripción abarcará desde la definición de la arquitectura diseñada para soportar la construcción del prototipo, hasta la especificación del plan de pruebas realizadas sobre el mismo, junto con los respectivos resultados obtenidos.

5.2. DESCRIPCIÓN DEL PROTOTIPO

El prototipo construido, para evaluar el mecanismo propuesto, consiste en una aplicación Java, implementada a partir de la adopción del Modelo de Construcción de Soluciones (Serrano, 2005), como referencia metodológica para soportar el proceso de desarrollo. Las posteriores secciones de este capítulo, se dedican a describir la aplicación de dos macrocomponentes fundamentales de dicho modelo: la Estructura para la Descripción del Sistema y el Modelo del Proceso de Desarrollo. El primer macrocomponente, involucra la especificación de artefactos esenciales, para el entendimiento adecuado de la funcionalidad y el comportamiento esperados del sistema/solución (Modelo de Casos de Uso, Arquitectura de Referencia, Modelo de Diseño (**Anexo A**), Modelo de Despliegue (**Anexo A**) y Plan de Pruebas); mientras que el segundo macrocomponente define cuatro fases de referencia, las cuales, ejecutadas en conjunto de forma iterativa/incremental, orientan el proceso de desarrollo de la solución propuesta.

5.2.1. Modelo de Casos de Uso del Sistema

La definición de los casos de uso del sistema, parte de la identificación de los requisitos funcionales del mismo. A continuación se detallan los requisitos funcionales abstraídos para el sistema implementado:

- i. **Comparar semánticamente, tareas de Procesos de Negocio de Telecomunicaciones, descritos con BPMO.** Teniendo en cuenta un conjunto de los atributos, que definen la funcionalidad de las tareas (Entradas, Salidas e Identificador), el sistema debe realizar una comparación basada en la semántica de los conceptos asociados a dichos atributos.
- ii. **Determinar la similitud semántica entre los atributos de las tareas.** El sistema debe estimar de manera cuantitativa, la similitud existente entre los atributos que definen la funcionalidad de dos tareas.
- iii. **Determinar la similitud semántica entre dos tareas.** A partir del análisis de la funcionalidad de las dos tareas comparadas, el sistema debe cuantificar el nivel de similitud entre ellas.
- iv. **Recuperar las tareas que más se adecúen a una solicitud específica.** Este requisito funcional se refiere, a la capacidad que debe tener el sistema de discriminar de un conjunto de tareas pertenecientes a procesos de negocio disponibles, las tareas semánticamente más similares a una tarea de consulta.
- v. **Configurar parámetros de comparación.** El sistema debe permitir ajustar el peso asignado a los atributo de las tareas a comparar, para ponderar su similitud semántica general.
- vi. **Clasificar tareas disponibles de acuerdo a su nivel de relevancia, respecto a una solicitud determinada.** A partir de la comparación entre una tarea de consulta y un conjunto de tareas pertenecientes a uno o varios procesos de negocio disponibles, el sistema debe clasificar dicho conjunto, de acuerdo al grado de similitud semántica existente entre cada una de sus tareas y la tarea solicitada.

A partir del análisis de los requisitos funcionales del sistema, se definieron ocho casos de uso los cuales se describen a continuación:

Caso de Uso No. 1: Comparar Tareas

Iniciador: Usuario (Agente Externo: Persona, Aplicación)

Propósito: Comparar semánticamente tareas pertenecientes a procesos de negocio de Telecomunicaciones.

Resumen: Este caso de uso satisface el requisito funcional (i), a partir del análisis del enriquecimiento semántico asociado a los parámetros que definen las tareas comparadas.

Caso de Uso No. 2: Recuperar Tareas más Relevantes

Iniciador: Usuario (Agente Externo: Persona, Aplicación)

Propósito: Entregar el conjunto de tareas semánticamente más relevantes para satisfacer una tarea de consulta.

Resumen: Este caso de uso obedece al requisito funcional (iv). Una vez ejecutadas las comparaciones entre una tarea de consulta y un conjunto de tareas publicadas, se determina cuales de estas últimas resultan más relevantes para satisfacer la petición definida por la primera.

Caso de Uso No. 3: Clasificar Tareas Publicadas

Iniciador: Usuario (Agente Externo: Persona, Aplicación)

Propósito: Clasificar las tareas pertenecientes a procesos de negocio disponibles, de acuerdo con su nivel de similitud respecto a una tarea de consulta.

Resumen: Este caso de uso satisface el requisito funcional (vi). Las tareas publicadas se ordenan de acuerdo a su valor de similitud semántica, respecto a una tarea de consulta específica, con la cual se ejecutó previamente el proceso de comparación.

Caso de Uso No. 4: Determinar Similitud General

Iniciador: Usuario (Agente Externo: Persona, Aplicación)

Propósito: Cuantificar el nivel de similitud semántica existente entre dos tareas comparadas.

Resumen: Este caso de uso soporta el requisito funcional (iii). A partir del análisis de los parámetros, que definen la funcionalidad de dos tareas comparadas, se calcula la similitud semántica existente entre estas.

Caso de Uso No. 5: Determinar Similitud de Entradas

Iniciador: Usuario (Agente Externo: Persona, Aplicación)

Propósito: Cuantificar el nivel de similitud semántica existente entre las entradas de las dos tareas comparadas.

Resumen: Este caso de uso se deriva del requisito funcional (ii). Mediante la comparación entre los conceptos del enriquecimiento semántico, asociado a las entradas de cada tarea comparada, se determina un valor de similitud para este parámetro de las dos tareas.

Caso de Uso No. 6: Determinar Similitud de Salidas

Iniciador: Usuario (Agente Externo: Persona, Aplicación)

Propósito: Cuantificar el nivel de similitud semántica existente entre las salidas de las dos tareas comparadas.

Resumen: Este caso de uso se origina a partir del requisito funcional (ii). Mediante la comparación entre los conceptos del enriquecimiento semántico, asociado a las salidas de cada tarea comparada, se determina un valor de similitud para este parámetro de las dos tareas.

Caso de Uso No. 7: Determinar Similitud de Identificadores

Iniciador: Usuario (Agente Externo: Persona, Aplicación)

Propósito: Cuantificar el nivel de similitud semántica existente entre los identificadores de las dos tareas comparadas.

Resumen: Este caso de uso satisface parte del requisito funcional (ii). Mediante la comparación entre los conceptos del enriquecimiento semántico asociado a los identificadores de cada tarea comparada, se determina un valor de similitud para este parámetro de las dos tareas.

Caso de Uso No. 8: Configurar Parámetros de Comparación

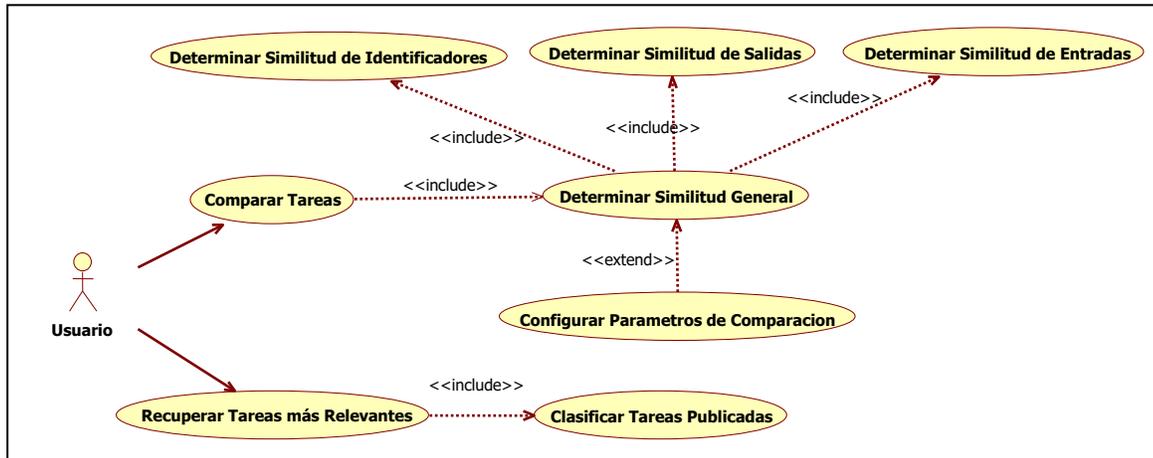
Iniciador: Usuario (Agente Externo: Persona, Aplicación)

Propósito: Ajustar los parámetros que intervienen en el cálculo de la similitud general entre dos tareas comparadas.

Resumen: Este caso de uso satisface el requisito funcional (v). Determinadas las similitudes correspondientes a cada atributo de dos tareas comparadas, se debe

establecer la contribución de cada uno de estos valores en el resultado de similitud general.

Figura 17. Diagrama de Casos de Uso del Sistema



Junto con los casos de uso es prioritario, definir una arquitectura de referencia, que proporcione una visión de los componentes más significativos del sistema y provea soporte a las actividades definidas para su desarrollo. En la siguiente sección, se describe la arquitectura establecida para el sistema implementado.

5.2.2. Descripción de la Arquitectura de Referencia

El sistema desarrollado para evaluar el mecanismo de comparación propuesto en el Capítulo 4 de este documento, adopta una arquitectura de aplicación en tres capas: la capa de Lógica de Presentación, la capa de Lógica de Negocio y una capa de Soporte. La Figura 18, presenta un diagrama de la arquitectura de referencia del sistema, en el cual, se observa como la capa de Lógica de Presentación, implementada mediante una aplicación web, recibe la petición del usuario representada a través de un proceso de negocio, modelado con BPMO (proceso de negocio de consulta). La petición es tramitada en la capa de Lógica de Negocio, donde inicialmente se llevan a cabo dos procedimientos necesarios para ejecutar la comparación de tareas:

1. Conversión de procesos BPMO a un modelo de Grafos (tanto para el proceso de consulta como para los procesos publicados).
2. Discriminación de los nodos correspondientes a las tareas de los procesos publicados y de consulta.

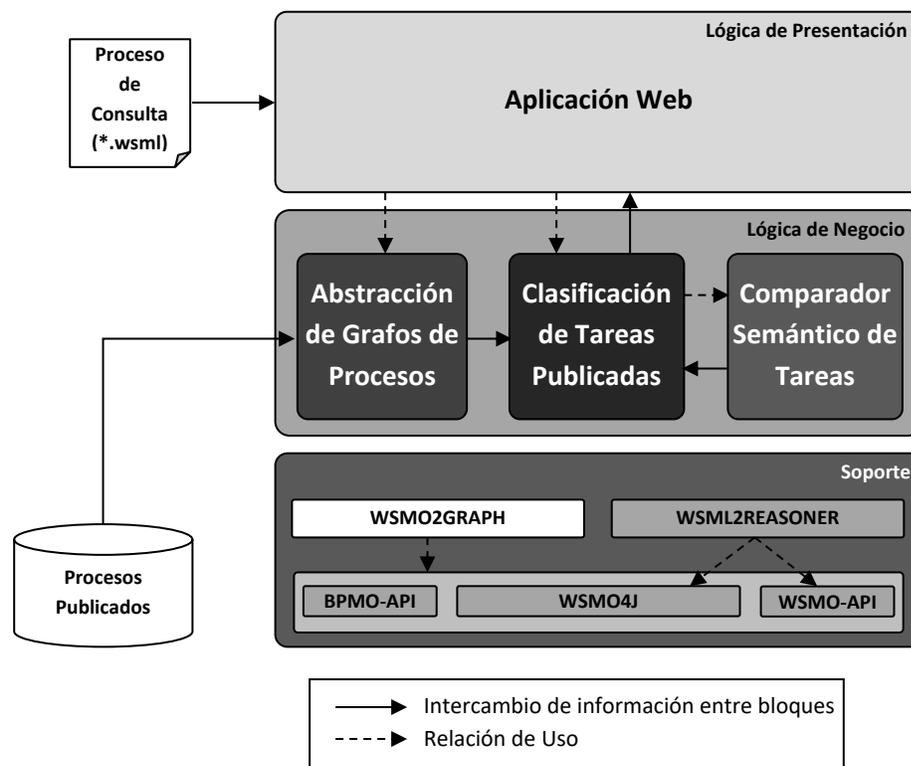
Posteriormente, las tareas de consulta y publicadas (pertenecientes a los procesos de consulta y publicados respectivamente) se comparan entre sí, determinándose el valor de similitud semántica entre ellas. De acuerdo con este resultado, se lleva a cabo la clasificación de las tareas publicadas, generando de esta manera, el conjunto de tareas más relevantes, para satisfacer la petición definida por cada tarea de consulta. La

información relacionada con este conjunto de tareas, es dirigida hacia la aplicación web donde se procesa y se despliega para el usuario.

Todo el comportamiento definido para las capas superiores de la aplicación, se sustenta en un conjunto de API de soporte.

Los procesos publicados, a los cuales se hace referencia en la arquitectura, corresponden a un conjunto de veinte procesos de negocio modelados como parte del trabajo previo realizado, para evaluar el comportamiento del sistema propuesto. En el **Anexo C** se presenta el diagrama de cada uno de los veinte procesos elaborados para dicho propósito.

Figura 18. Arquitectura de Referencia del Sistema



A continuación, se aborda una descripción detallada de los componentes de la arquitectura de referencia del sistema.

5.2.2.1. SOPORTE

Esta capa está constituida por un conjunto de herramientas software, que hacen posible el desarrollo de operaciones de conversión de procesos BPMO a Grafos, transformación de las ontologías a un modelo de objetos y procedimientos de inferencia sobre las ontologías de dominio, basados en los axiomas definidos dentro de estas y en las relaciones existentes entre sus conceptos. Las API que hacen parte de esta capa son:

- **WSMO2GRAPH.**

Constituye un módulo que implementa un mecanismo de transformación de procesos de negocio descritos en BPMO, a un completo modelo de grafos. Los grafos obtenidos de este módulo, están conformados por nodos y aristas que representan los componentes propios de la notación definida por BPMO, para especificar un proceso de negocio (Tareas, Eventos, *Graph Patterns*, *Block Patterns*, entre otros). El desarrollo de WSMO2GRAPH es producto del trabajo colaborativo de tres proyectos (incluido el presente), los cuales se están ejecutando dentro del Grupo de Ingeniería Telemática de la Universidad del Cauca. La solución desarrollada en el presente trabajo, hace parte de la arquitectura definida para el sistema propuesto en uno de estos proyectos: El trabajo de grado de maestría descrito en (Figuerola, et al., 2010), el cual pretende definir mecanismos para el descubrimiento de procesos de negocio, considerando requerimientos de búsqueda semántica y de comportamiento de los servicios. En (Figuerola, et al., 2010) se lleva a cabo un modelamiento matemático del requerimiento relacionado con el emparejamiento de procesos de negocio de manera tal que pueda abordarse, como un problema de emparejamiento de grafos. De ahí, la importancia del módulo WSMO2GRAPH y el porqué de su utilización en el desarrollo del presente proyecto.

- **BPMO-API, WSMO-API y WSMO4j.**

Conforman un conjunto de implementaciones de referencia, para construir aplicaciones basadas en las especificaciones de WSMO y BPMO. Estas API proveen soporte para el mapeo de todas las entidades definidas en estas dos iniciativas a un modelo de objetos, lo cual facilita el procesamiento y administración de la información y el conocimiento contenidos en artefactos, como ontologías y procesos de negocio.

- **WSML2Reasoner.**

Provee una plataforma para razonamiento sobre las variantes *flight* y *rule* de WSML, soportada en la infraestructura para inferencia sobre formalismos basados en reglas. La plataforma opera, a partir de una transformación sintáctica (preservando la semántica) de ontologías WSML, a programas en Datalog. De esta manera, las tareas razonamiento se realizan por medio de consultas Datalog aplicadas sobre una ontología transformada.

5.2.2.2. LOGICA DE NEGOCIO

En esta capa se implementan y articulan, todos los mecanismos necesarios para manejar las peticiones del usuario y satisfacer los casos de uso definidos para el sistema. Este nivel de la arquitectura, se descompone en tres unidades funcionales, las cuales interactúan con la capa de Lógica de Presentación y la capa de Soporte: el bloque de Abstracción de grafos de Procesos, el bloque de Clasificación de Tareas Publicadas y el bloque Comparador Semántico de Tareas.

- **Abstracción de Grafos de Procesos.**

En este bloque, se lleva a cabo la conversión de procesos BPMO a grafos, y la discriminación de los nodos correspondientes, a las tareas de los procesos publicados y de consulta. Estas operaciones emplean los métodos implementados por las entidades que componen el módulo WSMO2GRAPH de la capa de soporte. De esta manera, recibiendo como entrada el archivo WSML del proceso de negocio de consulta, y conociendo la ubicación del banco de procesos publicados, este bloque entrega al módulo de clasificación las tareas publicadas y de consulta en forma de dos conjuntos independientes.

- **Clasificación de Tareas Publicadas.**

Se constituye en un módulo elemental de ordenamiento de información. Este bloque, toma una a una las tareas de consulta y las compara con cada una de las tareas publicadas (empleando el módulo comparador semántico de tareas), para estimar su valor de similitud semántica. Con base en este valor, se determina para cada tarea de consulta el conjunto de tareas publicadas más relevantes, disponiéndolas en un arreglo en orden descendente de acuerdo con su similitud.

- **Comparador Semántico de Tareas.**

Este es el bloque más importante de la arquitectura de referencia del sistema. En él se implementan todos los Algoritmos descritos en el Capítulo 4 de este documento, los cuales definen fundamentalmente el mecanismo de comparación propuesto. El comparador semántico de tareas, es un subsistema complejo debido a la gran cantidad de operaciones que debe ejecutar, razón por la cual se definió una arquitectura para soportar su desarrollo(Figura 19):

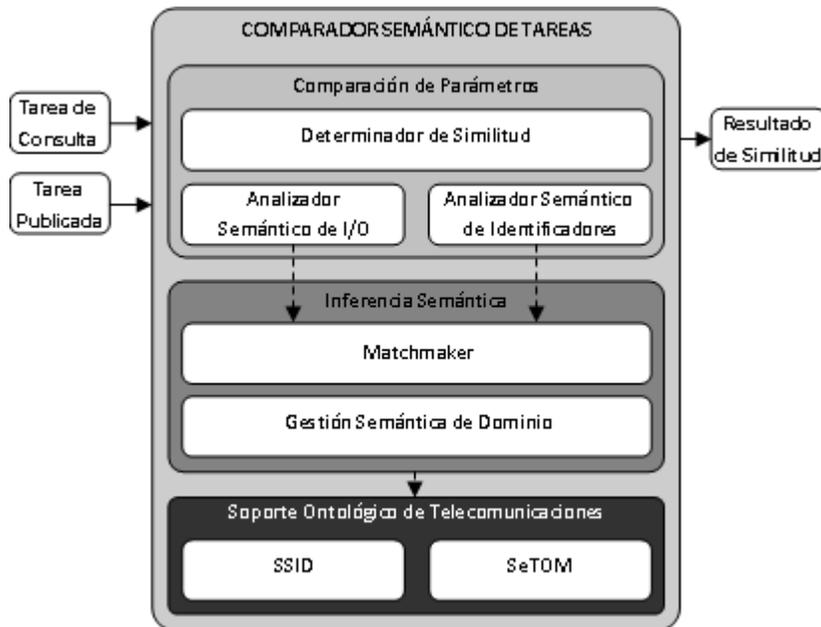
El proceso de comparación semántica, comienza con la recepción de una tarea de consulta, que será comparada con la tarea publicada. Las dos tareas pertenecientes a procesos de negocio de telecomunicaciones, deben estar enriquecidas semánticamente para establecer un valor general de similitud entre ellas.

El comparador está compuesto por los siguientes módulos:

- a. **Soporte Ontológico de Telecomunicaciones.**

En este módulo está dispuesto el conjunto de ontologías de dominio de telecomunicaciones, utilizadas para las operaciones de razonamiento llevadas a cabo sobre el enriquecimiento semántico de las tareas. De acuerdo con lo establecido en el Capítulo 3 de esta monografía, las ontologías que hacen parte de este componente son SeTOM y SSID.

Figura 19. Arquitectura para la Comparación Semántica de Tareas



b. **Gestión Semántica de Dominio.**

La función principal de este módulo, consiste en extraer de las ontologías de dominio los superconceptos (o componentes de nivel superior en el caso de SeTOM), asociados a los términos del enriquecimiento semántico de las tareas, los cuales son ordenados en un arreglo en forma ascendente (del más específico al más general), y son entregados al módulo Matchmaker. Para esto se emplean las API WSMO4J y WSMO, en conjunto con el razonador WSML2Reasoner.

c. **Matchmaker.**

Este módulo, calcula la Similitud Semántica (SS), de acuerdo a los grados de correspondencia y la Distancia Semántica (SD), entre los conceptos que enriquecen las tareas, a partir de la comparación de los arreglos de superconceptos entregados por el módulo Gestión Semántica de Dominio. Todas las operaciones llevadas a cabo dentro de este componente, se soportan mediante la implementación de los Algoritmos 2, 3, 4 y 5 previamente definidos.

Los resultados obtenidos del módulo Matchmaker, se entregan a los componentes superiores de la arquitectura, para determinar la similitud existente entre los atributos de las tareas comparadas.

d. **Analizador semántico de I/O.**

Este módulo cumple con la función de determinar dos valores de similitud generales, para los conjuntos de entradas y salidas de las dos tareas comparadas, a partir de los valores de similitud semántica entre conceptos (pertenecientes en este caso a la ontología SSID)

entregados por el módulo Matchmaker. Las entidades que conforma este componente, implementan los Algoritmos 6 y 7 propuestos y descritos en el Capítulo 4.

Los resultados de la comparación entre entradas y salidas suministrados por el módulo *Analizador Semántico de I/O*, se entregan posteriormente módulo *Determinador de Similitud*.

e. Analizador semántico de Identificadores.

Este módulo, se encarga de establecer el grado de correspondencia y la similitud semántica, entre los identificadores de las dos tareas comparadas. Cuando se ejecuta esta etapa de comparación, la ontología utilizada en el módulo Soporte de Dominio es SeTOM, la cual contiene un conjunto de axiomas, que permiten realizar inferencia sobre sus conceptos. La operación que ejecuta este módulo, se fundamenta en lo establecido en el Algoritmo 5 presentado en el Capítulo anterior.

f. Determinador de Similitud.

Finalmente, éste módulo analiza los resultados obtenidos de las comparaciones de entradas/salidas, y de identificadores suministradas por los dos módulos anteriores, y establece un valor total de similitud semántica para las dos tareas, de acuerdo con la operación definida el Capítulo 4 por medio de la ecuación 5.

Los submódulos *Matchmaker* y *Gestión Semántica de Dominio*, se encuentran separados en una agrupación denominada *Inferencia Semántica*. Asimismo, los bloques *Determinador de Similitud*, *Analizador Semántico de I/O*, y *Analizador Semántico de Identificadores*, se disponen en otra agrupación conocida como *Comparación de Parámetros*. Esta separación se realiza, con el fin de establecer, que estos submódulos operan en diferentes niveles del proceso de comparación. De esta manera, los bloques contenidos en *Inferencia Semántica*, son responsables de las operaciones realizadas directamente sobre los conceptos de las ontologías de dominio (comparación a nivel de conceptos), mientras que los bloques pertenecientes a *Comparación de Parámetros*, se encargan de interpretar los resultados obtenidos de la agrupación previa, para determinar la similitud en cuanto a funcionalidad de las tareas comparadas (comparación a nivel de los atributos de las tareas).

El diseño, implementación y verificación de los componentes que constituyen la capa de Lógica de Negocio, comprendieron dos iteraciones del proceso de desarrollo del sistema. La primera de ellas, se dedicó a los bloques de *Abstracción de Grafos de Procesos* y los submódulos *Matchmaker* y *Gestión Semántica de Dominio* del *Comparador Semántico de Tareas*. Los submódulos restantes de este componente, junto con el bloque de *Clasificación de Tareas Publicadas*, se abordaron en la segunda iteración, al final de la cual, se sometió el prototipo del sistema a un proceso de pruebas de rendimiento, el cual se describirá más adelante en este capítulo.

5.2.2.3. LÓGICA DE PRESENTACIÓN

Esta capa expone las funcionalidades ofrecidas por los componentes pertenecientes a la capa de Lógica de Negocio. Consiste en una aplicación web que dispone de un conjunto

de vistas, que satisfacen los casos de uso definidos en la sección anterior, desde la perspectiva del usuario. La aplicación cuenta con un total de cinco (5) interfaces, a través de las cuales, este puede transmitir su petición (en forma de un proceso de negocio de consulta), y configurar diferentes parámetros del procedimiento de comparación, para finalmente recibir el resultado que arroja el procesamiento de su solicitud por parte del sistema. Al acceder a la aplicación se despliega la interfaz de bienvenida, la cual suministra al usuario información relacionada con el sistema y le permite validarse (en caso de haber accedido previamente a la aplicación) o registrarse (caso de un usuario nuevo).

Posterior al registro (o validación), se despliega la interfaz de configuración de los parámetros de comparación (Figura 20), en la cual el usuario especifica su solicitud a partir de la carga de los archivos *.wsml y *.layout²⁷ del proceso de consulta, necesarios para extraer sus tareas respectivas y generar el diagrama del proceso de negocio. Esta interfaz también le permite al usuario, definir el porcentaje de contribución de la similitud correspondiente a cada uno de los atributos de las tareas a comparar, sobre el valor de similitud semántica general. Así mismo, el usuario puede establecer si desea someter todas las tareas de consulta al proceso de comparación o si prefiere omitir algunas de ellas. Si opta por la última opción, la aplicación le permite elegir cuales de las tareas no serían tenidas en cuenta en la comparación, por medio de la interfaz que se muestra en la Figura 21.

Una vez configurados todos los parámetros de comparación, el sistema procesa esta información y compara las tareas de consulta seleccionadas, con cada una de las tareas publicadas disponibles.

Finalmente, se despliega para el usuario una interfaz que presenta los resultados de la comparación (Figura 22). Esta interfaz contiene una tabla con información distribuida de la siguiente forma: en la primera columna, se consigna el nombre de cada una de las tareas de consulta seleccionadas en el proceso de configuración; en la segunda columna, se ordenan las k tareas publicadas más similares, para cada tarea de consulta (donde k puede tomar los valores de 3, 5 o 10, de acuerdo con el criterio del usuario), junto con el nombre del proceso de negocio al cual pertenecen; en la tercera columna, se indica el valor de similitud semántica general, calculado entre cada tarea publicada y de consulta presentes en la tabla; por último, en la cuarta columna, se dispone de un vínculo para el despliegue de un panel (Figura 23), que contiene información detallada sobre el resultado de similitud obtenido entre cualquier par de tareas (publicada y de consulta) registradas.

²⁷ Estos archivos son generados por la herramienta de modelado de procesos de negocio que proporciona WSMO Studio.

Figura 20. Configuración de Parámetros de Comparación

Semantic Task Comparator

A Business Process Reengineering Solution



Home
Comparison Set up
Contact Us
Log Out

Directions

Logged in as: admin

Please, follow the next instructions to setup the semantic task comparison procedure.

1. Upload the *.wsml and *.layout files of the query business process (there must be two wsml files of the business process: the enriched one and the simple one). Make sure that the three files belong to the same business process.
2. Once you upload the business process files, the application deploys the business process diagram. Later, you have to choose the type of comparison you want to execute. If you choose 'Compare Complete Process' (set as default), the application will take into account all of the query business process task in the comparison procedure, however if you choose 'Select Tasks to Compare' the application allows you to skip some of the query business process task in the in the comparison procedure.
3. The next step is to set the contribution percentages for each task parameter similarity (Inputs, Outputs and ID). The sum of these values must be 100.
4. Finally, clic on 'continue'

■ Comparison Set Up

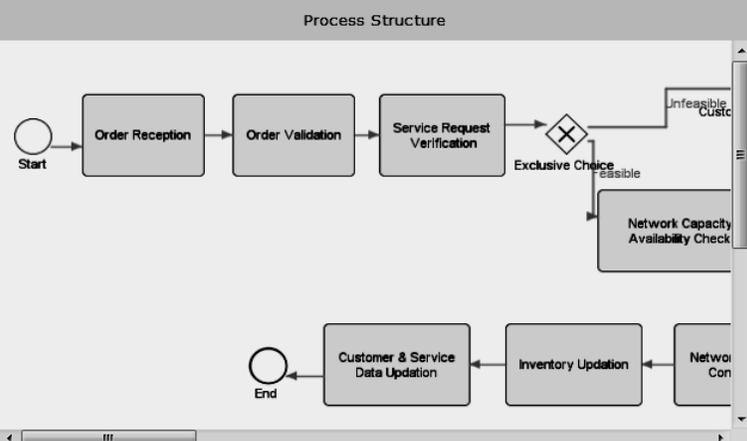
+ Upload the WSML Query Process files (*.wsml, *.layout)

✕ Clear All

ServiceProvisioningProcess.wsml
Clear

Done

Process Structure



Define Comparison Settings:

Select Tasks to Compare
 Compare Complete Process

Set Parameter Percentages

Input:	35
Output:	20
ID:	45

[Continue](#)

Contact Us | Terms of Use | Privacy Statement
Copyright © 2010 ---. All Rights Reserved.

Figura 21. Interfaz de Selección de las Tareas de Consulta a Comparar

Semantic Task Comparator
A Business Process Reengineering Solution

Home | Task Comparator | Contact Us | Log Out

Directions
Logged in as: admin

Please, follow the next instructions to complete the semantic task comparison procedure setup:

1. As default all the query business process task are selected. You only have to let selected those tasks that you want to compare.
2. Finally, clic on 'compare'

Task Comparator

Query Process Structure

```
graph LR; Start((Start)) --> OR[Order Reception]; OR --> OV[Order Validation]; OV --> SRV[Service Request Verification]; SRV --> EC{Exclusive Choice}; EC -- Infeasible Cus --> NCA[Network Capacity Availability Check]; EC -- Feasible --> NCA; NCA --> NetCon[Network Con]; NetCon --> IU[Inventory Updation]; IU --> CSUD[Customer & Service Data Updation]; CSUD --> End((End));
```

Select the task to compare:

- Customer & Service Data Updation
- Service Instance Creation
- Service Request Verification
- Order Validation
- Network Capacity Availability Check
- Order Reception
- Network Equipment Configuration
- Inventory Updation

Contact Us | Terms of Use | Privacy Statement
Copyright © 2010 ---. All Rights Reserved.

Figura 22. Interfaz de Resultados de Comparación y Ranking de Tareas.

Semantic Task Comparator
A Business Process Reengineering Solution

Home | Task Comparator | Contact Us | Log Out

Directions
Logged in as: admin
This page shows you the comparison procedure results:
The table in the right pane, depicts the task ranking for each query task. You can set the task ranking size (quantity of deployable target task) by clicking the 'top k' links above the table.
If you want, you can get the comparison details of each pair of tasks (query and target), by clicking the 'details' link on the fourth column of the results table.

Task Comparator

Task Ranking

| Top 3 | Top 5 | Top 10

Query Task	Target Task (from Process)	Similarity	Details
Order Reception	Order Reception (ServiceProvisioningProcess)	5.0	Details
	Establish Customer Contact (Customer Retention and Loyalty)	2.556648432064109	Details
	Contact the Customer (ManageCustomerRelationships)	2.467970680678519	Details
Service Instance Creation	Service Instance Creation (ServiceProvisioningProcess)	5.0	Details
	Enable Service Activation (Telecom Service Delivery)	4.067368421052632	Details
	Authorize Service Activation (Activate Service)	3.3607503607503606	Details
Network Capacity Availability Check	Resource Capability Check (Telecom Service Delivery)	5.0	Details
	Network Capacity Availability Check (ServiceProvisioningProcess)	5.0	Details
	Detect Customer Problem (Solve Customer Problem)	1.5114155251141552	Details

Back to Comparison Setup

Contact Us | Terms of Use | Privacy Statement
Copyright © 2010 ---. All Rights Reserved.

Figura 23. Interfaz de Resultados Detallados de Comparación.

Comparison Details

Query: Register Order

Query Inputs		Query Outputs	
infrastructureSetUpNotification	customerServiceRequest	serviceLevelAgreement	serviceRequestResponse

Target: Service Feasibility Check

Target Inputs		Target Outputs	
serviceOrder		requestAssessmentNotification	resourceRequirementRequest

Inputs Similarity: 0.6877283473028153	Outputs Similarity: 1.9519847606241418	Identifier Similarity: 5.0
Overall Similarity: 2.8811018736808136		

hide this panel

El trabajo realizado alrededor de esta aplicación web, se abordó como parte de la tercera iteración del proceso de desarrollo del sistema. Para su implementación, se empleó la tecnología Java Server Faces, la cual permite establecer una clara separación entre la Lógica de de Presentación y la Lógica de Negocio subyacente, de forma similar a la alcanzada aplicando el patrón de diseño MVC, en el desarrollo de una aplicación de escritorio. Como contenedor web, para desplegar la aplicación diseñada se empleó el servidor Apache Tomcat v6.0.20, el cual implementa las especificaciones *Java Servlet* y *Java Server Pages* (JSP) de *Sun Microsystems* y provee un potente entorno de ejecución. Vale la pena señalar, que la capa de Lógica de Presentación, fue desarrollada con fines puramente ilustrativos, es decir, no constituye una parte relevante de la funcionalidad del sistema, concentrada fundamentalmente en la capa de Lógica de Negocio. En consecuencia, es posible separar y empaquetar los componentes de la capa central de la arquitectura, como una aplicación independiente, la cual puede importarse y utilizarse en cualquier otra aplicación que lo requiera.

5.3. Pruebas del Prototipo

Una vez concluida la implementación del prototipo operacional, es necesario someterlo a un proceso de evaluación y puesta a punto que permita identificar y corregir deficiencias en la operación del sistema y garantizar su óptimo desempeño.

Con el fin de llevar a cabo actividades de evaluación sobre el prototipo desarrollado, se requirió inicialmente, el modelado de los procesos de negocio, que conforman el banco de procesos publicados. Estos procesos se organizan en cinco (5) dominios diferentes, de acuerdo con su ubicación dentro del modelo eTOM, y representan en conjunto un total de noventa y siete (97) tareas semánticamente enriquecidas, las cuales alimentan el mecanismo de comparación implementado.

Posterior al establecimiento del banco de procesos publicados, se desarrollo un módulo que permite extraer y almacenar en una base de datos, la información relacionada con los resultados de similitud obtenidos por el comparador de tareas. De esta manera, diez (10) de las tareas pertenecientes a los procesos publicados, fueron definidas como tareas de consulta, para luego compararlas con las 97 tareas publicadas, generando así una base de resultados de comparación del mecanismo, o lo que se ha denominado *benchmark del prototipo*.

Benchmark puede definirse como un punto de referencia, a partir del cual es posible evaluar comparativamente el rendimiento de un sistema respecto de otro, el cual por lo general representa las mejores prácticas de su dominio (Chillarege, 1999). Para el caso de interés, el benchmark está integrado por los resultados obtenidos de la ejecución del sistema sobre un conjunto de tareas preestablecidas. De esta manera, para evaluar el rendimiento del prototipo, particularmente, para determinar que tan adecuados son los resultados de similitud obtenidos por el mecanismo de comparación, se requiere precisamente conocer, cuáles son los resultados de similitud apropiados para cada una de las comparaciones ejecutadas por el prototipo, entendiendo estos resultados apropiados como aquellos que se ajustan al criterio de un experto del dominio. En este

sentido, debió desarrollarse una herramienta (la cual se describe en el **Anexo B**) que permitiera recoger las comparaciones realizadas manualmente por un grupo de personas con amplio conocimiento del dominio de los procesos de negocio de telecomunicaciones, con el objetivo de construir un *benchmark de referencia* para evaluar el rendimiento del prototipo implementado. El benchmark de referencia consiste en un conjunto resultados de similitud, obtenidos a partir de las comparaciones manuales entre tareas de consulta y publicadas, realizadas por un grupo de 8 evaluadores expertos.

En la herramienta desarrollada para este propósito, se realizaron en total 2910 comparaciones entre parejas de tareas (publicada y de consulta), garantizando que cada una de las 970 posibles parejas (10 tareas de consulta x 97 tareas publicadas) fuera comparada por al menos 3 de los evaluadores. En consecuencia, se contaba al final del proceso de evaluación con 3 resultados de similitud para cada par de tareas, los cuales fueron promediados en cada caso, conformando de esta manera la base de 970 resultados de similitud que constituyen el benchmark de referencia. Para determinar el nivel de rendimiento del sistema, se compararon los resultados consignados en este benchmark con aquellos registrados en el benchmark del prototipo, a partir de un análisis estadístico basado en la aplicación de un conjunto de medidas empleadas en la evaluación del desempeño de sistemas de recuperación de información. Dichas medidas se describen en la siguiente sección.

5.3.1. Medidas de Desempeño

Una funcionalidad básica del sistema desarrollado en el presente proyecto, consiste en generar un ranking de las tareas publicadas más relevantes (dispuestas por orden de similitud), para satisfacer una petición definida por medio de una tarea 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 caracterización del desempeño de sistemas de recuperación de información (p.ej. motores de búsqueda en internet). Estas medidas se denominan: *precision(p)*, *recall(r)*, *overall(o)* (Yatskevich, 2003), *top-k precision (P_k)* y *top-p precision (P_p)* (Lu, et al., 2005).

Las tres primeras medidas (*precision*, *recall* y *overall*), operan sobre la cardinalidad de tres conjuntos conformados por las tareas publicadas recuperadas por el prototipo, en respuesta a una tarea de consulta específica: el conjunto de *verdaderas positivas(TP)* (tareas publicadas recuperadas correctamente de acuerdo con los resultados registrados en el benchmark de referencia), el conjunto de *falsas positivas(FP)* (tareas publicadas incorrectamente recuperadas) y el conjunto de *falsas negativas(FN)* (tareas publicadas que no se recuperan, a pesar de ser relevantes, de acuerdo con el benchmark de referencia). De esta manera, las medidas *precision*, *recall* y *overall* se definen como:

Precision(p): se refiere a la proporción de tareas clasificadas como *verdaderas positivas*, respecto al número total de tareas publicadas recuperadas por el prototipo, esto es:

$$p = \frac{|TP|}{|TP|+|FP|} \quad \text{Ec. 6.}$$

Recall(*r*): identifica la proporción de tareas clasificadas como *verdaderas positivas*, respecto al número total de tareas publicadas consideradas como relevantes (es decir, todas aquellas tareas que deben ser recuperadas de acuerdo con el benchmark de referencia):

$$r = \frac{|TP|}{|TP|+|FN|} \quad \text{Ec. 7.}$$

Overall(*o*): relaciona las medidas de *precision* y *recall*, y se emplea para determinar la calidad del proceso de comparación y recuperación de tareas. Matemáticamente se calcula:

$$o = r * \left(2 - \frac{1}{p}\right) \quad \text{Ec. 8.}$$

De acuerdo con las dos primeras definiciones, la medida de *precision* determina el porcentaje de relevancia del conjunto de tareas recuperadas por el sistema, mientras que *recall* estima que porcentaje del total de tareas relevantes fue recuperado.

Por otra parte, las medidas de *top-k precision* (P_k) y *top-p precision* (P_p) operan con una orientación diferente a las tres previamente precisadas. Antes de abordar la descripción de estas dos medidas, es necesario introducir una serie de términos empleados en su definición:

Ranking del Prototipo: conjunto obtenido a partir de la clasificación de las tareas publicadas, según su valor de similitud, respecto a una tarea de consulta determinada, de acuerdo con el benchmark del prototipo.

Ranking Base: conjunto obtenido a partir de la clasificación de las tareas publicadas, según su valor de similitud, respecto a una tarea de consulta determinada, de acuerdo con el benchmark de referencia.

Tareas Relevantes de nivel *k* (TR_k): conjunto definido por las tareas compartidas entre los primeros *k* niveles del ranking del prototipo y del ranking base. Teniendo en cuenta que tanto en el ranking del prototipo como en ranking base, pueden existir tareas publicadas cuyo valor de similitud respecto a la tarea de consulta es el mismo, se considera apropiado que dicho conjunto de tareas sea clasificado en un mismo nivel, esto con el fin de evitar que el orden aleatorio en que están dispuestas estas tareas dentro de ambos rankings, influya con los resultados de la aplicación de las dos últimas medidas de desempeño definidas.

Teniendo en cuenta lo establecido en los términos anteriores, se definen las medidas *top-k precision* (P_k) y *top-p precision* (P_p):

top-k precision (P_k): evalúa la proporción de tareas relevantes (TR_k) respecto al número total de tareas clasificadas en los primeros *k* niveles del ranking del prototipo (RP_k):

$$P_k = \frac{|TR_k|}{|RP_k|} \quad \text{Ec. 9.}$$

top-p precision (P_p): determina el número de tareas relevantes (TR_p) cuya clasificación en cuanto a niveles coincide con la disposición del ranking base para un nivel k definido:

$$P_p = \frac{[TR_k]TR_p}{k} \quad \text{Ec. 10.}$$

Estas dos últimas medidas, especifican que tan adecuado es el criterio de clasificación de tareas implementado en el prototipo, para una tarea de consulta en particular.

Con el fin de evaluar el rendimiento general del sistema, en cuanto a la calidad del método para estimar la similitud y a su capacidad para recuperar tareas realmente relevantes, se llevó a cabo la aplicación de las medidas de desempeño descritas, sobre los resultados de similitud registrados en el benchmark del prototipo para cada una de las 10 tareas de consulta establecidas.

La siguiente sección detalla el plan de pruebas dispuesto para el sistema implementado, junto con los resultados obtenidos a partir de su ejecución.

5.3.2. Plan de Pruebas y Resultados Obtenidos

Todas las pruebas a las cuales se hace referencia a continuación, se llevaron a cabo sobre el prototipo generado, al concluir la segunda iteración del proceso de desarrollo de la solución propuesta. De esta manera, se realizó la evaluación del desempeño de las funcionalidades implementadas en la Capa de Lógica de Negocio de la Arquitectura descrita en secciones previas. Asimismo, la ejecución iterativa y la continua verificación de los resultados de este proceso de evaluación, soportaron la puesta a punto del sistema, a partir de la configuración de los parámetros de comparación definidos en los Algoritmos de Emparejamiento, Distancia y Similitud Semánticos presentados en el Capítulo 4.

A continuación se describen las pruebas de calidad y eficiencia ejecutadas sobre el prototipo desarrollado:

5.3.2.1. Plan de Pruebas.

Tabla 2. Plan de Pruebas

Transformación BPMO a Grafos de los Procesos Publicados Disponibles
Prueba de Rendimiento: permite medir el tiempo que le toma al sistema llevar a cabo el mapeo a un modelo de grafos, de los procesos de negocio disponibles en el Banco de Procesos Publicados, así como la discriminación de los nodos correspondientes a las tareas que hacen parte de estos procesos.
Comparación y Recuperación de Tareas
Prueba de Rendimiento: permite determinar el tiempo que toma el sistema en

recuperar las tareas publicadas más relevantes para satisfacer una tarea de consulta determinada.

Prueba de Calidad: permite evaluar la calidad de los resultados de similitud obtenidos por el sistema, al comparar una tarea de consulta con las tareas publicadas disponibles.

Los resultados de las pruebas de rendimiento están condicionados por las prestaciones del equipo donde se ejecuta el software del prototipo. En este caso, el equipo empleado para desplegar el plan de pruebas descrito cuenta con las siguientes especificaciones técnicas:

Tabla 3. Especificaciones Técnicas del Equipo empleado para Pruebas del Prototipo

Microprocesador	Intel Pentium Dual-Core Mobile T2390* de 1,86GHz
Caché del microprocesador	1MB de caché de nivel 2
Memoria	3072MB (2 dimm)
Disco duro	250GB (5400 RPM)
Sistema Operativo	Windows Vista Home Premium

5.3.2.2. Resultados

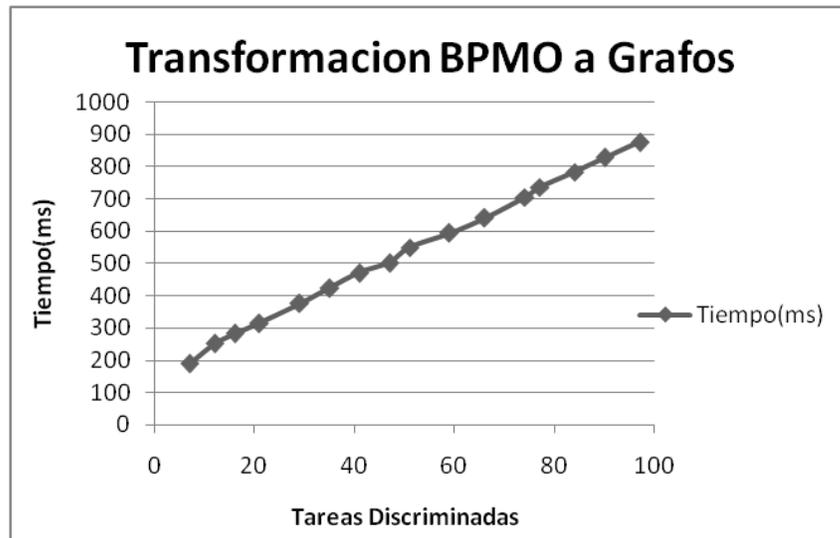
A partir de la información obtenida en la ejecución del plan de pruebas descrito, se generó el conjunto de gráficas que presenta esta sección, las cuales que permiten ilustrar y analizar el desempeño general del sistema desarrollado.

- **Transformación BPMO a Grafos de los Procesos Publicados Disponibles**

Prueba de Rendimiento: a partir de la ejecución de esta prueba, se determinó el tiempo que tarda el sistema en convertir todos los procesos de negocio disponibles en el banco de procesos publicados, a la vez que discrimina y aísla sus tareas componentes.

De acuerdo con, el comportamiento de la curva que se ilustra en la Figura 24, se observa como existe una relación de linealidad entre el número de tareas obtenidas de los Procesos de Negocio publicados, y el tiempo que toma el sistema en la ejecución de esta actividad. Por otra parte, se considera que el tiempo empleado en esta operación es aceptable (menos de 1 segundo para la obtención de alrededor de 100 tareas), teniendo en cuenta la complejidad del procedimiento de lectura y mapeo de las ontologías correspondientes a los procesos de negocio, y la gran cantidad de transacciones llevadas a cabo en la capa de *Soporte*.

Figura 24. Gráfica de rendimiento del sistema en la Transformación de Procesos de Negocio BPMS a Grafos y discriminación de sus tareas.



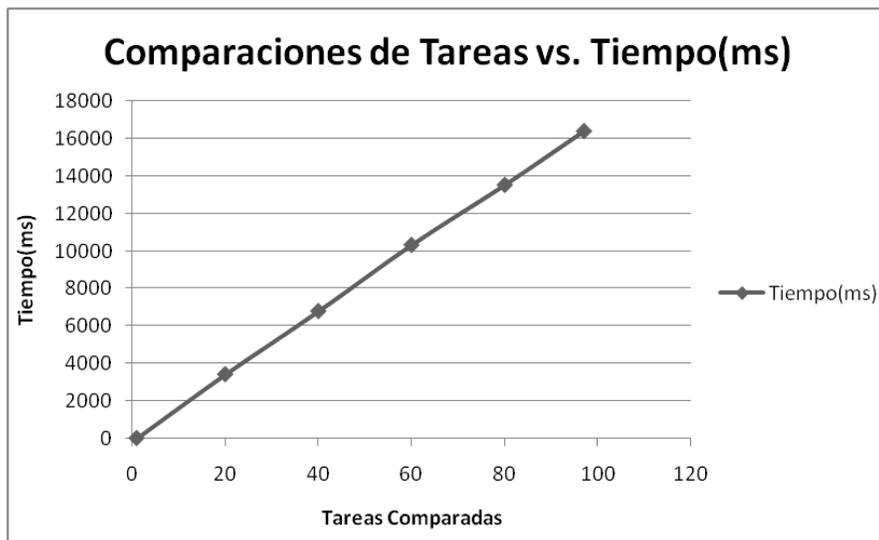
- **Comparación y Recuperación de Tareas**

Prueba de Rendimiento: por medio de esta prueba, se determinó la eficiencia del procedimiento de comparación, en cuanto al tiempo que tarda el sistema en llevar a cabo la evaluación de la similitud semántica general, entre una tarea de consulta determinada y cada una de las tareas pertenecientes a los procesos publicados. De acuerdo con la Figura 25, la duración de este procedimiento aumenta linealmente con el número de tareas disponibles en el banco de procesos.

El procedimiento de comparación de tareas, involucra un gran número de operaciones complejas y dispendiosas, relacionadas por ejemplo con el razonamiento sobre las ontologías de soporte y el manejo de arreglos de objetos, razón por la cual, tal como se muestra en la Figura 25, la ejecución de esta funcionalidad particular del sistema consume un tiempo considerable. Sin embargo, dado que no se tiene una referencia de otro sistema software, que permita llevar a cabo la función de comparación de tareas propuesto, y teniendo en cuenta que de otra manera, este procedimiento de comparación tendría que llevarse a cabo de forma manual, el sistema implementado facilita considerablemente las actividades de análisis y reingeniería de procesos de negocio, relacionadas por ejemplo, con la identificación de componentes (tareas) reutilizables para componer nuevos servicios dentro de un operador de telecomunicaciones.

La optimización en cuanto a rendimiento de la función de comparación del sistema implementado no se tiene en cuenta dentro del alcance del presente trabajo de grado, razón por la cual esta actividad puede considerarse para futuros proyectos de investigación.

Figura 25. Gráfica del rendimiento del sistema en la operación de Comparación de Tareas.



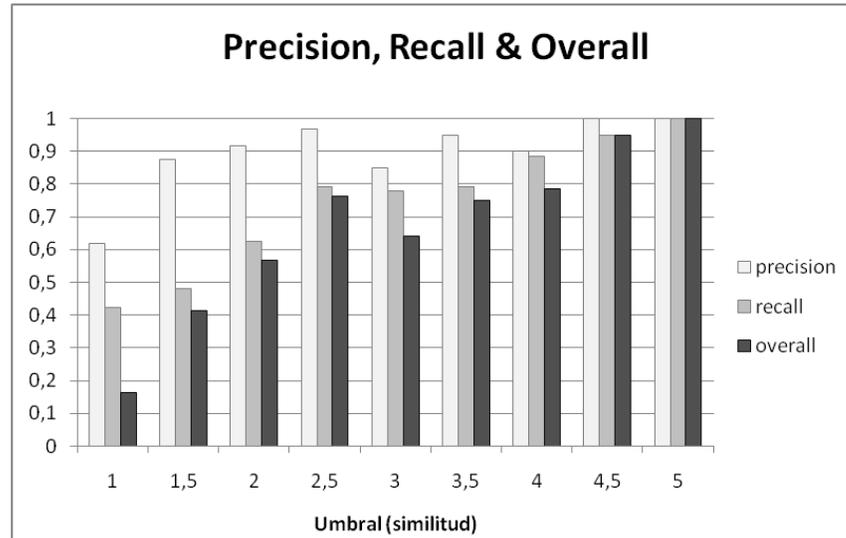
Prueba de Calidad: la ejecución de esta prueba, permite evaluar la calidad de los resultados obtenidos del mecanismo propuesto, en las actividades de comparación y clasificación de las tareas publicadas relevantes, para una petición particular especificada a través de una tarea de consulta. En las gráficas que se presentan a continuación se ilustran los resultados generales de las medidas de *precision*, *recall*, *overall*, *top-k precision* y *top-p precision*, evaluadas para diferentes valores de similitud y diferentes niveles del ranking de tareas recuperado por el sistema, considerando para las dos últimas medidas dos escenarios de pruebas:

- 1) **Dominio vs. Dominio:** evaluación de las medidas de *top-k precision* y *top-p precision* sobre los resultados de similitud de las comparaciones, realizadas entre tareas de consulta y tareas publicadas pertenecientes al mismo dominio.
- 2) **Dominio vs. Todos los Dominio:** evaluación de las medidas de *top-k precision* y *top-p precision* sobre todos los resultados de similitud de las comparaciones realizadas (sin discriminar el dominio al que pertenecen las tareas de consulta).

La gráfica de la Figura 26, presenta los resultados obtenidos de la aplicación de las medidas de *precision*, *recall* y *overall* para diferentes valores de similitud semántica (*umbral* en la gráfica).

De acuerdo con estos resultados, el desempeño del sistema es en general adecuado, en la medida en que gran parte de las tareas que recupera son tareas relevantes, conforme con el benchmark de referencia. Esto se refleja en el valor de la medida de *precision*, el cual es superior al 60% y es en promedio del 90%, para todo el rango de valores de similitud semántica. El mejor desempeño del sistema se obtiene para similitudes superiores a 3,5 de acuerdo con la medida de *overall*, la cual determina que en este intervalo, el sistema recupera tareas relevantes a la vez que descarta apropiadamente aquellas que no lo son.

Figura 26. Medidas de Desempeño: Precision, Recall y Overall para el Sistema.

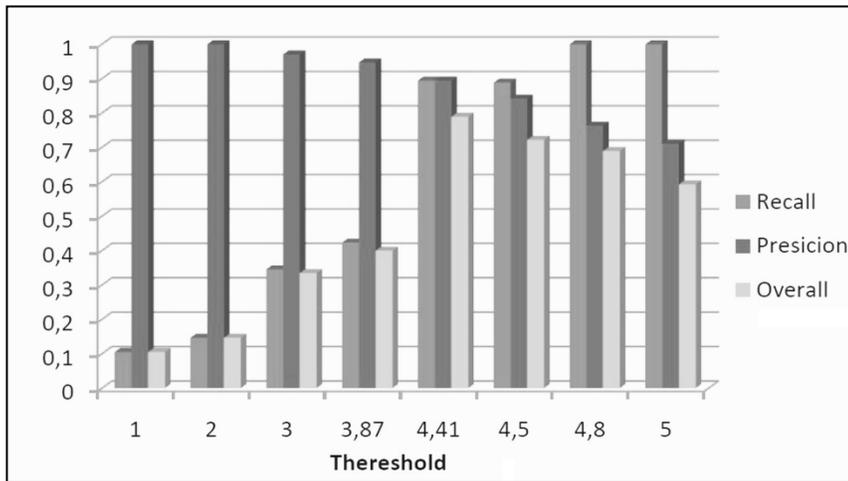


Por otra parte, el desempeño más pobre del sistema se presenta en los valores bajos de similitud (valores inferiores a 2), rango en el que se manifiesta la dificultad que presenta el sistema, para estimar la similitud semántica de tareas disímiles. En este rango, a pesar de que la mayor parte de las tareas recuperadas son tareas relevantes (*precision* aceptable), también se descartan muchas de ellas (bajo *recall*), lo cual se debe a que para estos valores de similitud, el mecanismo de comparación semántico evalúa relaciones conceptuales muy débiles, las cuales pueden resultarle ambiguas, subestimándolas o sobreestimándolas respecto al criterio de un evaluador experto. De esta manera el “ruido” recuperado para este intervalo particular es considerable.

Sin embargo, a pesar del inconveniente previamente mencionado, la calidad de los resultados obtenidos mediante la aplicación del mecanismo de comparación semántica, es en general, sustancialmente mejor respecto al desempeño de sistemas similares de recuperación de servicios basados en el emparejamiento sintáctico de sus interfaces (tales como el presentado en (Rojas, et al., 2010), cuya gráfica de desempeño se ilustra en la Figura 27), lo cual manifiesta la ventaja del desarrollo y la aplicación de enfoques como el propuesto en el presente trabajo.

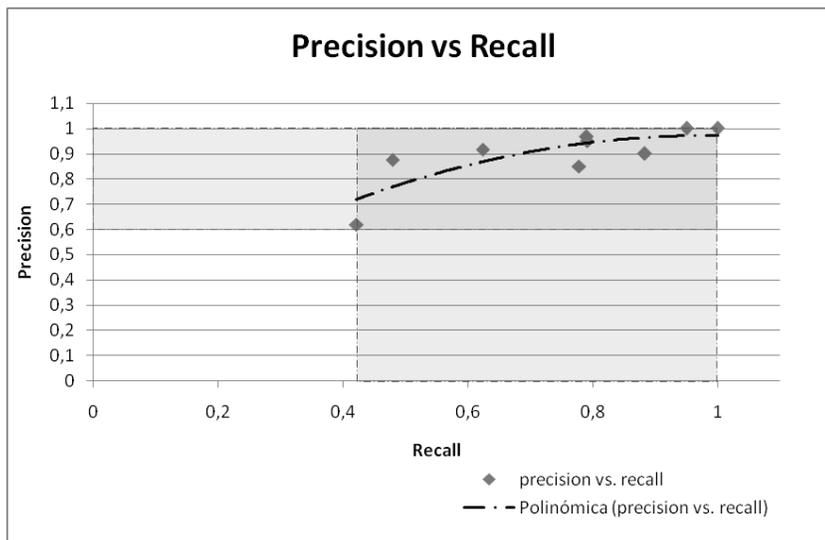
La Figura 28, ilustra la relación existente entre las medidas de *precision* y *recall*. En esta gráfica se visualiza como el desempeño del sistema está en general ubicado en una zona de alta *precision* (superior al 60%) y *recall* considerable (en promedio del 74% y superior al 40%), lo cual implica que el mecanismo implementado, a la vez que es significativamente selectivo, recupera un gran porcentaje de las tareas relevantes disponibles en el banco de procesos publicados.

Figura 27. Medidas de Desempeño: Precision, Recall y Overall para un sistema de recuperación de servicios basado en el emparejamiento sintáctico interfaces



Fuente: (Rojas, et al., 2010)

Figura 28. Gráfica de Precision vs. Recall



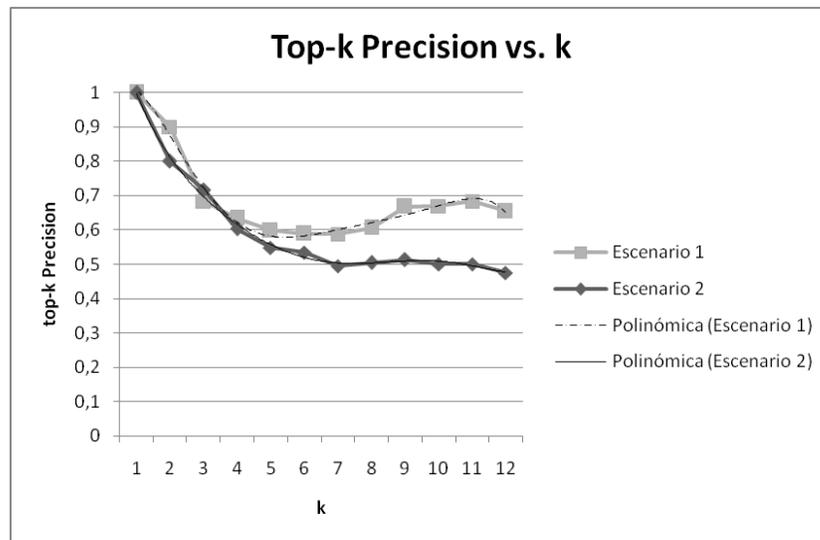
Las dos últimas medidas de desempeño: *top-k precision* y *top-p precision*, permiten evaluar la calidad del procedimiento de clasificación (ranking) de las tareas más relevantes, teniendo en cuenta los escenarios de prueba definidos previamente.

La precisión del sistema, medida en términos del porcentaje de tareas relevantes recuperadas en los primeros k niveles del ranking del prototipo, se ilustra en la Figura 29. De acuerdo con los resultados presentados en esta gráfica, la precisión del sistema disminuye conforme se incrementa el valor de k, llegando un valor mínimo alrededor del

cual se mantiene en los niveles restantes. Dicho valor mínimo, corresponde al nivel k del ranking para el cual el prototipo, recupera cerca del 50% del total de las tareas publicadas disponibles (es necesario recordar que en un nivel del ranking pueden posicionarse varias tareas con el mismo valor de similitud semántica).

De acuerdo con la Figura 29, el mejor comportamiento del sistema se obtiene para las condiciones establecidas en el escenario 1 de pruebas, en el cual se realiza un filtrado de las tareas publicadas, de tal manera que sólo se tiene en cuenta los resultados de similitud de aquellas que pertenecen al mismo dominio de las tareas de consulta. En este escenario la precisión del sistema es considerablemente alta entre los primeros 3 niveles del ranking, siendo ideal (100%) para el primer nivel y manteniéndose por encima del 60% hasta el doceavo.

Figura 29. Gráfica de top-k precision vs. k.



Para el escenario 2 de pruebas, la precisión es ligeramente inferior pero significativamente favorable, puesto que al igual que en el primer escenario, alcanza un valor máximo del 100% en el primer nivel del ranking, el cual disminuye hasta el 70% en el nivel 3 y se mantiene superior al 50% en los niveles restantes.

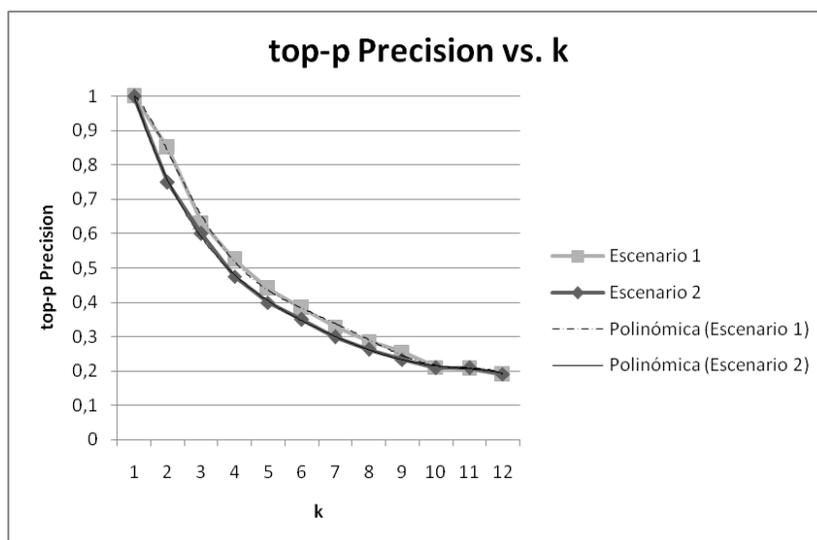
De esta manera, el desempeño del sistema respecto a la medida de *top-k precision* es óptimo, para valores de k inferiores a 3, y es en general aceptable para los demás niveles del ranking.

La medida de desempeño *top-p precision*, es más exigente que la anterior, en tanto que evalúa la precisión del sistema en la distribución de las tareas relevantes, dentro de los niveles del ranking del prototipo. Los resultados de la aplicación de esta medida, sobre el sistema implementado, se ilustran en la Figura 30. En esta gráfica, se observa como las curvas de *top-p precision* (para ambos escenarios de pruebas), tienen un comportamiento muy similar: las dos curvas alcanzan un valor máximo de precisión del 100% en el primer

nivel, el cual decrece en forma continua conforme el número k de niveles del ranking aumenta.

El primer escenario de pruebas, presenta una precisión ligeramente superior respecto al segundo, hecho que se atribuye al filtrado de tareas por dominio. Es así como, puede caracterizarse el desempeño del sistema (en relación con la medida de *top-p precision*) como óptimo en los 3 primeros niveles del ranking, en los cuales la precisión es superior al 60% y aceptable en hasta el nivel 4, donde la precisión decrece hasta el 50%.

Figura 30. Gráfica de *top-p precision* vs. k



El comportamiento del sistema, proyectado en esta última medida de desempeño, se justifica en el carácter semántico del mecanismo de comparación. Es así como, en determinados casos, el sistema puede llegar a estimar relaciones de similitud entre tareas, las cuales no son inmediatamente evidentes para una persona común, e incluso para un evaluador experto, debido a que los términos que identifican los atributos de dichas tareas, difieren significativamente desde el punto de vista sintáctico. De esta manera, el sistema puede clasificar en niveles primarios del ranking, tareas que semánticamente son muy semejantes, pero que sintácticamente no lo son tanto. Es posible ilustrar lo anterior con el siguiente ejemplo: considere las tareas cuyos atributos se consignan en la Tabla 5. Adviértase que, desde el punto de vista sintáctico, los parámetros que definen la funcionalidad de ambas tareas (a excepción del caso de las salidas), difieren significativamente, en consecuencia, para un evaluador estas tareas pueden resultar poco semejantes. Sin embargo, en esencia (semánticamente) ambas tareas desempeñan la misma operación, por lo que el mecanismo propuesto estima una similitud considerable y ubica la tarea identificada como “*Order Validation*” en uno de los primeros niveles del ranking de la tarea “*Receive Request*”.

Tabla 4. Ejemplo Comparación de Tareas

	Identificador (Enriquecimiento)	Entradas (Enriquecimiento)	Salidas (Enriquecimiento)
Tarea de Consulta	<i>Receive Request (etom#orderHandling)</i>	<i>serviceRequest (sid#AccessServiceRequest)</i>	<i>serviceOrder (sid#ServiceOrder)</i>
Tarea Publicada	<i>Order Validation (etom#orderHandling)</i>	<i>customerOrder (sid#CustomerOrder)</i>	<i>serviceOrder (sid#ServiceOrder)</i>
Similitud Semántica	1.0	0.586	1.0

Finalmente, es necesario señalar que todos los resultados a los cuales se ha hecho referencia en esta sección, obedecen a la configuración de parámetros del sistema para la cual se obtuvo las mejores medidas de calidad, en las pruebas de comparación y recuperación de tareas. Esta configuración se especifica a continuación en la Tabla 5:

Tabla 5. Configuración de parámetros del prototipo.

<i>wld</i>	<i>wln</i>	<i>wO</i>	$\beta(m>n)$		$\beta(n>m)$	
			<i>Comparación de entradas</i>	<i>Comparación de salidas</i>	<i>Comparación de entradas</i>	<i>Comparación de salidas</i>
45%	35%	20%	0.8	1	1	0.8

5.4. RESUMEN

En este capítulo se abordó una descripción detallada del proceso de desarrollo llevado a cabo, alrededor de la construcción de un sistema software que implementa en el mecanismo de comparación propuesto. De esta manera, se identificaron inicialmente los requisitos funcionales, a partir de los cuales fue posible definir los casos de uso del sistema y la arquitectura de referencia que permitió soportar el desarrollo del mismo.

Asimismo, se realizó una descripción del benchmark de referencia generado para soportar el proceso de evaluación y puesta a punto del sistema. Posteriormente, se definieron las medidas de desempeño, empleadas para determinar la calidad del procedimiento de comparación y recuperación de tareas adoptado por el mecanismo propuesto.

De acuerdo con estas medidas, el desempeño general del sistema es satisfactorio en cuanto a su capacidad para estimar la similitud entre dos tareas, ajustándose de manera aproximada al criterio de un evaluador con conocimiento del dominio de los procesos de negocio de telecomunicaciones. De esta manera, se demostró que un enfoque semántico de comparación, efectivamente disminuye la ambigüedad propia de sistemas de emparejamiento basados en criterios sintácticos, al mismo tiempo que permite optimizar el consumo de tiempo empleado en la identificación de relaciones de similitud, entre tareas de procesos de negocio de telecomunicaciones, lo cual evidencia la viabilidad y favorabilidad de la propuesta presentada en el presente trabajo de grado.

6. CONCLUSIONES Y TRABAJO FUTURO

En la actualidad los operadores del sector de telecomunicaciones deben enfrentarse a uno de los mercados más exigentes del mundo, requiriendo estar a la vanguardia de las herramientas y tecnologías, que les permitan satisfacer las cambiantes necesidades de los usuarios. Para lograr este objetivo, sus procesos de negocio están en frecuente evolución, por lo que necesitan ser gestionados y evaluados constantemente, para determinar los posibles cambios que pueden realizarse para mejorar su desempeño en cuanto a costos, calidad, servicio y tiempo.

En el presente trabajo de grado se ha abordado esta necesidad, proponiendo un mecanismo de comparación de tareas entre procesos de negocio de telecomunicaciones, basado en el cálculo de su similitud semántica. Este mecanismo permite al operador realizar la identificación de tareas, que pueden adaptarse y reutilizarse para satisfacer nuevas necesidades, con mayor rapidez y con precisión aceptable, contribuyendo en gran medida con la optimización del consumo de recursos involucrado en el despliegue de procesos de negocio.

El mecanismo aquí propuesto, está basado en algunos trabajos de investigación relacionados con el descubrimiento de servicios web semánticos, los cuales fueron adaptados al ámbito de las tareas pertenecientes a procesos de negocio de telecomunicaciones, y permitieron generar una herramienta de comparación automática (Semantic Task Comparator) que favorece el análisis, gestión y reingeniería de procesos dentro del operador.

Este capítulo describe inicialmente las principales contribuciones del trabajo realizado, posteriormente presenta las conclusiones a las que se llegó durante su desarrollo, y finalmente propone los trabajos futuros.

6.1. CONTRIBUCIONES

Entre las principales contribuciones de éste proyecto de grado se destacan las siguientes:

- La adaptación de dos ontologías del dominio de telecomunicaciones, basadas en los modelos eTOM y SID de la NGOSS, denominadas SeTOM y SSID, las cuales fueron analizadas y aplicadas al proyecto para realizar el enriquecimiento de las tareas de procesos de negocio de telecomunicaciones. La adaptación se realizó a partir de definición de la relación existente entre el conocimiento dispuesto en dichas ontologías y los atributos que definen la funcionalidad de las tareas (identificador, entradas y salidas).
- La definición de un mecanismo de comparación semántica de tareas, entre procesos de negocio del dominio de telecomunicaciones descritos con BPMO, el cual está basado en el cálculo de la similitud semántica, entre los parámetros que describen dichas tareas. Este mecanismo se compone de una serie de algoritmos,

que comprenden procesos de emparejamiento y cálculo de distancia y similitud entre conceptos ontológicos, los cuales fueron propuestos a partir del estudio de algoritmos aplicados en el descubrimiento de servicios web semánticos.

- El prototipo *Semantic Task Comparator*, que constituye una herramienta automática para la comparación de tareas de procesos de negocio de telecomunicaciones, y fue definida a partir del mecanismo antes mencionado. Esta herramienta permite realizar comparaciones entre una o varias tareas de un proceso de consulta, con todas las tareas disponibles en el operador, permitiendo establecer no solo la similitud semántica entre las tareas comparadas, sino también un ranking de tareas, de acuerdo a dichos valores de similitud. Los resultados de esta herramienta, pueden contribuir a la disminución del tiempo de despliegue de nuevas soluciones en las empresas del sector, en la medida en que disminuyen el tiempo que requiere la ejecución de dichas comparaciones, independizando esta actividad de la interpretación subjetiva del personal técnico que realiza las búsquedas.
- La plataforma *Business Process Task Comparison Tool*, dedicada a la evaluación y comparación manual de tareas de procesos de negocio, a partir del análisis y estimación de la similitud entre sus atributos (identificador, entradas y salidas). Esta plataforma implementa una metodología de *Benchmarking* en la cual se genera manualmente un *Benchmark de Referencia*, a partir del criterio de varios expertos del dominio de telecomunicaciones. Con los resultados obtenidos de esta herramienta, es posible realizar un análisis estadístico basado en medidas como *recall*, *precision*, *overall*, *k-Precision* y *P-Precision*.
- El modelado en BPMO, de un banco de 20 procesos de negocio de telecomunicaciones, los cuales fueron posteriormente enriquecidos con conceptos dispuestos en dos ontologías propias este dominio. Estos procesos en conjunto proporcionan 97 tareas, y sirvieron de base para las comparaciones realizadas por parte de los expertos, en la herramienta *Business Process Task Comparison Tool*, y en la plataforma *Semantic Task Comparator*.
- Finalmente el aporte de la plataforma *Semantic Task Comparator* al trabajo de grado de maestría “Descubrimiento Automático de Procesos de Negocio Basado en Semántica del Comportamiento” del Ing. Cristhian Figueroa Martínez, desarrollado en la Universidad del Cauca. Esta herramienta sirve de soporte para el módulo Analizador Semántico, definido en su arquitectura. De esta manera se contribuye a la investigación en torno a la temática del descubrimiento de procesos de negocio, soportado en la aplicación de ontologías de dominio específicas del sector de las telecomunicaciones, con el objetivo de incentivar futuros trabajos dentro del grupo Ingeniería Telemática (GIT) de la Universidad del Cauca, integrando las líneas de investigación relacionadas con aplicaciones y servicios en internet y servicios avanzados de telecomunicaciones.

6.2. CONCLUSIONES

Dentro del proyecto realizado, se ha abordado la temática de comparación de tareas entre procesos de negocio de telecomunicaciones descritos con BPMO, la cual fue realizada, a partir del análisis semántico, no solo de sus entradas y salidas, como usualmente se realiza en actividades de descubrimiento de servicios, sino añadiendo un atributo importante, que permitió obtener una visión más completa de las operaciones o funcionalidades que se llevan a cabo en dichas tareas: el identificador. Este análisis, permitió proponer un mecanismo de comparación basado en el cálculo de la similitud semántica, entre conceptos que definen el enriquecimiento semántico de las tareas, el cual se soporta en la utilización de ontologías del dominio de telecomunicaciones, que fueron adaptadas al proyecto para realizar dicho enriquecimiento.

Al finalizar este proyecto se entregó una plataforma denominada *Semantic Task Comparator*, la cual se basa en el mecanismo antes mencionado, permitiendo realizar comparaciones entre una o varias tareas de un proceso de consulta, con todas las que componen los procesos disponibles en el operador, con el fin de establecer una clasificación o ranking, de acuerdo al valor de similitud semántica existente entre las tareas comparadas. Esta plataforma permite también definir el porcentaje de contribución de cada uno de los atributos de las tareas, sobre el valor de similitud semántica general, y entrega información, no solo de dicho resultado, sino de los valores obtenidos en la comparación de cada atributo, y de las entradas o salidas que podrían modificarse para aproximarse más a los requerimientos de la tarea de consulta.

Lo anterior, permite concluir que el trabajo realizado, facilita el descubrimiento de procesos de negocio de un operador de telecomunicaciones, a partir del enriquecimiento y comparación de las tareas que los componen, teniendo en cuenta criterios semánticos que permiten establecer de manera fiable su similitud.

Finalmente, a continuación se describen las principales conclusiones que surgieron durante la ejecución del presente proyecto:

- Para facilitar el entendimiento y gestión eficiente de los procesos de negocio, dentro de las empresas de telecomunicaciones, es importante contar con un lenguaje que permita enriquecer semánticamente las descripciones de dichos procesos, con elementos que los hagan comprensibles, tanto para el personal técnico experto en TI, como para los profesionales del negocio. Una alternativa muy interesante, es utilizar BPMO para la definición de los procesos, ya que esta plataforma, permite no solo modelar dichos procesos (al estar fundamentada en BPMN), sino también enriquecerlos con una ontología, que favorece un entendimiento común de las necesidades y los objetivos del negocio en todos los niveles de la empresa, y permite reducir la ambigüedad en su interpretación.
- Un requisito fundamental para asegurar la fiabilidad en los resultados obtenidos de las comparaciones, es garantizar que el enriquecimiento semántico de las tareas, sea realizado de manera adecuada.

- Para esto, es necesario tener en cuenta que dicho enriquecimiento, debe estar soportado en herramientas computacionalmente comprensibles como las ontologías, que proporcionan un vocabulario común dentro de un dominio, y aseguran la interpretación correcta del conocimiento, reduciendo ambigüedades conceptuales y terminológicas. Para favorecer la interoperabilidad entre las aplicaciones y sistemas de gestión, estas ontologías, en lo posible deben estar fundamentadas en modelos ampliamente reconocidos. En el caso particular del sector de las telecomunicaciones, una buena alternativa son las propuestas de la NGOSS: eTOM y SID, que constituyen la base fundamental para las ontologías SeTOM y SSID de YATOSP.
- Otro aspecto importante, consiste en asegurar que los conceptos ontológicos se hayan asociado fielmente a los atributos que definen la funcionalidad de las tareas, en otras palabras, que los conceptos de la ontología, asociados al identificador, entradas y salidas, reflejen efectivamente el sentido o la esencia de dichos atributos.
- El estudio de los modelos eTOM y SID, permitió definir la relación existente entre el conocimiento dispuesto en las ontologías de dominio elaboradas a partir de ellos (SeTOM y SSID), y los atributos que definen la funcionalidad de las tareas que conforman procesos de negocio del dominio de las telecomunicaciones. De esta manera, se hizo posible la adaptación de dichas ontologías al mecanismo de enriquecimiento semántico de tareas, mediante la asociación de los conceptos de la ontología SeTOM a sus identificadores (dado que ambos representan actividades que se llevan a cabo dentro de los procesos de operador), y los conceptos de la ontología SSID a sus entradas y salidas (en tanto que corresponden a la información que se comparte entre las tareas de un proceso de negocio, es decir a los objetos del negocio que se intercambian entre dichas tareas).
- Dada la escasez de documentación respecto a mecanismos de comparación semántica de tareas, fue necesario recurrir a otras líneas de investigación afines, con el propósito de adaptar tales propuestas al campo de aplicación de interés. Una alternativa adecuada, se identificó al tener en cuenta la correspondencia existente entre los principales atributos que describen la funcionalidad de las tareas, y los parámetros definidos para los servicios web (identificador o nombre, entradas y salidas). De esta manera, resultó de gran interés el estudio del estado actual de los trabajos de investigación, relacionados con el descubrimiento de servicios web semánticos, haciendo especial énfasis en los algoritmos de emparejamiento, distancia y similitud semánticos aplicados sobre los conceptos que enriquecen sus descripciones, los cuales se soportan en la utilización de ontologías.
- En cuanto a los mecanismos empleados para el emparejamiento de servicios web, se identificaron tres enfoques (descritos en la sección 2.3.1.1. de esta monografía), de los cuales se consideró adecuado aplicar, en primera instancia, un algoritmo de emparejamiento basado en categorías, que permitiera caracterizar cualitativamente la correspondencia existente entre

tareas. Y posteriormente, cuantificar estas relaciones, y emplear criterios de clasificación (ranking), que permitieran reflejar la capacidad de las tareas para satisfacer un requerimiento de otra, mediante la definición de similitudes semánticas entre estas.

- A partir del estudio de los mecanismos para cuantificar la distancia y similitud entre conceptos (sección 2.3.1.2), se logró identificar dos enfoques: uno basado en la información contenida en un documento descriptor (cuerpo de conocimiento), y otro basado en el conocimiento codificado en ontologías. Dadas las características del presente proyecto, se consideró que el enfoque más adecuado para determinar la similitud semántica entre tareas, es aquel que está basado solo en la ontología, en tanto es el que más se aproxima a los resultados de la opinión de los expertos del dominio, dada su simplicidad, escalabilidad, y a que es independiente del tamaño, la calidad y la disponibilidad del documento de texto. Dentro de este enfoque, el mejor mecanismo para calcular la similitud, se trata de aquel que considera la contribución de cada segmento del trayecto que hay entre los dos conceptos comparados, el cual además de considerar las relaciones dadas por la jerarquía de la ontología, considera también la influencia de la profundidad de la localización de los conceptos, en el resultado de similitud.
- Las pruebas de rendimiento realizadas sobre la plataforma *Semantic Task Comparator* permiten concluir, que el tiempo que toma realizar las comparaciones entre tareas es considerable, debido a que dicho procedimiento involucra un gran número de operaciones complejas y dispendiosas, relacionadas principalmente con el razonamiento sobre las ontologías de soporte, y el manejo de arreglos de objetos. Sin embargo, en contraste al tiempo que toma realizar dicho procedimiento de forma manual, los resultados de estas pruebas son muy pequeños, lo cual puede considerarse como una mejora en la disminución del tiempo de despliegue de nuevas soluciones en las empresas del sector.
- Los resultados entregados por la plataforma *Business Process Task Comparison Tool* permiten concluir (de acuerdo con las medidas de *Precision*, *Recall*, *Overall*, *k-precision* y *P-Precision*) que el mecanismo propuesto para determinar la similitud semántica de tareas, entre procesos de negocio de telecomunicaciones (implementado a través de la prototipo *Semantic Task Comparator*), se ajusta de manera aproximada al criterio de un evaluador con conocimiento de este dominio. Lo anterior demuestra que dicho mecanismo, garantiza un cálculo de similitud de gran calidad, lo que evidencia la viabilidad y favorabilidad de la propuesta presentada, en la comparación semántica de tareas, y por ende en el descubrimiento de procesos de negocio del dominio de los servicios de telecomunicaciones.

6.3. TRABAJOS FUTUROS

Esta tesis puede considerarse como un estudio inicial de mecanismos de comparación semántica de tareas, que aporta soluciones al problema de descubrimiento de servicios y procesos de negocio del dominio de telecomunicaciones. De esta manera, en el campo de investigación del presente proyecto de grado, se proponen los siguientes trabajos futuros:

Diseño e implementación de una herramienta para el enriquecimiento de procesos de negocio

A partir de la necesidad de garantizar un adecuado enriquecimiento semántico de las tareas, de procesos de negocio de telecomunicaciones, como requisito para asegurar la fiabilidad en los resultados de la comparación, se propone desarrollar una herramienta que permita modelar y enriquecer los procesos dentro del sistema implementado, que facilite la asociación de los conceptos ontológicos, con todos los atributos que describen las tareas. Esto considerando que las herramientas existentes para este fin, se enfocan únicamente en el enriquecimiento de ciertos parámetros (como las entradas y salidas), lo que dificulta el enriquecimiento de otros componentes, útiles para el razonamiento sobre las tareas.

Análisis e inclusión de otros parámetros al mecanismo de comparación propuesto

El mecanismo de comparación propuesto en el presente proyecto, tiene en cuenta los principales atributos que definen la funcionalidad de las tareas (identificador, entradas y salidas), sin embargo, podría resultar de gran interés el estudio de otros parámetros, en la estimación de la similitud semántica, que permitan satisfacer aun mejor los requerimientos de las solicitudes (por ejemplo el tipo, recursos, tecnología empleada, etc). Y en este mismo sentido, ampliar el campo de investigación, hacia otros componentes de procesos de negocio, aparte de las tareas, con el fin de proponer una solución más completa al problema de descubrimiento de procesos.

Estudio y definición de un mecanismo de comparación de tareas que involucre criterios sintácticos y semánticos

Algunos trabajos de investigación en torno a descubrimiento de servicios web, proponen el uso de ambas técnicas en el proceso de emparejamiento, tanto la evaluación de las descripciones a partir de sus palabras clave, como el razonamiento semántico sobre los conceptos del enriquecimiento. Si bien se ha demostrado que de ellas, la más recomendada es la que involucra criterios semánticos, es posible adicionar el análisis sintáctico al mecanismo de comparación de tareas propuesto, con el fin de reforzar la precisión en los resultados obtenidos.

Evolución del mecanismo de comparación propuesto hacia el uso de múltiples ontologías

El mecanismo propuesto en este trabajo de grado, se basa en la comparación y razonamiento sobre conceptos pertenecientes a la misma ontología de dominio, en el caso de los identificadores, a términos dispuestos en SeTOM, y en las entradas y salidas

a SSID. Sin embargo, podría considerarse la posibilidad de enriquecer esos atributos con múltiples ontologías, ya que en algunos casos, los operadores prefieren utilizar ontologías propietarias, que se adaptan a sus modelos de negocio, para lo cual, se propone ampliar el campo de investigación al uso de mediadores de ontologías, que soporten este tipo de razonamiento.

Evolución del mecanismo de comparación propuesto hacia otros dominios

El conjunto de algoritmos propuestos dentro del presente proyecto, pueden ser evaluados y adaptados a otros dominios de aplicación (especialmente los algoritmos 1, 2, 5, 6 y 7), con el fin de proponer un mecanismo de comparación de tareas, que pueda ser implementado en distintos campos del conocimiento.

Experimentación de la plataforma Semantic Task Comparator en un ambiente real

Por último, se propone realizar la experimentación de la plataforma en un ambiente real, para determinar su desempeño, y validar su utilidad en actividades de gestión y reingeniería de procesos de un operador de telecomunicaciones.

REFERENCIAS

- Al-Mubaid, H and Nguyen, H. 2006.** *A Cluster-based Approach for Semantic Similarity in the Biomedical Domain.* In Proc. of IEEE Eng Med Biol Soc. 2006.
- Ayomi, B., et al. 2007.** *A Semantic Approach for Service Matching in Pervasive Environments.* s.l. : Universidad de Southampton, 2007.
- Bechhofer, S., Goble, C. and Horrocks, I. 2001.** *DAML+OIL is not Enough.* University of Manchester, Oxford. 2001.
- Bellur, U. and Vadodaria, H. 2008.** *Semantic Matchmaking Algorithms.* Department of Computer Science and Engineering, Indian Institute of Technology, Bombay. 2008.
- Benatallah, B., et al. 2003.** *Request Rewriting-based Web Service Discovery.* In Proc. of the International Symposium on Wearable Computers (ISWC 2003). 2003.
- Bianchini, D., De Antonellis, V. and Melchiori, M. 2006.** *Evaluating Similarity and Difference in Service Matchmaking.* Workshop on Enterprise Modeling and Ontologies for Interoperability (INTEROP-EMOI'06). Luxemburgo : s.n., 2006.
- Bianchini, D., et al. 2006.** *Ontology-based Methodology for e-Service discovery.* J. Information Systems (JIS), vol. 31, nos. 4/5, Special Issue on Semantic Web and Web Services. 2006. pp. 361–380.
- Bishop, B. and Fischer, F. 2008.** *IRIS - Integrated Rule Inference System.* University of Innsbruck. Austria : s.n., 2008.
- Blázquez, J., et al. 2006.** *Ordenación de Resultados en Consultas Semánticas sobre Bases de Datos.* Universidad Carlos III de Madrid. España : s.n., 2006.
- BPMI. 2008.** *Business Process Management Initiative (BPMI.org).* <http://www.bpmi.org/>. (Última actualización: Enero de 2008). 2008.
- Brickley, D. and Guha, R. 2004.** *RDF Vocabulary Description Language 1.0: RDF Schema.* W3C Recommendation. Disponible en <http://www.w3.org/TR/rdf-schema>. 2004.
- Budanitsky, A. and Hirst, G. 2006.** *Evaluating WordNet-based Measures of Semantic Distance.* Computational Linguistics. 2006.
- Burrows, M. and Pendlebury, M. 2008.** *eTOM, ITIL & ISO/IEC20000 Harmonization.* The IT service Management Forum (ITSMF UK) - BSM impact Limited. Reino Unido : s.n., 2008.

- Cabral, L., Norton, B. and Domingue, J. 2009.** *The Business Process Modelling Ontology*. Workshop: Semantic Business Process Management (SBPM 2009) at European Semantic Web Conference (ESWC 2009). Grecia : s.n., 2009.
- Castells, P. 2004.** *La web semántica* . Escuela Politécnica Superior, Universidad Autónoma de Madrid. Disponible en <http://www.ii.uam.es/~castells>. 2004.
- Çelik, D. and Elçi, A. 2006.** *Discovery and Scoring of Semantic Web Services based on Client Requirement(s) through a Semantic Search Agent*. IEEE International Workshop on Engineering Semantic Agent Systems (ESAS). 2006.
- Chillarege, R. 1999.** *Software Testing Best Practices*. Center for Software Engineering, IBM Research. 1999. IBM Technical Report.
- Choi, O., Han, S. and Abraham, A. 2005.** *Extended Semantic Web Services Model for Automatic Integrated Framework*. Dept. of Comput. Sci. & Eng., Chung Ang University. Seoul, Korea del Sur : s.n., 2005.
- Cimpian, E., Komazec, S., Lintner, D., Blamauer, C., Evenson, M. 2007.** *Business Process Modeling Ontology BPMO*. Sembiz Deliverable D1.3. Disponible en <http://www.sembiz.org/attach/D1.3.pdf>. 2007.
- Corcho, O. and Gómez-Pérez, A. 2000.** *A Roadmap to Ontology Specification Languages*. EKAW. 2000. pp. 80-96.
- Cuenca Grau, B., et al. 2009.** *OWL 2 Web Ontology Language: Profiles*. W3C Recommendation. Disponible en <http://www.w3.org/TR/2009/REC-owl2-profiles-20091027>. 2009.
- d'Amato, C. 2007.** *Similarity-based Learning Methods for the Semantic Web*. PhD Thesis, University of Bari. 2007.
- Dawidziuk, R. and Cieslak, P. 2007.** *A proof-of-concept ontology of telecommunication services*. SIMS Deliverable D3.5. final versión. 2007.
- de Bruijn, J., et al. 2005.** *The Web Service Modeling Language WSML*. WSML Deliverable D16.1v0.2. Disponible en <http://www.wsmo.org/TR/d16/d16.1/v0.2/>. 2005.
- de Francisco, D. and Grenon, P. 2009.** *Enhancing Telecommunication Business Process Representation and Integration with Ontologised Industry Standards*. Workshop: 4th International Workshop on Semantic Business Process Management at 6th Annual European Semantic Web Conference (ESWC 2009). Creta, Grecia. 2009.
- Delgado, A. and García-Rodríguez, I. 2009.** *Ontología para relacionar procesos de negocio y su realización como servicios*. Actas de los Talleres de las Jornadas de Ingeniería del Software y Bases de Datos, Vol. 3, No. 3. 2009.

- Devitt, A., Danev, B. and Matusikova, K. 2006.** *Ontology-driven Automatic Construction of Bayesian Networks for Telecommunication Network Management.* En The Second Formal Ontologies Meet Industry Workshop (FOMI 2006). 2006.
- Dinos, J. 2004.** *Arquitectura de un sistema basado en agentes para la recuperación de metadatos RDF en base a una ontología de documentos.* Universidad de Puerto Rico. Disponible en <http://grad.uprm.edu/tesis/dinosrojas.pdf>. 2004.
- Domingue, J. 2008.** *Deliverable 4.5 SUPER. sBPMN and sEPC to BPMO Translation.* <http://www.ip-super.org/res/Deliverables/M24/D4.5.pdf>. . 2008.
- Feier, C. and Domingue, J. 2005.** *WSMO Primer.* WSMO Working Draft D3.1v0.2. 2005.
- Fensel, D., et al. 2005.** *OWL DL vs. OWL Flight: Conceptual Modeling and Reasoning for the Semantic Web.* Fourteenth International World Wide Web Conference (WWW2005). 2005.
- Figuroa, C. and Corrales, J. 2010.** *Automatic Business Process Discovery Based on Behavioral Semantics. [Unpublished].* Grupo de Ingeniería Telemática, Universidad del Cauca. 2010.
- Fleck, J. 2003.** *Overview of the Structure of the NGOSS Architecture.* Disponible en <http://www.tmforum.org/browse.aspx?catID=884&linkID=28267&docID=2052>. 2003.
- Frankowski, J., Rubach, P. and Szczekocka, E. 2007.** *Collaborative Ontology Development in Real Telecom Environment.* Disponible en <http://subs.emis.de/LNI/Proceedings/Proceedings116/gi-proc-116-007.pdf> . 2007.
- García, F. 2007.** *Sistema basado en Tecnologías del Conocimiento para Entornos de Servicios Web Semánticos.* Departamento de Ingeniería de la Información y las Comunicaciones, Universidad de Murcia. 2007.
- Ge, J. and Qiu, Y. 2008.** *Concept Similarity Matching Based on Semantic Distance.* In Proc. of SKG2008. Southwest University. 2008.
- Golbreich, C. and Wallace, E. 2009.** *OWL 2 Web Ontology Language: New Features and Rationale.* W3C Recommendation. Disponible en <http://www.w3.org/TR/2009/REC-owl2-new-features-20091027>. 2009.
- Gomez-Perez, A., Fernandez-Lopez, M. and Corcho, O. 2004.** *Ontological Engineering: with examples from the areas of Knowledge Management, E-Commerce And The Semantic Web.* Springer. Londres : s.n., 2004.
- Grimm, S., et al. 2006.** *A Reasoning Framework for Rule-Based WSML.* Digital Enterprise Research Institute (DERI), University of Innsbruck. Austria : s.n., 2006.

- Gruber, T. 1993.** *A translation approach to portable ontology specifications.* Knowledge Acquisition. 1993.
- Hardman, L., Ossenbruggen, J and Rutledge, L. 2002.** *Hypermedia and the semantic web: a research agenda.* En: Journal of digital information. 2002.
- Hau, J., Lee, W. and Darlington, J. 2005.** *A Semantic Similarity Measure for Semantic Web Services.* Web Service Semantics Workshop. London eScience Centre. 2005.
- Hernández, L. and Villa, G. 2009.** *Limitaciones de la Web Actual - Futuro de la Web.* Vol. 2, No. 3. 2009.
- Heymans, S. 2007.** *Business Process Ontology Framework.* SUPER Deliverable 1.1. Disponible en <http://www.ip-super.org/res/Deliverables/M12/D1.1.pdf>. 2007.
- Horridge, M. and Patel-Schneider, P. 2009.** *OWL 2 Web Ontology Language: Manchester Syntax.* W3C Working Group Note. Disponible en <http://www.w3.org/TR/2009/NOTE-owl2-manchester-syntax-20091027/>. 2009.
- IEEE, Standard Upper Ontology Working Group.** IEEE P1600.1 de SUO WG. Disponible en <http://suo.ieee.org/>.
- Inkpen, D. and Désilets, A. 2005.** *Semantic Similarity for Detecting Recognition Errors in Automatic Speech Transcripts.* University of Ottawa. 2005.
- Jędrzejczak, A., et al. 2008.** *sBPMN and sEPC to BPMO Translation.* SUPER Deliverable 4.5. Disponible en <http://www.ip-super.org/res/Deliverables/M24/D4.5.pdf>. 2008.
- Keller, U., Lara, R. and Polleres, A. 2004.** *WSMO Web Service Discovery.* WSMO deliverable D5.1. Disponible en <http://www.wsmo.org/2004/d5/d5.1/v0.1>. 2004.
- Kwiatkowski, K. and D., Gourdelier. 2007.** *Convergence in Telecommunication Whitepaper.* Comarch Information Technology. Inglaterra : s.n., 2007.
- Lambea, J. 2009.** *Business SLA Management.* Deliverable D.A2. Proyecto SLA@SOI. 2009.
- Lausen, H., et al. 2005.** *Wsmi - A Language Framework for Semantic Web Services.* Digital Enterprise Research Institute (DERI). 2005.
- Lee, W., et al. 2008.** *Comparison of Ontology-based Semantic-Similarity Measures.* Center for Biomedical Informatics and Research, Stanford University. 2008.
- Li, L. and Horrocks, I. 2003.** *A software Framework for Matchmaking based on Semantic Web Technology.* In Proc. of WWW'03. Budapest : s.n., 2003.

- Lin, D. 1998.** *An information-theoretic definition of similarity.* In Proc. of (ICML-98). 296–304. 1998.
- Lord, P., et al. 2003.** *Semantic Similarity Measures as Tools for Exploring the Gene Ontology.* University of Manchester. Pacific Symposium on Biocomputing. 2003.
- Lu, J., Wang, S. and Wang, J. 2005.** *An Experiment on the Matching and Reuse of XML Schemas.* In Proc. of ICWE-2005. 2005.
- Manola, F. and Miller, E. 2004.** *RDF Primer.* W3C Recommendation. Disponible en <http://www.w3.org/TR/2004/REC-rdf-primer-20040210>. 2004.
- Martinez, J. and Pérez, N. 2008.** *YATOSP: Marco de referencia semántico para el sector Telco.* Telecom I+D Conference, Bilbao. 2008.
- McGuinness, D.L. and van Harmelen, F. 2004.** *OWL Web Ontology Language: Overview.* Disponible en <http://www.w3.org/TR/owl-features/>. 2004.
- Motik, B., Patel-Schneider, P. and Parsia, B. 2009.** *OWL 2 Web Ontology Language: Structural Specification and Functional-Style Syntax.* W3C Recommendation. Disponible en <http://www.w3.org/TR/2009/REC-owl2-syntax-20091027/>. 2009.
- NGOSS. 2002.** *NGOSS Architecture Technology Neutral Specification, TMF 053.* Public Evaluation Version 2.5, The TeleManagement Forum. 2002.
- Okkyung, C. and Sangyong, H. 2008.** *Flexible Rule-based Web Services System for users' preferences.* In Proceedings of the 2008 4th International Conference on Next Generation Web Services Practices NWESP'08. 2008.
- OMG. 2008.** *Business Process Modeling Notation.* OMG Available Specification. V1.1. Disponible en <http://www.omg.org/spec/BPMN/2.0>. 2008.
- Paolucci, M., et al. 2002.** *Semantic Matching of Web Service Capabilities.* In: Proc. Of Int. Semantic Web Conference (ISWC,2002). Italia : s.n., 2002.
- Pérez, N., et al. 2007.** *Gestión de Procesos de Negocio Semánticos.* Telefónica Investigación y Desarrollo. 2007.
- Qiao, X. and Li, X. 2010.** *Semantic Web Application for Telecommunication Services.* Project "Service Intelligence for Convergent Network Environment". Institute of Network Technology in Beijing University of Posts and Telecommunications. Disponible en <http://int.bupt.edu.cn/jsp/centers/bupt506/intro.htm>. 2010.
- Resnik, P. 1993.** *Selection and Information: A Class-Based Approach to Lexical Relationships.* Ph.D Dissertation. 1993.

- Resnik, P. 1995.** *Using information content to evaluate semantic similarity.* In Proc. of IJCAI'95. Montreal, Canada : s.n., 1995.
- Rodriguez, R. and Garcia, L. 2008.** *La Gestión de los Procesos de Negocio en las Empresas de Telecomunicaciones.* Santa Clara, Cuba : s.n., 2008.
- Rój, M. 2007.** *Techniques for ontologydriven semantic interface artefacts.* SIMS Deliverable D3.4. final version. 2007.
- Rojas, L and Suarez, L. 2010.** *Descubrimiento de Servicios en Ambientes Ubicuos. [Trabajo de Grado en proceso de evaluación].* Grupo de Ingeniería Telemática, Universidad del Cauca. 2010.
- Samper, J. 2005.** *Ontologías para Servicios Web Semánticos de Información de Tráfico: Descripción y Herramientas de Explotación.* Universidad De Valencia. 2005.
- Serrano, C. 2005.** *Modelo de Construcción de Soluciones.* Universidad del Cauca. 2005.
- Silva, L. 2002.** *Representación de ontologías en la web semántica.* Instituto de Informática, Universidade Federal do Rio Grande do Sul. Brasil. 2002.
- TeLQAS.** *Proyecto TeLQAS: Telecommunication Litration Question Answering System.* Disponible en <http://homepages.mcs.vuw.ac.nz/~kouros/projects/telqas>. Última actualización: 11 de Septiembre de 2009.
- TMF - eTOM. 2005.** *Enhanced Telecom Operations Map (eTOM) – The business process framework.* International Telecommunication Union. ITU-T Recommendation M.3050.1. Ginebra, Suiza : s.n., 2005.
- TMF - SID. 2005.** *Information Framework (SID) In Depth.* Disponible en <http://www.tmforum.org/InformationFramework/6647/home.html>. 2005.
- Van Heijst, G., Schreiber, A. and Wielinga, B. 1997.** *Using explicit ontologies in KBS development.* International Journal of Human-Computer Studies, 45: 183-292. 1997.
- Villa, G. and Hernández, L. 2009.** *Limitaciones de la Web Actual Futuro de la Web.* Instituto Politecnico Naciona. Distrito Federal, México : s.n., 2009.
- W3C. 2010.** W3C. "Guía Breve de Servicios Web". [En línea] 06 de 05 de 2010. [Citado el: 13 de 05 de 2010.] <http://www.w3c.es/Divulgacion/GuiasBreves/ServiciosWeb>.
- W3C, OWL Working Group. 2004.** *W3C Recommendation: OWL Web Ontology Language: Overview.* <http://www.w3.org/TR/2004/REC-webont-req-20040210/>. 2004.
- Wahle, S. 2008.** *Network Domain Federation - An Architectural View on How to Federate Testbeds.* Future Internet Research and Experimentation. FIRE Launch Event, Paris City Hall. 2008.

Watkins, S., et al. 2005. *WP 8: Case Study B2B in Telecommunications: Prototype Platform Design*. Project DIP. Deliverable D8.3. . 2005.

Weske, M. 2007. *BPM Concepts, Languages, Architectures*. Springer, ISBN: 978-3-540-73521-2. 2007.

WfMC. 1999. *Terminologia & Glosario*. WfMC. U.K. : s.n., 1999.

Xu, B., et al. 2006. *A Semantic Matchmaker for Ranking Web Services*. *Journal of Computer Science & Technology*. 21(4):574-581. 2006.

Yatskevich, M. 2003. *Preliminary Evaluation of Schema Matching Systems*. Department of Information and Communication Technology, University of Trento. 2003.

ANEXO A

ANEXO A

MODELO DE DESPLIEGUE Y MODELO DE DISEÑO DEL SISTEMA

Este anexo aborda la documentación correspondiente a la definición y descripción de un conjunto de diagramas que especifican la Vista de Diseño y Despliegue del Sistema. De acuerdo con la referencia metodológica adoptada, este documento forma parte del macrocomponente denominado *Estructura para la Descripción del Sistema*, el cual involucra la especificación de artefactos esenciales (Modelo de Casos de Uso, Arquitectura de Referencia, Modelo de Diseño, Modelo de Despliegue y Plan de Pruebas) para el entendimiento adecuado de la funcionalidad y el comportamiento esperados del sistema/solución.

A.1. Diagrama de Paquetes del Sistema

La Figura 31 define el conjunto de paquetes en los cuales se distribuye el software del sistema desarrollado. Tal como se observa en la gráfica, dichos paquetes se encuentran dispuestos de acuerdo con la separación establecida en la arquitectura de referencia descrita en el Capítulo 5 de la monografía. La definición de cada uno de los paquetes que conforman la aplicación del sistema se aborda a continuación:

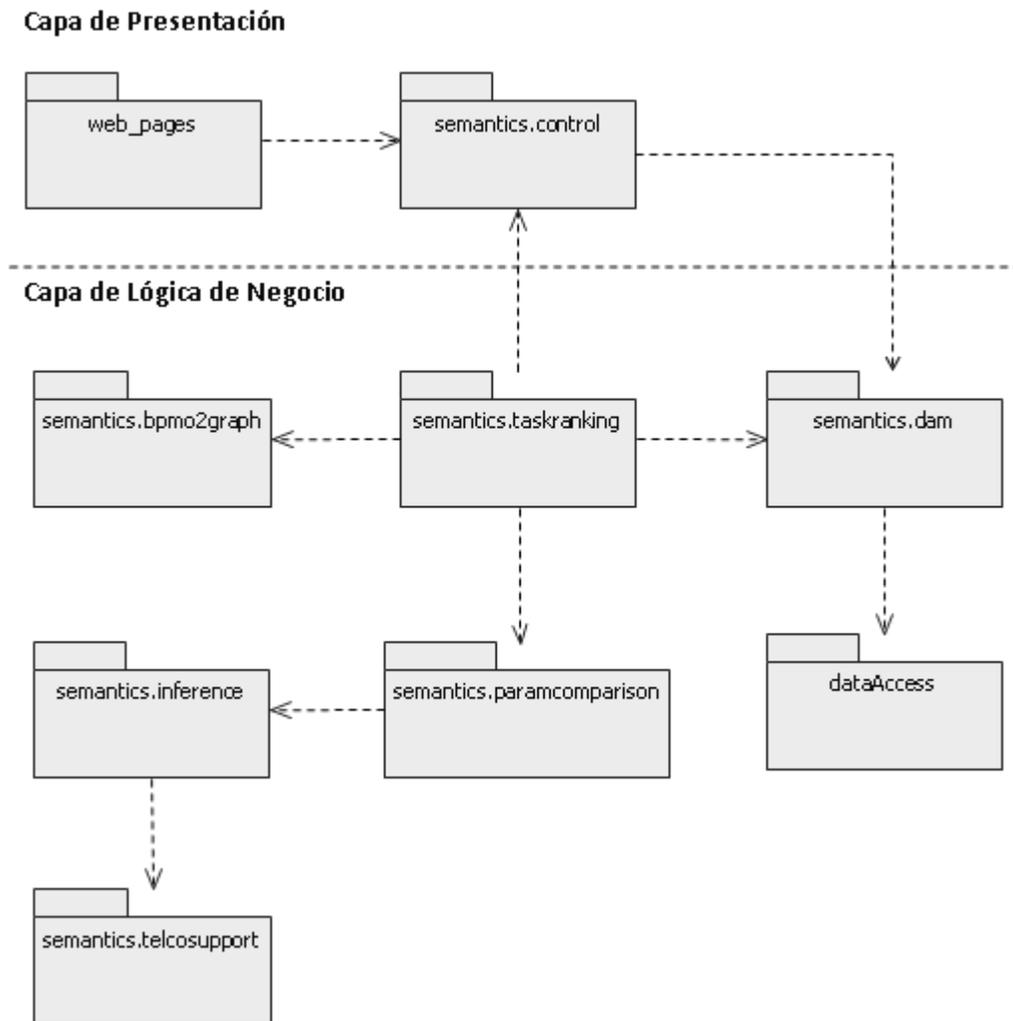
A.1.1. Capa de Presentación.

Esta capa de la arquitectura de referencia se implementa mediante los siguientes paquetes:

- *web_pages*: implementa las interfaces gráficas de usuario que componen la aplicación web. Fundamentalmente, consiste en un conjunto de páginas *jsp* y dos archivos xml de configuración: el archivo descriptor de despliegue de la aplicación (*web.xml*) y el archivo que especifica el control de la navegación (*faces-config.xml*).
- *semantics.control*: constituye un elemento mediador entre la Lógica de Presentación y la Lógica de Negocio, el cual garantiza la captura adecuada de los datos suministrados por el usuario en su interacción con la aplicación web, y de la

misma manera, que la información que se despliega ante él corresponda apropiadamente con sus peticiones.

Figura 31. Diagrama de Paquetes del Sistema



A.1.2. Capa de Lógica de Negocio.

Los paquetes que implementan esta capa de la arquitectura de referencia, representan las funcionalidades más importantes del sistema propuesto. En esta sección se realiza una descripción de cada uno de ellos:

- *semantics.bpmo2graph*: este paquete contiene el software encargado de llevar a cabo la transformación a un modelo de grafos de los procesos de negocio BPMO (publicados y de consulta), así como la discriminación de los nodos correspondientes a las tareas que los constituyen.
- *semantics.taskranking*: este componente se ocupa de articular el ranking de tareas publicadas más relevantes para cada una de las tareas que hacen parte del proceso de negocio de consulta, partiendo de la estimación del valor de su similitud semántica general.
- *semantics.inference*: en este paquete se implementan los componentes de la agrupación denominada “*Inferencia Semántica*”, la cual hace parte de la arquitectura definida para el comparador semántico de tareas, descrita en el Capítulo 5 de la monografía. Dichos componentes se encargan de determinar la similitud semántica entre los conceptos que enriquecen las tareas comparadas, soportándose en operaciones de razonamiento sobre las ontologías de dominio de telecomunicaciones *SeTOM* y *SSID*.
- *semantics.telcosupport*: este paquete comprende los archivos en formato wsml que describen las ontologías de dominio de soporte para llevar a cabo la comparación semántica de tareas entre procesos de negocio del dominio de las telecomunicaciones.
- *semantics.dam*: (*dam*: **d**ata **a**ccess **m**anagement) el software que integra esta aplicación se encarga de gestionar la conexión con la base de datos de usuarios y soportar las operaciones consulta y actualización de la información que en ella se registra.
- *dataAccess*: este componente provee un mecanismo de mapeo de los registros de la base de datos de usuarios a un modelo de objetos, facilitando de esta manera su gestión.

A.2. Diagrama de Clases del Sistema

La Figura 32 presenta el Diagrama de Clases del sistema desarrollado. Este diagrama provee la vista estática de la aplicación, definiendo además las relaciones existentes entre las entidades que la componen. En esta sección se realiza la descripción de cada una de las clases ilustradas, junto con sus atributos y métodos más representativos.

Figura 32. Diagrama de Clases del Sistema

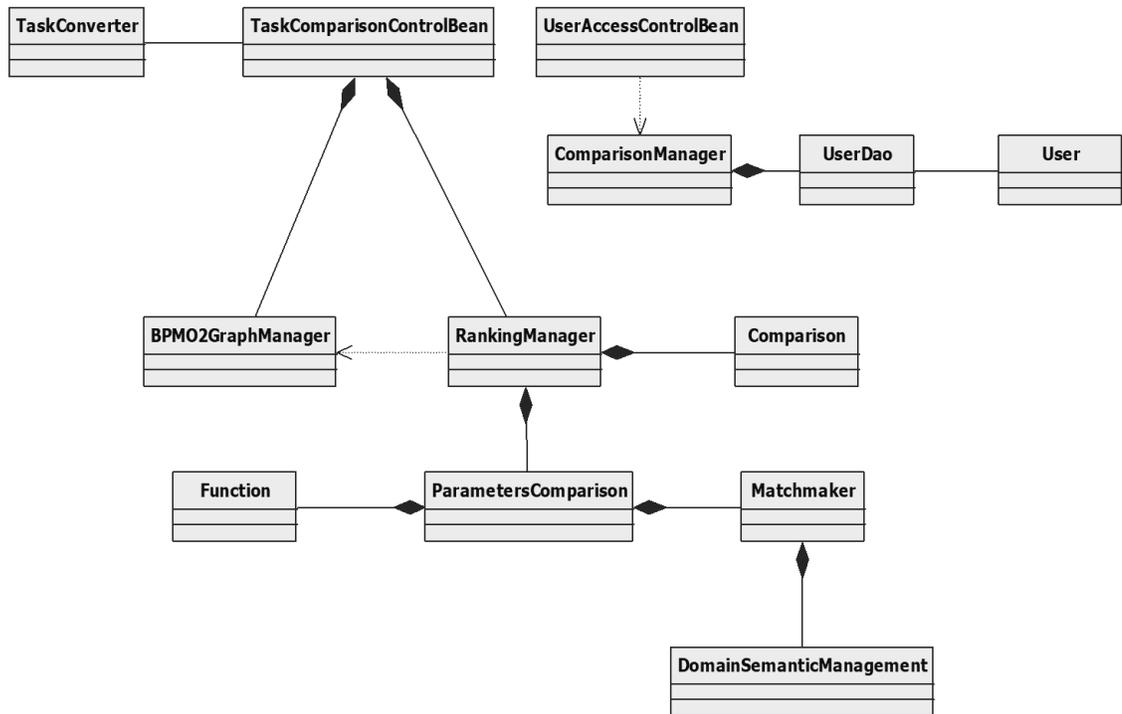


Tabla 6. Descripción de la Clase TaskConverter

Nombre
<i>TaskConverter</i> (paquete: <i>semantics.control</i>)
Descripción
Esta clase implementa dos métodos que permiten tratar tareas de un proceso de negocio como componentes desplegados (objetos de tipo String) dentro de la aplicación web.
Atributos
Ninguno
Métodos
<ul style="list-style-type: none"> • <i>getAsObject()</i>: a partir de una cadena de caracteres correspondiente al nombre de una tarea obtiene un objeto de tipo <i>Function</i> (Clase que representa la abstracción de una tarea en el modelo de grafos). • <i>getAsString()</i>: a partir de un objeto de tipo <i>Function</i> obtiene una cadena de caracteres desplegable en cualquiera de las vistas de la aplicación web.

Tabla 7. Descripción de la Clase TaskComparisonControlBean

Nombre
<i>TaskComparisonControlBean</i> (paquete: <i>semantics.control</i>)
Descripción
Esta Clase representa el Bean que lleva a cabo el control de las operaciones necesarias para la captura de las peticiones del usuario y el llamado a los métodos de las clases encargadas de realizar la comparación semántica entre tareas publicadas y de consulta.
Atributos
<ul style="list-style-type: none"> • <i>_WSMLQueryPath</i>: Cadena que especifica la ubicación del archivo <i>wsmi</i> del proceso de consulta, en el sistema de archivos del equipo servidor. • <i>dsm</i>: Objeto de tipo <i>DomainSemanticManagement</i> empleado para el soporte de las operaciones de inferencia sobre las ontologías de dominio. • <i>rankingManager</i>: Encargado de gestionar la generación del ranking de tareas relevantes para cada tarea de consulta. • <i>bpmo2GraphManager</i>: Encargado de realizar la conversión de los procesos BPMO a un modelo de grafos. • <i>inputPercentage</i>: variable de tipo entero que identifica el porcentaje de contribución de la similitud de las entradas de las tareas en el cálculo de la similitud general. • <i>outputPercentage</i>: variable de tipo entero que identifica el porcentaje de contribución de la similitud de las salidas de las tareas en el cálculo de la similitud general. • <i>idPercentage</i>: variable de tipo entero que identifica el porcentaje de contribución de la similitud de identificadores de las tareas en el cálculo de la similitud general. • <i>queryTasks</i>: conjunto de tareas del proceso de consulta. • <i>wsmiQFile</i>: referencia del archivo <i>wsmi</i> del proceso de consulta sin enriquecimiento semántico. • <i>wsmiEnQFile</i>: referencia del archivo <i>wsmi</i> del proceso de consulta enriquecido semánticamente. • <i>wsmiLayoutQFile</i>: referencia del archivo <i>layout</i> del proceso de consulta (contiene información para desplegar el diagrama del proceso de negocio en la aplicación web). • <i>comparisonType</i>: cadena que captura el tipo de comparación que el usuario desea realizar. El usuario elige entre comparar todas las tareas del proceso de consulta u omitir algunas de ellas. • <i>selectedQueryTasks</i>: conjunto de tareas de consulta seleccionadas por el usuario para someterlas al proceso de comparación semántica. • <i>taskRanking</i>: ranking de las tareas publicadas más relevantes para cada tarea de consulta. • <i>topK</i>: variable de tipo entero que especifica el número de niveles del ranking de tareas que se despliegan en la aplicación web.
Métodos

<ul style="list-style-type: none"> • <i>extractQueryTasks()</i>: empleando los métodos del atributo <i>bpmo2GraphManager</i> extrae las tareas de consulta del proceso entregado por el usuario. • <i>getSelectedQueryTasks()</i>: obtiene las tareas de consulta seleccionadas por el usuario para someterlas al proceso de comparación. • <i>extractTargetTasks()</i>: empleando los métodos del atributo <i>bpmo2GraphManager</i> extrae las tareas pertenecientes a los procesos de negocio publicados. • <i>setComparisonParameters()</i>: captura los parámetros de comparación definidos por el usuario en la interacción con las interfaces de la aplicación web. • <i>setRanking()</i>: establece el ranking de tareas que se despliega para el usuario de acuerdo con el valor del atributo <i>topk</i>. • <i>uploadProcessListener()</i>: captura los archivos correspondientes al proceso de negocio de consulta proporcionados por el usuario.

Tabla 8. Descripción de la Clase *UserAccessControlBean*

Nombre
<i>UserAccessControlBean</i> (paquete: <i>semantics.control</i>)
Descripción
Esta Clase representa el Bean que ejecuta las operaciones de validación y registro de usuarios en el sistema.
Atributos
<ul style="list-style-type: none"> • <i>user</i>: atributo de tipo <i>User</i> que contiene la información del usuario que desea acceder o registrarse en el sistema. • <i>cm</i>: atributo de tipo <i>ComparisonManager</i> que permite realizar consultas sobre la base de base de datos del sistema.
Métodos
<ul style="list-style-type: none"> • <i>validateUser()</i>: ejecuta una consulta para verificar la validez de los datos de <i>login</i> y <i>password</i> proporcionados por el usuario. • <i>addUser()</i>: adiciona un registro en la base de datos con la información de un usuario que desea darse de alta en el sistema. • <i>logout()</i>: gestiona el cierre de sesión en la aplicación web.

Tabla 9. Descripción de la Clase *ComparisonManager*

Nombre
<i>ComparisonManager</i> (paquete: <i>semantics.dam</i>)
Descripción
Esta Clase gestiona la conexión con la base de datos del sistema y permite la ejecución de operaciones de consulta y actualización de la información que en ésta se registra.
Atributos
<ul style="list-style-type: none"> • <i>uDao</i>: atributo de tipo <i>UserDao</i> que permite el mapeo de los registros de la tabla de usuarios de la base de datos, a objetos de tipo <i>User</i> (empleando

<p>el patrón DAO: <i>Data Access Object</i>).</p> <ul style="list-style-type: none"> • <i>cx</i>: atributo de tipo <i>Connection</i> que habilita la conexión con la base de datos del sistema.
Métodos
<ul style="list-style-type: none"> • <i>searchUser()</i>: retorna un objeto de tipo <i>User</i> correspondiente a uno de los usuarios registrados en el sistema, cuya información coincida con los datos proporcionados en el argumento de éste método. • <i>addUser()</i>: adiciona un registro en la base de datos con la información de un usuario que desea darse de alta en el sistema. • <i>deleteUser()</i>: elimina un registro de la base de datos correspondiente al usuario que se pasa como argumento. • <i>getUsers()</i>: obtiene un arreglo de objetos de tipo <i>User</i> correspondiente al conjunto de usuarios registrados en el sistema.

Tabla 10. Descripción de la Clase UserDao

Nombre
<i>UserDao</i> (paquete: <i>semantics.dataAccess</i>)
Descripción
Esta Clase implementa el patrón DAO que permite el mapeo de registros de la tabla de usuarios de la base de datos del sistema a objetos java, facilitando de esta manera la gestión de la información de las personas que acceden a la aplicación.
Atributos
Ninguno
Métodos
<ul style="list-style-type: none"> • <i>load()</i>: ejecuta una consulta en la base de datos para retornar el registro de un usuario particular del sistema. posteriormente mapea los datos de dicho registro sobre un objeto de tipo <i>User</i>. • <i>loadAll()</i>: obtiene todos los registros de los usuarios del sistema, los mapea sobre objetos de tipo <i>User</i> y retorna un arreglo conformado por los objetos mapeados. • <i>create()</i>: adiciona un registro en la base de datos con la información de un usuario que desea darse de alta en el sistema. • <i>delete()</i>: elimina un registro de la base de datos correspondiente al usuario que se pasa como argumento.

Tabla 11. Descripción de la Clase User

Nombre
<i>User</i> (paquete: <i>semantics.dataAccess</i>)
Descripción
Esta Clase representa la abstracción de un usuario del sistema. Los atributos la clase <i>User</i> corresponden con los campos definidos en la tabla de usuarios de la base de datos.
Atributos
<i>id</i> : atributo de tipo entero que representa el ID del usuario registrado en la base

<p>de datos del sistema. <i>login</i>: cadena de caracteres que identifica el nombre con el que un usuario accede al sistema a través de la aplicación web. <i>password</i>: cadena de caracteres que identifica la contraseña con la que un usuario accede al sistema a través de la aplicación web. <i>name</i>: cadena de caracteres correspondiente al nombre del usuario registrado. <i>email</i>: cadena de caracteres que especifica el correo electrónico del usuario registrado.</p>
Métodos
<ul style="list-style-type: none"> • <i>getters y setters</i>.

Tabla 12. Descripción de la clase BPMO2GraphManager

Nombre
<i>BPMO2GraphManager</i> (paquete: <i>semantics.bpmo2graph</i>)
Descripción
Esta Clase implementa las operaciones necesarias para realizar la transformación a un modelo de grafos de los procesos publicados y de consulta. A partir de los grafos obtenidos, extrae las nodos correspondientes a las tareas y los ordena en arreglos separados.
Atributos
<ul style="list-style-type: none"> • <i>queryProcessGraph</i>: grafo del proceso de negocio de consulta. • <i>_WSMLQueryPath</i>: cadena que especifica la ubicación del archivo <i>wsmi</i> del proceso de consulta, en el sistema de archivos del equipo servidor. • <i>availableProcessPath</i>: cadena que especifica la ubicación del directorio de proceso publicados, en el sistema de archivos del equipo servidor. • <i>processDomains</i>: arreglo de cadenas de caracteres que especifican los nombres de los dominios en los cuales se encuentran clasificados los procesos de negocio publicados. • <i>queryTasks</i>: conjunto de tareas del proceso de negocio de consulta. • <i>targetTasks</i>: conjunto integrado por de todas las tareas de los procesos de negocio publicados. • <i>dsm</i>: Objeto de tipo <i>DomainSemanticManagement</i> empleado para verificar la validez del enriquecimiento semántico de las tareas.
Métodos
<ul style="list-style-type: none"> • <i>getProcessGraph()</i>: obtiene el grafo de un proceso de negocio, a partir de la especificación de la ubicación del archivo <i>wsmi</i> que lo describe. • <i>extractQueryTasks()</i>: extrae las tareas del grafo del proceso de consulta. • <i>extractTargetTasks()</i>: extrae las tareas pertenecientes a los grafos de los procesos de negocio publicados.

Tabla 13. Descripción de la Clase RankingManager

Nombre

<i>RankingManager</i> (paquete: <i>semantics.taskranking</i>)
Descripción
Esta Clase ejecuta las operaciones necesarias para configurar el ranking de las tareas publicadas más similares a las tareas de consulta.
Atributos
<ul style="list-style-type: none"> • <i>queryTasks</i>: conjunto de objetos de tipo <i>Función</i> correspondientes a last areas de consulta. • <i>targetTasks</i>: conjunto de objetos de tipo <i>Función</i> correspondientes a last areas de publicadas. • <i>parametersComparator</i>: instancia de la clase <i>ParametersComparator</i> que permite estimar la similitud semántica entre las tareas comparadas. • <i>dsm</i>: Objeto de tipo <i>DomainSemanticManagement</i> empleado para el soporte de las operaciones de inferencia sobre las ontologías de dominio. • <i>comparisons</i>: arreglo de objetos de tipo <i>comparison</i> que almacena las comparaciones entre tareas de consulta y publicadas en orden ascendente de acuerdo a su valor de similitud (almacena las 10 comparaciones por cada tarea de consulta).
Métodos
<ul style="list-style-type: none"> • <i>setComparisonParameters()</i>: establece el valor de los porcentajes de contribución de las similitudes de entradas, salidas e identificadores en el cálculo de la similitud general de las tareas comparadas. • <i>compare()</i>: por medio del llamado a los métodos del atributo <i>parametersComparator</i> realiza la comparación semántica entre las tareas de consulta (cada uno de los objetos que componen el arreglo <i>queryTasks</i>) y las tareas publicadas (objetos que conforman el arreglo <i>targetTasks</i>). Al mismo tiempo, con base en los resultados de similitud de las comparaciones genera el ranking de tareas publicadas para cada tarea de consulta (registrando las parejas de tareas más similares en el arreglo <i>comparisons</i>).

Tabla 14. Descripción de la Clase Comparison

Nombre
<i>Comparison</i> (paquete: <i>semantics.taskranking</i>)
Descripción
Esta Clase permite registrar los detalles de la comparación semántica realizada entre dos tareas.
Atributos
<ul style="list-style-type: none"> • <i>queryTask</i>: instancia de la clase <i>Función</i> que representa tarea de consulta. • <i>targetTask</i>: instancia de la clase <i>Función</i> que representa tarea de publicada. • <i>overallSim</i>: variable de tipo <i>double</i> que especifica el valor de similitud semántica general entre la tarea de consulta y la tarea publicada. • <i>idSim</i>: variable de tipo <i>double</i> que especifica el valor de similitud semántica de identificadores existente entre la tarea de consulta y la tarea publicada. • <i>inputsSim</i>: variable de tipo <i>double</i> que especifica el valor de similitud

<p>semántica de entradas existente entre la tarea de consulta y la tarea publicada.</p> <ul style="list-style-type: none"> • <i>outputsSim</i>: variable de tipo <i>double</i> que especifica el valor de similitud semántica de salidas existente entre la tarea de consulta y la tarea publicada.
Métodos
<ul style="list-style-type: none"> • <i>getters</i> y <i>setters</i>.

Tabla 15. Descripción de la Clase ParametersComparator

Nombre
<i>ParametersComparator</i> (paquete: <i>semantics.paramcomparison</i>)
Descripción
Esta Clase implementa los algoritmos propuestos en el Capitulo 4 de la monografía para determinar los valores de similitud semántica de entradas, salidas e identificadores existentes entre dos tareas.
Atributos
<ul style="list-style-type: none"> • <i>match</i>: instancia de la clase <i>Matchmaker</i> que permite determinar la similitud entre conceptos que enriquecen los parámetros de las tareas comparadas. • <i>queryTask</i>: instancia de la clase <i>Function</i> que representa tarea de consulta. • <i>targetTask</i>: instancia de la clase <i>Function</i> que representa tarea de publicada. • <i>maximumsSims</i>: arreglo de objetos de tipo <i>Double</i> empleados en el cálculo de la similitud semántica de entradas y salidas. • <i>minimumsSims</i>: arreglo de objetos de tipo <i>Double</i> empleados en el cálculo de la similitud semántica de entradas y salidas. • <i>inputsQ</i>: arreglo de cadenas de caracteres correspondientes a los nombres de las entradas de la tarea de consulta. • <i>outputsQ</i>: arreglo de cadenas de caracteres correspondientes a los nombres de las salidas de la tarea de consulta. • <i>inputsT</i>: arreglo de cadenas de caracteres correspondientes a los nombres de las entradas de la tarea publicada. • <i>outputsT</i>: arreglo de cadenas de caracteres correspondientes a los nombres de las salidas de la tarea publicada. • <i>inputSimMatrix</i>: matriz de similitud empleada en el cálculo de la similitud general de entradas. • <i>outputSimMatrix</i>: matriz de similitud empleada en el cálculo de la similitud general de salidas. • <i>totalInputSim</i>: variable de tipo <i>Double</i> que especifica el valor de similitud de entradas existente entre la tarea de consulta y la tarea publicada. • <i>totalOutputSim</i>: variable de tipo <i>Double</i> que especifica el valor de similitud de salidas existente entre la tarea de consulta y la tarea publicada. • <i>totalIdSim</i>: variable de tipo <i>Double</i> que especifica el valor de similitud de identificadores existente entre la tarea de consulta y la tarea publicada. • <i>totalTaskSim</i>: variable de tipo <i>Double</i> que especifica el valor de similitud semántica general entre la tarea de consulta y la tarea publicada.

<ul style="list-style-type: none"> • <i>idSimWeigth</i>: variable de tipo <i>Double</i> que identifica el porcentaje de contribución de la similitud de identificadores de las tareas en el cálculo de la similitud general. • <i>inputSimWeigth</i>: variable de tipo <i>Double</i> que identifica el porcentaje de contribución de la similitud de entradas de las tareas en el cálculo de la similitud general. • <i>outputSimWeigth</i>: variable de tipo <i>Double</i> que identifica el porcentaje de contribución de la similitud de salidas de las tareas en el cálculo de la similitud general.
Métodos
<ul style="list-style-type: none"> • <i>compareTaskInputs()</i>: permite determinar la similitud semántica en cuanto a entradas existente entre las dos tareas que recibe como argumento. Para esto, implementa el procedimiento definido el Algoritmo 6 presentado en el Capítulo 4 de la monografía. • <i>compareTaskOutputs()</i>: estima la similitud semántica en cuanto a salidas existente entre las dos tareas que recibe como argumento. Para esto, implementa el procedimiento definido el Algoritmo 7 presentado en el Capítulo 4 de la monografía. • <i>compareTaskId()</i>: establece la similitud semántica en cuanto a identificadores existente entre las dos tareas que recibe como argumento. Para esto, implementa el procedimiento definido el Algoritmo 5 presentado en el Capítulo 4 de la monografía. • <i>determineTotalTaskSimilarity()</i>: pondera la similitud general de las tareas comparadas de acuerdo con los valores de los atributos <i>idSimWeigth</i>, <i>inputSimWeigth</i> y <i>outputSimWeigth</i>.

Tabla 16. Descripción de la Clase Matchmaker

Nombre
<i>Matchmaker</i> (paquete: <i>semantics.inference</i>)
Descripción
Esta Clase implementa los algoritmos propuestos en el Capitulo 4 de la monografía para determinar el grado de correspondencia, distancia y similitud semánticas entre conceptos pertenecientes a las ontologías de dominio de soporte (<i>SeTOM</i> y <i>SSID</i>), los cuales enriquecen los parámetros de las tareas comparadas.
Atributos
<ul style="list-style-type: none"> • <i>dsm</i>: Instancia de la clase <i>DomainSemanticManagement</i> empleada para el soporte de las operaciones de inferencia sobre las ontologías de dominio. • <i>semanticDistance</i>: variable de tipo <i>double</i> que representa el valor de la distancia semántica existente entre dos conceptos determinados. • <i>semanticSimilarity</i>: variable de tipo <i>double</i> que especifica el valor de la similitud semántica existente entre dos conceptos determinados. • <i>matchingDegree</i>: cadena de caracteres que identifica el grado de correspondencia en el cual se clasifican dos conceptos comparados. • <i>p</i>: variable de tipo <i>double</i> que representa el factor que determina la contribución del valor de distancia semántica en el cálculo de la similitud semántica para dos conceptos específicos.

<ul style="list-style-type: none"> • <i>superConceptsQ</i>: arreglo de cadenas de caracteres correspondientes a los nombres de los superconceptos del término perteneciente al enriquecimiento semántico de la tarea de consulta. • <i>superConceptsT</i>: arreglo de cadenas de caracteres correspondientes a los nombres de los superconceptos del término perteneciente al enriquecimiento semántico de la tarea de publicada.
Métodos
<ul style="list-style-type: none"> • <i>findMatchingDegree()</i>: permite determinar el grado de correspondencia entre dos conceptos dentro de una ontología de dominio, implementando los procedimientos establecidos en los Algoritmos 2 y 4 definidos en el Capítulo 4 de la monografía. • <i>findSemanticDistance()</i>: calcula la distancia semántica existente entre dos conceptos pertenecientes a la ontología de <i>SSID</i>, de acuerdo con el procedimiento establecido en el Algoritmo 1 propuesto en el Capítulo 4 de la monografía. • <i>findSemanticSimilarity()</i>: implementa el procedimiento establecido en el Algoritmo 5 para determinar el valor de similitud semántica existente entre dos conceptos pertenecientes a una de las ontologías de dominio. • <i>registerIdOntology()</i>: permite registrar la ontología de <i>SeTOM</i> en el razonador empleado para realizar operaciones de inferencia sobre esta ontología de dominio. • <i>doesExist()</i>: permite determinar si un término (cadena de caracteres que recibe como argumento) existe como concepto o instancia dentro de la ontología de <i>SeTOM</i>. Este método se implementa para validar el enriquecimiento de identificadores de las tareas.

Tabla 17. Descripción de la Clase DomainSemánticManagement

Nombre
<i>DomainSemanticManagement</i> (paquete: <i>semantics.inference</i>)
Descripción
Esta Clase se encarga de llevar a cabo toda la gestión y configuración necesarias para habilitar las operaciones de razonamiento sobre las ontologías de dominio de <i>SeTOM</i> y <i>SSID</i> .
Atributos
<ul style="list-style-type: none"> • <i>leFactory</i>: instancia de la clase <i>LogicalExpressionFactory</i> perteneciente al API de soporte <i>WSMO4j</i>, la cual habilita la creación de consultas descritas en sintaxis <i>WSML</i> para operaciones de inferencia sobre las ontologías de dominio. • <i>wsmoManager</i>: instancia de la clase <i>WSMO4JManager</i> perteneciente al API de soporte <i>WSML2Reasoner</i>, la cual habilita la creación de <i>Fabricas</i> de objetos que permiten implementar las diferentes entidades definidas dentro de <i>WSMO</i>. • <i>wsmIparser</i>: instancia de la clase <i>Parser</i>, la cual hace parte del API de soporte <i>WSML-API</i>. Este atributo hace posible el mapeo de las ontologías de dominio descritas en sintaxis <i>wsml</i>, a un modelo de objetos. • <i>sid</i>: referencia del objeto que representa la ontología <i>SSID</i>.

<ul style="list-style-type: none"> • <i>etom</i>: referencia del objeto que representa la ontología <i>SeTOM</i>. • <i>reasoner</i>: instancia de la clase <i>DatalogBasedWSMLReasoner</i> perteneciente al API de soporte <i>WSML2Reasoner</i>, la cual permite la ejecución de consultas (razonamiento) sobre las ontologías de dominio. • <i>sidLoc</i>: cadena de caracteres que especifica la ubicación del archivo <i>wsml</i> de la ontología de dominio <i>SSID</i>. • <i>etomLoc</i>: cadena de caracteres que especifica la ubicación del archivo <i>wsml</i> de la ontología de dominio <i>SeTOM</i>.
Métodos
<ul style="list-style-type: none"> • <i>parseAndLoadOntology()</i>: permite obtener un modelo de objetos de las ontologías de dominio descritas en <i>wsml</i>. • <i>getSuperConcepts()</i>: permite obtener el arreglo de superconceptos para un concepto particular perteneciente a una de las ontologías de dominio . • <i>sortArray()</i>: implementa un algoritmo para ordenar un arreglo de conceptos (de acuerdo con su ubicación en la jerarquía de las ontologías de dominio) del más específico al más general. • <i>isSubConcept()</i>: permite determinar si un concepto A es subconcepto de un concepto B (A y B se reciben como argumento). • <i>getReasoner()</i>: ejecuta la configuración del razonador empleado para las operaciones de inferencia. • <i>runQuery ()</i>: mediante el llamado de los métodos del razonador (<i>reasoner</i>) ejecuta una consulta sobre la ontología de dominio que se encuentre registrada en éste. • <i>registerOntologyId()</i>: permite registrar la ontología <i>SeTOM</i> en el razonador. • <i>registerOntologyI_O()</i>: permite registrar la ontología <i>SSID</i> en el razonador. • <i>doesExist()</i>: permite determinar si un término (cadena de caracteres que recibe como argumento) existe como concepto o instancia dentro de la ontología de <i>SeTOM</i>. Este método se implementa para validar el enriquecimiento de identificadores de las tareas. • <i>getNumComponents()</i>: obtiene el número de componentes que hacen parte de un concepto perteneciente a la ontología <i>SeTOM</i>.

La clase **Function** ilustrada en el diagrama de clases de la Figura 32 hace parte del API de soporte **WSMO2GRAPH**. Esta clase representa la abstracción de una tarea perteneciente a un proceso de negocio descrito en BPMO. Sus métodos y atributos se describen en la Tabla 18.

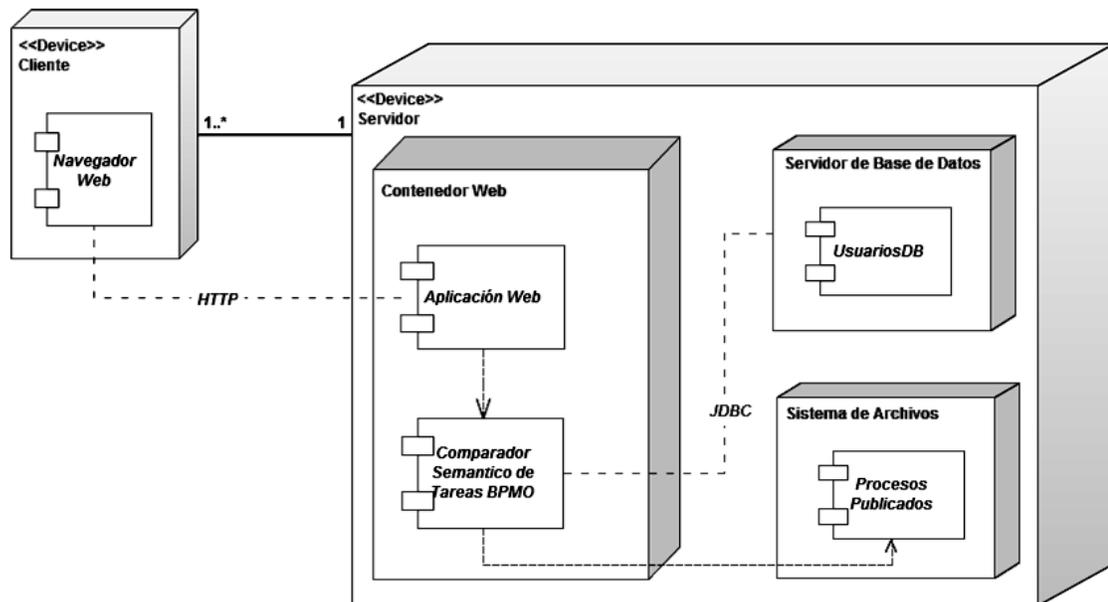
Tabla 18. Descripción de la Clase Function

Nombre
<i>Function</i> (paquete: <i>bpmo.graph.model</i>)
Atributos
<ul style="list-style-type: none"> • <i>name</i>: cadena de caracteres que identifica la tarea. • <i>iri</i>: identificador de la tarea. • <i>idEnTask</i>: cadena de caracteres que especifica el enriquecimiento semántico de identificador de la tarea. • <i>inEnTask</i>: arreglo de cadenas de caracteres que contiene el

<p>enriquecimiento semántico de entradas de la tarea.</p> <ul style="list-style-type: none"> • <i>outEnTask</i>: arreglo de cadenas de caracteres que contiene el enriquecimiento semántico de salidas de la tarea. • <i>inputs</i>: arreglo de cadenas de caracteres que contiene las salidas de la tarea. • <i>outputs</i>: arreglo de cadenas de caracteres que contiene las salidas de la tarea. • <i>homeProcess</i>: cadena de caracteres que identifica el proceso al cual pertenece la tarea.
Métodos
<ul style="list-style-type: none"> • <i>getters</i> y <i>setters</i>.

A.3. Diagrama de Despliegue del Sistema

Figura 33. Diagrama de Despliegue del Sistema



El diagrama de despliegue del sistema, presentado en la Figura 33, ilustra el conjunto de nodos en los cuales se distribuyen físicamente los artefactos implementados para soportar su operación. Tal como se muestra en el diagrama, el sistema está dispuesto en una configuración típica de Cliente/Servidor. De esta manera, tanto la lógica de negocio como la lógica de presentación están desplegadas dentro del Contenedor Web presente en el nodo correspondiente al equipo Servidor. En este nodo también se encuentran disponibles el banco de procesos publicados, el cual no es más que un directorio donde

se almacenan los archivos de estos procesos de negocio, y un gestor de base de datos para administrar la información de los usuarios registrados en el sistema. En cuanto al equipo Cliente, es necesario que esté en red con el Servidor y que cuente con un navegador web, por medio del cual el usuario pueda acceder al portal del sistema.

ANEXO B

ANEXO B

DESCRIPCIÓN DE LA HERRAMIENTA PARA LA COMPARACIÓN MANUAL DE TAREAS DE PROCESOS DE NEGOCIO.

Para generar el benchmark de referencia empleado en la evaluación del desempeño del sistema desarrollado, fue necesaria la construcción de una herramienta que permitiera, en primer lugar, llevar a cabo la comparación entre tareas a través de una interfaz gráfica de usuario, en segundo lugar, soportar el acceso simultáneo de varios evaluadores y en tercer lugar, almacenar los resultados de las comparaciones en una base de datos. Con base en estos requisitos se determinó construir una aplicación web, la cual se desarrolló a partir de la adaptación del prototipo implementado en el marco del trabajo de grado: **“Plataforma para la Evaluación de Sistemas de Recuperación de Servicios Basados en Comportamiento”** (Sandino, y otros, 2010) que se lleva a cabo en la Universidad del Cauca, el cual permite, entre otras funcionalidades, realizar comparaciones manuales entre procesos de negocio completos.

La herramienta finalmente adaptada, provee dos interfaces a través de las cuales se realiza el proceso de comparación manual de tareas. La Figura 34 muestra la interfaz de validación que permite a un evaluador registrado en el sistema, identificarse y acceder a la herramienta. Si el proceso de validación es exitoso, el evaluador es dirigido a la Interfaz que se observa en la Figura 35. En esta interfaz se dispone de un formulario por medio del cual el evaluador califica con un valor entre cero (0) y cinco (5) la similitud existente entre los términos que identifican los atributos de dos tareas (una de consulta y otra publicada).

En esta herramienta se realizaron en total 2910 comparaciones entre parejas de tareas (publicada y de consulta), garantizando que cada una de las 970 posibles parejas (10 tareas de consulta x 97 tareas publicadas) fuera comparada por al menos 3 de los evaluadores. En consecuencia, se contaba al final del proceso de evaluación con 3 resultados de similitud para cada par de tareas, los cuales fueron promediados en cada caso, conformando de esta manera la base de 970 resultados de similitud que constituyen el benchmark de referencia requerido para evaluar el rendimiento del sistema desarrollado.

Figura 34. Interfaz de Validación de la Herramienta de Comparación

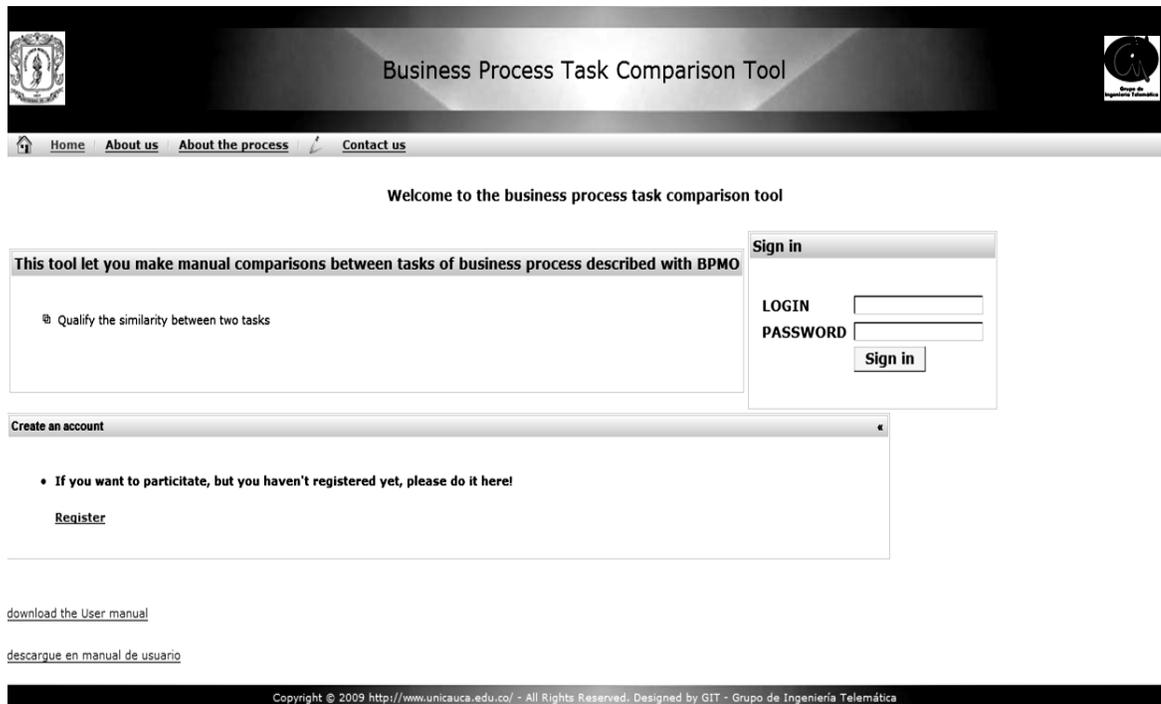
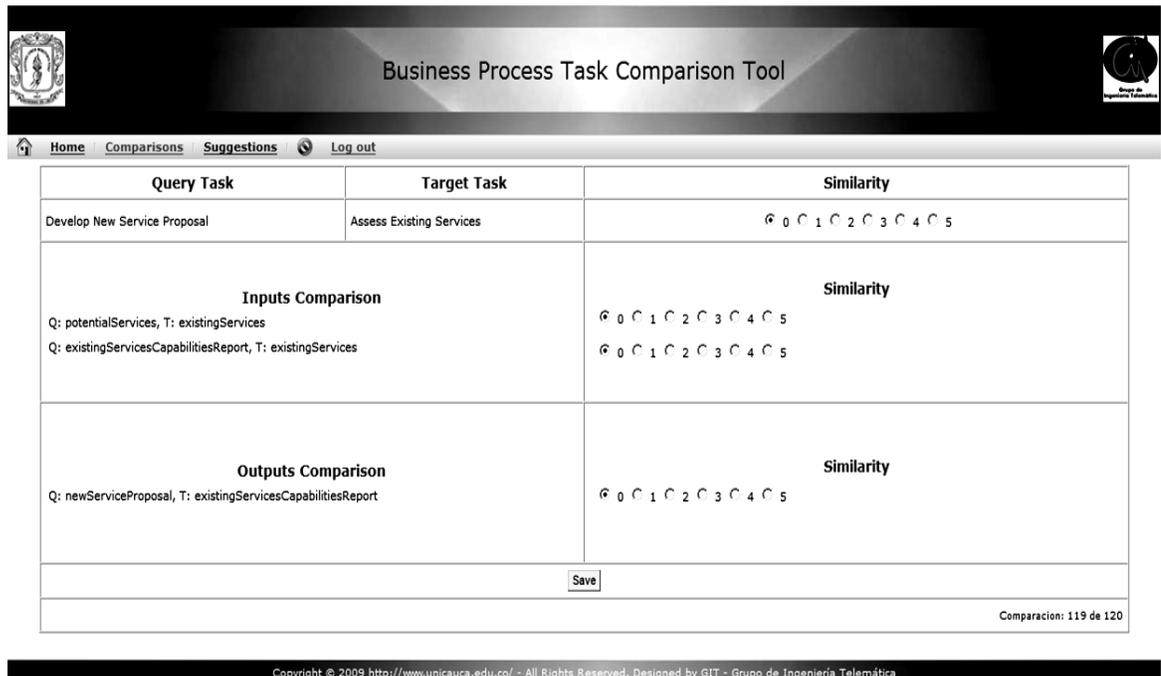


Figura 35. Interfaz de Comparación de la Herramienta



ANEXO C

ANEXO C

BANCO DE PROCESOS PUBLICADOS

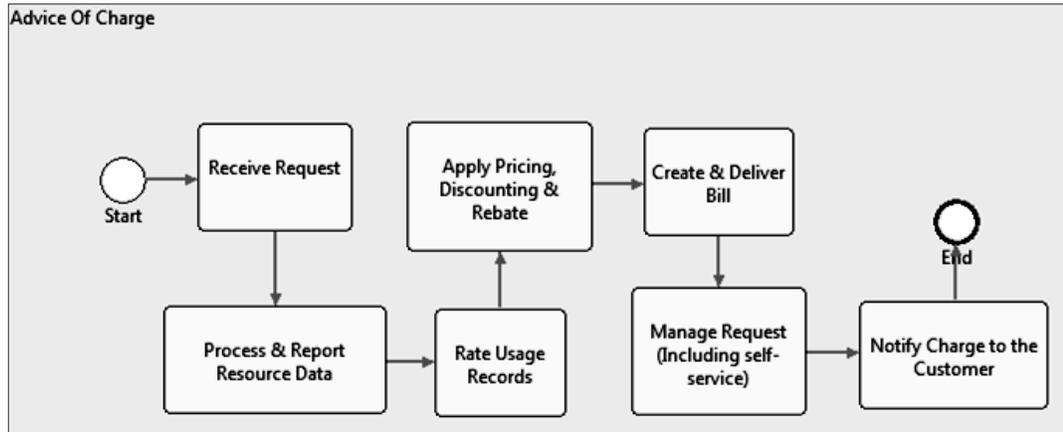
Este anexo abarca la descripción de los procesos de negocio de telecomunicaciones diseñados en la ejecución del presente proyecto con el objetivo de someter el mecanismo de comparación semántica propuesto a un procedimiento de evaluación y puesta a punto. Los procesos diseñados se basan fundamentalmente en el estudio de las relaciones existentes entre los componentes del marco de referencia para el sector de las telecomunicaciones eTOM y se han clasificado en 5 dominios, los cuales corresponden a 5 diferentes áreas funcionales definidas dentro de este estándar formulado dentro de la iniciativa NGOSS (ITU-T Study Group 4, 2005):

C.3. Gestión de las Relaciones con el Cliente (*CRM, Customer Relationships Management*):

Esta agrupación de operaciones comprende todas las actividades llevadas a cabo dentro de un operador de telecomunicaciones relacionadas con establecimiento, mantenimiento y retención de las relaciones con los clientes. Entre los procesos de negocio diseñados pertenecientes a esta área funcional se encuentran:

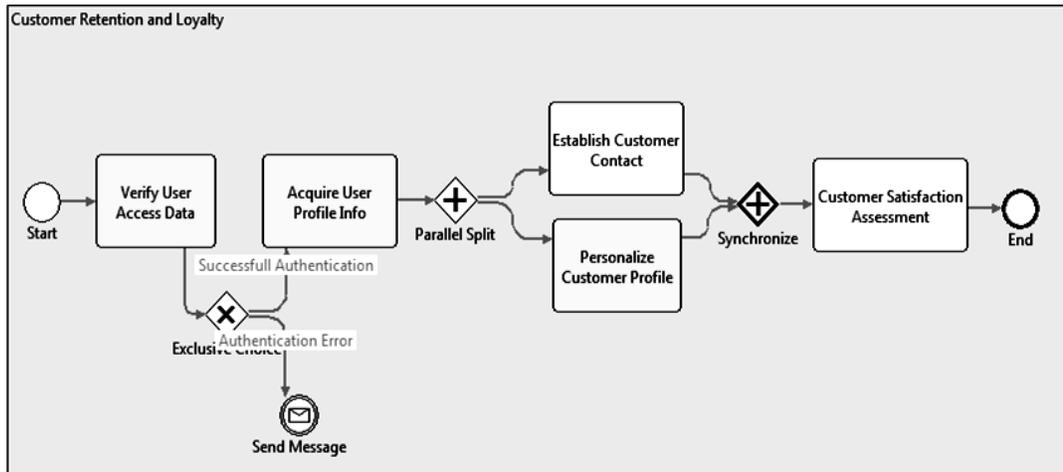
- **Proceso de Notificación de Cobro (*Advice of Charge*):** permite conocer el costo que representa para un usuario el acceso a un servicio de valor agregado que no hace parte del contrato convenido con el operador y que generalmente implica diferentes reglas de facturación (Figura 36).

Figura 36. Diagrama del proceso Advice of Charge



- **Proceso de Retención y Fidelidad del Cliente (*Customer Retention and Loyalty*):** este proceso permite llevar a cabo la evaluación de la satisfacción del cliente respecto a las condiciones de prestación de los servicios contratados con el operador (Figura 37).

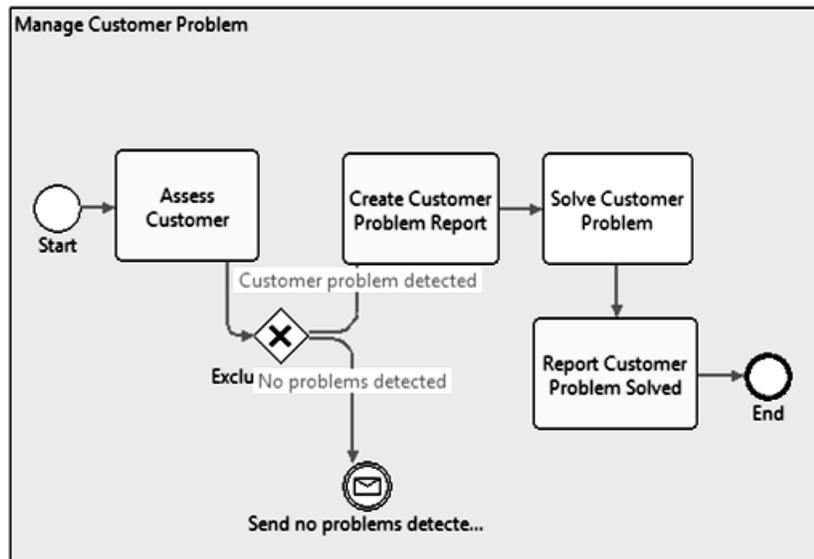
Figura 37. Diagrama del proceso Customer Retention and Loyalty



- **Proceso de Gestión de Problemas del Cliente (*Manage Customer Problem*):** este proceso dispone de un conjunto de actividades a partir de las cuales es posible detectar la existencia de problemas en la provisión de servicios para un

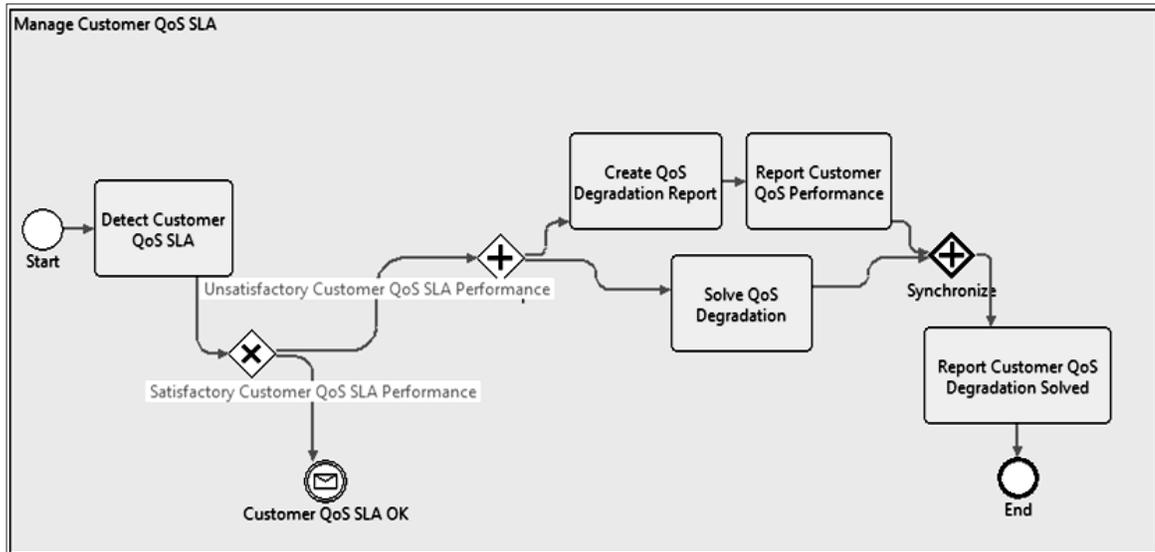
cliente determinando, así como ejecutar las medidas necesarias para restablecer el normal desempeño de dichos servicios (Figura 38).

Figura 38. Diagrama del proceso Manage Customer Problem



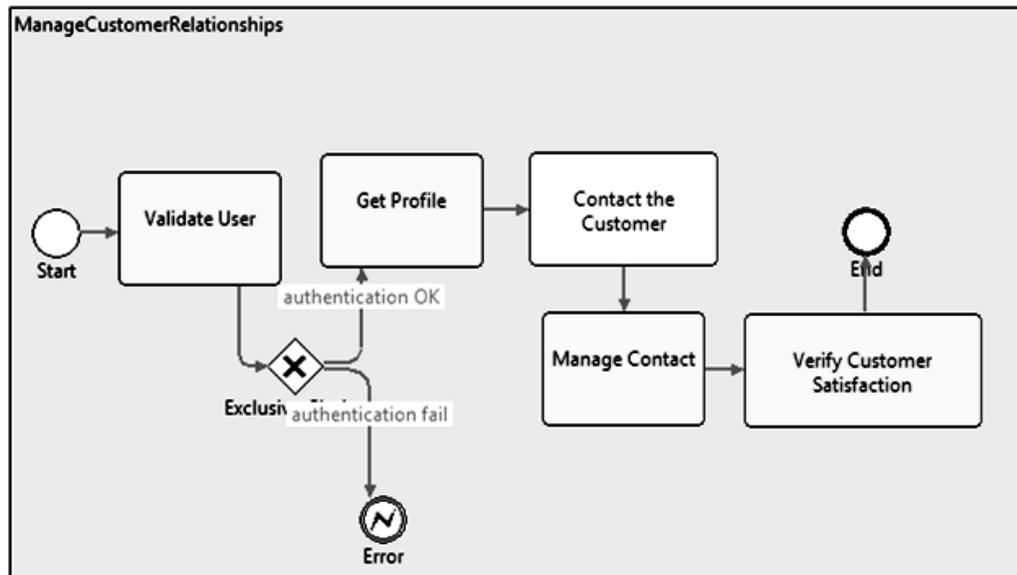
- **Proceso de Gestión de QoS y Acuerdo de Nivel de Servicio del Cliente (*Manage Customer QoS/SLA*):** define un flujo de trabajo que permite determinar si existe alguna inconsistencia en cuanto a la Calidad de Servicio percibida por el cliente y aquella estipulada en el Acuerdo de Nivel de Servicio, caso en el cual el proceso reporta y provee solución a dicho inconveniente (Figura 39).

Figura 39. Diagrama del proceso Manage Customer QoS/SLA



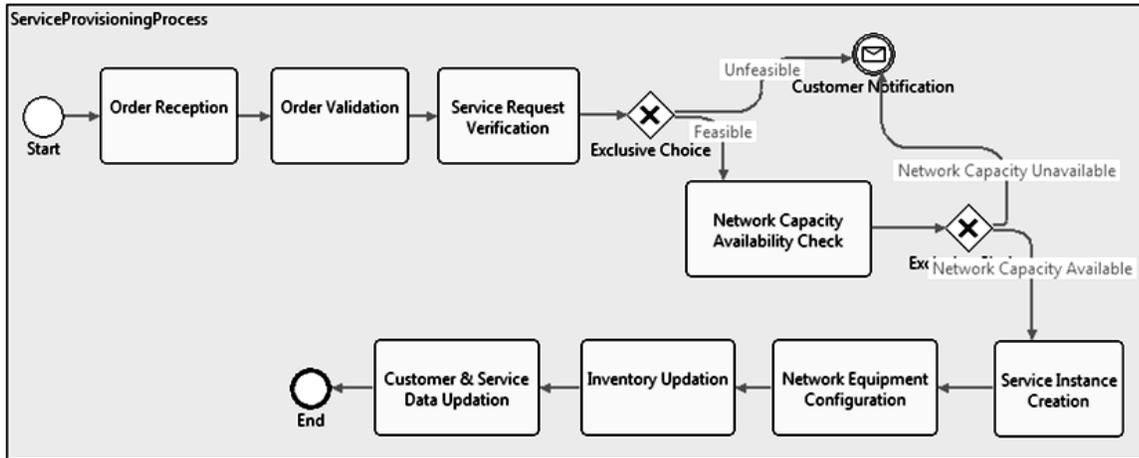
- **Proceso de Gestión de la Relación con el Cliente (*Manage Customer Relationships*):** define un proceso sencillo que permite evaluar la satisfacción del cliente respecto al nivel de desempeño del servicio que le provee el operador (Figura 40).

Figura 40. Diagrama del proceso Manage Customer Relationships



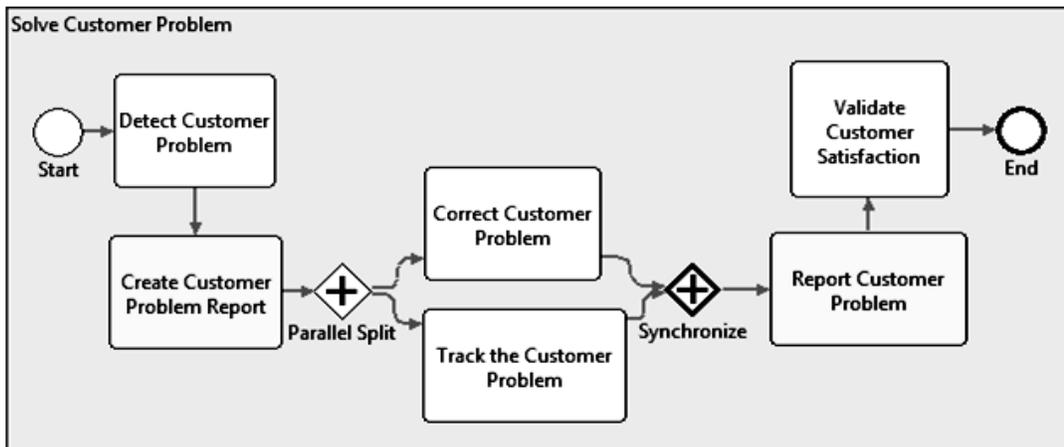
- **Proceso de Provisión de Servicio (Service Provisioning Process):** especifica el flujo de trabajo necesario para llevar a cabo la prestación de un servicio, a partir de la recepción y verificación de la orden del cliente, así como la evaluación de la disponibilidad de los recursos necesarios para satisfacer la dicha orden (Figura 41).

Figura 41. Diagrama del proceso Service Provisioning Process



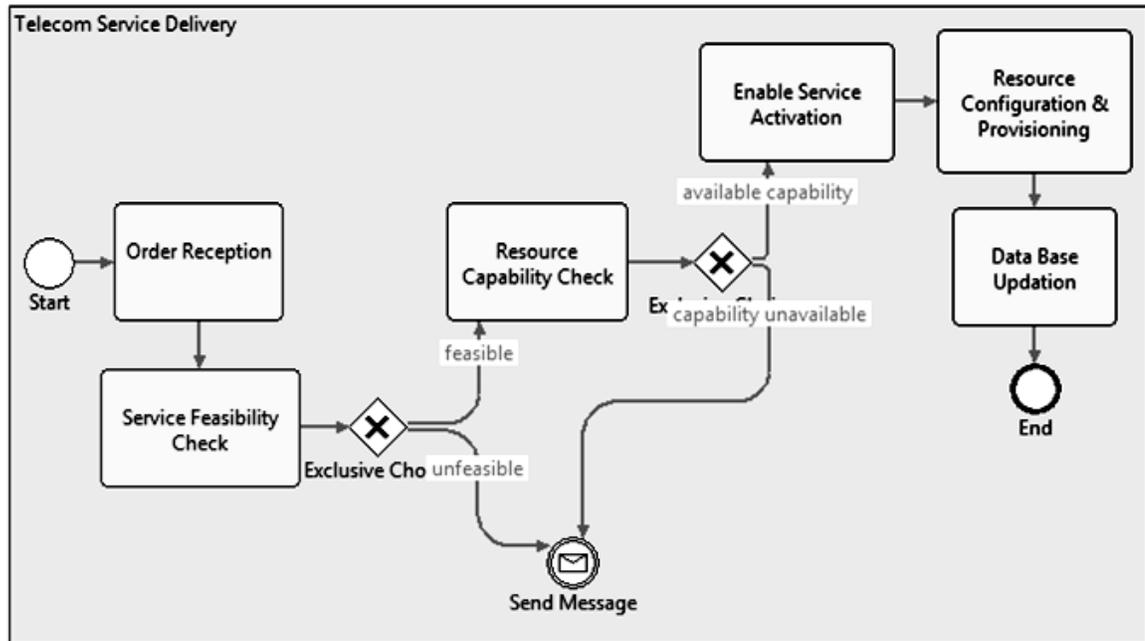
- **Proceso para Resolver Problemas del Cliente (Solve Customer Problem):** define un conjunto de actividades cuya ejecución permite detectar, reportar, rastrear y proveer solución a los problemas que se presentan en la prestación de servicio para un cliente determinado (Figura 42).

Figura 42. Diagrama del proceso Solve Customer Problem



- **Proceso de Entrega de Servicios (*Telecom Service Delivery*):** este proceso provee soporte para la realización de las actividades involucradas en la activación de un servicio de telecomunicaciones (Figura 43).

Figura 43. Diagrama del proceso Telecom Service Delivery

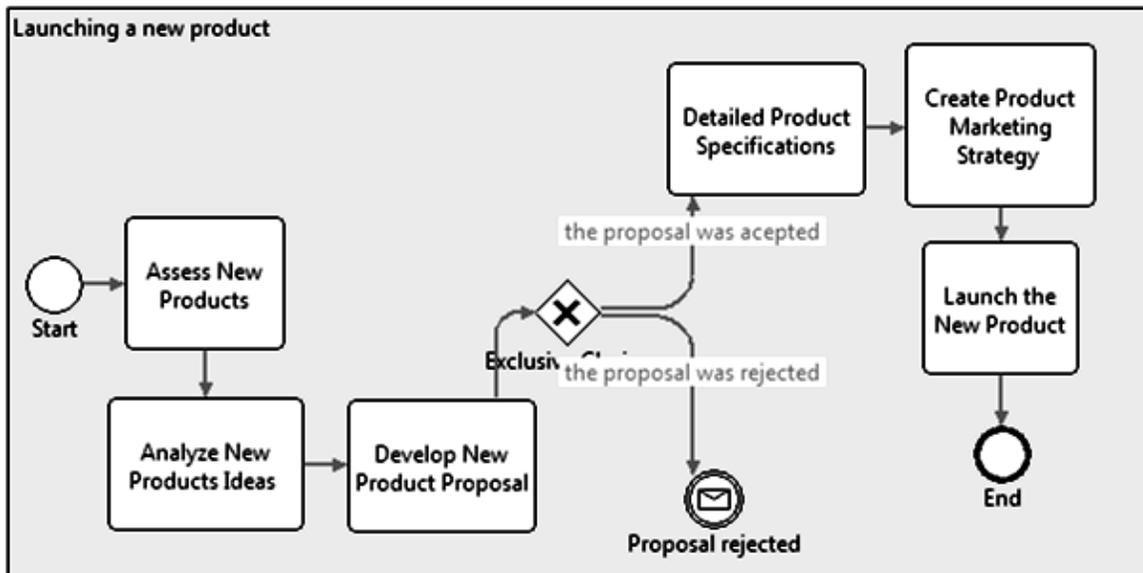


C.2. Gestión de Mercadeo y Ofertas (*Marketing & Offer Management*):

Dentro de esta agrupación funcional se consideran las actividades realizadas por los operadores, destinadas a la definición e implementación de estrategias de mercadeo y oferta, el desarrollo de nuevos productos y la gestión de productos existentes. Los procesos de negocio diseñados y clasificados dentro de este dominio son:

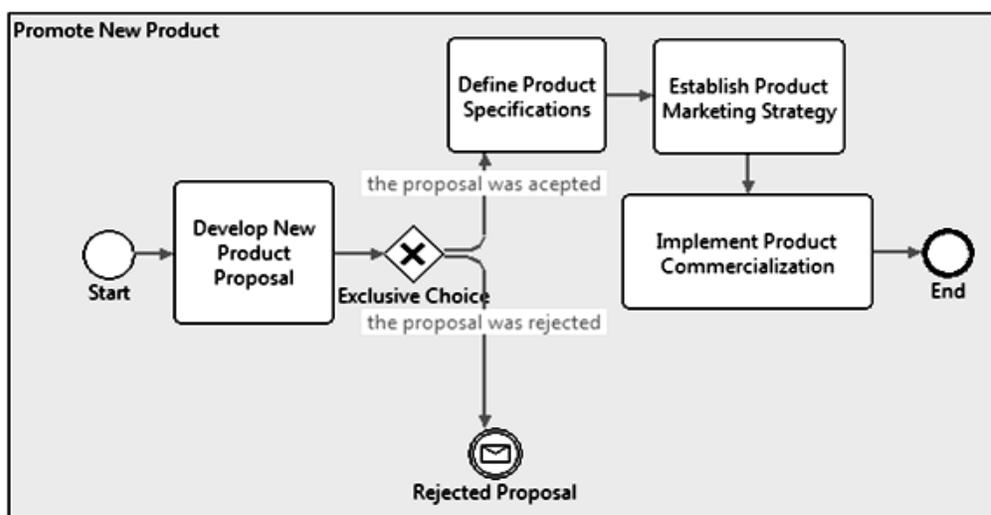
- **Proceso para Lanzar un Producto Nuevo al Mercado (*Launching a New Product*):** las tareas que constituyen este proceso se encargan de llevar a cabo la evaluación de nuevos productos que satisfagan la demanda de servicios de telecomunicaciones, así como generar estrategias de comercialización que soporten el despliegue de dichos productos dentro del mercado (Figura 44).

Figura 44. Diagrama del proceso Launching a New Product



- **Proceso para Promover un Nuevo Producto en el Mercado (*Promote a New Product*):** básicamente este proceso se deriva del anterior, en la medida en que provee una funcionalidad similar, pero en este caso, partiendo del desarrollo de la propuesta del nuevo producto, sin considerar la etapa inicial de evaluación de que se ejecuta en el proceso anterior (Figura 45).

Figura 45. Diagrama del proceso Promote a New Product

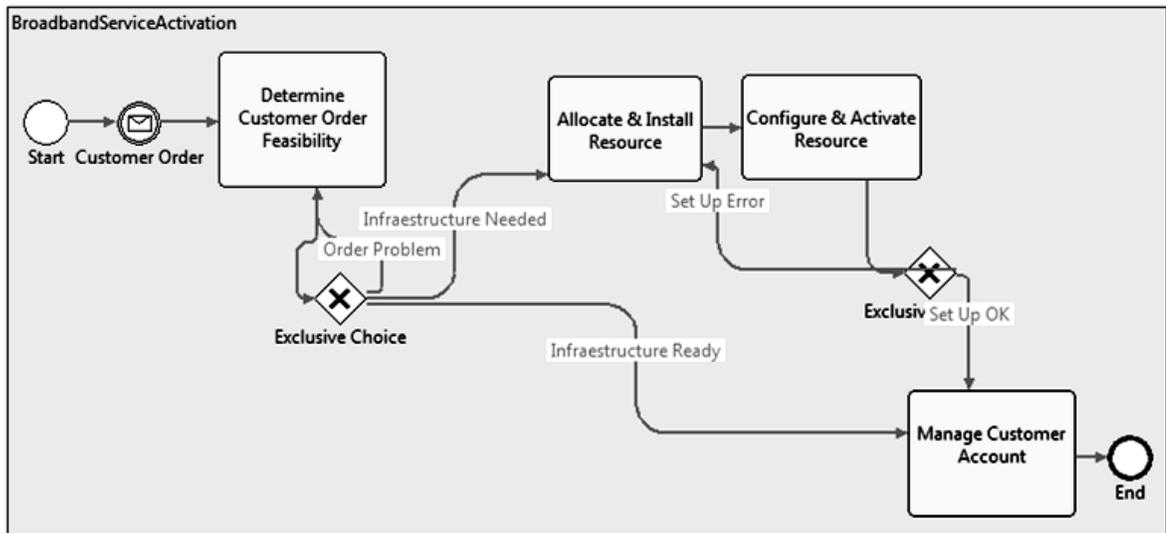


C.3. Gestion y Operaciones de Recursos (*Resource Management & Operations*):

Esta agrupación de operaciones involucra todas las actividades necesarias para administrar los recursos (sistemas de TI, redes, servidores, routers, etc.) con los que cuenta el operador para soportar la prestación de los servicios requeridos por los clientes y al interior de la empresa. A este dominio pertenecen los procesos a los cuales se hace referencia a continuación:

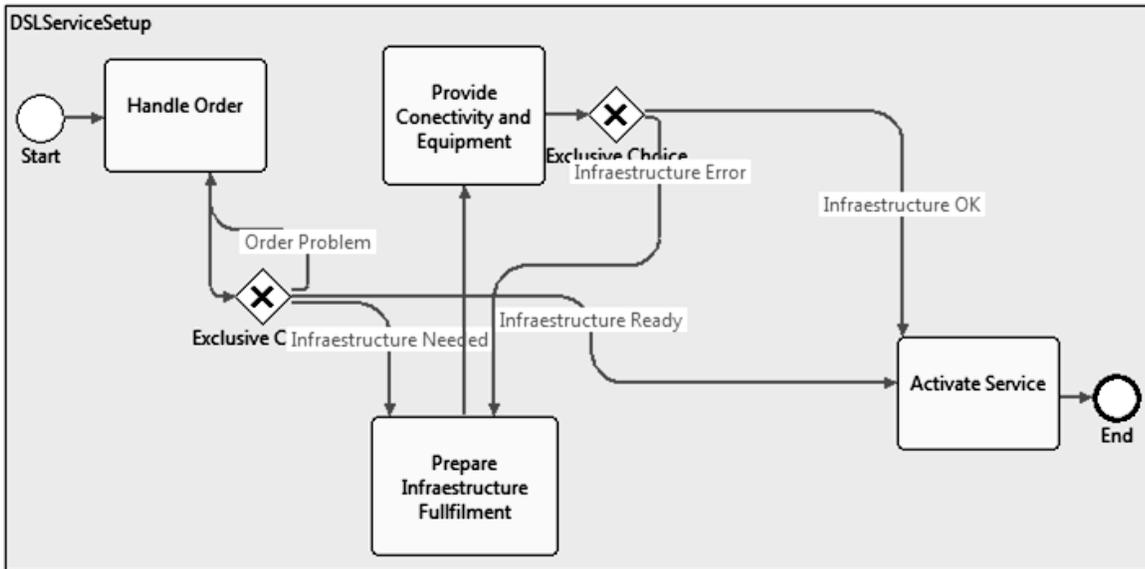
- **Proceso de Activación de Servicios de Banda Ancha (*Broadband Service Activation*):** modela el flujo de trabajo que ejecuta un operador de telecomunicaciones para llevar a cabo la activación de un servicio de banda ancha, para lo cual tiene en cuenta la evaluación previa de la viabilidad de la orden del cliente y la ubicación, instalación y configuración de los recursos que soportan la prestación de dicho servicio (Figura 46).

Figura 46. Diagrama del Proceso Broadband Service Activation



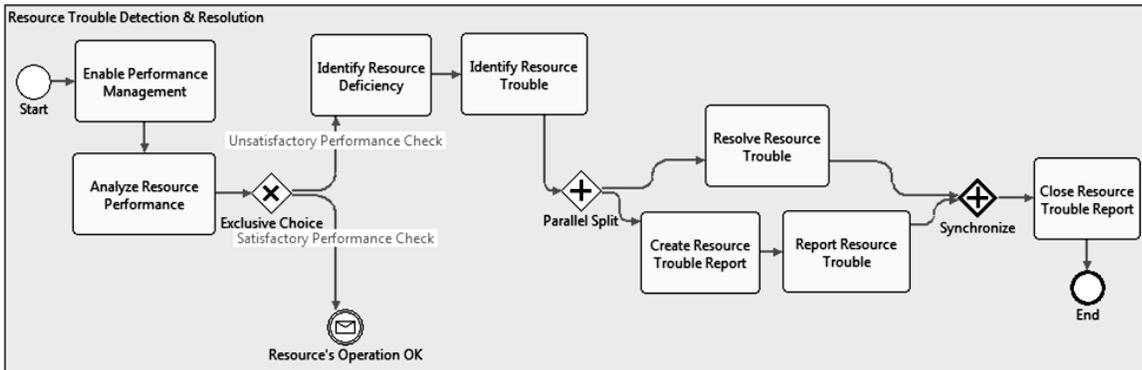
- **Proceso de Configuración de Servicios DSL (*DSL Service Setup*):** constituye una variación del proceso anterior, la cual fundamentalmente cumple la misma función, esta vez aplicada a un servicio específico de acceso de banda ancha mediante *línea de abonado digital* (Figura 47).

Figura 47. Diagrama del Proceso DSLServiceSetup



- Proceso de Detección y Resolución de Problemas de Recursos (*Resource Trouble Detection and Resolution*):** permite llevar a cabo la evaluación del rendimiento de los recursos involucrados en la prestación de los servicios del operador. A partir de esta evaluación determina la posible existencia de deficiencias en el desempeño de dichos recursos e identifica, reporta y provee solución a los problemas que limitan su operación adecuada (Figura 48).

Figura 48. Diagrama del Proceso Resource Trouble Detection and Resolution

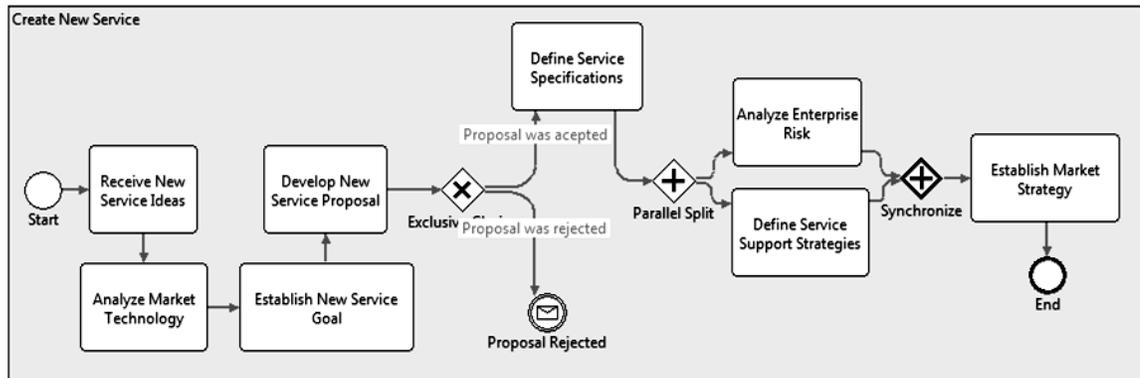


C.4. Desarrollo y Gestion de Servicios (*Service Development & Management*):

Las actividades que se consideran dentro de esta agrupación funcional están dirigidas hacia la definición de estrategias para el diseño y creación de servicios, la gestión de servicios existentes y el aseguramiento de las capacidades para satisfacer la demanda futura de servicios. Dos de los procesos publicados se clasifican en este dominio:

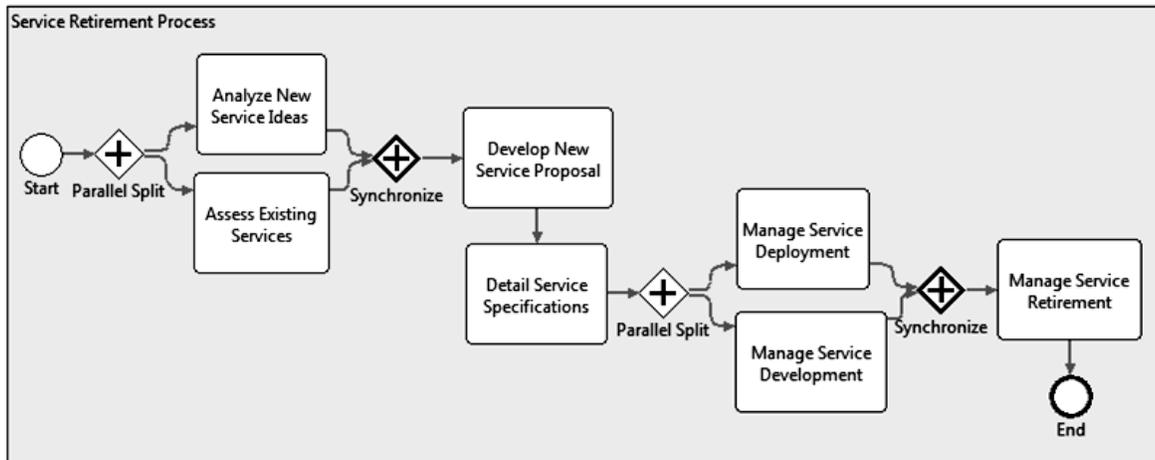
- **Proceso de Creación de un Nuevo Servicio (*Create New Service*):** representa un conjunto de actividades relevantes necesarias para la creación de un servicio, considerando la demanda del mercado y definiendo las especificaciones y requerimientos necesarios para satisfacer la nueva necesidad (Figura 49).

Figura 49. Diagrama del proceso Create New Service



- **Proceso de Retiro de un Servicio (*Service Retirement Process*):** este proceso permite analizar los servicios existentes en el mercado para determinar si se debe realizar el retiro de algún servicio del proveedor, dependiendo de los nuevos requerimientos de los usuarios y de la posibilidad de reemplazar el servicio a liquidar (Figura 50).

Figura 50. Diagrama del proceso Service Retirement Process

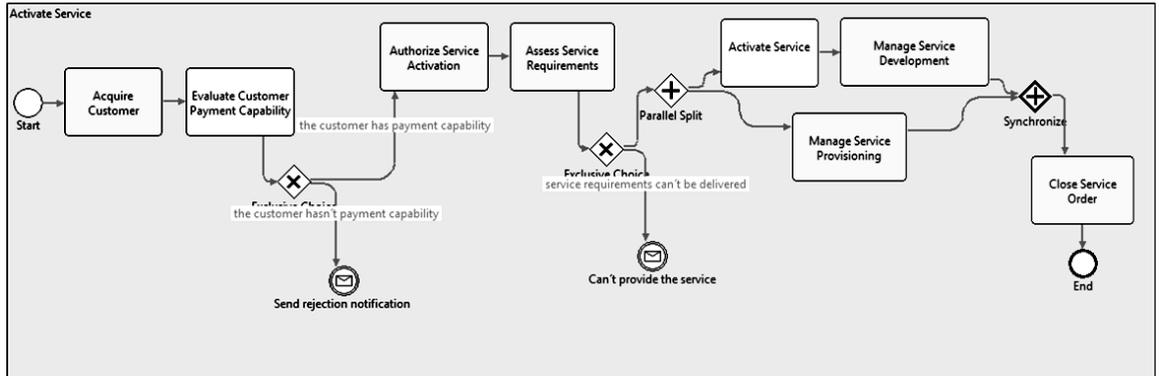


C.5. Gestión y Operaciones de Servicios (*Service Management & Operations*):

Ésta agrupación de operaciones está directamente relacionada con la experiencia diaria del cliente. Comprende las actividades llevadas a cabo dentro del operador, orientadas hacia la administración y la operación de servicios de información y comunicaciones demandados por los clientes. Éstas actividades se enfocan en la prestación y gestión del servicio, mas no en la administración de la red y la tecnología de información subyacente. Los procesos de negocio diseñados pertenecientes a esta agrupación funcional son:

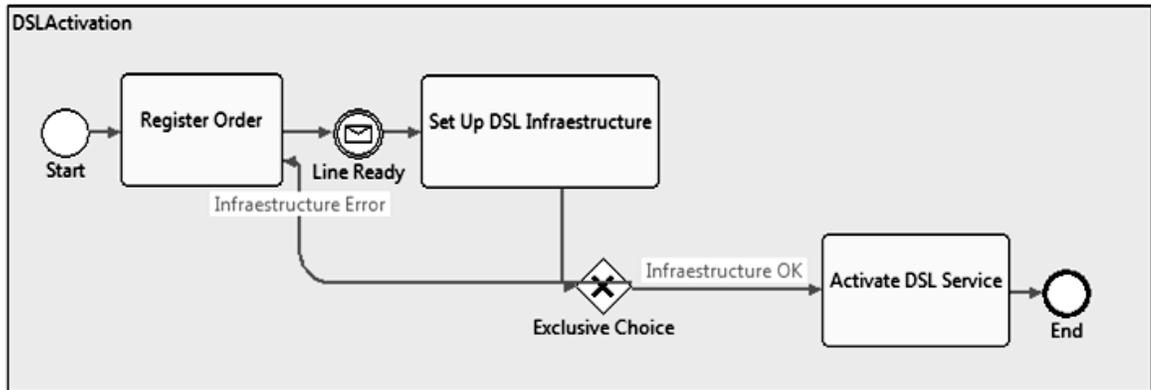
- **Proceso de Activación de Servicio (*Activate Service*):** esta compuesto por un conjunto de operaciones que permiten evaluar una solicitud de servicio del cliente para verificar la viabilidad de la prestación del servicio. Además gestiona los elementos necesarios para la activación del servicio en cumplimiento con la orden recibida (Figura 51).

Figura 51. Diagrama del proceso Activate Service



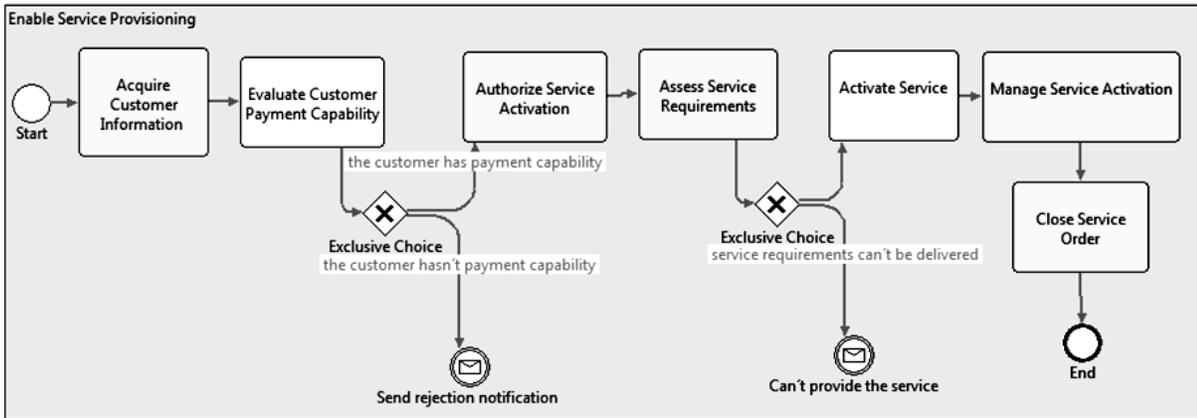
- **Proceso de Activación de Servicio DSL (DSL Activation):** modela las actividades necesarias para llevar a cabo la activación de un servicio soportado en una conexión digital sobre línea de abonado (Figura 52).

Figura 52. Diagrama del proceso DSL Activation



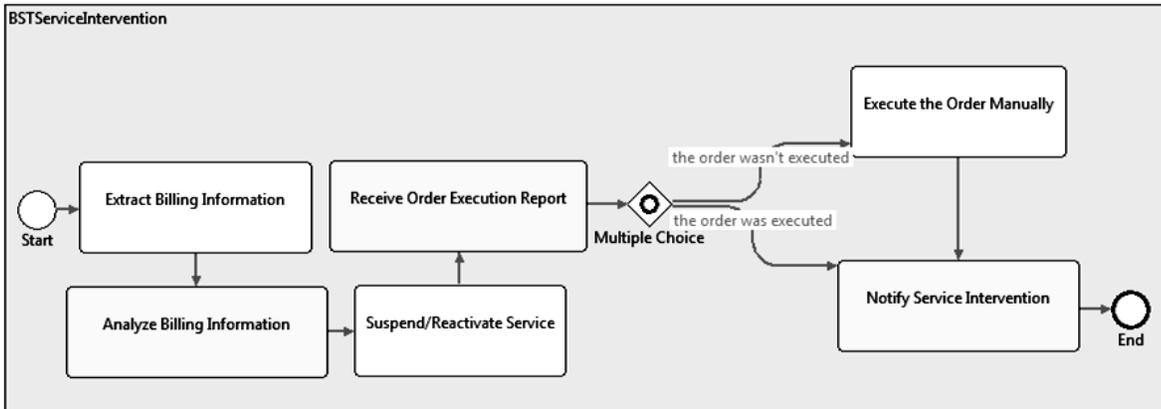
- **Proceso para Habilitar la Prestación de un Servicio (Enable Service Provisioning):** este proceso define las operaciones necesarias para autorizar la activación de un servicio de telecomunicaciones, a partir del análisis de la capacidad de pago del cliente y de la disponibilidad de los recursos necesarios para la activación (Figura 53).

Figura 53. Diagrama del proceso Enable Service Provisioning



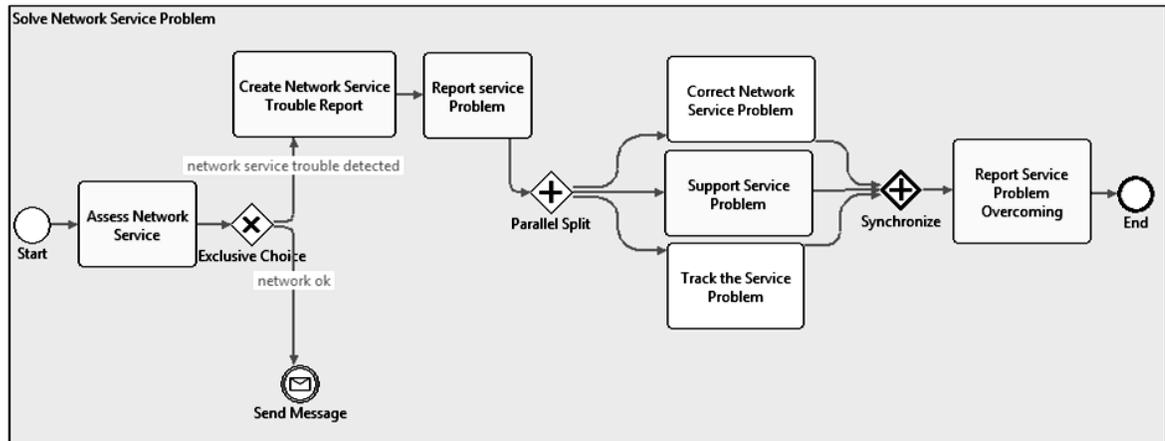
- Proceso de Intervención del Servicio de Telefonía Básica Conmutada (BST Service Intervention):** especifica el flujo de trabajo que lleva a cabo un operador de telecomunicaciones para reactivar o suspender el servicio de telefonía básica conmutada, teniendo en cuenta la información recibida del departamento de facturación (Figura 54).

Figura 54. Diagrama del proceso BST Service Intervention



- Proceso para Solucionar Problemas de Red del Servicio (Solve Network Service Problem):** describe las operaciones necesarias para detectar y corregir un problema en la red del operador de telecomunicaciones, incluye actividades de reporte y seguimiento del problema hasta asegurar su solución (Figura 55).

Figura 55. Diagrama del proceso Solve Networ Service Problem



Referencias

ITU-T Study Group 4. 2005. *ITU-T Recommendation M.3050.1, Enhanced Telecom Operations Map (eTOM) – The business process framework*. International Telecommunication Union. Ginebra, Suiza : s.n., 2005.

Sandino, L., Figueroa, C. y Corrales, J. 2010. *Platform for the Evaluation of Web Processes Retrieval Systems. [Unpublished]*. Grupo de Ingeniería Telemática, Universidad del Cauca. 2010.