



# Universidad Nacional de La Plata

Facultad de Informática

Tesis presentada para obtener el grado de Magister en Ingeniería de  
Software

## **Normalización de Información de Servicios Requeridos para Sistemas de Información Geográficos**

Tesista: *Viviana Ester Saldaño*

Directora: *Silvia Gordillo*

Codirectora: *Alejandra Cechich*

Marzo 2013



## **PREFACIO**

Esta Tesis es presentada como parte de los requisitos para optar al grado académico de Magister en Ingeniería de Software, de la Universidad Nacional de La Plata y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otras. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en la Unidad Académica Caleta Olivia de la Universidad Nacional de la Patagonia Austral, bajo la dirección de la Dra. Silvia Gordillo, Profesora de la Universidad Nacional de La Plata, la co-dirección de la Dra. Alejandra Cechich, Profesora de la Universidad Nacional del Comahue y la asesoría científica de la Dra. Agustina Buccella, Profesora de la Universidad Nacional del Comahue.

Viviana Saldaño

FACULTAD DE INFORMÁTICA

UNIVERSIDAD NACIONAL DE LA PLATA



## **AGRADECIMIENTOS**

*Un enorme agradecimiento a Alejandra Cechich y Agustina Buccella, quienes me guiaron y acompañaron incondicionalmente durante todos estos años. Sin su constante dedicación, seguimiento y aliento, esta tesis no se hubiera concluido.*

*A mi compañera de investigación, Gabriela Gaetán con quien compartimos tanto alegrías como frustraciones, desvaríos y descubrimientos, gracias por su aporte crítico y siempre tan preciso y acertado.*

*Y finalmente, agradezco inmensamente a mi familia, esposo e hijos que debieron soportar tantas postergaciones a causa de “la Tesis”.*



## **RESUMEN**

La amplia difusión del proceso de desarrollo basado en componentes, en particular en el dominio de aplicaciones geográficas, ha generado el surgimiento de una gran cantidad y variedad de componentes software orientados a satisfacer las necesidades de los desarrolladores. En este contexto, uno de los principales inconvenientes encontrados al desarrollar aplicaciones mediante reuso de componentes, consiste en la localización de aquellos más adecuados a un requerimiento o conjunto de requerimientos en particular. A esto se denomina proceso de selección de componentes, que incluye dos actividades principales: la oferta de componentes disponibles para reuso y la demanda de servicios a ser satisfechos por la interface de los componentes ofertados.

Nuestra propuesta consiste en una mejora al proceso de demanda mediante la estandarización de la especificación de requerimientos o servicios orientándolos a la búsqueda de componentes, en particular para el dominio de aplicaciones geográficas. Para ello se trabaja sobre la normalización de los requerimientos, categorizándolos según información estándar para servicios geográficos.

Así, en esta Tesis se define en primer lugar, una Taxonomía de Servicios Geográficos. Luego, se propone una metodología para extraer conocimiento a partir de los casos de uso y categorizar esos requerimientos en clases de servicios geográficos. Por último, se describe la herramienta ReqGIS, la cual automatiza la metodología señalada anteriormente.





## ABSTRACT

The wide-spread use of component based development processes, particularly in geographic application domain, has caused the appearance of a great number and variety of software components oriented to satisfying developers' needs. In this context, one of the main difficulties when developing applications with software reuse is discovering the most suitable components to a particular requirement. This activity is known as *selection process*, which includes two main activities: *offering* reusable components and *requiring* services to be satisfied by the components' interfaces.

Our proposal suggests improving the *requiring process* by standardizing the software requirement specification, and making requirements more suitable for searching suitable components, particularly for the geographic application domain. To do so, we work on requirement normalization, classifying requirements according to standard information for geographic services.

In this Thesis, we firstly define a Geographic Service Taxonomy. Then, we propose a methodology to extracting knowledge from use cases, so requirements could be classified as normalized geographic services. Finally, we describe the ReqGIS tool, which supports the normalization process.



# Indice general

<b>1. INTRODUCCIÓN.....</b>	<b>1</b>
1.1. Contexto .....	1
1.2. Objetivos .....	3
1.3. Metodología .....	4
1.4. Estructura .....	4
<b>2. FUNDAMENTOS DEL TRABAJO DE TESIS.....</b>	<b>7</b>
2.1. Desarrollo de sistemas basados en componentes .....	7
2.1.1. Selección de Componentes .....	9
2.1.2. Enfoques para la Selección de Componentes.....	11
2.1.2.1. OTSO (Off-The-Shelf Option) .....	11
2.1.2.2. PORE (Procurement Oriented REquirements).....	14
2.1.2.3. Sistema DesCOTS (Description, Evaluation and Selection of COTS components) .....	16
2.1.2.4. CARE (COTS-Aware Requirements Engineering).....	18
2.1.2.5. Propuesta de Lucena.....	20
2.2. Especificación de Requerimientos para COTS .....	22
2.2.1. Conocimiento a partir de Casos de Uso.....	24
2.2.2. Conceptos Lingüísticos.....	28
2.2.3. Algunas Herramientas Lingüísticas .....	33
2.2.4. Extracción de Conocimiento en Casos de Uso.....	38
2.3. Resumen .....	39
<b>3. CATEGORIZACIÓN Y SELECCIÓN DE SERVICIOS SIG: ESTADO DEL ARTE .....</b>	<b>41</b>
3.1. Características del Dominio Geográfico.....	41
3.2. Categorización de Servicios Geográficos .....	44

3.3. Estándares para Definición y Categorización de Servicios SIG.....	47
3.3.1. Categorización de Servicios del estándar ISO 19119.....	50
3.4. Selección de Componentes SIG .....	52
3.5. Resumen .....	54
<b>4. PROPUESTA DE MEJORA PARA EL PROCESO DE DEMANDA DE SERVICIOS</b>	
<b>SIG.....</b>	<b>55</b>
4.1. Marco Contextual .....	55
4.2. Propuesta para Estandarizar el Proceso de Demanda de Servicios SIG .....	58
4.2.1. Definición de una Taxonomía de Servicios Geográficos .....	59
4.2.2. Método Propuesto para la Extracción de Conocimiento en Casos de Uso de Aplicaciones SIG .....	64
4.3. ReqGIS: Herramienta de Clasificación de Requerimientos Geográficos .....	68
4.3.1. Descripción de la Arquitectura .....	68
4.3.2. Descripción de la Funcionalidad.....	72
4.3.3. Interfaces de Usuario .....	81
4.4. Resumen .....	83
<b>5. CASO DE ESTUDIO .....</b>	<b>85</b>
5.1. Especificación de Requerimientos.....	85
5.1.1. Alcance.....	85
5.1.2. Caso de Uso Ejemplo.....	87
5.2. Clasificación de Requerimientos con ReqGIS .....	88
5.3. Validación de los resultados obtenidos.....	90
5.3.1. La Encuesta .....	90
5.3.2. Diseño de la Encuesta .....	92
5.3.3. Resultados de la Encuesta.....	96
5.4. Resumen .....	100
<b>6. CONCLUSIÓN Y TRABAJO FUTURO .....</b>	<b>103</b>
6.1. Análisis de la Consecución de Objetivos .....	103
6.2. Principales Aportes .....	104

6.3. Publicaciones derivadas de esta Tesis .....	105
6.4. Trabajos Futuros .....	107
<b>APÉNDICE A. CATEGORIZACIÓN ISO 19119.....</b>	<b>109</b>
<b>APÉNDICE B. CASOS DE USO.....</b>	<b>125</b>
<b>APÉNDICE C. RESULTADOS DEL PROCESAMIENTO DE REQGIS .....</b>	<b>131</b>
<b>APÉNDICE D. RESPUESTAS A ENCUESTA .....</b>	<b>133</b>



# Índice de figuras

<b>Figura 2.1:</b> Actividades del proceso de desarrollo basado en componentes .....	8
<b>Figura 2.2:</b> Fases del método de selección OTSO .....	12
<b>Figura 2.3:</b> Procesos genéricos para lograr los objetivos esenciales.....	15
<b>Figura 2.5:</b> Actividades del proceso CARE.....	19
<b>Figura 2.6:</b> Principales actividades de los procesos de Publicación y Selección de componentes .....	20
<b>Figura 2.7:</b> Plantilla de Caso de Uso Textual .....	26
<b>Figura 2.8:</b> Árbol sintáctico de la oración “ <i>The children ate the cake</i> ” . .....	31
<b>Figura 2.9:</b> Definiciones para “tree” y “flower” .....	35
<b>Figura 2.10:</b> Solapamiento entre definiciones de “Tree” y “Flower” .....	35
<b>Figura 2.11:</b> Hipónimos e hiperónimos de las palabras “tree” y “flower” y análisis de solapamientos que calcula el algoritmo Lesk Adaptado. ....	36
<b>Figura 4.1:</b> Procedimiento propuesto para la búsqueda y selección de componentes. A la derecha se observa el proceso de demanda, el cual es objeto de desarrollo en esta tesis.....	56
<b>Figura 4.3:</b> Diagrama de componentes de la herramienta ReqGIS. ....	69
<b>Figura 4.4:</b> Algoritmo que calcula el valor más alto de relación semántica. ....	71
<b>Figura 4.5:</b> Caso de uso textual.....	72
<b>Figura 4.6:</b> Aplicación de la metodología a un caso de uso. ....	73
<b>Figura 4.7:</b> Árbol sintáctico correspondiente a: “ <i>User modifies coordinate attribute</i> ”. .....	74
<b>Figura 4.8:</b> Formación del Token de Evento. ....	75
<b>Figura 4.9:</b> Medición de la relación semántica entre el verbo modify del token de evento y verbos de la categoría Human Interaction. ....	77
<b>Figura 4.10:</b> Archivo XML que almacena la categoría de servicios elegida. ....	81
<b>Figura 4.11:</b> Interfaz de la herramienta ReqGIS – Ingreso de Información. ....	82

<b>Figura 4.12:</b> Interfaz de la herramienta ReqGIS – Presentación de Resultados. ....	82
<b>Figura 5.1:</b> Caso de Uso “Find city facilities” .....	88
<b>Figura B.2:</b> Caso de Uso “Assess property” .....	126
<b>Figura B.3:</b> Caso de Uso “Find school district” .....	126
<b>Figura B.5:</b> Caso de Uso “Create landmarks” .....	127
<b>Figura B.6:</b> Caso de Uso “Find schools” .....	128
<b>Figura B.7:</b> Caso de Uso “Measure distance” .....	128
<b>Figura B.8:</b> Caso de Uso “Create formatted maps and reports” .....	129
<b>Figura B.9:</b> Caso de Uso “Turn on/off layers” .....	129
<b>Figura B.10:</b> “Run statistics” .....	130
<b>Figura D.1.:</b> Cuestionario 1 .....	133
<b>Figura D.2.:</b> Cuestionario 2 .....	134
<b>Figura D.3.:</b> Cuestionario 3 .....	135
<b>Figura D.4.:</b> Cuestionario 4 .....	136
<b>Figura D.5.:</b> Cuestionario 5 .....	137
<b>Figura D.6.:</b> Cuestionario 6 .....	138
<b>Figura D.7.:</b> Cuestionario 7 .....	139
<b>Figura D.8.:</b> Cuestionario 8 .....	140
<b>Figura D.9.:</b> Cuestionario 9 .....	141
<b>Figura D.10.:</b> Cuestionario 10 .....	142



# Índice de tablas

<b>Tabla 2.1:</b> Etiquetas correspondientes a las distintas categorías sintácticas .....	29
<b>Tabla 3.1:</b> Operaciones SIG Universales .....	45
<b>Tabla 4.1:</b> Esquema de clasificación normalizado para el proceso de Oferta .....	58
<b>Tabla 4.2:</b> Taxonomía de servicios geográficos. ....	62
<b>Tabla 4.4:</b> Valores de relación semántica entre objetos representativos de la categoría “Human Interaction” y objeto “attribute” del token de evento. ....	79
<b>Tabla 4.5:</b> Valores promedio de relación semántica para todas las categorías de la Taxonomía SIG.....	80
<b>Tabla 5.1:</b> Resultados de procesamiento del caso de uso “Find city facilities” con ReqGIS.....	89
<b>Figura 5.2:</b> Cuestionario de la encuesta.....	95
<b>Tabla 5.2:</b> Conocimiento previo sobre Categorización de Servicios .....	96
<b>Tabla 5.3:</b> Tabulación de Cuestionarios .....	97
<b>Figura 5.3:</b> Gráfico circular “Categoría Seleccionada por ReqGIS vs Categoría Seleccionada en las Encuestas” .....	98
<b>Tabla 5.4:</b> Tabulación de cuestionarios teniendo en cuenta el conocimiento previo de ISO 19119.....	99
<b>Figura 5.4:</b> Gráfico de columnas comparativo según conocimiento previo de ISO 19119.....	99
<b>Tabla A.1.:</b> Categorización de servicios geográficos del estándar ISO 19119 .....	123
<b>Tabla C.1:</b> Resultados de la estandarización de los casos de uso .....	132



# Capítulo 1

## Introducción

En este capítulo, se presenta una breve descripción sobre el contexto de la tesis, su objetivo, la metodología utilizada y su estructura.

### 1.1. Contexto

El avance en la tecnología en cuanto a dispositivos de colección de datos más sofisticados y el desarrollo de gran cantidad de nuevos sistemas de información espaciales han generado un alto volumen de datos sobre nuestra tierra. Por ejemplo, hoy en día el uso de dispositivos GPS es tan común que está disponible incluso en los relojes de pulsera.

Los Sistemas de Información Geográficos (SIG), son los responsables de almacenar y manipular toda la información espacial. A diferencia de los sistemas de información tradicionales, debido a la naturaleza de los datos y a las características exclusivas de la información geográfica, los SIG poseen requerimientos funcionales especiales. En consecuencia, es muy importante determinar esos requerimientos especiales y la funcionalidad que deben proveer para poder diseñar e implementar SIG con capacidades apropiadas para el modelado, colección, consulta y visualización de la información geográfica. Así, la construcción de un SIG no es una tarea simple.

Para desarrollar un Sistema de Información Geográfico se puede utilizar un proceso de desarrollo de software tradicional en el cual todas las tareas como análisis, diseño, implementación y verificación se realizan desde el principio, o se puede hacer uso de componentes ya existentes diseñados especialmente para dar soluciones específicas. Ésta segunda opción permite el reuso del análisis, diseño e implementación del componente evitando gran cantidad de esfuerzo extra si debiéramos construirlo como indica la primera opción. Esta metodología se conoce como Desarrollo de Software Basado en Componentes (DSBC).

El desarrollo de software basado en componentes, es el proceso de construir sistemas de software mediante el ensamblado e integración de soluciones prefabricadas en la forma de componentes OTS (Off-The-Shelf). Existen distintos criterios sobre lo que constituye un componente OTS. En general, es un software que provee una funcionalidad específica o servicio<sup>1</sup>, disponible en el mercado para ser integrado en otros sistemas de software [73]. Más precisamente: “Un componente es un producto software que: (1) se vende o licencia al público en general; (2) es comercializado por un vendedor; (3) es mantenido y actualizado por un vendedor quien tiene los derechos de propiedad intelectual; (4) está disponible en múltiples copias idénticas; y (5) se utiliza sin modificaciones al código fuente por parte del cliente” [50].

Los beneficios de aplicar esta metodología de desarrollo consisten principalmente en la reducción de costos y tiempo de desarrollo, debido a que los componentes pueden adquirirse en vez de desarrollarse desde cero. Este factor sumado al creciente tamaño y complejidad de los sistemas, ha convertido la aplicación del paradigma de Desarrollo de Software Basado en COTS (DSBC) en un estándar de desarrollo de software.

Por otra parte, se han detectado algunas dificultades en la utilización de este paradigma, principalmente en lo que concierne a las actividades de identificación y recuperación de componentes existentes.

En el DSBC intervienen por un lado el desarrollador del componente y por otro, el usuario del componente, definiéndose así dos modelos, de *demanda* y de *oferta*, representando requerimientos especiales y soluciones de acuerdo a las restricciones y problemática que involucra cada uno de ellos.

El modelo de demanda hace referencia a los requerimientos o servicios solicitados por los clientes de un sistema, por ejemplo empresas que desarrollan SIG, los cuales, en un desarrollo basado en componentes, buscan en la Web o en repositorios equivalentes, componentes que brinden servicios que se adecuen a las

---

<sup>1</sup> El concepto de servicio que se utiliza en esta Tesis es diferente a aquel de “servicio web” en una arquitectura basada en SOA o SOAP. A diferencia de un servicio web, en esta Tesis se asocia servicio con una funcionalidad de un único componente [63].

necesidades de sus aplicaciones. El modelo de oferta se refiere a los requerimientos o servicios que brindan los desarrolladores de componentes para SIG, para lo cual publican en la Web los componentes desarrollados junto a la información técnica y funcional asociada a los mismos.

Por lo tanto, para que ambas partes logren entenderse, es fundamental la forma en que los componentes son publicados, llamado *proceso de publicación*, y la forma en que los mismos son seleccionados, llamado *proceso de selección* [44]. Además, para integrar la funcionalidad de ambos procesos surge la necesidad de utilizar un *modelo de mediación* que facilite tanto la recuperación de la información como la selección automatizada de componentes candidatos [32].

## 1.2. Objetivos

El objetivo de este trabajo consiste en mejorar el modelo de demanda de componentes para SIG, al definir herramientas que estandaricen los servicios requeridos. Para ello será necesario:

a) *Clasificar Servicios SIG*: proponer una clasificación de servicios SIG en base a la bibliografía existente. La misma debe contemplar tanto los estándares internacionales para arquitecturas de servicios geoespaciales (creados por el Consorcio OpenGis y la ISO TC211) como las necesidades del mercado de usuarios y desarrolladores de SIG. Además de poseer requerimientos en cuanto a aspectos funcionales (servicios necesitados), la clasificación debe contemplar el conjunto de atributos no técnicos solicitados y de interés para este grupo de usuarios.

b) *Construir el Modelo de Demanda*: proponer un modelo de demanda considerando la información real obtenida del estudio de los estándares y de los usuarios y desarrolladores de SIG. El modelo instanciará la clasificación teórica de los requerimientos funcionales, en base a la demanda actual existente de los servicios SIG y a sus características específicas. Se aplicarán mejoras para la identificación automática de servicios.

c) *Validar el Modelo*: validar la viabilidad del modelo de demanda en base a un caso de estudio real.

### **1.3. Metodología**

En este trabajo se utilizará el método de Investigación-Acción (I-A) que ha obtenido una amplia aceptación y aplicación en la investigación en ingeniería del software en los últimos años [5][75]. Este método presenta como principales características [8]: orientación a la acción y al cambio, focalización en un problema, un modelo de proceso “orgánico” que engloba etapas sistemáticas y algunas veces iterativas.

El método permite generar un beneficio al “cliente” de la investigación y, al mismo tiempo, generar “conocimiento de investigación” relevante. Identificamos para ello durante la investigación los siguientes tipos de roles:

- El investigador: rol desempeñado por el tesista.
- El objeto investigado: clasificación de requerimientos de usuarios según categorías de servicios geográficos.
- El grupo crítico de referencia: constituido por desarrolladores de aplicaciones geográficas que buscan componentes software que brinden los servicios requeridos.
- El beneficiario: empresas desarrolladoras de aplicaciones geográficas, que utilicen el paradigma de Desarrollo Basado en Componentes.

Nuestro trabajo se centra en identificar problemas en el objeto investigado desde el punto de vista del grupo crítico de referencia y en proponer mejoras para el objeto investigado – la categorización de los requerimientos geográficos.

### **1.4. Estructura**

El documento de Tesis está organizado en los siguientes capítulos:

- En el Capítulo 2 se describe el proceso de desarrollo basado en componentes y en particular se analiza la fase de selección de componentes. Además, se presentan casos de uso como medio para especificar requerimientos de componentes y se estudian distintas herramientas para extraer automáticamente el conocimiento de los casos de uso.
- En el Capítulo 3, se describen las principales características de los sistemas de información geográficos y se detallan distintas clasificaciones existentes para servicios geográficos. También se resumen trabajos relacionados en el área de selección de componentes para SIG.
- El Capítulo 4 se centra específicamente en nuestro trabajo de investigación. Allí describimos la propuesta de mejora para el proceso de demanda y la herramienta de software que implementa nuestra propuesta.
- En el Capítulo 5 se presenta un caso de estudio para describir el funcionamiento de nuestra propuesta. El mismo está acompañado de una encuesta realizada a expertos del dominio lo que nos permite evaluar el resultado obtenido con la herramienta.
- El Capítulo 6 presenta las conclusiones y los trabajos futuros destacando las contribuciones y trabajos futuros de esta tesis.
- En el Apéndice A se presenta la categorización completa propuesta por el estándar ISO 19119, en la que se detallan las categorías de servicios, los servicios ejemplo de cada una, y las descripciones de cada servicio ejemplo.
- En el Apéndice B se detalla la totalidad de casos de uso que especifican la funcionalidad de la aplicación geográfica ejemplo.
- En el Apéndice C se muestran los resultados de la clasificación de los requerimientos especificados en los casos de uso detallados en el Apéndice B mediante la utilización de la herramienta ReqGIS.

- En el Apéndice D se exhiben las respuestas a las encuestas realizadas a desarrolladores de aplicaciones geográficas.

## **1.5. Marco de desarrollo**

El trabajo de esta tesis se realizó en el marco de los siguientes proyectos de investigación:

- 29/B090 – “Identificación de Soluciones Off-The-Shelf para Sistemas de Información Geográficos”- Universidad Nacional de la Patagonia Austral. 2008-2009
- 29/B107 – “Mejora del Proceso de Selección de Componentes para Sistemas de Información Geográficos”- Universidad Nacional de la Patagonia Austral. 2010-2011



## Capítulo 2

# FUNDAMENTOS DEL TRABAJO DE TESIS

En este capítulo se analizan diversos temas vinculados al mejoramiento del desarrollo de aplicaciones SIG mediante la agilización del proceso de selección de componentes. Para ello, se estudian los casos de uso como herramienta para especificar requerimientos de componentes, se analizan distintos enfoques para la selección de componentes adecuados y se describen técnicas del procesamiento del lenguaje natural que serán utilizadas en la propuesta de mejora.

### 2.1. Desarrollo de sistemas basados en componentes

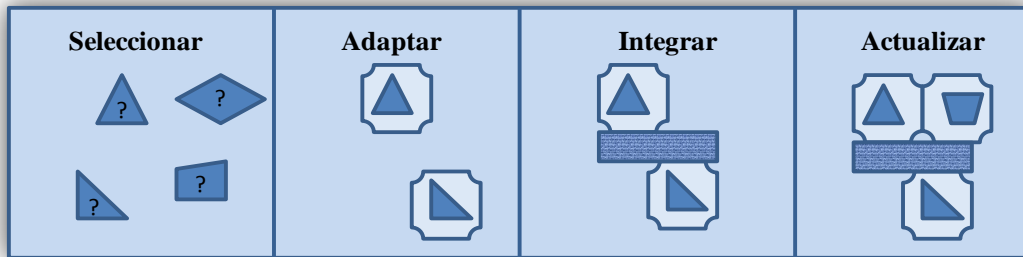
La reutilización de software se ha convertido paulatinamente en una práctica común para la construcción de productos software. La reducción de costo, tiempo y esfuerzo en los procesos de elaboración son algunos de los motivos que han llevado a los desarrolladores de software a considerar técnicas para la reutilización de partes en prácticamente cualquier fase del ciclo de vida del producto (análisis, diseño e implementación) [53]. En consecuencia se ha incrementado el reuso de componentes software para la elaboración de estos sistemas, los analistas utilizan metodologías, procesos y técnicas específicas. El sistema a desarrollar estará compuesto por una o más aplicaciones software, e incluso puede que algunas de estas aplicaciones hayan sido construidas mediante la composición de otras partes (componentes) durante el desarrollo del sistema [31].

Los sistemas de componentes OTS (Off-The-Shelf) se construyen mediante la integración a gran escala de componentes desarrollados por terceros. Sin embargo, dado que los componentes son diseñados para satisfacer las necesidades de un mercado genérico y no para satisfacer los requerimientos particulares de una organización, no se puede asegurar que alguno de los componentes disponibles satisfaga todos los requerimientos especificados y se adapte a la arquitectura del software. Por lo tanto, la

selección de componentes necesita que la especificación de requerimientos de usuario haya sido definida previamente.

En consecuencia, el desarrollo de sistemas basados en componentes introduce cambios fundamentales en algunas de las actividades ya conocidas del desarrollo tradicional. El diseño de la arquitectura puede realizarse al mismo tiempo que la evaluación del componente, y se incorporan nuevas actividades tales como la adaptación e integración del producto. Asimismo, el proceso de ingeniería de requerimientos en DSBC tiene características específicas inherentes al paradigma, como la priorización y negociación de requerimientos [3].

Existen diversas propuestas de procesos de desarrollo de sistemas basados en componentes (DSBC). Sin embargo la mayoría de ellos coincide con un proceso compuesto por un conjunto de etapas [11] que pueden variar en el nombre según el proceso al que pertenecen, pero que esencialmente son las siguientes (Figura 2.1): Selección de componentes, Adaptación, Integración y Actualización.



**Figura 2.1:** Actividades del proceso de desarrollo basado en componentes

A continuación se describe cada una de las etapas que componen el proceso de desarrollo basado en componentes.

**Selección de Componentes.** Esta etapa debe realizarse lo antes posible en el ciclo de vida de desarrollo, dado que si un componente seleccionado no es el adecuado, se requerirá de mucho esfuerzo para adaptarlo e integrarlo en el sistema desarrollado. En un proceso completo de selección se debe considerar no solamente la funcionalidad requerida, sino también los aspectos no funcionales como rendimiento, confiabilidad, flexibilidad.

**Adaptación de Componentes.** Esta actividad incluye el desarrollo de todos los adaptadores de software necesarios. Un enfoque muy utilizado es el de envoltorio (“wrappers”) de componentes. Los adaptadores o envoltorios pueden filtrar entradas y/o salidas de los componentes. Este enfoque restringe la exhibición de la funcionalidad no deseada interponiendo una barrera alrededor de los componentes, limitando su funcionalidad.

**Integración de Componentes.** Esta fase abarca todos los esfuerzos de desarrollo necesarios para interconectar los distintos componentes seleccionados como una única aplicación. En esta fase también se construyen partes adicionales del sistema deseado que no son soportadas por ninguno de los componentes seleccionados y también se realizan las pruebas del sistema completo.

**Actualización de Componentes.** La actividad de actualización de componentes es una tarea complicada que consume mucho tiempo. Sin la planificación adecuada, un cambio a un componente puede tener repercusiones imprevistas en muchos otros componentes. Para evitar estos problemas, es necesario definir las interfaces de los componentes y controlar la interacción entre los mismos.

La selección de componentes COTS (Commercial OTS) suele ser una tarea no trivial, donde hay que considerar diferentes aspectos de los componentes comerciales y de la arquitectura de software [47]. En la actualidad representa el principal centro de interés en el área de los sistemas basados en componentes comerciales, y en la literatura podemos encontrar diversos enfoques que serán analizados en la próxima sección.

### **2.1.1. Selección de Componentes**

La selección de componentes es un proceso que determina qué componentes ya desarrollados pueden ser utilizados. Es una tarea compleja que requiere de mucho tiempo y esfuerzo, donde gran parte del tiempo se utiliza para buscar y analizar los posibles componentes candidatos [12].

En un proceso de desarrollo basado en COTS, la evaluación temprana de productos software candidatos es un aspecto clave del ciclo de vida de desarrollo del sistema. El éxito depende fundamentalmente del entendimiento preciso sobre la funcionalidad y limitaciones de cada uno de los productos individualmente.

Diferentes autores [53] [51] describen el proceso de selección como un conjunto de fases y estrategias. En general, todos ellos coinciden en que el proceso de selección está compuesto básicamente por tres actividades principales superpuestas e iterativas [6]:

- 1) Búsqueda de componentes candidatos disponibles en el mercado
- 2) Evaluación de los mismos con respecto a los requerimientos del sistema
- 3) Decisión sobre los COTS a utilizar

En la primera actividad, búsqueda de componentes candidatos, se realiza la exploración de los distintos segmentos del mercado para encontrar los componentes y la información existente sobre los servicios ofrecidos (y requeridos) que puedan coincidir con los requerimientos de usuario definidos. Se identifican las propiedades de un componente, como por ejemplo, la funcionalidad del componente (qué servicios proporciona) y otros aspectos relativos a la interfaz de un componente (como el uso de estándares), aspectos de calidad que son difíciles de aislar y aspectos no técnicos, como la cuota de mercado de un vendedor o el grado de madurez del componente dentro de la organización. Esta fase de búsqueda es tediosa, donde hay mucha información difícil de cuantificar, y en algunos casos, difícil de obtener. Debemos notar que el término “servicios” en este contexto se asocia al comportamiento del componente y se utiliza para identificar la adecuación con respecto a los requerimientos provistos por una aplicación destino, que será la anfitriona en el momento de componer o integrar.

A continuación, en la actividad de evaluación, se efectúa la medición de distintas características de los componentes candidatos determinándose el grado de aptitud del componente. Esta evaluación puede realizarse mediante alguna de las siguientes estrategias [54]:

**Filtrado progresivo**, donde se comienza con un gran número de COTS y progresivamente se definen criterios discriminantes a través de iteraciones sucesivas de ciclos de evaluación de productos.

**Armado de rompecabezas**, el cual asume que un sistema basado en componentes requiere que varios componentes se integren como piezas en un rompecabezas, es decir que se deben considerar los requerimientos de cada componente y simultáneamente tener en cuenta los requerimientos de los demás componentes del rompecabezas.

**Identificación de piedra angular**, donde se comienza identificando un requerimiento clave y luego se realiza la búsqueda de productos que satisfagan ese requerimiento. Esto permite eliminar rápidamente una gran cantidad de componentes que no cumplen con un determinado requisito.

Por último, en la actividad de toma de decisión, teniendo en cuenta la información proporcionada por la actividad de evaluación y también otra información relevante como por ejemplo restricciones del sistema, se resuelve cuál es el componente a utilizar [51].

## **2.1.2. Enfoques para la Selección de Componentes**

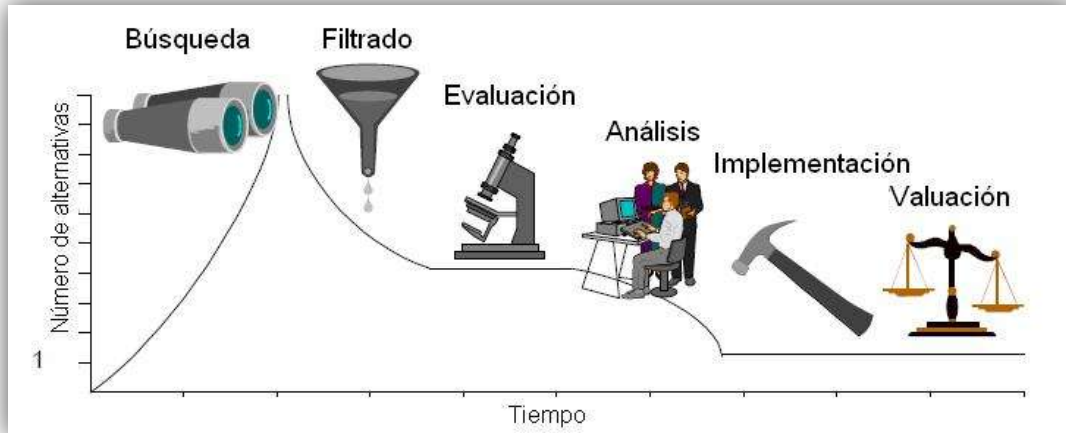
Durante el transcurso de las últimas dos décadas se han propuesto distintos enfoques para la selección de componentes. A continuación se describen algunos de ellos.

### **2.1.2.1. OTSO (Off-The-Shelf Option)**

El primer método más difundido fue OTSO (Off-The-Shelf Option), propuesto por Kontio en 1995 [38]. El método define la estructura básica del método de selección de COTS y sirve como base para otros enfoques. Los principios fundamentales del método son: la definición del proceso, la definición del criterio de evaluación y el enfoque de estimación costo-valor.

a) *Definición del proceso de selección*

El método OTSO incluye la definición de un proceso general para la selección de componentes reusables, compuesto por seis fases que se muestran en la Figura 2.2.



**Figura 2.2:** Fases del método de selección OTSO

El eje horizontal de la Figura 2.2 representa el progreso de la evaluación y el eje vertical la cantidad de alternativas consideradas en cada fase. En la fase de *búsqueda*, el número de alternativas posibles crece rápidamente, luego los candidatos potenciales son *filtrados* y las alternativas seleccionadas son *evaluadas* más detalladamente para determinar el grado de satisfacción del criterio de evaluación. Estos resultados se documentan sistemáticamente.

A continuación, durante la fase de *análisis*, se interpretan los datos provistos por la fase de evaluación. A veces es posible obtener conclusiones directas si una de las alternativas es claramente superior a otras. Sin embargo, en la mayoría de los casos es necesario usar técnicas de decisión de múltiples criterios para tomar una decisión acerca de la alternativa a seleccionar y utilizar.

Finalmente, para optimizar el proceso de selección y proveer retroalimentación para un futuro reuso del componente, es necesario realizar la *valuación* del éxito del componente reusado en el proyecto.

*b) Definición del criterio de evaluación*

La definición del criterio de evaluación se realiza en paralelo junto a las fases de búsqueda, filtrado y evaluación. El criterio de evaluación juega un rol central en el proceso de selección. La selección de un componente OTS adecuado requiere la consideración de distintos factores.

El proceso de definición del criterio de evaluación descompone los objetivos de reuso en un conjunto jerárquico de criterios. Cada rama de la jerarquía finaliza en una “prueba”: medición, observación o prueba bien definida que es realizada para determinar cómo una alternativa satisface el criterio asociado a la prueba. Es posible identificar cuatro fases en la definición del criterio de evaluación: definición del criterio de búsqueda, definición de la línea base, definición detallada del criterio de evaluación, y ponderación del criterio.

*c) Enfoque de estimación costo/valor*

La decisión sobre qué componente OTS se va a utilizar, se basa en los costos y valores estimados que aportarán las distintas alternativas al proyecto. El método OTSO utiliza el enfoque AHP (Analytic Hierarchy Process) [67] para consolidar la evaluación de datos para la toma de decisión. La técnica AHP se basa en la idea de descomponer un problema de múltiples criterios en una jerarquía de alternativas (o conjunto de soluciones factibles) para la consecución de la meta marcada. En la metodología AHP las alternativas son evaluadas bajo unos criterios que se tienen en cuenta en la decisión (como por ejemplo: calidad, cualidad, tamaño, durabilidad, etc), pudiendo ser éstos cualitativos o cuantitativos, lo que constituye una de las grandes virtudes de este método. El método AHP consiste, a grandes rasgos, en los cuatro siguientes pasos: i) Definir una jerarquía de decisión con niveles, incluyendo el objetivo en el nivel 1, los criterios implicados en la toma de decisiones en el nivel 2 y las alternativas posibles en el nivel 3; ii) Estimar pesos o preferencias de los criterios (nivel 2) comparándolos por pares; iii) Comparar las alternativas (nivel 3) por parejas utilizando cada criterio (nivel 2) por separado; y iv) Seleccionar la mejor alternativa.

La emisión de juicios para el establecimiento de los pesos se realiza por comparaciones pareadas, esto es, se determina primero el peso o importancia de cada

uno de los criterios mediante su enfrentamiento por pares en todas las posibles combinaciones, estableciéndose la preferencia en cada par mediante una escala discreta, como por ejemplo la propuesta por Saaty [67] con valores entre 1 (misma importancia) y 9 (absoluta importancia).

Una vez establecido el peso de los criterios dentro de una decisión, se estima el peso de cada alternativa realizando el mismo proceso y utilizando la misma escala anteriormente mencionada, pero esta vez se enfrentan las alternativas por separado bajo cada criterio. A partir de todas esas comparaciones pareadas, una serie de cálculos matriciales permiten obtener el peso o prioridad final de las alternativas consideradas, además de un grado de inconsistencia, esto es, un valor que indica si la calificación establecida es lógica, o si por el contrario tiene cierta inconcordancia. Esta inconsistencia se calcula en base a la valoración de la homogeneidad de las matrices normalizadas con las matrices de comparación, y según Saaty [68], una matriz de comparaciones se muestra consistente, y con ello válida, si este grado de inconsistencia es menor de 0,1.

### **2.1.2.2. PORE (Procurement Oriented REquirements)**

El enfoque PORE (Procurement Oriented Requirements) [46] destaca la importancia de un proceso de ingeniería de requerimientos apropiado para el DSBC.

Este enfoque se basa en un proceso iterativo y paralelo de elicitación de requerimientos y evaluación y selección de COTS. Su modelo identifica cuatro objetivos esenciales para seleccionar/rechazar productos candidatos:

- requerimientos de cliente esenciales atómicos (Objetivo 1);
- requerimientos de cliente no esenciales atómicos (Objetivo 2);
- requerimientos de cliente complejos no atómicos (Objetivo 3);
- requerimientos del cliente usuario (Objetivo 4).

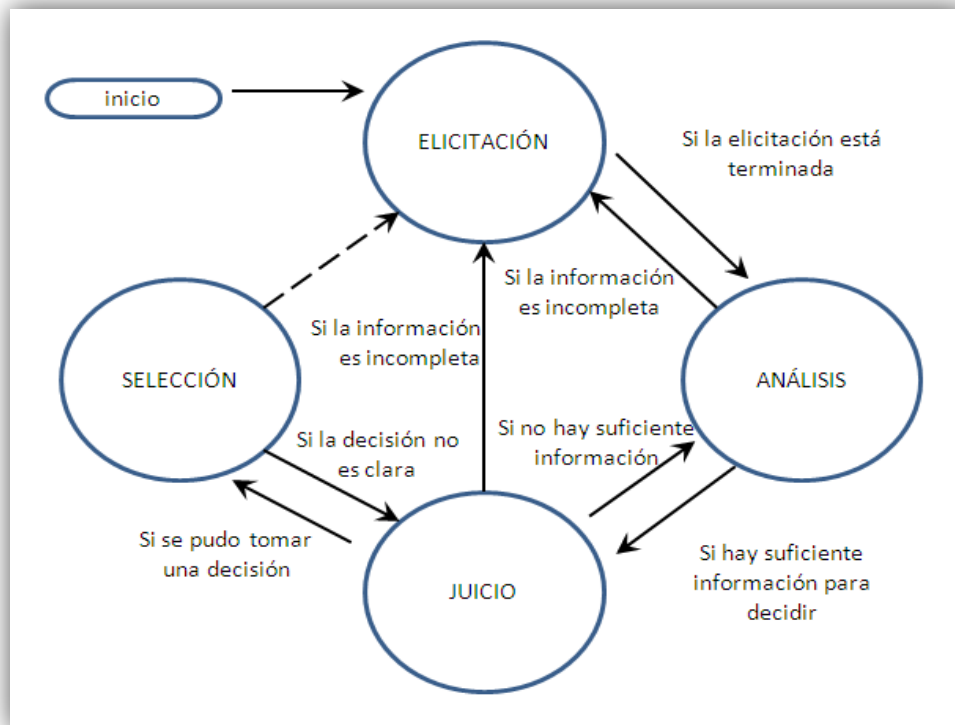
Estos objetivos deben ser alcanzados en secuencia. Para ello, PORE determina los siguientes cuatro procesos genéricos que permiten lograr cada uno de los cuatro



objetivos esenciales, tal como se ilustran en la Figura 2.3: *elicitación, análisis, juicio, y selección.*

En el proceso de *Elicitación* se obtienen los requerimientos de los clientes, luego en el proceso de *Análisis*, se estudia la información recolectada para determinar si está completa y correcta. A continuación en el proceso de *Juicio*, si la información obtenida es suficiente, se decide sobre el cumplimiento de los requerimientos del producto y finalmente durante el proceso de *Selección*, se eligen uno o más candidatos COTS.

Este método, tiene la ventaja que no asume la existencia de un conjunto de requerimientos predefinidos, sino que realiza en paralelo la elicitación de requerimientos y la selección de los componentes, facilitando la retroalimentación entre ambas actividades. Sin embargo no establece claramente un proceso que guíe la negociación de las diferencias entre los requerimientos y los componentes disponibles.



**Figura 2.3:** Procesos genéricos para lograr los objetivos esenciales

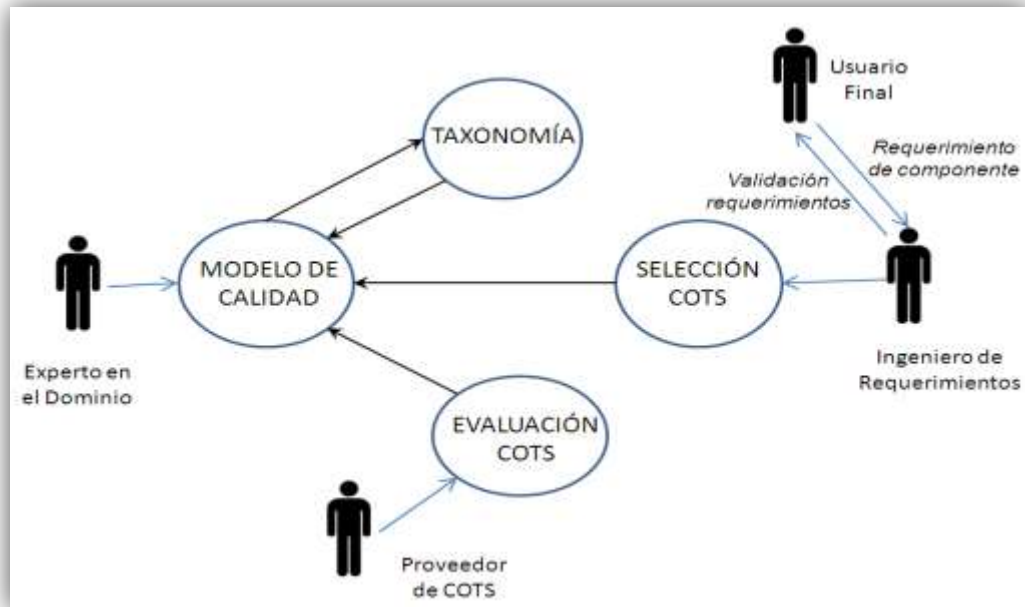
El proceso iterativo PORE realiza la selección por descarte, es decir aquellos productos que no cumplen los requerimientos del cliente son eliminados de la lista de candidatos. Sin embargo, este método no establece cómo se recolectan los COTS candidatos para la lista inicial.

### **2.1.2.3. Sistema DesCOTS (Description, Evaluation and Selection of COTS components)**

En este enfoque [28], se define un proceso para selección de componentes basado en la utilización de modelos de calidad. El proceso comprende distintas actividades, cada una de las cuales es soportada por una herramienta software. Las herramientas que componen el sistema DesCOTS son:

- Herramienta Modelo de Calidad, la cual permite definir los modelos de calidad,
- Herramienta Evaluación de COTS, para evaluar los componentes,
- Herramienta Selección de COTS, permite definir los requerimientos que guiarán la selección de componentes COTS,
- Herramienta Taxonomía, permite organizar los dominios de los COTS en forma de una taxonomía que soporta el reuso de los modelos de calidad.

En la Figura 2.4 se muestra la arquitectura del sistema DesCOTS. En este modelo se describen los actores intervinientes y las relaciones o dependencias entre ellos. Los actores pueden ser actores software (herramientas del sistema DesCOTS) o actores humanos (experto en el dominio, proveedor de componentes, usuario final, ingeniero de requerimientos):



**Figura 2.4:** Arquitectura del sistema DesCOTS. Relación de cada herramienta del sistema con los distintos tipos de usuarios definidos

Así, la herramienta “Modelo de Calidad” provee funcionalidad para definir los factores de calidad del software, para reusarlos en diferentes modelos de calidad, establecer relaciones entre ellos, diseñar métricas para su evaluación futura y también definir patrones de requerimientos.

La herramienta “Evaluación COTS” brinda soporte a la tarea de evaluación de componentes candidatos COTS usando un modelo de calidad definido por la herramienta Modelo de Calidad y apropiado al dominio de los componentes. Sus funciones principales son la administración de catálogos de componentes COTS y la gestión de sus evaluaciones. Así, la herramienta permite la registración, modificación y eliminación de los distintos componentes COTS y la gestión de sus evaluaciones.

La herramienta “Selección COTS” es la responsable de mapear dos procesos distintos. Brinda soporte para la definición de los requerimientos para la selección, y además analiza los requerimientos definidos y las evaluaciones de componentes COTS para informar sobre la selección de un componente perteneciente a un dominio específico. Ambos procesos están relacionados y pueden aplicarse de manera cíclica

tantas veces como sea necesario para dar soporte efectivo a la negociación de requerimientos.

La herramienta “Taxonomía” está integrada junto a la herramienta Modelo de Calidad para soportar el reuso de los modelos de calidad. Además la herramienta Taxonomía también es utilizada por la herramienta Selección COTS, cuando el Ingeniero de Requerimientos al determinar los requerimientos para la selección, recorre la taxonomía para encontrar cuál es el tipo de componente COTS que se necesita.

Este método, a diferencia de otros enfoques considera las características de calidad y puede ser adaptado a distintos dominios. Sin embargo, tampoco provee un procedimiento definido para el manejo de las diferencias entre los requerimientos y las características de los componentes.

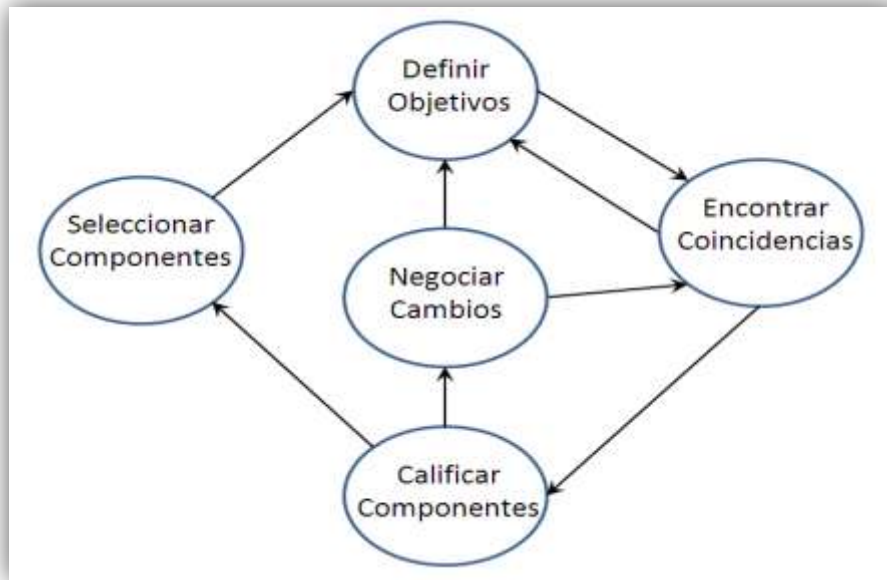
#### **2.1.2.4. CARE (COTS-Aware Requirements Engineering)**

Otro enfoque interesante, es CARE (COTS-Aware Requirements Engineering) [13], el cual intenta definir un enfoque de selección de componentes más completo. Define dos tipos de requerimientos: nativos (requerimientos de clientes) y foráneos (requerimientos implementados por COTS existentes). CARE intenta reducir la brecha entre los requerimientos nativos y foráneos mediante un proceso de negociación entre clientes que pueden modificar sus objetivos y vendedores que pueden adaptar sus COTS. Las especificaciones de los COTS son almacenadas en una base de conocimiento o repositorio, al cual se realizan las consultas para encontrar los componentes requeridos.

En este proceso se definen las siguientes actividades, que se ejecutan de forma iterativa, según se ilustra en la Figura 2.5:

- a) **Definir objetivos:** se crea un modelo de los objetivos del sistema,
- b) **Encontrar componentes coincidentes:** se buscan los componentes COTS que concuerden con las necesidades del sistema en desarrollo,

- c) **Calificar componentes:** se ponderan los componentes según el grado de coincidencia con los requerimientos definidos,
- d) **Negociar cambios:** se decide sobre la modificación de los requerimientos definidos o modificación a los componentes disponibles,
- e) **Seleccionar componentes:** se eligen los componentes más adecuados.



**Figura 2.5:** Actividades del proceso CARE

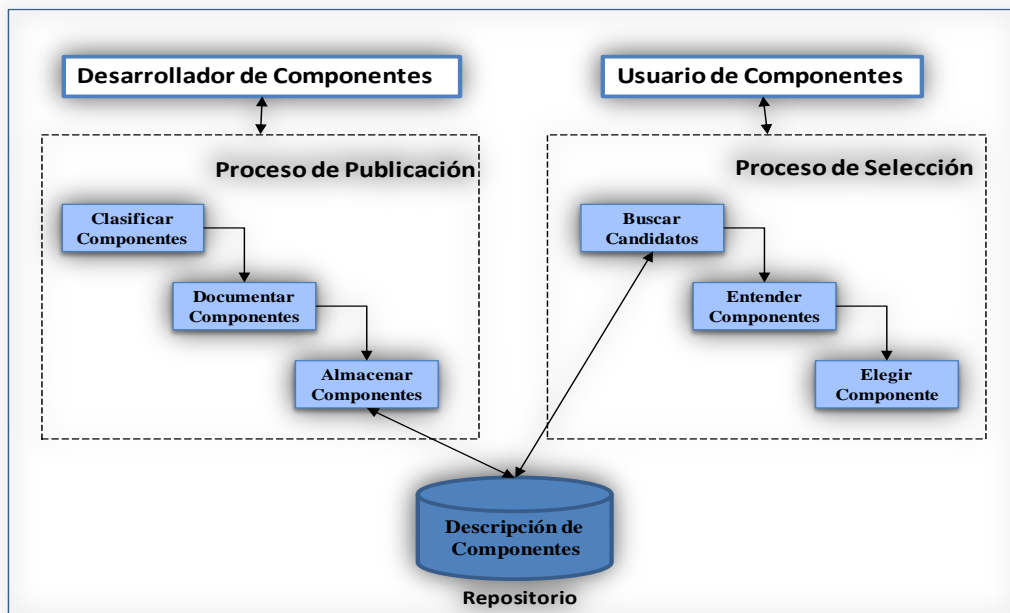
A pesar de que el método intenta cubrir el vacío dejado por sus métodos predecesores, también presenta algunas dificultades: 1) Asume que el repositorio de COTS existe, pero no especifica cómo construir dicho repositorio, 2) No establece cómo obtener la información para describir los COTS en el repositorio, 3) El proceso de búsqueda no es muy eficiente teniendo que buscar componentes dentro de un rango muy amplio de descripciones.

### 2.1.2.5. Propuesta de Lucena

Un enfoque distinto para el desarrollo de aplicaciones basadas en componentes es el presentado en [44]. En ese trabajo, Lucena propone un proceso de desarrollo basado en componentes reusables en el que intervienen los siguientes tres factores:

- **Desarrollador del Componente:** es el encargado de investigar la funcionalidad requerida, desarrollar el producto y luego almacenar el componente en un repositorio para ser usado posteriormente.
- **Usuario del Componente:** es el desarrollador de aplicaciones que busca construir sus aplicaciones mediante la utilización de componentes previamente desarrollados.
- **Repositorio de Componentes:** es el sitio donde se almacenan los componentes desarrollados, junto con la documentación que los describe.

En la Figura 2.6 se pueden observar las distintas actividades que integran los procesos de publicación y selección de componentes.



**Figura 2.6:** Principales actividades de los procesos de Publicación y Selección de componentes

Los componentes desarrollados para reuso, deben construirse para satisfacer distintas demandas futuras, por lo cual deben ser componentes estándares lo suficientemente generales como para ser fácilmente utilizados en proyectos futuros.

Luego de concluir el desarrollo del componente, éste debe ser almacenado en el repositorio. El desarrollador debe documentar el componente con información de suma importancia para la posterior búsqueda, entendimiento y toma de decisión sobre la utilidad del nuevo componente desarrollado. Este procedimiento se efectúa en tres pasos, la clasificación del componente, su documentación completa y por último el almacenamiento de toda la información en el repositorio. La denominación adoptada para este procedimiento es “Proceso de Publicación”.

Desde la perspectiva de los usuarios, el desarrollo de aplicaciones basadas en componentes reusables comienza con la recepción de los requerimientos del sistema. Los desarrolladores realizan el análisis teniendo en cuenta los componentes disponibles, es decir durante el análisis buscan los componentes existentes para guiar su trabajo. Para esto se realiza un procedimiento llamado “Proceso de Selección”.

El primer paso para la selección de componentes consiste en la búsqueda de componentes candidatos que cumplan determinadas características funcionales. El objetivo es proveer a los desarrolladores de aplicaciones una lista de componentes que satisfagan sus requerimientos.

De esta manera, en el segundo paso del proceso de selección, el número de candidatos posibles se ha reducido. Los usuarios de componentes deben estudiar, entender y finalmente verificar si estos posibles candidatos realmente satisfacen sus necesidades. Esta evaluación técnica es un paso excluyente pero no es suficiente para decidir si un componente es el más adecuado para una aplicación.

En el último paso, los usuarios de componentes trabajan con los componentes aprobados en el paso anterior; el objetivo es determinar qué componente será integrado a su aplicación.

La forma en que se realiza la búsqueda de componentes puede diferir. Por ejemplo, se puede comenzar la búsqueda con información relacionada a características

funcionales, así como también a características operacionales del componente deseado. Es decir que se puede limitar el número de posibles candidatos mediante un filtrado por términos clave y características técnicas.

El resultado final esperado del proceso de selección es la elección de un componente para ser usado en la aplicación a desarrollar. En el mejor de los casos, este componente será el mejor de entre un conjunto de componentes con características y rendimientos similares. Sin embargo, esta situación no ocurre en todos los casos. A veces no es posible seleccionar un componente que cumpla con todos los requerimientos esperados. Esto puede suceder cuando el componente buscado no existe, o cuando un componente no funciona como se esperaba o no cumple con los requerimientos no técnicos.

Además de los enfoques analizados en secciones anteriores, existen también otros métodos. Sin embargo, todos utilizan procedimientos manuales, que involucran la participación activa de uno o más expertos en el dominio para guiar el proceso de selección y así poder decidir cuál es el componente más adecuado a utilizar. Así, estos métodos brindan un marco para la selección de componentes, pero no se puede estimar el tiempo y esfuerzo que demandará la selección del componente buscado, dependiendo en gran medida del conocimiento del experto en el dominio y de la información con que se cuente de los componentes analizados.

## **2.2. Especificación de Requerimientos para COTS**

En un entorno de desarrollo tradicional, el objetivo principal del ingeniero de requerimientos es asegurar que la especificación de requerimientos satisface los deseos de los usuarios y que representa una descripción clara y concisa del sistema a desarrollar. De esta manera, los requerimientos especificados se transforman en una arquitectura de software y por último se implementan. El proceso de ingeniería de requerimientos para desarrollo basado en componentes es afectado por problemas muy distintos a los de sistemas tradicionales [3]:



**Flexibilidad de requerimientos:** En desarrollos basados en componentes, los requerimientos deben ser más flexibles y menos específicos. Por ejemplo, suponga que el rendimiento es un requerimiento crítico de un sistema de base de datos pero ninguno de los productos evaluados satisface el tiempo de respuesta deseado. Esta es una situación típica para decidir si comprar o desarrollar. Si la decisión final es comprar un producto, el cliente debe aceptar las limitaciones y los requerimientos que no pueden ser satisfechos por ningún componente disponible. Cuando se desarrollan sistemas con el objetivo de maximizar el uso de COTS, los requerimientos especificados no deben ser tan estrictos que excluyan el uso de COTS o que requieran grandes modificaciones al producto para satisfacer esos requisitos. De hecho, un enfoque interesante consiste en permitir que las características de los COTS disponibles determinen los requerimientos. Consecuentemente, es necesario lograr el mejor balance entre precisión y flexibilidad en los requerimientos.

**Reducción del control:** los productos COTS son desarrollados en base a requerimientos que los vendedores creen que serán útiles a la mayor cantidad de posibles clientes. Los desarrolladores de componentes tratan de satisfacer las necesidades del mercado en vez de satisfacer los requerimientos de una organización particular. Por lo tanto, los COTS son diseñados para satisfacer requerimientos genéricos. Esto implica que el cliente debe poseer un entendimiento preciso de las características del producto para decir qué partes deben ser adaptadas para cumplir con sus necesidades particulares. Además, el desarrollador/vendedor tiene control total sobre las versiones y actualizaciones del producto, exponiendo a los clientes a situaciones inesperadas sobre las que no tienen control.

**Proceso de requerimientos continuo:** En el desarrollo tradicional de sistemas, los requerimientos evolucionan a la par de los cambios que se producen en el ambiente en el cual los sistemas operan. Entre los cambios típicos a las especificaciones de requerimientos se encuentran agregar o eliminar requerimientos y corregir errores. La evolución de los requerimientos puede causar una inestabilidad temporal, pero tan pronto como los cambios son administrados y los requerimientos acordados, la situación es controlada. Sin embargo, en desarrollos basados en COTS, los

requerimientos son extremadamente volátiles principalmente debido a los rápidos cambios del mercado de COTS. Los vendedores solicitan a los clientes aceptar nuevas versiones con nuevas características que pueden no ser deseadas o conflictivas para los requerimientos planteados. Esto lleva a un proceso continuo de negociación. Las presiones de la competencia en el mercado fuerzan a los vendedores a innovar y a diferenciar características de los productos más que a estandarizarlos.

En consecuencia, en el entorno de DSBC la fase de requerimientos consiste en la generación de un conjunto de requerimientos flexibles seguidos por la exploración del mercado para seleccionar aquellos componentes que mejor se ajusten a estos requerimientos.

### **2.2.1. Conocimiento a partir de Casos de Uso**

Los casos de uso capturan sencillamente los requerimientos funcionales, es decir describen el comportamiento del sistema, como éste interactúa con sus usuarios y otros sistemas para hacer algo requerido por sus usuarios. Los casos de uso pueden utilizarse también para capturar requerimientos no funcionales específicos a un caso de uso. Los requerimientos no funcionales se describen mejor usando requerimientos declarativos. Estos son luego adjuntados a los casos de uso a los que aplican. En resumen, los requerimientos funcionales pueden ser capturados como casos de uso, y muchos de los requerimientos no funcionales pueden ser asociados a los mismos. Así, se espera que el modelo de casos de uso esté complementado con documentación de requerimientos adicional que contenga los requerimientos no funcionales. Estos requerimientos no funcionales son capturados en las especificaciones suplementarias e incluyen requerimientos de plataforma del sistema, restricciones de diseño, requerimientos de rendimiento y otros requerimientos que no pueden expresarse como una secuencia de acciones que el sistema realiza [4].

En general, el modelado de casos de uso se basa en una técnica formal. Esto representa tanto una fortaleza como una debilidad. Es una fortaleza dado que se puede utilizar el formalismo subyacente para mejorar la precisión y proveer profundidad adicional y rigor a las actividades de modelado. Es una debilidad porque es fácil de

asumir (erróneamente) que el modelado de casos de uso está limitado al dibujo de diagramas y porque la formalidad y la terminología pueden confundir a las personas nuevas en casos de usos. La realidad es que el formalismo provee una estructura para la creación de modelos y los diagramas proveen una buena visión general del sistema. Pero el verdadero valor de los casos de uso se encuentra en las descripciones textuales de los mismos. Es en las descripciones de los casos de uso que se encuentra la mayor parte del contenido del modelo, y es en la confección de las descripciones de casos de uso en que se requiere más esfuerzo.

Los casos de uso han logrado una amplia aceptación como medio para describir las interacciones realizadas entre el sistema y su entorno. Actualmente el Lenguaje Unificado de Modelado (UML) [58] es un estándar altamente adoptado que define los conceptos centrales para el modelado de casos de uso. Al aplicar este estándar, los casos de uso se identifican y estructuran mediante los diagramas de casos de uso de UML y luego se especifican con descripciones textuales.

Un caso de uso describe cómo una entidad (sistema) coopera con otras entidades (actores), comunicándose y realizando acciones internas para lograr un objetivo particular. Un caso de uso se especifica como una secuencia de pasos que conforman el escenario principal, con escenarios alternativos que detallan condiciones excepcionales a la ejecución normal del caso de uso. Las acciones descritas en cada paso se formulan en lenguaje natural para facilitar su lectura a todo tipo de audiencia.

Existen distintas propuestas en cuanto a la notación utilizada para la especificación de casos de uso. La diferencia entre las mismas está relacionada con el nivel de formalismo empleado para realizar la especificación. Existen notaciones formales [35][72][74], que proveen la posibilidad de inspeccionar y analizar las descripciones de casos de uso automáticamente, pero cuentan con la desventaja de la dificultad de interpretación para usuarios no técnicos. Existen también enfoques [4][14][27] que proponen la utilización de formas restringidas de lenguaje natural, mediante la utilización de plantillas para detallar el comportamiento de los casos de uso, de las cuales la más difundida es la propuesta de Cockburn [14], la cual se puede ver en la Figura 2.7.

Nombre Caso de Uso	<El nombre debe ser el objetivo del caso de uso, descrito como una frase corta>	
Descripción	<Una descripción más detallada del objetivo, si es necesario>	
Actores	<Listar los actores involucrados en el caso de uso>	
Precondiciones	<Lo que se considera como estado actual del sistema>	
Postcondiciones	<Las condiciones de finalización exitosa del caso de uso, tras haberse satisfecho los objetivos de los actores>	
Escenario Principal	1	<Describir los pasos que narran la interacción entre los actores y el sistema>
	2	
	...	
Extensiones	1	<Describir las extensiones, haciendo referencia al paso del escenario principal que extienden>
	2	
	...	
Sub-Variaciones	1	<Describir las sub-variaciones, que pueden causar bifurcación en el escenario principal >
	2	
	...	

**Figura 2.7:** Plantilla de Caso de Uso Textual

Esta plantilla es utilizada para realizar la descripción textual de un caso de uso, y así aclarar de forma literal el diagrama que representa gráficamente al caso de uso. A continuación se detallan los ítems que componen esa descripción textual:

- **Nombre del caso de uso:** El nombre debe ser conciso, debe estar orientado al resultado del caso de uso. Debe reflejar las tareas que el usuario necesita llevar a cabo para usar el sistema. El nombre debe llevar un verbo pues es una acción
- **Descripción:** Proporcionar una breve descripción de la razón y los resultados del caso de uso, o una descripción de alto nivel de la secuencia de acciones y los resultados de la ejecución del caso de uso.

- **Actores:** En esta parte se ubican los nombres de los actores implicados en el caso de uso, teniendo en cuenta que en un caso de uso o un conjunto de ellos pueden interactuar uno o muchos actores.
- **Precondiciones:** Lista de todas las actividades que deben ser tenidas en cuenta o todas las condiciones que deben ser cumplidas, antes de que el caso de uso empiece. Las precondiciones deben ser numeradas.
- **Postcondiciones:** Describen el estado del sistema una vez terminada la ejecución del caso de uso. Deben ir numeradas.
- **Escenario Principal:** Provee una descripción detallada de las acciones del usuario y las respuestas del sistema durante la ejecución normal del caso de uso, son las condiciones esperadas cuando el caso de uso se ejecuta sin contratiempos. Esta secuencia de diálogo en última instancia conduce a lograr el objetivo expuesto en el nombre y la descripción de casos de uso. Esta descripción puede ser escrita como una respuesta a la pregunta hipotética: ¿Cómo se pueden hacer cumplir las tareas que sugiere el nombre del caso de uso? Para lograr una buena conformación del flujo normal se recomienda numerar las acciones de los usuarios y las respuestas del sistema a manera de pasos para llegar a la meta del caso de uso.
- **Extensiones:** Es donde se establece en qué momento, del hilo de ejecución del caso de uso se va a extender a otro.
- **Sub-Variaciones:** Documenta los escenarios que pueden ocurrir cuando el caso de uso no puede ser ejecutado completamente. Se enumeran igual que el flujo normal de eventos con las acciones del usuario y las respuestas del sistema.

Otra de las alternativas dentro de este enfoque es la utilización de lenguajes controlados, en la cual se restringe la gramática del lenguaje natural al uso de sólo ciertas palabras y estructuras de oración. Una de las propuestas de este enfoque, es la utilización del patrón SVDPI (Subject, Verb, Direct object, Preposition, Indirect

object): “La estructura de la oración debe ser simple”...”Sujeto ... verbo ... objeto directo...preposición...objeto indirecto”[27].

En [30], se propone un metamodelo para descripciones textuales de casos de uso, que permite especificar el comportamiento de los mismos en forma de flujo. Este metamodelo define una representación textual del comportamiento de los casos de uso que preserva los beneficios del lenguaje natural, como la facilidad de lectura y entendimiento, para sus lectores contando además con el grado de formalidad requerido para realizar actividades de chequeo de consistencia entre ambas partes del modelo de casos de uso de UML -- el modelo propiamente dicho y las descripciones textuales del comportamiento de los mismos.

Una alternativa interesante es la presentada por Bittner y Spence [9], en la cual se introduce cierto grado de formalidad para las descripciones de casos de uso y se permite a la vez la utilización del lenguaje natural sin restricciones. Esta propuesta consiste en describir los casos de uso en forma de flujos en lenguaje natural, los cuales definen el comportamiento a través de secuencias de eventos, contando con la ventaja de legibilidad para usuarios no técnicos. Veamos el uso de técnicas de procesamiento del lenguaje natural en más detalle.

### **2.2.2. Conceptos Lingüísticos**

La especificación de casos de uso está expresada en oraciones de lenguaje natural. Para extraer y procesar automáticamente la información de los casos de uso se utilizan herramientas lingüísticas que realizan análisis morfológicos, sintácticos y semánticos sobre el texto que compone cada caso de uso. A continuación se realiza una breve descripción de los conceptos más relevantes empleados por dichas herramientas [34][48].

Las palabras del lenguaje se agrupan en clases que muestran un comportamiento sintáctico similar denominadas categorías gramaticales o sintácticas, conocidas tradicionalmente como *partes del lenguaje* o POS (Parts Of Speech).

Dos categorías importantes del lenguaje son: *sustantivo* y *verbo*. Los sustantivos hacen referencia generalmente a personas, animales, conceptos y cosas. Los verbos son usados para expresar acciones en las oraciones. Cada una de estas categorías tiene asociado un conjunto de abreviaturas que se denominan etiquetas.

Para cada idioma, se han elaborado distintos compendios o cuerpos del lenguaje con variados ejemplos de uso del lenguaje. Según el cuerpo *Brown* para el idioma inglés americano, uno de los cuerpos más ampliamente difundidos, algunas de las etiquetas correspondientes para las distintas categorías sintácticas son los que pueden verse en la Tabla 2.1.

Categoría	Etiqueta	Descripción
SUSTANTIVO	NN	sustantivo singular
	NNP	sustantivo propio
	NNR	sustantivo adverbial
	NNS	sustantivo plural
	NNPS	sustantivo propio plural
	NNRS	sustantivo adverbial plural
VERBO	VB	forma base (take)
	VBZ	tercera persona singular (takes)
	VBD	tiempo pasado (took)
	VBG	gerundio y participio presente (taking)
	VBN	participio pasado (taken)
	MD	auxiliar modal (can, may, must, could, might,...)

**Tabla 2.1:** Etiquetas correspondientes a las distintas categorías sintácticas

Las palabras se organizan en frases, agrupamientos de palabras que conforman una unidad. La sintaxis es el estudio de las regularidades y restricciones en

el ordenamiento de palabras y estructura de la frase. Los tipos de frases más importantes son:

- **Frase sustantiva.** Este tipo de frase es la unidad sintáctica en la que se reúne información sobre el sustantivo. El sustantivo es la cabeza de la frase sustantiva. En el cuerpo *Brown*, la etiqueta asociada es NP (Noun Phrase). Están constituidas por un determinante opcional, cero o más frases adjetivas, un sustantivo y luego pueden incluirse algunos modificadores, tales como frases preposicionales o modificadores de cláusulas también denominados cláusulas relacionadas.
- **Frase preposicional.** Este tipo de frase aparece dentro de otros tipos de frases como las frases sustantivas y las frases verbales, expresando generalmente ubicación espacial o temporal y otros atributos. Son encabezadas por una preposición y se etiquetan como PP (Prepositional Phrase).
- **Frase verbal.** En este tipo de frase, el verbo es la cabeza de la frase, que se etiqueta como VP (Verb Phrase). La frase verbal organiza todos los elementos de la oración que dependen sintácticamente del verbo.
- **Frase adjetiva.** Las frases adjetivas complejas son poco comunes. Se etiquetan como AP (Adjective Phrase).

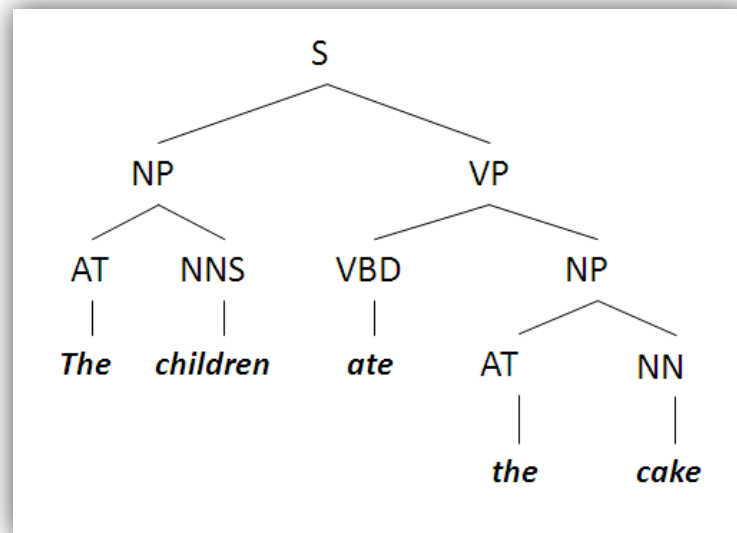
Los lenguajes tienen restricciones en cuanto al ordenamiento de las palabras. Las regularidades de los distintos ordenamientos se expresan mediante *reglas de reescritura*. Una regla de reescritura tiene la forma ‘categoría → categoría\*’ y estipula que el símbolo del lado izquierdo puede reescribirse con la secuencia de símbolos del lado derecho. De esta manera, una categoría sintáctica puede reescribirse como una o más categorías y/o palabras. Las posibles reescrituras dependen únicamente de la categoría y no del contexto de esa categoría, por lo que a este tipo de gramáticas se las conoce como *gramáticas libres de contexto*.

La forma más intuitiva de representar la estructura de frases es con un árbol. Los nodos hoja del árbol son los *nodos terminales* y corresponden a palabras del lenguaje mientras que los nodos internos son los *nodos no terminales* y corresponden a



frases sintácticas y/o partes del lenguaje (POS). En la Figura 2.8, se muestra un ejemplo de árbol sintáctico.

En esta Figura podemos observar que el árbol consta de un único nodo raíz que consiste en el símbolo inicial de la gramática ‘S’ (de sentencia). El nodo raíz tiene dos nodos hijos, NP y VP que corresponden a una frase sustantiva y una frase verbal respectivamente. El nodo NP tiene dos nodos hijos, que corresponden a dos categorías sintácticas o POS: AT un determinante y NNS un sustantivo plural. Estos nodos tienen cada uno un hijo que corresponde a un nodo terminal y que por lo tanto contienen propiamente palabras del lenguaje. Por otra parte, el nodo VP tiene también dos hijos: VBD y NP. El nodo VBD corresponde a la categoría sintáctica de “verbo pasado” y tiene como nodo hijo un nodo terminal que contiene la palabra del idioma Inglés “ate”. El nodo NP, es un nodo que contiene una frase sustantiva, y tiene como hijos a los nodos AT y NN, que corresponden a los POS determinante y sustantivo singular respectivamente, los cuales tienen como hijos a los nodos terminales “the” y “cake”.



**Figura 2.8:** Árbol sintáctico de la oración “*The children ate the cake*”.

Otro concepto importante de organización es la dependencia. La dependencia se entabla entre conceptos o argumentos que dependen de un evento. Generalmente las frases sustantivas son argumentos de los verbos. Los argumentos se pueden clasificar

de distintas maneras. Según los roles semánticos que cumplen, pueden ser **agentes** de una acción (persona o cosa que realiza una actividad), **pacientes** (persona o cosa para quien se realiza una acción) y otros roles como **instrumento** y **objetivo** que describen otra clase de relaciones semánticas. Por otra parte, los argumentos se pueden describir sintácticamente según las relaciones gramaticales. Todos los verbos tienen un **sujeto**, que consiste en la frase sustantiva que aparece antes del verbo. Muchos verbos tienen una o dos frases sustantivas que normalmente aparecen después del verbo a las que se denomina **objeto directo** (la que aparece inmediatamente después del verbo) y **objeto indirecto** (que describe el objeto o persona que indirectamente obtiene algo).

La semántica es el estudio del significado de las palabras: construcciones y expresiones. Se puede dividir la semántica en dos partes -- el estudio del significado de las palabras individuales (semántica léxica) y el estudio de cómo los significados de las palabras pueden combinarse dentro del significado de las oraciones o unidades más grandes.

Una forma de enfocar la semántica léxica es estudiar cómo se relacionan los significados de las palabras. Las palabras se pueden organizar en una jerarquía léxica como por ejemplo en WordNet, una base de datos léxica del idioma Inglés. A diferencia de los diccionarios tradicionales que están ordenados alfabéticamente, WordNet está ordenada semánticamente, creando una base de datos léxica de sustantivos, verbos, adjetivos y adverbios. Los sinónimos se agrupan formando conjuntos de sinónimos. Una palabra es *polisémica* si se encuentra en varios conjuntos de sinónimos, donde cada uno representa una acepción posible de la palabra. Por ejemplo la palabra *leaf* aparece en tres conjuntos: (leaf, leafage, foliage), (leaf, folio) y (leaf).

Cada conjunto tiene asociada una definición o *glosa*, consistente en una explicación breve describiendo el concepto representado por el conjunto de sinónimos. Por ejemplo la glosa del conjunto (leaf, leafage, foliage) es “the main organ of photosynthesis and transpiration in higher plants” y la glosa del segundo conjunto es “a sheet of any written or printed material (especially in a manuscript or book)”. Cada conjunto puede referenciarse por un identificador único, conocido como etiqueta.

Los conjuntos de sinónimos están conectados unos a otros a través de distintas relaciones semánticas. Para los sustantivos las relaciones más importantes son **hiponimia/hiperonimia** y **meronimia/holonimia**. Las relaciones de hiponimia/hiperonimia se establecen cuando dos conceptos se vinculan mediante la relación “es un”. Por ejemplo “loro es una clase de ave”, entonces “loro” es el hipónimo de “ave” y “ave” es el hiperónimo de “loro”. Las relaciones de meronimia/holonimia, se establecen cuando dos conceptos se vinculan mediante la relación “es parte de”. Por ejemplo, “raíz”, “hoja”, “rama”, “tronco” son parte de “árbol”, entonces raíz, hoja, rama y tronco son merónimos de árbol y árbol es holónimo de raíz, hoja, rama y tronco.

Los verbos se relacionan mediante relaciones de troponimia/hiperonimia, cuando se vinculan mediante la relación “una manera de”. Por ejemplo “freir” es una manera de “cocinar”, entonces freir es un tropónimo de cocinar y cocinar es un hiperónimo de freir.

### **2.2.3. Algunas Herramientas Lingüísticas**

En esta sección se describe la funcionalidad de tres herramientas lingüísticas, las cuales son componentes de nuestra herramienta ReqGIS. Estas herramientas realizan análisis morfológicos, sintácticos y semánticos sobre el texto y son utilizadas en ReqGIS para extraer y procesar automáticamente la información de los casos de uso.

- **Freeling**

FreeLing [60] es una herramienta multilingüe de código abierto que provee una amplia variedad de analizadores del lenguaje. Provee servicios a los desarrolladores de aplicaciones de procesamiento de lenguaje natural tales como tokenización, analizadores morfológicos, lematización, etiquetado PoS, análisis de frases. Esta herramienta es adaptable y extensible, pudiendo utilizarse los recursos lingüísticos (diccionarios, lexicons, gramáticas, etc.) directamente o adaptándolos a dominios específicos.

- **Algoritmo Lesk Adaptado.**

El algoritmo Lesk original [41] desambigua palabras en frases cortas. Para ello, cada sentido o definición de una palabra de la frase se compara con los sentidos de las demás palabras de la frase, y se le asigna a esa palabra el sentido que más concuerde con los de las demás palabras de la frase. Las definiciones que se tienen en cuenta son las de diccionarios tradicionales.

Este algoritmo se modificó, aprovechando las ventajas provistas por las relaciones altamente vinculadas entre sinónimos que ofrece WordNet, dando origen al Algoritmo Lesk Adaptado [7].

La idea principal de este algoritmo es desambiguar palabras encontrando el solapamiento entre sus definiciones. Así, dadas dos palabras,  $W_1$  y  $W_2$ , cada una con sus respectivas definiciones  $N_{w1}$  y  $N_{w2}$ , para cada par de definiciones  $W_1^i$  y  $W_2^j$ ,  $i=1..N_{w1}$ ,  $j=1..N_{w2}$ , se determina el solapamiento entre las correspondientes definiciones contando el número de palabras que tienen en común. El par de definiciones con el mayor solapamiento es seleccionado.

Una característica de este algoritmo es que realiza el procedimiento de desambiguación para palabras del idioma inglés contenidas en el diccionario WordNet. En consecuencia, el ejemplo que se detalla a continuación estará constituido por palabras de dicho idioma.

Un ejemplo representativo de este algoritmo sería el siguiente: consideremos que queremos desambiguar las palabras “tree” y “flower”. El diccionario WordNet provee tres definiciones para “tree” y tres definiciones para “flower”, tal y como muestra la Figura 2.9.

<b>Tree</b>	
1	a tall perennial woody <b>plant having</b> a main trunk and branches forming a distinct elevated crown; includes both <b>gymnosperms</b> and <b>angiosperms</b>
2	a figure that branches from a single root; "genealogical tree"
3	English actor and theatrical producer noted for his lavish productions of Shakespeare (1853-1917)
<b>Flower</b>	
1	a <b>plant</b> cultivated for its blooms or blossoms
2	reproductive organ of <b>angiosperm plants</b> especially one <b>having</b> showy or colorful parts
3	the period of greatest prosperity or productivity

**Figura 2.9:** Definiciones para “tree” y “flower”

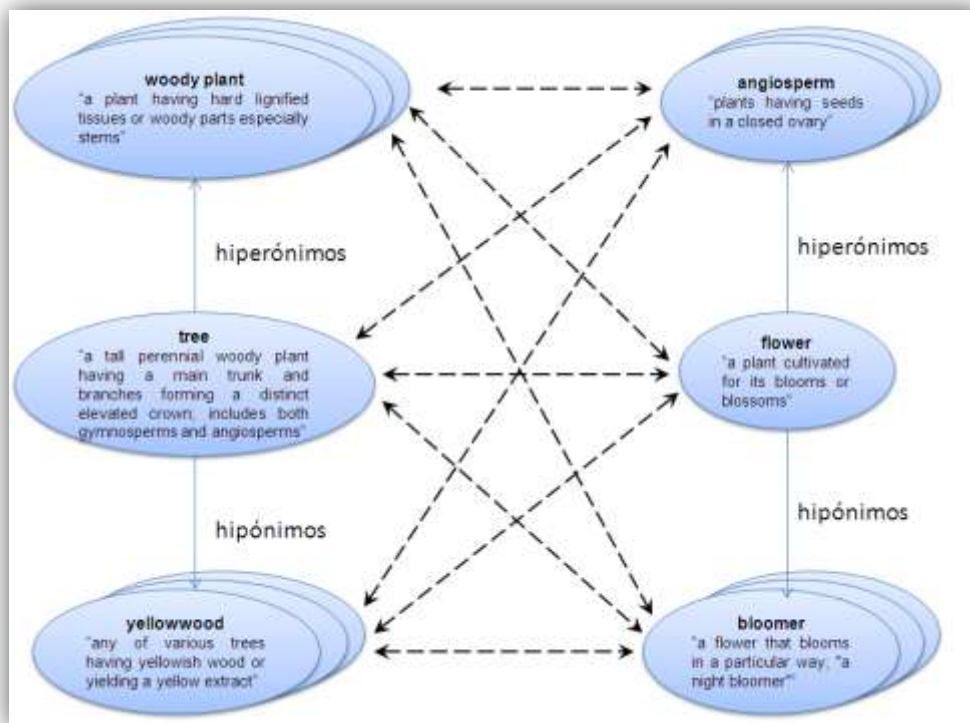
Tree#1 $\cap$ Flower#1 = 1 (plant)
Tree#2 $\cap$ Flower#1 = 0
Tree#3 $\cap$ Flower#1 = 0
Tree#1 $\cap$ Flower#2 = 3 (have, plant, angiosperm)
Tree#2 $\cap$ Flower#2 = 0
Tree#3 $\cap$ Flower#2 = 0
Tree#1 $\cap$ Flower#3 = 0
Tree#2 $\cap$ Flower#3 = 0
Tree#3 $\cap$ Flower#3 = 0

**Figura 2.10:** Solapamiento entre definiciones de “Tree” y “Flower”

En la Figura 2.10 podemos ver el solapamiento existente entre cada definición de “tree” y “flower”. La primera definición de “tree” y la segunda de “flower” tienen el máximo solapamiento entre todas las posibles combinaciones de definiciones.

Además de los solapamientos entre las definiciones, el algoritmo Lesk Adaptado compara también los solapamientos entre las definiciones de otras palabras relacionadas que están conectadas a través de las distintas relaciones semánticas definidas en WordNet, como hponimia, hiperonimia, holonimia y meronimia (para sustantivos) e hiperonimia y troponimia (para verbos).

Para las palabras utilizadas en el ejemplo, “Tree” y “Flower”, se observan en la Figura 2.11, los hipónimos y los hiperónimos de “tree” y “flower”, y además se observa también que cada palabra puede tener más de una definición (por ejemplo la palabra “tree” tiene tres definiciones) y que a su vez cada una de esas definiciones puede no tener ninguno o muchos hiperónimos e hipónimos.



**Figura 2.11:** Hipónimos e hiperónimos de las palabras “tree” y “flower” y análisis de solapamientos que calcula el algoritmo Lesk Adaptado.

Además, también se distinguen las múltiples comparaciones que el algoritmo realiza entre las definiciones de palabras y los hiperónimos e hipónimos. Así, el algoritmo tiene en cuenta tanto los solapamientos entre las definiciones de las palabras a desambiguar, como los solapamientos entre las palabras vinculadas a ellas mediante otras relaciones semánticas. Esto otorga al cálculo un análisis más profundo de la semántica del lenguaje.

- **WordNet::Similarity.**

Es un componente de software disponible en forma gratuita que permite medir la similitud y la relación semántica existente entre dos conceptos. Este software provee seis medidas para calcular la similitud y tres medidas para calcular la relación semántica, todas ellas basadas en la base de datos léxica WordNet.

Las medidas de similitud determinan en cuánto se parecen dos conceptos, basándose en la información contenida en una jerarquía “es-un”. Por ejemplo, automóvil debería considerarse más parecido a un bote que a un árbol, dado que automóvil y bote comparten a vehículo como un ancestro común en una jerarquía es-un.

La base de datos léxica WordNet se adapta particularmente a medidas de similitud, ya que organiza los sustantivos y verbos en relaciones de jerarquía *es-un*. Además, los conceptos pueden relacionarse de otras maneras aparte de ser semejantes entre ellos. En consecuencia, WordNet establece también entre los conceptos otras relaciones como tiene-parte, está-compuesto-por, es-un-atributo-de, etc. Las medidas de relación semántica se basan en estas relaciones.

Este componente implementa las distintas medidas de similitud y relación semántica como módulos Perl que toman como entrada dos conceptos, y devuelven un valor numérico que representa el grado en que se asemejan o relacionan ambos conceptos. Una de las medidas de relación semántica utilizadas por este componente, es la que calcula el Algoritmo Lesk Adaptado y es la que se aplicará en este trabajo de Tesis para encontrar la categoría de servicio geográfico más relacionada con un determinado requerimiento funcional.

## 2.2.4. Extracción de Conocimiento en Casos de Uso

Con respecto a la extracción de conocimiento en casos de uso, en [17][49] se propone un método para procesamiento de requerimientos textuales, el cual se basa en la identificación de los atributos principales de las acciones descritas en cada paso de un caso de uso, mediante la utilización de herramientas lingüísticas disponibles que permiten obtener un árbol sintáctico. Habiendo identificado las acciones, permite conocer las operaciones aceptadas y solicitadas por una entidad, lo que facilita la definición de interfaces de un componente, o servicios requeridos en este trabajo de Tesis

Este método parte de una especificación de casos de uso, construida mediante la utilización de las plantillas sugeridas por Cockburn [14] y de las recomendaciones para escritura de casos de uso [39], donde cada uno de los pasos que conforman el caso de uso debe cumplir con las siguientes premisas:

- Premisa 1: Un paso de un caso de uso textual describe: (a) la comunicación entre un actor y el sistema, o (b) una acción interna.
- Premisa 2: La acción se describe mediante una oración simple, según el patrón SVDPI (“Sujeto...Verbo...objeto Directo...Preposición...objeto Indirecto”).

Además el método consta de los siguientes pasos:

- a) Determinación del POS (part-of-speech) de cada palabra, es decir el tipo sintáctico de la palabra (SS, VB, etc.).
- b) Generación del árbol sintáctico de cada oración, es decir de cada paso que conforma el Escenario Principal del caso de uso, obteniendo así la estructura sintáctica de la oración (NP, VP, PP, etc.).
- c) Obtención de los Tokens de Eventos. Estos tokens de eventos se construyen a partir del verbo principal de la oración (elemento que encabeza la frase verbal, VP) y el objeto representativo (objeto directo, elemento que encabeza la frase sustantiva NP posterior al verbo).



- d) A partir de los resultados obtenidos en el paso anterior, se convierten los casos de uso textuales en una especificación de comportamiento utilizando Pro-cases [64]. Pro-cases es una técnica formal que permite especificar el comportamiento como un conjunto de secuencias finitas de eventos y que en consecuencia permite razonar sobre la consistencia de su especificación.

Los tokens de eventos se construyen para posibilitar la representación de las acciones en una especificación de comportamiento procesable automáticamente. Además de los tokens de eventos, se obtienen también a partir de los casos de uso, los actores involucrados y el tipo de operación (envío, recepción o interna) que representa cada paso del escenario principal de un caso de uso. Luego, con toda esta información y mediante la utilización de una serie de guías de transformación sencillas se traducen los casos de uso textuales a pro-cases.

### **2.3. Resumen**

En este capítulo se analizaron los temas fundamentales sobre los que se basa nuestra propuesta de mejora para la selección de componentes SIG. Se evaluaron distintos enfoques utilizados para la selección de componentes, se analizaron los casos de uso como herramienta para especificar requerimientos de componentes y se detallaron distintas herramientas lingüísticas que permiten obtener información desde el contenido de los casos de uso.



# Capítulo 3

## Categorización y Selección de Servicios

### SIG: Estado del Arte

En este capítulo se presenta una breve descripción del dominio de aplicación al cual hace referencia esta tesis: Sistemas de Información Geográficos (SIG). Se analizan las características del dominio geográfico y en particular diversas clasificaciones de servicios SIG. También se presentan distintos estándares referidos al dominio geográfico.

#### 3.1. Características del Dominio Geográfico

En los últimos años, la información sobre aspectos geográficos ha sido ampliamente considerada como parte esencial de todo sistema de información. Con las nuevas tecnologías emergentes, tales como dispositivos GPS o técnicas de teledetección remota, la tarea de capturar información geográfica es cada vez más fácil. Además, existen varias aplicaciones de software de código libre (open source) y propietario para almacenar este tipo de información de una manera eficiente. Por lo tanto, varias organizaciones están interesadas en implementar portales web interactivos o sistemas stand-alone que les permita consultar y administrar este tipo de información.

A estos sistemas espaciales se los denomina Sistemas de Información Geográficos. En la literatura existen varias definiciones para este término, cada una considerando la funcionalidad del sistema desde diferentes perspectivas. Igualmente, todas ellas se centran en tres aspectos en común:

1. *un SIG es una colección de herramientas para realizar simulaciones y análisis geográficos* [10][40][65],

2. *un SIG posee un conjunto de estructuras de datos y algoritmos para representar, recuperar y manipular información geográfica [65][76] y*

3. *un SIG es una utilidad que ayuda a las personas a tomar decisiones en tareas relacionadas con la geografía [29][43].*

Considerando estos aspectos, podemos concluir en la definición dada en [66], *donde un SIG es un sistema basado en computadora diseñado para modelar, capturar, almacenar, manipular, consultar, recuperar, analizar y visualizar información eficientemente y donde parte de la misma es de naturaleza geográfica.*

Una característica importante de los sistemas de información geográficos es que pueden ofrecer más que una simple herramienta para producir mapas. Mientras que en la cartografía tradicional el mapa es la base de datos, en un SIG el mapa es sólo una proyección de una vista particular de una base de datos geográfica en un momento dado. Esto permite que el usuario final del SIG pueda obtener un número ilimitado de alternativas de análisis y la capacidad de crear mapas desde distintos puntos de vista de acuerdo a aspectos diferentes de la información [66].

A diferencia de los sistemas de información tradicionales, debido a la naturaleza de los datos y a las características exclusivas de la información geográfica, los SIG poseen requerimientos funcionales especiales. Algunos de ellos son:

- Se necesitan tipos de datos y operaciones especiales para manipular la información geográfica. Se requieren tipos de datos como puntos, curvas, superficies, etc., junto con operaciones para manipularlos como distancia, dirección, intersección, etc.
- La información geográfica requiere diferentes procedimientos de visualización y análisis. Se requieren nuevas técnicas de análisis y visualización para mostrar los valores que retornan los nuevos tipos de datos y operaciones.
- La información geográfica posee un gran volumen con una estructura jerárquica impuesta. La representación de la información geográfica requiere gran cantidad de espacio de almacenamiento. Por ejemplo, para

representar un país se necesitan miles de pares de números reales. Además, la naturaleza del espacio geográfico genera relaciones y una jerarquía entre los valores espaciales que no existe en otro tipo de sistemas. Por ejemplo, las rutas se relacionan por una relación de contención con la ciudad en la cual se encuentran.

- El procesamiento de la información geográfica se caracteriza por transacciones que son más largas que las que existen en un sistema relacional estándar. Esto se debe a la característica anterior ya que al tener la información un gran volumen, la creación y modificación de los datos geográficos requieren mucho más tiempo.
- Existen dos vistas conceptuales diferentes del espacio geográfico. Estas son: modelos basados en campos y modelos basados en objetos. Los modelos basados en campos representan la información geográfica como colecciones de distribuciones espaciales donde cada una de ellas puede formalizarse como una función matemática desde un marco de trabajo espacial a un atributo del dominio. Aquí, el espacio geográfico se considera como un elemento existente en sí mismo, el cual posee propiedades asociadas a cada ubicación. El clima, la vegetación y los mapas geológicos son ejemplos de características geográficas modeladas como campos. Por otro lado, los modelos basados en objetos representan la información geográfica como entidades discretas e identificables que consisten de una colección de atributos para describir las propiedades de un fenómeno o hecho del mundo real utilizando valores geométricos y alfanuméricos. Objetos construidos por humanos, como caminos y edificios, son modelados como objetos. Los valores geométricos definen la forma del objeto plasmada en un plano, mientras que los valores alfanuméricos hacen referencia a los atributos temáticos asociados al mismo (por ejemplo nombre, altura en pisos, capacidad, etc. en el caso de un edificio). Ambos modelos consideran el espacio geográfico como un contenedor de objetos con propiedades asociadas. Cada vista es diferente

desde el punto de vista conceptual y necesita representarse en un modelo conceptual con abstracciones diferentes.

- Cada vista conceptual del espacio puede representarse en varias formas diferentes en una computadora. Los modelos lógicos deben extenderse o deben crearse nuevos para soportar la representación de datos geográficos. Además, cada uno de los modelos conceptuales (previamente explicados) pueden representarse utilizando un modelo lógico en muchas formas diferentes.

Estos y otros requerimientos impactan sobre la creación y utilización de los SIG. Por lo tanto, es muy importante determinar los requerimientos especiales y la funcionalidad que deben proveer. Así, se podrán diseñar e implementar SIG con capacidades apropiadas para el modelado, colección, consulta y visualización de la información geográfica. Además, la capacidad de asociación de bases de datos temáticas junto con la descripción espacial precisa de objetos geográficos y las relaciones entre los mismos (topología) es lo que diferencia a un SIG de otros sistemas informáticos de gestión de información.

## **3.2. Categorización de Servicios Geográficos**

Existen en la literatura distintos enfoques que categorizan las operaciones o servicios que una aplicación SIG debe brindar. Estos trabajos representan las tareas desde la perspectiva del usuario de un sistema geográfico, quien no necesariamente cuenta con conocimientos sobre conceptos espaciales abstractos o de un usuario sin experiencia en el dominio geográfico. De esta manera, estas categorizaciones pretenden brindar al usuario un conjunto de operaciones de referencia que le sirvan de soporte para definir sus requerimientos y evaluar los servicios brindados por las distintas aplicaciones SIG. A continuación se presentan algunas de las propuestas de categorización de servicios geográficos.

Albretch [1] presenta un conjunto de 20 operaciones SIG universales o tareas, independientes de la estructura de los datos subyacentes, las cuales cubren la totalidad

de las capacidades analíticas ofrecidas por el software geográfico del mercado actual. Las mismas surgieron a partir de un relevamiento realizado a usuarios para determinar sus expectativas respecto de la funcionalidad de un software SIG. Los resultados obtenidos revelaron un amplio rango de complejidad, desde operaciones elementales a tareas compuestas, por lo que se debió realizar un proceso de “normalización” de operaciones SIG. El conjunto de operaciones encontrado, según puede verse en Tabla 3.1, se ha agrupado en seis categorías: Búsqueda, Análisis de Ubicación, Análisis de Terreno, Distribución Vecindario, Análisis Espacial y Mediciones. Cada una de estas categorías reúne las operaciones relacionadas; por ejemplo en la categoría ‘Análisis de Terreno’ encontramos las operaciones: Pendientes, Zona de Influencia, Sistema de Drenaje y Análisis de Visibilidad, operaciones que hacen referencia al análisis del terreno.

CATEGORÍA	OPERACIONES
BÚSQUEDA	Interpolación Búsqueda temática Búsqueda espacial (Re-)clasificación
ANÁLISIS DE UBICACIÓN	Límites Rutas Solapamiento Polígonos de Thiessen/Voronoi
ANÁLISIS DE TERRENO	Pendientes Zona de influencia Sistema de drenaje Análisis de visibilidad
DISTRIBUCIÓN VECINDARIO	Costo/Difusión/Propagación Proximidad Vecino más cercano
ANÁLISIS ESPACIAL	Análisis multivariado Patrón de dispersión Centralización/Conexión Forma
MEDICIONES	Mediciones

**Tabla 3.1:** Operaciones SIG Universales

CATEGORÍA	OPERACIONES
PARAMETRIZACIÓN DEL PROYECTO	<ul style="list-style-type: none"> <li>• Inicialización de software SIG</li> <li>• Inicialización motor base de datos</li> <li>• Ajuste parámetros visualización</li> </ul>
ENTRADA DE DATOS	<ul style="list-style-type: none"> <li>• Digitalización de mapas</li> <li>• Escaneo de mapas</li> <li>• Carga por lotes</li> <li>• Asignación de atributos a cada característica</li> </ul>
CONVERSIÓN DE DATOS	<ul style="list-style-type: none"> <li>• Raster a vector</li> <li>• Vector a raster</li> </ul>
VALIDACIÓN DE DATOS	<ul style="list-style-type: none"> <li>• Identificación/corrección de errores topológicos</li> <li>• Identificación y corrección de datos en tablas</li> </ul>
VISUALIZACIÓN / RENDERIZACIÓN	<ul style="list-style-type: none"> <li>• Zoom/Pan/cambiar vista</li> <li>• Refrescar pantalla</li> <li>• Simbolización de características</li> <li>• Filtrado de características</li> <li>• Gestión de imágenes de fondo</li> </ul>
GESTIÓN DE BASES DE DATOS DE MAPAS	<ul style="list-style-type: none"> <li>• Georeferenciamiento</li> <li>• Unión de hojas/mosaicos de mapas</li> <li>• Rectificación/fusión de capas</li> <li>• Cambio de proyección</li> </ul>
GESTIÓN DE ATRIBUTOS DE BASES DE DATOS	<ul style="list-style-type: none"> <li>• Enlace de atributo básico a base de datos principal</li> <li>• Definición de relaciones complejas</li> <li>• Definición de conexiones (SQL, middleware) a sistemas o bases de datos secundarios</li> </ul>
ANÁLISIS / PROCESAMIENTO DE DATOS	<ul style="list-style-type: none"> <li>• Creación/guardado de vistas</li> <li>• Análisis de proximidad</li> <li>• Superposición de mapas</li> <li>• Análisis de redes (rutas óptimas, asignación de recursos)</li> <li>• Análisis raster (álgebra de mapas)</li> <li>• Generalización/suavizado</li> <li>• Generación de vistas 3D – DTM (Modelado Digital de Terreno )</li> <li>• Análisis secundarios de DTM</li> </ul>
SALIDA: PRODUCCIÓN DE MAPAS / REPORTES	<ul style="list-style-type: none"> <li>• Generación de estadísticas resumen</li> <li>• Generación de reportes de texto a partir de atributos en base de datos</li> <li>• Generación de mapas simples</li> <li>• Generación de composición de mapas</li> </ul>

**Tabla 3.2:** Operaciones Clave de SIG



La Unión Europea a través de sus proyectos ESPRIT [18], también presenta una lista de operaciones clave para SIG. La confección de esta lista, está motivada por la idea que usuarios mejor informados pueden tomar mejores decisiones respecto a sus requerimientos y que pueden describir mejor esas necesidades a sus proveedores SIG y a sus consultores durante las fases de selección e implementación. Por lo tanto, se propone la lista de funciones u operaciones comunes mostradas en la Tabla 3.2 orientada a usuarios inexpertos en el dominio.

En esta lista, las funciones se encuentran agrupadas en nueve categorías: ‘Parametrización del Proyecto’, ‘Entrada de Datos’, ‘Conversión de Datos’, ‘Validación de Datos’, ‘Visualización/Renderización’, ‘Gestión de Bases de Datos de Mapas’, ‘Gestión de Atributos de Bases de Datos’, ‘Análisis/Procesamiento de Datos’ y ‘Salida: Producción de Mapas/Reportes. Cada categoría agrupa todas las operaciones vinculadas con el tópico de esa categoría. Por ejemplo la categoría ‘Entrada de Datos’ reúne a las operaciones: Digitalización de Mapas, Escaneo de Mapas, Carga por Lotes y Asignación de atributos a cada característica, todas operaciones referidas al tema “Entrada de datos” al sistema.

Se puede observar que ambas categorizaciones difieren en el alcance; mientras que la primera está orientada a brindar orientación específica sobre la funcionalidad SIG para una interfaz de usuario, la segunda no sólo tiene en cuenta lo anterior (con un nivel mayor de abstracción), sino que considera los aspectos inherentes a la funcionalidad del software en su totalidad.

### **3.3. Estándares para Definición y Categorización de Servicios SIG**

Una importante iniciativa para lograr la interoperabilidad SIG ha sido desarrollada por el OpenGis Consortium (OGC). Este organismo está conformado por una asociación de agencias gubernamentales, organizaciones de investigación y academia, buscando definir un conjunto de requerimientos, estándares y especificaciones que soporten la interoperabilidad SIG. Las especificaciones provistas

por el Consorcio OGC permiten la interoperabilidad sintáctica y la catalogación de información geográfica. Sin embargo, no se ha definido ningún método para resolver el problema de la heterogeneidad semántica. Este problema presenta un gran desafío para el descubrimiento y recuperación de datos y servicios geográficos en un ambiente geoespacial abierto y distribuido. A continuación, se presenta una breve descripción de los estándares definidos por el consorcio OGC, vinculados a definición y catalogación de servicios.

El Consorcio OpenGis y la ISO TC211 desarrollaron conjuntamente un estándar internacional para arquitecturas de servicios geoespaciales denominado ISO 19119/OGC Service Architecture, [57][62]. Esta arquitectura provee a los desarrolladores un framework para crear software que permite a los usuarios acceder y procesar datos geográficos almacenados en diferentes fuentes de datos a través de una interfaz genérica. Esta arquitectura se basa en el Modelo de Referencia para Procesamiento Distribuido Abierto (RM-ODP) [33] y está conformada por un conjunto de componentes, conexiones y topologías definidas a través de una serie de cuatro vistas:

- *Vista Computacional*: describe patrones de interacción entre entidades e interfaces de un sistema distribuido independientemente de su implementación y contenido semántico. En especial detalla la combinación de servicios, es decir de funcionalidades provistas por esas entidades a través de sus interfaces, para lograr una tarea específica a través del encadenamiento de servicios y datos de maneras no previstas por los proveedores de datos y servicios.
- *Vista de Información*: hace referencia a la semántica del procesamiento de la información. Esta sección contiene la descripción de una taxonomía de servicios, en la cual se categorizan los servicios geográficos en seis clases, *Interacción Humana*, *Gestión de Modelo/Información*, *Gestión de Workflow/Tarea*, *Procesamiento (espacial, temático, temporal, metadatos)*, *Comunicación* y *Gestión de Sistema*.

- *Vista de Ingeniería*: tiene relación con el diseño de aspectos de distribución de servicios, ya sea en una aplicación monolítica o en distintas arquitecturas cliente-servidor.
- *Vista tecnológica*: concierne la infraestructura subyacente en un sistema distribuido. Describe los componentes de hardware y software usados.

Otro documento generado por el consorcio OpenGIS es “OpenGIS Web Services Architecture” [56], donde se describen los aspectos más significativos de la Arquitectura de Servicios Web propuesta por el OGC. Esta arquitectura, se basa en los roles fundamentales de proveedor y consumidor de servicios dentro de un sistema distribuido. La definición de componentes se enfoca en provisión y/o consumo de servicios definidos. Según la definición provista por esta arquitectura, los servicios se organizan en cuatro capas: *cliente, aplicación, procesamiento y gestión de información*. Cada capa tiene un propósito general, independiente de datos y servicios geográficos e incluye distintos tipos de servicios específicos, entre los cuales se encuentran algunos adaptados a servicios y datos geográficos.

Haciendo una comparación entre ambas propuestas del OGC, podemos observar que existe una correspondencia entre los conceptos de capas de la arquitectura de servicios web y categorías de ISO 19119:

- Las capas Aplicación y Clientes conjuntamente proveen los servicios de la categoría Interacción Humana.
- La capa de Procesamiento, provee los servicios de la categoría Procesamiento.
- La capa de Gestión de Información provee los servicios de la categoría Gestión de modelo/información.

### **3.3.1. Categorización de Servicios del estándar ISO 19119**

En el estándar ISO 19119, se define una categorización de servicios geográficos según se mencionara en la sección anterior. Esta taxonomía establece el nombre de las categorías y las definiciones correspondientes a cada una de ellas:

- Interacción Humana, incluye servicios para administración de interfaces de usuario, gráficos, multimedia y presentación de documentos compuestos.
- Gestión de Modelo/Información, abarca servicios para la gestión del desarrollo, manipulación y almacenamiento de metadatos, esquemas conceptuales y estructuras de datos.
- Gestión de Workflow/Tarea, son servicios para soportar tareas específicas o relacionadas con el flujo de trabajo realizado por individuos. Soportan el uso de recursos y desarrollo de productos que involucran una secuencia de actividades o pasos que pueden ser ejecutados por distintas personas.
- Procesamiento (espacial, temático, temporal, metadatos), incluye servicios que ejecutan cálculos a gran escala con grandes cantidades de datos.
- Comunicación, abarca servicios de codificación y transferencia de datos entre redes de comunicación.
- Gestión de Sistema, son servicios para la gestión de componentes del sistema, aplicaciones, redes y gestión de usuarios.

Además, para cada una de las categorías definidas, brinda ejemplos de servicios específicos. Por ejemplo, para la categoría Interacción Humana, los servicios ejemplo son (Tabla 3.3): Visor de catálogo, Visor geográfico, Visor de hoja de cálculo geográfica, Editor de servicios, Editor de encadenamientos, Gestor de ejecución de workflow, Editor de características geográficas, Editor de símbolos geográficos, Generalizador de características, Visor de estructuras de datos geográficos.

CAPA	SERVICIO
HUMAN INTERACTION	Catalogue Viewer
	Geographic Viewer
	Geographic spreadsheet viewer
	Service editor
	Chain definition editor
	Workflow enactment manager
	Geographic feature editor
	Geographic symbol editor
	Feature generalization editor
	Geographic data-structure viewer

**Tabla 3.3:** Servicios ejemplo de la categoría Interacción Humana - Taxonomía de Servicios de ISO 19119

Cada ejemplo de servicio, está acompañado por una breve descripción, donde se detalla la funcionalidad brindada por el servicio. En el caso del ejemplo de servicio denominado “Visor geográfico”, la descripción que lo acompaña es “Permite al usuario visualizar una o más colecciones o capas de características. Permite también interactuar con datos de mapas.”.

En la Tabla 3.4 se muestran algunas de las descripciones de los ejemplos de servicio de la categoría Interacción Humana.

En el Apéndice A, se puede consultar la categorización del estándar ISO19119 completa, incluyendo los ejemplos de servicios de cada categoría y sus respectivas descripciones.

CAPA	SERVICIO	DESCRIPCIÓN
HUMAN INTERACTION	Catalogue Viewer	Client service that allows a user to interact with a catalogue to locate, browse, and manage metadata about geographic data or geographic services.
	Geographic Viewer	Client service that allows a user to view one or more feature collections or coverages. This viewer allows a user to interact with map data.
	Geographic Feature Editor	Geographic viewer that allows a user to interact with feature data, e.g., displaying, querying. Supports feature annotation.

**Tabla 3.4:** Descripciones de algunos servicios ejemplo de la categoría Interacción Humana.

La exigencia de cumplimiento de la taxonomía de este estándar reside en que si un sistema provee un servicio denominado de la misma manera que un servicio ejemplo de esta taxonomía, entonces deberá estar clasificado con la misma categoría del servicio ejemplo. Además, cualquier catálogo de servicios geográficos que adhiera a este estándar deberá clasificar sus servicios según las categorías de esta taxonomía.

### 3.4. Selección de Componentes SIG

Existen algunos grupos de investigación que han trabajado sobre el tema de oferta y demanda de servicios SIG. En estos trabajos se definen ontologías para servicios SIG, que intentan por un lado solucionar el problema de la interoperabilidad semántica y por otro, facilitar el descubrimiento de componentes en un repositorio o catálogo determinado.

Lutz [45] propone una metodología para mejorar el descubrimiento de servicios SIG. A diferencia de los enfoques basados en palabras clave, los cuales sufren de las limitaciones impuestas por las ambigüedades del lenguaje natural, esta metodología utiliza ontologías que describen las operaciones geoespaciales para crear descripciones de los requerimientos y capacidades de los servicios. En este contexto, la ontología hace uso del catálogo de servicios conjuntamente con conceptos de Lógica Descriptiva (DL) y capacidades de razonamiento de Lógica de Primer Orden (FOL)

para optimizar el descubrimiento de servicios específicos dentro del conjunto de servicios web espaciales.

Li y otros [42] proponen un enfoque semántico que permite a los clientes seleccionar los servicios que mejor se adapten a sus requerimientos en una Organización Virtual GGS (GIS Grid Services) dinámica. Los servicios son seleccionados de acuerdo a los requerimientos funcionales y a los requerimientos dinámicos de calidad de servicio (QoS). Este framework utiliza una ontología para definir la terminología de las funciones GGS y la semántica de los requerimientos QoS dinámicos.

O'Dea y otros [55] realizan un estudio de los distintos avances realizados sobre Web Semántica Geoespacial. La web semántica permite mejorar la capacidad de recuperación de información mediante una mejor expresión del contexto y significado de una consulta. Sin embargo, los servicios involucrados en este trabajo de Tesis son distintos repositorios de datos espaciales, y no componentes software que brinden funcionalidad SIG.

Yue y otros [77], proponen una mejora al Catálogo de Servicios para la Web (CSW), mediante la representación semántica de datos y servicios geoespaciales, permitiendo realizar búsquedas en CSW teniendo en cuenta las relaciones semánticas definidas en OWL/OWL-S [15][59]. Esto permite optimizar las búsquedas de datos y servicios, utilizando además de la concordancia directa de palabras clave con metadatos que proveen la funcionalidad clásica de búsqueda, la información semántica implícitamente embebida en los metadatos, tal como las relaciones jerárquicas entre las entidades.

En todos estos trabajos mencionados, se estudia la manera de generar ontologías para servicios SIG, pero no se presenta ni se define ningún proceso que facilite la identificación de servicios SIG solicitados.

### **3.5. Resumen**

En este capítulo se describieron las características del dominio geográfico y en particular diversas clasificaciones de servicios SIG. También se presentaron distintos estándares referidos al dominio geográfico y a la categorización de servicios SIG. Además, se evaluaron distintos enfoques utilizados para la selección de componentes.

El análisis del dominio geográfico realizado en este capítulo será de utilidad para la definición de nuestra propuesta de mejora que presentamos en el capítulo siguiente.



## Capítulo 4

# PROPUESTA DE MEJORA PARA EL PROCESO DE DEMANDA DE SERVICIOS SIG

En el desarrollo de software basado en componentes, los procesos de búsqueda y selección de componentes ofrecen soluciones que constituyen el núcleo del paradigma. Sin embargo, las metodologías de desarrollo no automatizan la ejecución de dichos procesos.

En este capítulo describimos nuestra propuesta [26], la cual constituye una parte del procedimiento integral construido especialmente para facilitar los procesos de búsqueda y selección de componentes. Así, nuestra propuesta plantea una mejora a dichos procesos consistente en la estandarización de los criterios de búsqueda de componentes. Para ello se definen:

- una taxonomía de servicios geográficos,
- un proceso para extracción de conocimiento en casos de uso y
- una aplicación que permite clasificar un requerimiento según las categorías de servicios geográficas propuestas en la taxonomía.

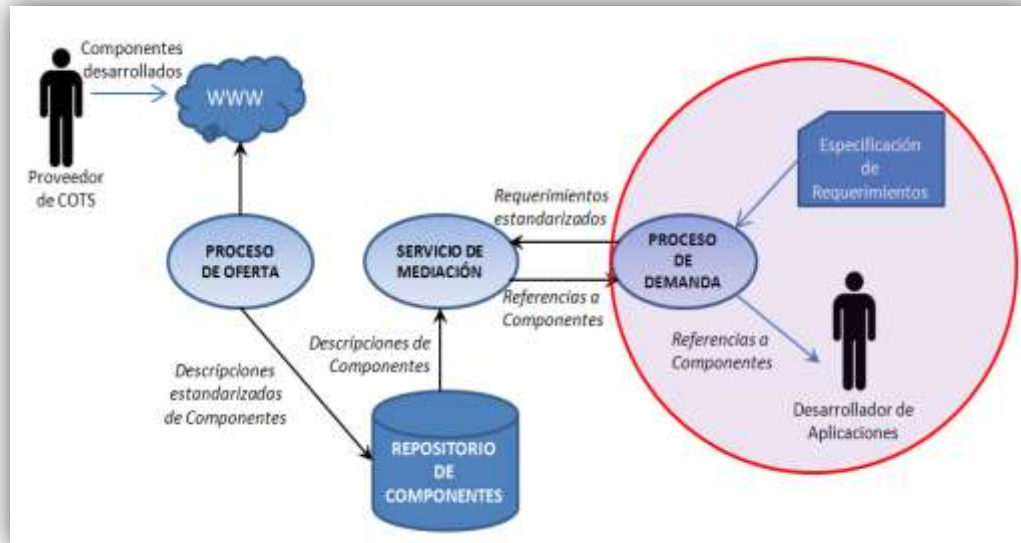
Cada uno de estos ítems que integran nuestra propuesta de mejora se detalla en las secciones 4.2 y 4.3.

### 4.1 Marco Contextual

Uno de los principales inconvenientes que se presentan al momento de realizar los procesos de búsqueda y selección consiste en la carencia de documentación suficientemente expresiva de los componentes que contribuyan a una selección efectiva del mismo. Otra dificultad consiste en que los criterios de búsqueda y

selección no tienen la precisión esperada y además las actuales metodologías de desarrollo de software no cuentan con procesos intermediarios que agilicen la tarea de localización de componentes que brinden los servicios requeridos.

Teniendo en cuenta estas limitaciones, hemos elaborado un procedimiento que contribuye a la estandarización del proceso de demanda de componentes, Figura 4.1.



**Figura 4.1:** Procedimiento propuesto para la búsqueda y selección de componentes. A la derecha se observa el proceso de demanda, el cual es objeto de desarrollo en esta tesis.

Este procedimiento se basa en la *Propuesta de Lucena* [44] detallada en el Capítulo 2, Sección 2.1.2.5 y en la incorporación de un *Modelo de Mediación de Componentes* [32], que facilita tanto la recuperación de información como la selección automatizada de componentes candidatos. Así, este procedimiento involucra dos actividades que representan dos perspectivas distintas del desarrollo basado en componentes, la del desarrollador de componentes y la del desarrollador de aplicaciones, las cuales reciben el nombre de *Proceso de Oferta* y *Proceso de Demanda* respectivamente.

El Proceso de Oferta, tiene que ver con la recopilación y almacenamiento de la información de los componentes desarrollados en un repositorio determinado, mientras

que el Proceso de Demanda, involucra la generación de la consulta para la posterior búsqueda y selección de componentes en uno o más repositorios. El procedimiento incluye además un *Servicio de Mediación*, el cual provee el vínculo entre los Procesos de Oferta y de Demanda. Así, ante una petición del proceso de demanda, requiriendo el servicio de un componente particular, el servicio de mediación busca en las descripciones de componentes almacenadas por el proceso de oferta en el repositorio determinado y le retorna las referencias de aquellos componentes que proporcionan la clase de servicio requerido.

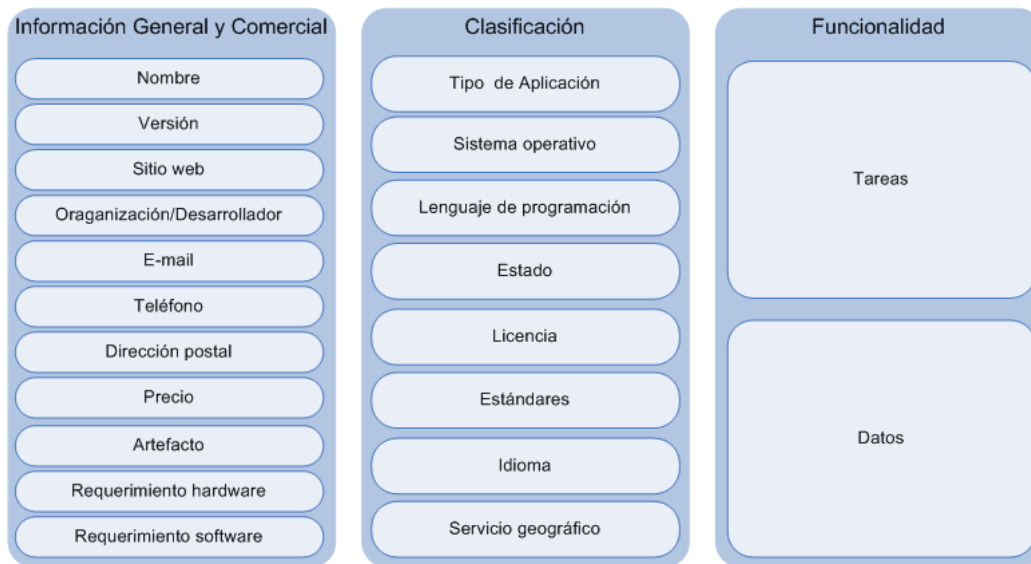
Para agilizar la tarea de búsqueda del servicio de mediación, surge la necesidad de contar con información estándar tanto de los componentes como de los requerimientos. Así, la oferta de servicios debe estandarizar la información de los componentes que almacena y *la demanda de servicios debe expresar los requerimientos en términos estándar para facilitar la búsqueda.*

Como antecedentes de este trabajo, para la estandarización del proceso de oferta, en [21][22][23] se define una taxonomía que contiene la información necesaria para describir un componente, y así facilitar el proceso de publicación. En esos trabajos se analiza qué información está realmente presente en la documentación de los catálogos de componentes para saber con qué datos contar a la hora de implementar el modelo de mediación. En [24][25], se describe el proceso de Extracción de información enriquecido con tecnologías asociadas al Procesamiento de Lenguaje Natural y a la Web Semántica. La idea principal consiste en poblar automáticamente un esquema de clasificación normalizado de componentes SIG (Tabla 4.1) con la información de oferta disponible en portales Web especializados.

En este esquema de clasificación normalizado se distinguen aspectos funcionales y no funcionales, organizados en 21 categorías distribuidas en tres grupos correspondientes con la clasificación propuesta en el Framework USCS [52]: Información general y comercial, Clasificación y Funcionalidad.

Así, el proceso de oferta almacenará en el repositorio la información de los componentes existentes utilizando la estructura de este esquema de clasificación normalizado.

La estandarización del proceso de demanda, es lo que precisamente constituye uno de los objetivos de esta tesis, en particular para la búsqueda y selección de componentes que provean servicios geográficos. La tarea del proceso de demanda consiste entonces en estandarizar los requerimientos recibidos en el documento de especificación, lo cual significa “traducir” esos requerimientos a categorías de servicios geográficos que se correspondan con la categoría Servicio Geográfico del esquema de clasificación normalizado (Tabla 4.1).



**Tabla 4.1:** Esquema de clasificación normalizado para el proceso de Oferta

Finalmente, al contar con información estandarizada de los componentes en el repositorio y recibir peticiones de servicio estandarizadas, el servicio de mediación podrá realizar una búsqueda más precisa y en menor tiempo retornando al proceso de demanda las referencias a los componentes que provean la funcionalidad requerida.

## 4.2. Propuesta para Estandarizar el Proceso de Demanda de Servicios SIG

Los desarrolladores de aplicaciones geográficas tienen requerimientos que satisfacer según la funcionalidad de la aplicación que deben construir. Los mismos están expresados en un Documento de Especificación de Requerimientos (DER) y

contienen un vocabulario muy variado y amplio. Debido a esa diversidad de lenguajes en que se expresan los requerimientos, surge la necesidad de estandarizar la especificación.

En nuestro trabajo, proponemos realizar esta estandarización de requerimientos, encontrando la categoría de servicio correspondiente a cada requerimiento de la especificación. De esta forma, la búsqueda de componentes consistirá en encontrar aquellos que brinden una categoría de servicios particular. Con este propósito, hemos trabajado en la generación de una taxonomía de servicios geográficos y un proceso para extracción de conocimiento a partir de casos de uso. A continuación describimos cada uno de estos ítems.

### **4.2.1. Definición de una Taxonomía de Servicios Geográficos**

Como hemos visto, existen diversas propuestas para categorizar la funcionalidad de los sistemas geográficos. La taxonomía propuesta en esta Tesis se elaboró tras un análisis de los trabajos previos descritos en el capítulo 3, los estándares del OGC (ISO19119 [57], OpenGIS Web Services Architecture [56]) y un relevamiento efectuado a usuarios SIG pertenecientes a organizaciones de la zona, acerca de sus expectativas sobre la funcionalidad de un sistema SIG. También se tuvo en cuenta la taxonomía definida en el proceso de oferta [21][22], la cual está detallada en la Sección 4.1 de este documento.

Para la elaboración de nuestra Taxonomía [69], se tuvieron en cuenta los niveles o capas definidos por el estándar ISO 19119 [57] y los servicios ejemplo de cada una de ellas. Se seleccionó esa categorización debido a que tras el estudio de los trabajos existentes sobre definición de taxonomías de servicios SIG, no se encontró otra clasificación que exhibiera mayor completitud. Además, se encontraron diferencias entre los niveles de abstracción de las categorías de las distintas clasificaciones, pero se constató que los servicios definidos por la ISO 19119 son lo suficientemente abarcativos para nuestro trabajo. Además el uso de estándares

constituye un paso más hacia la interoperabilidad y la definición de un lenguaje común.

La Taxonomía, tal como se muestra en la Tabla 4.2, está compuesta por las siguientes categorías de servicios:

- **Human Interaction:** incluye servicios para administración de interfaces de usuario, gráficos, multimedia y presentación de documentos compuestos
- **Model/Information Management Services:** abarca servicios para la gestión del desarrollo, manipulación y almacenamiento de metadatos, esquemas conceptuales y estructuras de datos
- **Workflow/Task Management Services:** son servicios para soportar tareas específicas o relacionadas con el flujo de trabajo realizado por individuos. Soportan el uso de recursos y desarrollo de productos que involucran una secuencia de actividades o pasos que pueden ser ejecutados por distintas personas
- **Geographic Processing Services:** incluye servicios que ejecutan cálculos a gran escala con grandes cantidades de datos. Esta categoría se subdivide según los tipos de propiedades de sus características en:
  - *Processing Spatial-Services*
  - *Processing Thematic-Services*
  - *Processing Temporal Services*
  - *Processing Metadata Services*
- **Communication Services:** abarca servicios de codificación y transferencia de datos entre redes de comunicación

Como se puede observar, esta Taxonomía está definida en idioma inglés. Esto se debe a que la herramienta que efectúa la estandarización de servicios, realiza los análisis lingüísticos sobre palabras de ese idioma, y esta Taxonomía tal como se detalla más adelante provee vocabulario que es utilizado por dicha herramienta para clasificar los requerimientos.

En el estándar ISO 19119 se describen además distintos ejemplos de servicios comprendidos por cada categoría. Su importancia radica en que mediante el discernimiento de la información provista por las descripciones de los servicios, es posible determinar a qué categoría pertenece un requerimiento particular.

Así, la estructura de nuestra Taxonomía, cuenta con dos columnas: *Categoría* y *Descripción del Servicio*. La primera columna detalla el nombre de cada categoría y la segunda (propia del desarrollo de esta tesis), provee información adicional que describe la funcionalidad brindada por cada categoría. La columna Descripción del Servicio, consiste en una lista de palabras clave que fue extractada de las descripciones de los ejemplos de servicios incluidos en cada categoría del estándar ISO 19119. La Descripción del Servicio es un elemento clave para el proceso de categorización de los servicios geográficos, dado que las herramientas toman estos vocablos como referencia para realizar los análisis lingüísticos.

Dentro de la lista de palabras clave asociadas a cada categoría, se pueden distinguir dos clases: verbos y sustantivos. Los verbos son muy significativos, ya que describen las acciones desarrolladas por los servicios incluidos en cada categoría. De igual manera, los sustantivos, y en particular los que cumplen el rol sintáctico de objeto directo, representan el dato procesado por el servicio. Así, ambas clases de palabras clave nos brindan información fundamental para el entendimiento de la funcionalidad provista por cada categoría de la Taxonomía.

CATEGORÍA	DESCRIPCIÓN DEL SERVICIO	
	Verbo Principal	Objeto Representativo
HUMAN INTERACTION	interact locate browse manage view display overlay query animate calculate edit	catalogue metadata feature coverage map spreadsheet service chain workflow view perspective texture symbol structure dataset
MODEL / INFORMATION MANAGEMENT SERVICES	access manage query describe discover order disseminate	feature store map graphic coverage sensor catalogue metadata gazetteer order
PROCESSING-SPATIAL SERVICES	convert transform rectify remove adjust extract tile measure manipulate correct verify	coordinate coverage vector image space input sample tiling dimension feature route position proximity distance
WORKFLOW/TASK MANAGEMENT SERVICES	define enact validate notify	chain workflow subscriptin event

**Tabla 4.2:** Taxonomía de servicios geográficos.



CATEGORÍA	DESCRIPCIÓN DEL SERVICIO	
	Verbo Principal	Objeto Representativo
PROCESSING THEMATIC SERVICES	register match generalize determine analyze calculate classify generalize extract count detect process convolve compress filter reduce manipulate change undestand synthesize create transform parse scan code	geoparameter theme region classification feature subset dataset terrain image resolution object text reference
PROCESSING TEMPORAL SERVICES	transform change extract analyze	time input sample subset proximity
PROCESSING METADATA SERVICES	calculate annotate add	statistics information label link description
COMMUNICATION SERVICES	encode decode transfer compress convert collaborate manage	rule protocol transfer communication file storage

**Tabla 4.2:** Taxonomía de servicios geográficos (cont.).

La columna *Descripción del Servicio* se encuentra a su vez subdividida en las columnas *Verbo Principal* y *Objeto Representativo*, cada una de las cuales incluye una lista de palabras clave asociadas, que consisten en los verbos y objetos directos extractados de los servicios brindados por cada categoría.

Por ejemplo, para la categoría *Human Interaction*, una de las ocho categorías de la taxonomía, en la columna de *Verbo Principal*, encontramos los verbos que representan las acciones posibles dentro de esa categoría: interact, locate, browse, manage, view, display, overlay, query, animate, calculate y edit. En la columna *Objeto Representativo*, encontramos los objetos directos: catalogue, metadata, feature, coverage, map, spreadsheet, service, chain, workflow, view, perspective, texture, symbol, structure y dataset. Estos objetos son los objetos directos que aparecen en las definiciones de los ejemplos de servicios de la categoría.

Así, la Taxonomía propuesta en esta tesis, consta de tres columnas: *Categoría*, que contiene los nombres de categorías propuestos por el estándar ISO 19119 y *Descripción del Servicio*, subdividido en *Verbo Principal* y *Objeto Representativo*. Estas dos últimas columnas contienen vocablos representativos de cada categoría que seleccionamos cuidadosamente para facilitar la clasificación de los servicios geográficos.

#### **4.2.2. Método Propuesto para la Extracción de Conocimiento en Casos de Uso de Aplicaciones SIG**

En esta sección se describe la metodología que proponemos [69][70][71] para estandarizar los requerimientos detallados en el DER.

Esta metodología constituye el núcleo del Proceso de Demanda. Recibe como entrada la especificación de requerimientos, expresada mediante casos de uso textuales y los procesa para obtener los requerimientos estandarizados, en forma de categorías genéricas de servicios geográficos, según la clasificación provista por la Taxonomía de Servicios Geográficos. La finalidad de este procesamiento es proveer esos

requerimientos estandarizados al Servicio de Mediación para que éste efectúe la búsqueda en el repositorio y nos devuelva las referencias a los componentes que satisfacen la funcionalidad requerida.

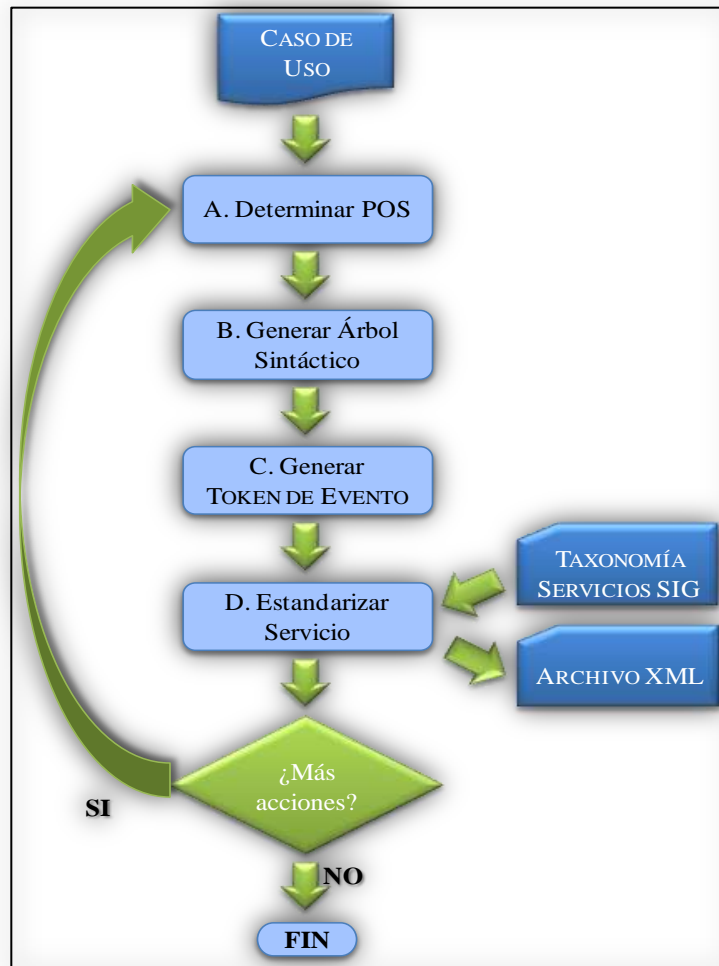
Esta metodología es también aplicable a otros dominios de información además de los servicios geográficos. Para reutilizar la metodología en cualquier otro dominio, solo sería necesario generar una nueva Taxonomía de Servicios, propia del dominio al que se quiere aplicar la metodología. Así, si se quisiera normalizar requerimientos por ejemplo del dominio contable, entonces se debería crear una taxonomía de servicios contables con sus respectivas categorías y descripciones de servicios.

En nuestro caso, dado que la información de los componentes ofrecidos también está categorizada según las clases de servicios del estándar ISO 19119, la consulta será más efectiva.

Para estandarizar o normalizar los requerimientos, el método busca en la Taxonomía de Servicios Geográficos (Tabla 4.2) la categoría que más se relaciona con el requerimiento especificado en el caso de uso textual. A fin de determinar el grado de relación existente, se analiza la relación entre las palabras claves asociadas a cada categoría de la Taxonomía y las palabras claves que se extractan del caso de uso: *verbos* y *objetos directos*. Así, para calcular el grado de relación existente, se comparan los verbos del caso de uso y los verbos de cada categoría de la Taxonomía (Verbos Principales), y además se compara también la relación entre los objetos directos del caso de uso y los objetos directos de cada categoría de la Taxonomía (Objetos Representativos).

Para comparar dos conceptos y determinar el grado de relación existente, se utiliza el Algoritmo Lesk Adaptado, detallado anteriormente en el Capítulo 2, Sección 2.2.3.

En la Figura 4.2 se muestran los pasos principales de esta metodología. Esta metodología está basada en un método para extracción de conocimiento en casos de uso textuales [17][49] detallado en el Capítulo 2, Sección 2.2.4.



**Figura 4.2:** Pasos para la extracción de servicios geográficos desde casos de uso

Como se puede observar en la Figura 4.2, los casos de uso constituyen la entrada de la metodología. El desarrollador proporciona un caso de uso textual en el cual se describe la funcionalidad más relevante. En nuestro enfoque, para aprovechar las ventajas del lenguaje natural y asimismo evitar ambigüedades, se utilizan casos de uso que usan lenguaje natural restringido. En particular, se emplea la propuesta presentada por Cockburn [14], detallada en el Capítulo 2 Sección 2.2.1., en la cual se utilizan plantillas para especificar el comportamiento en un caso de uso. Además, el lenguaje empleado dentro de los casos de uso es un lenguaje natural controlado, restringido mediante la estructura propuesta en [27], donde se aplica el patrón SVDPI

(Subject, Verb, Direct object, Preposition, Indirect object) : “La estructura de la oración debe ser simple”... “Sujeto... verbo... objeto directo... preposición... objeto indirecto”. De esta forma, estas dos propuestas [14][27] se combinan para facilitar la comprensión de los casos de uso a los usuarios comunes y además proporcionar una notación que pueda ser analizada y validada automáticamente.

Además, se puede observar también en la Figura 4.2 al componente Taxonomía de Servicios Geográficos explicado en la Sección 4.2.1.

Otro componente presente es el que corresponde al Archivo XML, el cual es utilizado para almacenar el resultado del servicio clasificado.

Resumiendo, tal como se muestra en la Figura 4.2, los pasos principales de la extracción de servicios son:

#### A) **Determinación del POS (part-of-speech)**

En este paso, se toma cada oración del Escenario Principal del caso de uso y se analizan sus palabras, etiquetándolas con su categoría gramatical correspondiente: (artículo, sustantivo, adjetivo, verbo, etc.).

#### B) **Generación del Árbol Sintáctico**

Para cada oración procesada en el paso anterior se analiza su sintaxis delimitando, separando y clasificando los constituyentes fundamentales (sujeto, predicado, objeto directo, etc.)

#### C) **Generación del Token de Evento**

En el tercer paso, se contruye el *token de evento*. El token de evento es el elemento fundamental de este proceso, ya que será el representante del requerimiento especificado en el caso de uso. Está integrado por el verbo principal y el objeto directo de la oración, que se obtienen del árbol sintáctico generado en el paso anterior.

#### D) **Estandarización del servicio requerido**

Por último, cada token de evento es procesado para obtener la categoría geográfica correspondiente según la Taxonomía de Servicios Geográficos. Para ello, es

necesario establecer con qué categoría de la Taxonomía el token de evento está más relacionado.

Con ese propósito, se realiza la comparación del *token de evento* con la *descripción de servicio* de cada categoría de la Taxonomía, calculando la relación existente mediante el Algoritmo Lesk Adaptado. En particular, se mide la relación existente entre el verbo del token de evento y cada uno de los verbos principales de las categorías de la Taxonomía. De igual manera se procede con el objeto directo del token de evento y cada uno de los objetos representativos de las categorías de la Taxonomía.

La categoría geográfica seleccionada, será aquella que obtenga el valor más alto de todas las categorías. Esto significa que el requerimiento del caso de uso está más relacionado con esa categoría, por lo cual a los efectos de consultar al Servicio de Mediación, se utilizará la categoría geográfica equivalente, la cual provee el nombre de servicio estandarizado.

### **4.3. ReqGIS: Herramienta de Clasificación de Requerimientos Geográficos**

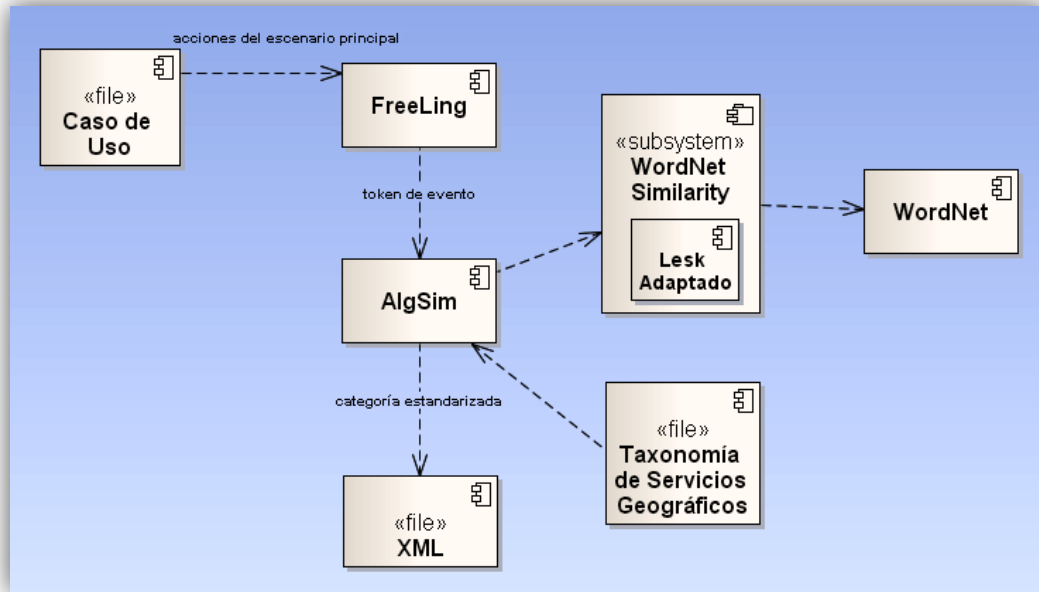
Para automatizar la metodología propuesta en la Sección 4.2.2., se desarrolló una herramienta software denominada ReqGIS.

El objetivo principal de la herramienta ReqGIS es automatizar el proceso de estandarización de los requerimientos geográficos (Figura 4.2) de los desarrolladores, clasificándolos según las categorías de la Taxonomía de Servicios Geográficos.

#### **4.3.1. Descripción de la Arquitectura**

Esta herramienta de clasificación de requerimientos se construyó mediante la integración de componentes disponibles en Internet y de un componente que se desarrolló específicamente para este proyecto, denominado AlgSim

Dicho componente es el encargado de efectuar el proceso de estandarización, objetivo de este trabajo. En la Figura 4.3 se muestran los componentes utilizados que de forma conjunta soportan la metodología propuesta en la Sección 4.2.2.



**Figura 4.3:** Diagrama de componentes de la herramienta ReqGIS.

Cada uno de estos componentes está detallado en el Capítulo 2, Sección 2.2.3. A continuación, se describe la contribución que cada uno de ellos realiza en el contexto de la herramienta ReqGIS y la interacción existente entre ellos.

- **Componente FreeLing.**

Dentro de la herramienta ReqGIS, el componente FreeLing efectúa los pasos A y B especificados en la metodología propuesta. Este módulo recibe como entrada un paso del escenario principal de un caso de uso, el cual consiste en una oración con formato SVDPI, y genera un árbol sintáctico de la oración. Este árbol sintáctico se utiliza luego para construir el *token de evento*, tomando las palabras etiquetadas como *top* y *direct object* para constituir *verbo principal* y *objeto representativo* del token de evento respectivamente.

- **Componente WordNet::Similarity**

WordNet::Similarity implementa las distintas medidas de similitud y relación semántica [61], que toman como entrada dos conceptos, y devuelven un valor numérico que representa el grado en que se asemejan o relacionan ambos conceptos. Una de las medidas de relación semántica utilizadas por este componente, es la que calcula el Algoritmo Lesk Adaptado y es la que el componente AlgSim utiliza para encontrar la categoría de servicio geográfico más relacionada con un token de evento.

- **Componente Algoritmo Lesk Adaptado**

Este algoritmo [7], se utiliza para medir la relación semántica entre el Token de Evento y cada categoría de la Taxonomía.

Para efectuar el cálculo de la relación semántica entre dos conceptos, este algoritmo realiza la comparación de las definiciones de ambos conceptos, las cuales son provistas por el componente WordNet. El cálculo se basa en la cantidad de palabras en común que tienen ambas definiciones. Así, una mayor cantidad de palabras en común determina un mayor grado de relación semántica.

- **Componente WordNet**

WordNet es una base de datos léxica del idioma Inglés [19]. A diferencia de los diccionarios tradicionales que están ordenados alfabéticamente, WordNet está ordenada semánticamente, creando una base de datos léxica de sustantivos, verbos, adjetivos y adverbios los cuales están organizados mediante distintas relaciones como: es-un, tiene-parte, está-compuesto-por, etc.

Así, este componente provee información semántica al componente WordNet::Similarity para calcular medidas de similitud y relación semántica.

- **Componente AlgSim**

Para satisfacer el objetivo de estandarizar los servicios requeridos por los clientes de componentes, desarrollamos el componente AlgSim en lenguaje Perl.



Este componente implementa el paso D de la metodología propuesta. Toma como entrada un token de evento y lo procesa, generando como resultado la categoría geográfica correspondiente al servicio requerido.

Para realizar la clasificación del servicio, el módulo ejecuta un algoritmo iterativo cuyo pseudocódigo se muestra en la Figura 4.4., accediendo a información almacenada en la Taxonomía de Servicios Geográficos y utilizando los servicios provistos por los componentes WordNet::Similarity, Algoritmo Lesk Adaptado y WordNet.

```

given eventToken (verb, object)
for each taxonomy category {
  for each category verb {
    calculate SIMILARITY(categoryVerb, tokenVerb)
  }
  calculate categoryVerbsSimilarityAverage
  for each category object {
    calculate SIMILARITY(categoryObject, tokenObject)
  }
  calculate categoryObjectSimilarityAverage
  calculate categorySimilarityAverage
}
category = category with highest categorySimilarityAverage
store(verb, object, category)
return category

```

**Figura 4.4:** Algoritmo que calcula el valor más alto de relación semántica.

El objetivo principal del algoritmo consiste en calcular el valor de relación semántica entre el token de evento y cada una de las categorías de la taxonomía, para así poder seleccionar aquella cuya relación semántica con el token de evento sea la mayor.

Para calcular el valor de relación semántica de cada categoría, se evalúa la relación semántica (mediante la aplicación del Algoritmo Lesk Adaptado) entre los elementos del token de evento y los elementos de cada categoría de la Taxonomía, es

decir pares de verbos (verbo de la categoría, verbo del token de evento) y pares de objetos (objeto de la categoría, objeto del token de evento).

Luego de encontrar la categoría más adecuada, el algoritmo almacena el resultado en un archivo XML.

### 4.3.2. Descripción de la Funcionalidad

En esta sección se detalla el funcionamiento de la herramienta ReqGIS mediante su aplicación al procesamiento de un caso de uso ejemplo.

En la Figura 4.5 se muestra el caso de uso utilizado como ejemplo, el cual describe la modificación de las coordenadas de ubicación de una línea eléctrica. El caso de uso está especificado en idioma inglés, dado que la herramienta lingüística que realiza el procesamiento trabaja con vocablos de ese idioma.

<b>Nombre</b>	Modify location coordinates of an electric line	
<b>Servicio</b>	Update information of electric lines	
<b>Descripción</b>	Assign new geographic coordinate	
<b>Actores</b>	System Administrator	
<b>Alcance</b>	Line layer, representing electric lines	
<b>Precondición</b>	Electric line layer has been created in the system	
<b>Escenario Principal</b>	<b>1</b>	The user selects an electric line
	<b>2</b>	The user changes coordinate attribute
	<b>3</b>	System displays electric line with location updated
<b>Postcondición</b>	System has updated location of electric line	

**Figura 4.5:** Caso de uso textual

Tal como se ilustra en la Figura 4.6, el caso de uso constituye el elemento de entrada de ReqGIS. Así, la herramienta aplica los cuatro pasos de la metodología a cada uno de los ítems que integran el “Escenario Principal” del caso de uso.

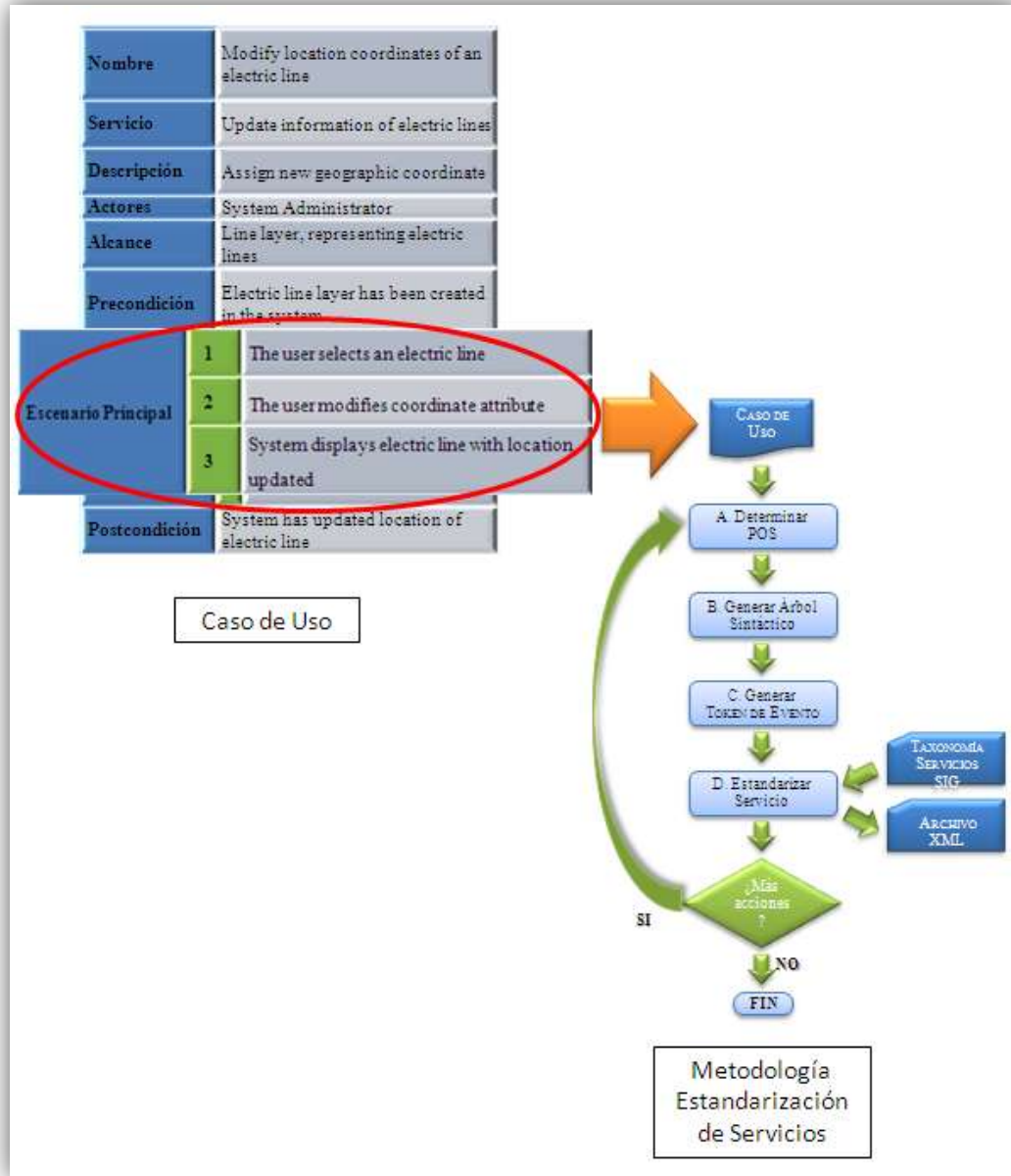
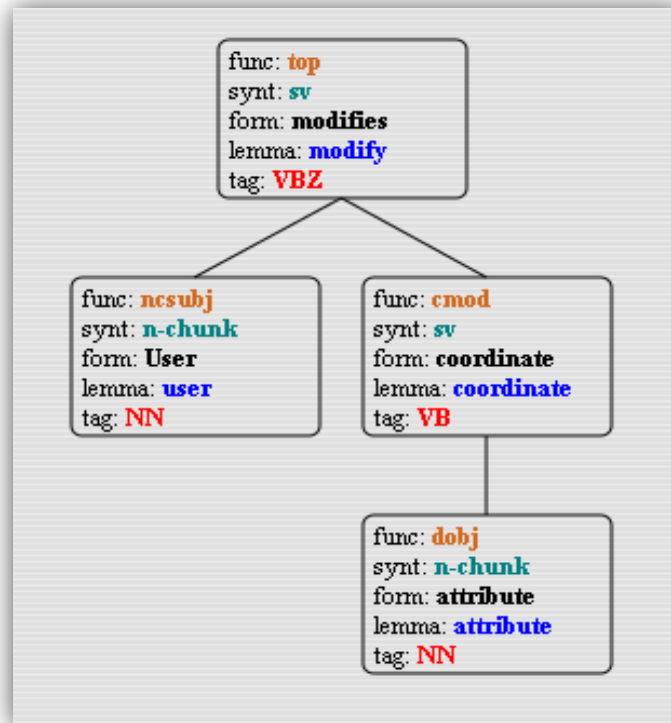


Figura 4.6: Aplicación de la metodología a un caso de uso.

Como resultado de aplicar los pasos A y B de la metodología, la herramienta genera un árbol sintáctico del caso de uso. Por ejemplo, para el ítem 2 del Escenario Principal del caso de uso: “User modifies coordinate attribute”, se genera el siguiente árbol sintáctico que clasifica cada una de las palabras de la oración, Figura 4.7:



**Figura 4.7:** Árbol sintáctico correspondiente a: “User modifies coordinate attribute”.

En ese árbol se puede ver que el nodo raíz etiquetado ‘top’ contiene la palabra ‘modifies’ la que según el cuerpo de lenguaje *Brown* para el idioma inglés se clasifica como VBZ, es decir que corresponde a un verbo en tercera persona singular cuya forma base es ‘modify’. Otro nodo del árbol es el que contiene la palabra ‘User’, la cual cumple con la función de sujeto de la oración y está clasificada como sustantivo singular (NN según el cuerpo *Brown*).

Continuando con el paso C de la metodología, ReqGIS genera el Token de Evento. Para ello, necesita encontrar el *verbo principal* y el *objeto representativo* de la oración. Así, la herramienta selecciona del árbol sintáctico generado previamente, el

nodo raíz etiquetado como “top” para representar el elemento “verbo principal” y el nodo etiquetado como “dobj” para representar el elemento “objeto representativo” del Token de Evento. En este ejemplo, la herramienta selecciona como verbo principal al verbo “modify” y como objeto representativo al sustantivo “attribute”. Por lo tanto, el Token de Evento resultante es “modify attribute”, tal como se muestra en la Figura 4.8.

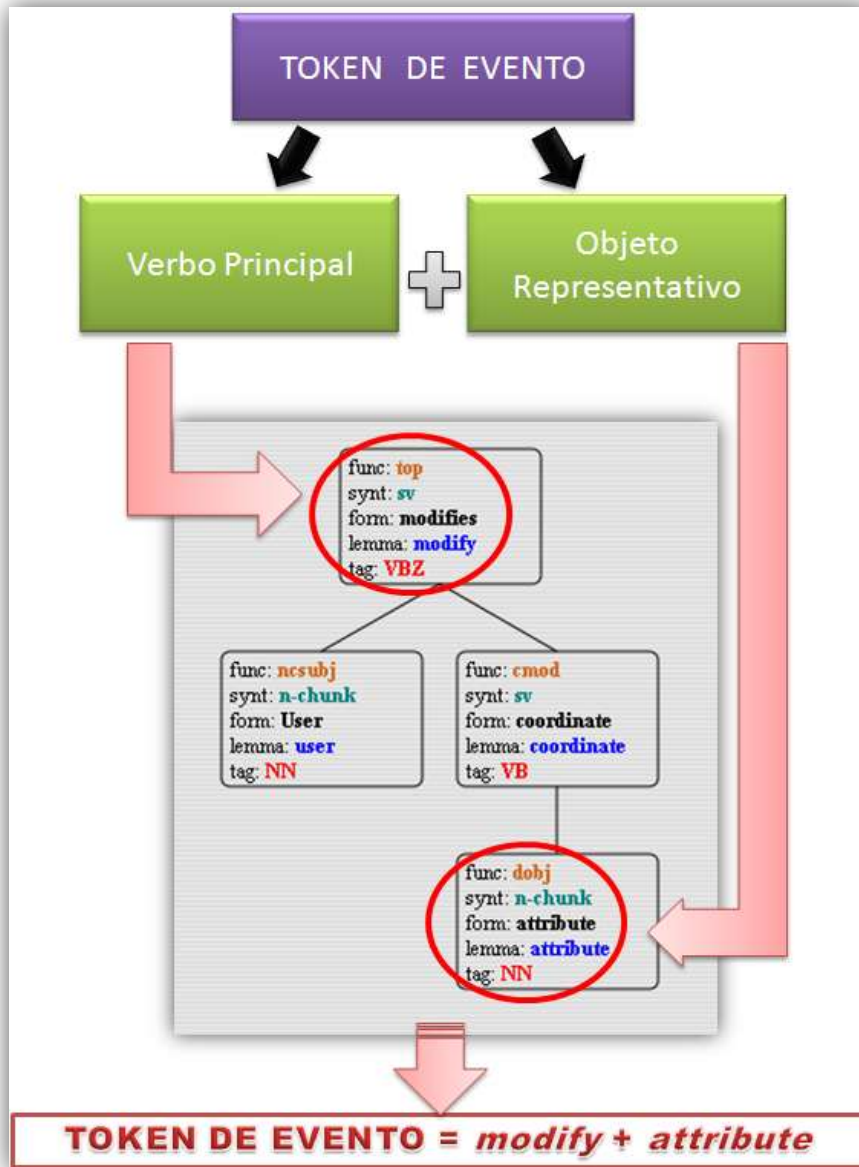


Figura 4.8: Formación del Token de Evento.

Finalmente, para el paso D de la metodología la herramienta procesa el token de evento, y obtiene el nombre estandarizado de la categoría de servicio que le corresponde.

Para efectuar este procesamiento, ReqGIS debe “razonar” cuál es la categoría de servicio que más se adecua al servicio requerido por el token de evento. Para ello, la herramienta realiza la medición de la relación semántica entre las palabras que componen el Token de Evento y las palabras clave definidas en cada categoría de la Taxonomía de Servicios Geográficos. Así, la categoría de la Taxonomía cuyo valor de relación semántica con el Token de Evento es el más alto, es la categoría seleccionada y constituye el nombre de servicio estandarizado provisto por la herramienta ReqGIS.

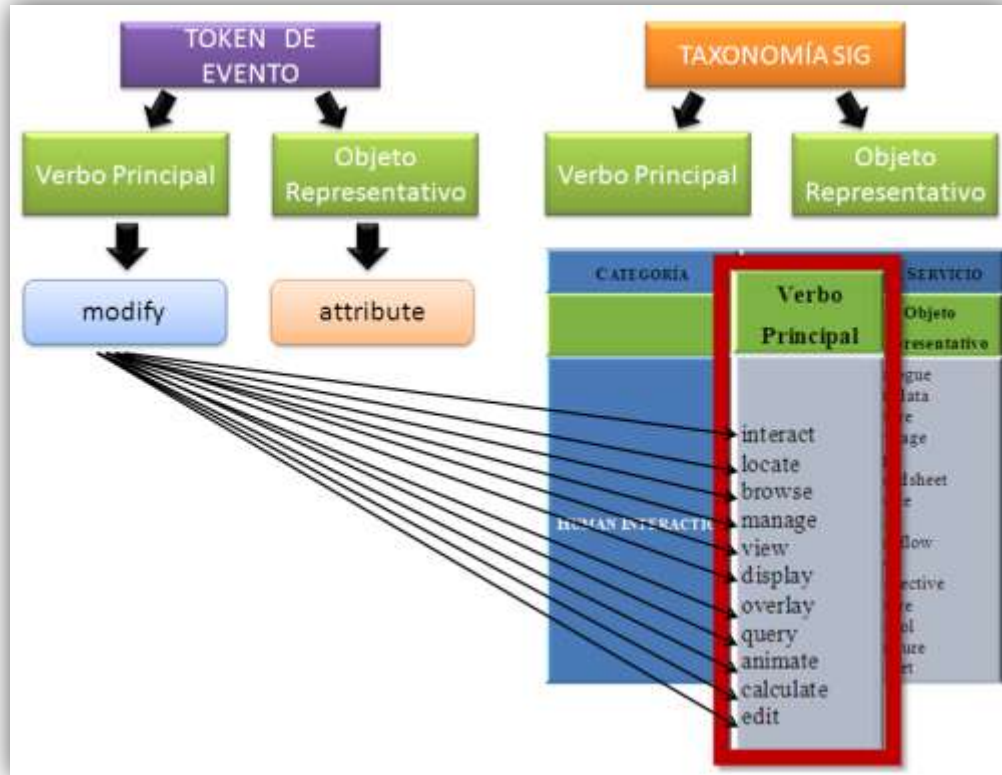
El algoritmo de cálculo de relación semántica utilizado por ReqGIS es el Algoritmo Lesk Adaptado, el cual se detalla en el Capítulo 2 Sección 2.2.1.2.

A continuación, se ilustra la medición de la relación semántica entre verbos. En particular, la medición entre el verbo *modify* del token de evento y cada uno de los verbos de la categoría *Human Interaction*. En la Figura 4.9 se observan todas las combinaciones posibles tenidas en cuenta para efectuar las mediciones de relación semántica.

En la Tabla 4.3, se pueden observar los valores de relación semántica resultantes para cada par (verbo categoría, verbo token evento) y en la última columna, el valor de relación semántica promedio para los verbos de la categoría *Human Interaction*.

Por ejemplo, si analizamos el cálculo de relación semántica entre los verbos “interact” y “modify”, encontramos que el valor de relación semántica existente entre ambos verbos es de 158, lo cual indica que “interact” es uno de los verbos más relacionados semánticamente con el verbo “modify”, así como también los verbos “calculate” y “edit” que registran valores de 122 y 213 respectivamente. Sin embargo, para el procesamiento desarrollado por ReqGIS, el valor que se toma como referencia es el promedio de todos los valores de relación semántica de los verbos de la categoría de la Taxonomía, es decir 72,54, valor que encontramos en la columna **Valor Promedio Categoría** de la Tabla 4.3. Se calcula el valor promedio de relación

semántica para disminuir el sesgo de una relación muy alta del verbo analizado con alguno de los verbos de la categoría y ninguna o muy poca relación con los demás verbos de esa misma categoría.



**Figura 4.9:** Medición de la relación semántica entre el verbo modify del token de evento y verbos de la categoría Human Interaction.

De igual manera se realiza la medición de relación semántica entre el objeto representativo que integra el token de evento y cada uno de los objetos representativos de la categoría Human Interaction.

Categoría	Verbos Categoría	Verbo Token Evento	Relación Semántica	Valor Promedio Categoría
HUMAN INTERACTION	interact	modify	158	72,54
	locate	modify	30	
	browse	modify	23	
	manage	modify	12	
	view	modify	95	
	display	modify	70	
	overlay	modify	30	
	query	modify	33	
	animate	modify	12	
	calculate	modify	122	
	edit	modify	213	

**Tabla 4.3:** Valores de relación semántica entre verbos de la categoría “Human Interaction” y verbo “modify” del token de evento.

En la Tabla 4.4, se pueden observar los valores de relación semántica resultantes para cada par (objeto representativo categoría, objeto representativo token evento) y en la última columna, el valor de relación semántica promedio para los objetos representativos de la categoría Human Interaction.



Luego de evaluar la relación semántica de verbos y objetos entre el token de evento y cada elemento de una categoría de la taxonomía, se calcula el valor general de relación semántica asociado a la categoría, promediando los valores promedio obtenidos previamente para los verbos y objetos de la categoría. Así, en nuestro ejemplo de la categoría *Human Interaction*, se promedian los valores promedio de verbo principal (72,54) y objeto representativo (54) obteniendo el valor 63,27, que respresenta la relación semántica de la categoría Human Interaction con el Token de Evento “modify attribute”

Categoría	Objetos Representativos Categoría	Objeto Representativo Token Evento	Relación Semántica	Valor Promedio Categoría
HUMAN INTERACTION	catalogue	attribute	35	54
	metadata	attribute	8	
	feature	attribute	127	
	coverage	attribute	8	
	map	attribute	76	
	spreadsheet	attribute	13	
	service	attribute	38	
	chain	attribute	31	
	workflow	attribute	9	
	view	attribute	64	
	perspective	attribute	64	
	texture	attribute	60	
	symbol	attribute	98	
	structure	attribute	179	

**Tabla 4.4:** Valores de relación semántica entre objetos representativos de la categoría “Human Interaction” y objeto “attribute” del token de evento.

La herramienta realiza estos cálculos para todas las categorías de la Taxonomía de Servicios Geográficos. La categoría que se selecciona es aquella que obtiene el valor más alto de relación semántica, lo cual indica que es la categoría cuyos verbos y objetos están más relacionados semánticamente con el token de evento.

Para el ejemplo analizado, la herramienta selecciona la categoría “Processing Metadata Services”, dado que su valor promedio de relación semántica es el mayor de todas las categorías. En la Tabla 4.5, se muestran los valores obtenidos por cada categoría.

CATEGORÍA	VALORES RELACIÓN SEMÁNTICA		
	Promedio Verbo	Promedio Objeto	Promedio Categoría
HUMAN INTERACTION	72,54	54,00	63,27
MODEL / INFORMATION MANAGEMENT SERVICES	37,14	45,50	41,32
WORKFLOW / TASK MANAGEMENT SERVICES	25,75	35,00	30,37
PROCESSING-SPATIAL SERVICES	119,18	76,50	97,84
PROCESSING THEMATIC SERVICES	118,85	58,53	88,69
PROCESSING TEMPORAL SERVICES	400,00	20,00	210,00
<b>PROCESSING METADATA SERVICES</b>	<b>476,25</b>	<b>45,60</b>	<b>260,92</b>
COMMUNICATION SERVICES	76,14	65,16	70,65

**Tabla 4.5:** Valores promedio de relación semántica para todas las categorías de la Taxonomía SIG.

La herramienta ReqGIS, luego de completar el procesamiento, muestra la categoría seleccionada y almacena también este resultado en un archivo XML, tal como se muestra en la Figura 4.10.

```

<service>
  <object> attribute </object>
  <verb> modify </verb>
  <category> Processing-Metadata_Services </Category>
</service>

```

**Figura 4.10:** Archivo XML que almacena la categoría de servicios elegida.

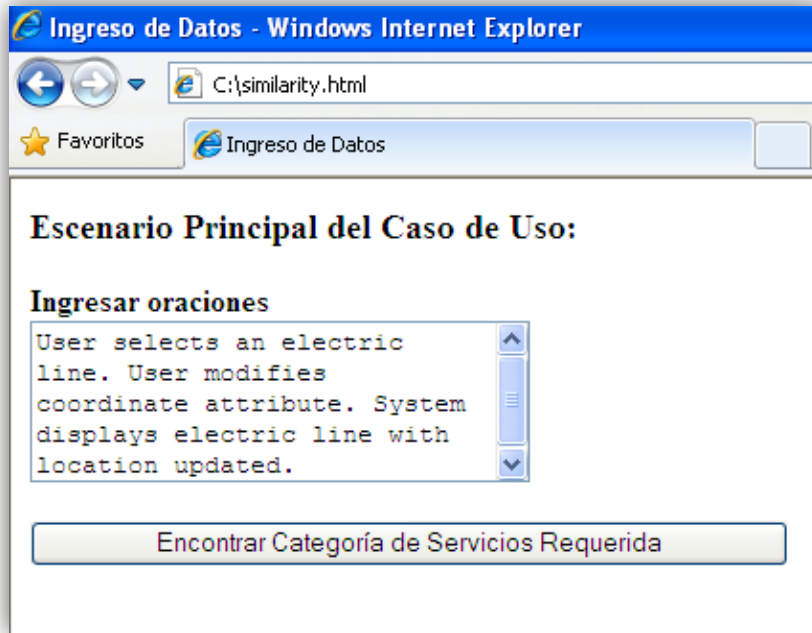
De esta manera, el modelo de demanda se encuentra estandarizado y podrá invocar a un servicio de mediación para que busque en el repositorio las referencias a componentes que brinden los servicios requeridos.

### 4.3.3. Interfaces de Usuario

ReqGIS es una aplicación web que cumple con el propósito de estandarizar los requerimientos geográficos de los desarrolladores de aplicaciones que construyen sus sistemas mediante la reutilización de componentes.

Así, esta herramienta recibe como entrada un caso de uso especificando ciertos requerimientos para un sistema SIG, y luego de efectuar el procesamiento de la información recibida aplicando la metodología de estandarización propuesta, detallada anteriormente en la Sección 4.2, devuelve al usuario el nombre del servicio estandarizado que corresponde al requerimiento especificado mediante el caso de uso.

A continuación se muestran las interfaces de usuario de la herramienta ReqGIS. En la primera, Figura 4.11, el usuario ingresa las oraciones que componen el escenario principal del caso de uso, que en nuestro ejemplo son cada uno de los pasos del caso de uso “*Modify location coordinates of an electric line*”.



**Figura 4.11:** Interfaz de la herramienta ReqGIS – Ingreso de Información.

En la segunda, Figura 4.12, se visualiza el resultado del procesamiento de las oraciones ingresadas, indicando la categoría estandarizada de servicio geográfico correspondiente, que en este ejemplo equivale a *Processing-Metadata Services*.



**Figura 4.12:** Interfaz de la herramienta ReqGIS – Presentación de Resultados.

Luego, el Servicio de Mediación podrá efectuar la búsqueda del servicio requerido en el repositorio de componentes, utilizando como parámetro de búsqueda la categoría de servicio proporcionada por ReqGIS.

#### **4.4. Resumen**

En este capítulo se describió nuestra propuesta de mejora al Proceso de Demanda. En primer lugar se detalló la Taxonomía de Servicios, la cual contiene los nombres de todas las posibles categorías de servicios geográficos así como también un conjunto de palabras seleccionadas que describen cada categoría de servicio geográfico y permiten discernir sobre la clasificación de un requerimiento geográfico. Además se detalló el procedimiento utilizado para estandarizar los requerimientos de servicios geográficos especificados mediante casos de uso. Por último se describió la herramienta ReqGIS que automatiza el procedimiento de estandarización de requerimientos y se ilustró su funcionamiento mediante el procesamiento de un ejemplo de caso de uso.

En el capítulo siguiente, se realizará el procesamiento de un caso de estudio completo para evaluar la aplicabilidad del procedimiento utilizado.



# Capítulo 5

## Caso de Estudio

En este capítulo aplicamos nuestra herramienta ReqGIS a un conjunto de requerimientos elicitados para el desarrollo de un Sistema de Información Geográfico para Información Pública. Luego, realizamos una experiencia con expertos en el dominio SIG para evaluar la validez del resultado proporcionado por ReqGIS.

### 5.1. Especificación de Requerimientos

Los requerimientos elicitados están especificados con casos de uso textuales, para lo cual se utilizó la plantilla propuesta por Cockburn [14]. Dicha plantilla propone describir los casos de uso mediante las siguientes secciones: Nombre del caso de uso, Descripción, Actores, Precondiciones, Postcondiciones, Escenario principal, Extensiones y Sub-variaciones, según se describe en el Capítulo 2, Sección 2.2.1.

#### 5.1.1. Alcance

El siguiente es un caso de estudio real basado en el Proyecto SIT SantaCruz [16]. Este proyecto está radicado actualmente en la Subsecretaría de Planeamiento de la provincia de Santa Cruz y consiste en una integración de herramientas y procedimientos de trabajo orientados a la generación, actualización y gestión de información relacionada con el territorio provincial, sus recursos, ocupación y uso.

Entre los objetivos propuestos se persigue alcanzar una multiplicidad de beneficios, entre los cuales se encuentran:

- Mejorar el acceso de los ciudadanos al Estado, al Gobierno y a sus acciones

- Mejorar la accesibilidad organizada a la información, lo cual permite hacer más eficientes los procesos de toma de decisiones por parte de las diferentes áreas técnicas del Estado
- Mejorar la precisión de la información al permitir la integración de múltiples fuentes en diversos formatos y escalas, según las necesidades específicas de usuarios y situaciones, en forma instantánea
- Permitir la consulta instantánea, el análisis, visualización o modelización de escenarios útiles, presentes y futuros, como soporte de decisiones
- Permitir la transferencia eficiente y ágil de datos a organismos de investigación, reguladores o controladores, incluso público en general
- Permitir ahorrar tiempos significativos en el acceso y utilización de la información, optimizando recursos humanos, tiempo y dinero
- Permitir la simplificación en la oferta de información a los ciudadanos, mediante la implementación de tecnologías y procedimientos de representación de datos, de fácil adopción y comprensión
- Permitir la mejora en los métodos de procesamiento de datos, lo que redundará en información más precisa para necesidades específicas de una organización
- Permitir la mejora en la utilización y gestión de recursos al asegurar una búsqueda más eficiente y una localización más rápida de éstos

La información actualmente se genera y procesa parcialmente mediante software bajo licencia aunque entre los objetivos de mediano plazo se encuentra la migración por completo (de procesos) a software libre, de acuerdo a los lineamientos establecidos en la Ley Provincial Nro 2787/05 de Uso de Software Libre en el Estado.

El proyecto SIT SantaCruz, no obstante encontrarse radicado en la Subsecretaría de Planeamiento (dependiente del Ministerio Secretaría General de la Gobernación), consiste en una iniciativa transversal en la estructura del Estado provincial. Con poco más de 50 acuerdos y protocolos específicos de trabajo,



acordades desde 2006 hasta la actualidad, SIT SantaCruz mantiene una veintena de líneas directas e indirectas de trabajo con diversos organismos públicos nacionales, provinciales y gobiernos locales, habiéndose alcanzado numerosos productos hasta el presente que incluyen diagnósticos territoriales, diagnósticos sectoriales, elaboración de cartografía temática y asistencias técnicas en materia de planificación y ordenamiento territorial.

En este marco, se plantea la necesidad de implementar un sistema de información geográfico (SIG) mediante un portal público, accesible por internet a todos los ciudadanos de la provincia que deseen realizar consultas. El sistema debe brindar información sobre:

- Propiedades y parcelas
- Barrios
- Proveedores de servicios públicos (municipales, provinciales, etc.)
- Seguridad y servicios de emergencia
- Servicios comunitarios (parques, comedores, centros comunitarios, centros asistenciales, escuelas, etc.)

Los requerimientos que expresan estas necesidades, están especificados mediante un conjunto de casos de uso textuales que se detallan en el Apéndice B de este documento.

### **5.1.2. Caso de Uso Ejemplo**

En esta sección se presenta un caso de uso textual, integrante de la especificación detallada en el Apéndice B. Este caso de uso se utiliza para ejemplificar el funcionamiento de la herramienta ReqGIS.

Este caso de uso, al igual que todos los que integran la Especificación de Requerimientos, está redactado en idioma inglés, dado que la herramienta ReqGIS con la que se estandarizarán los requerimientos especificados en ellos, opera en ese idioma.

En la Figura 5.1, se muestra el caso de uso “Find City Facilities”. Su finalidad es que el ciudadano pueda localizar las dependencias de servicios: policía, bomberos, hospitales, así como también plazas, zonas de recreación, etc.

<b>Nombre Caso de Uso</b>	<b>FIND CITY FACILITIES</b>
<b>Objetivo</b>	Citizen wants to know where all the parks, recreation facilities and trails are located around his neighborhood
<b>Escenario Principal</b>	1 Citizen chooses kind of facility from menu
	2 System shows selected facilities on map
	3 Citizen zooms map
	4 Citizen pans map to see facilities out of screen

**Figura 5.1:** Caso de Uso “Find city facilities”

## 5.2. Clasificación de Requerimientos con ReqGIS

En esta sección se muestran los resultados de procesar los requerimientos mediante la utilización de la herramienta ReqGIS. Recordemos que este software realiza la estandarización de los requerimientos solicitados, retornando las categorías de servicios geográficos correspondientes. Esas categorías son las categorías de servicios geográficos propuestas en la Taxonomía detallada en el Capítulo 4, Sección 4.2.1.

Así, al procesar el caso de uso “Find city facilities” (mostrado anteriormente en la Figura 5.1), los resultados obtenidos son los siguientes:

NOMBRE CASO DE USO	PASO DEL ESCENARIO PRINCIPAL	CATEGORÍA DE SERVICIO GEOGRÁFICO	VALOR DE RELACIÓN SEMÁNTICA
<i>“Find city facilities”</i>	1. Citizen chooses kind of facility from menu	Metadata	69
	2. System shows selected facilities on map	Human Interaction	308
	3. Citizen zooms map	Human Interaction	291
	4. Citizen pans map to see facilities out of screen	Human Interaction	292

**Tabla 5.1:** Resultados de procesamiento del caso de uso “Find city facilities” con ReqGIS

Como puede verse en la Tabla 5.1, la herramienta ReqGIS realiza la clasificación de servicios geográficos para cada uno de los pasos que componen el Escenario Principal del caso de uso “Find city facilities”. Así, para el primer paso “*Citizen chooses kind of facility from menu*”, el procesamiento determina que la categoría de servicios geográfica correspondiente es “*Metadata*” con un valor de relación semántica de 69 puntos. De la misma manera se procesan los pasos siguientes que componen el caso de uso, obteniendo una categoría y un valor de relación semántica para cada uno de ellos.

Así, puede observarse que la herramienta ha encontrado dos categorías de servicios geográficos relacionadas semánticamente con los pasos de este caso de uso: *Metadata* y *Human Interaction*. A fin de determinar cuál es la categoría que representa al caso de uso, ReqGIS evalúa los valores de relación semántica obtenidos para cada una de las categorías y selecciona aquella cuyo valor promedio es mayor, que representa a la categoría más relacionada semánticamente con el caso de uso. Así, para el caso de uso *Find city facilities* la herramienta selecciona la categoría *Human Interaction* con un valor de 297 puntos. En este caso, 297 es el valor promedio de los valores obtenidos por las distintas ocurrencias de la categoría *Human Interaction* en el caso de uso.

En el Apéndice C, se pueden consultar las categorías de servicios geográficos relacionados con los restantes casos de uso de la Especificación de Requerimientos.

## **5.3. Validación de los resultados obtenidos**

En esta sección se analizan los resultados obtenidos mediante la utilización de la herramienta ReqGIS. Para ello se confeccionó una encuesta que se detalla a continuación en la Sección 5.3.1 y luego se muestran las respuestas a dichas encuestas (Sección 5.3.2) proporcionadas por desarrolladores expertos en el dominio.

### **5.3.1. La Encuesta**

En cualquier investigación científica, en la etapa de recolección de datos, se usa un grupo de técnicas y herramientas a través de las cuáles podemos obtener y medir la información recopilada sobre un grupo de parámetros que queremos determinar, partiendo del diseño de la investigación, la muestra adecuada en concordancia con el problema científico a resolver y la hipótesis planteada, teniendo muy en cuenta las variables seleccionadas.

Existen varias técnicas e instrumentales para la recopilación de datos que se usan en las investigaciones científicas. En este trabajo específicamente se utiliza la técnica de la encuesta [36]. La encuesta es una técnica de investigación muy estandarizada que produce información cuantificable para ser analizada mediante técnicas estadísticas con fines descriptivos o para poner a prueba una hipótesis.

El cuestionario es un instrumento utilizado por la encuesta para la recolección de información, y es diseñado para poder cuantificar y universalizar la información. Su finalidad es conseguir la comparabilidad de la información.

Una encuesta no consiste sólo en el instrumento de recolección de información sino que además permite describir, comparar o explicar comportamientos, conocimientos y posturas. Por lo tanto, la encuesta es parte de un proceso mayor en el que se definen claramente las siguientes actividades:

1. Definir objetivos específicos medibles
2. Planificar y programar la entrevista

3. Asegurar la disponibilidad de recursos adecuados
4. Diseñar la encuesta
5. Preparar el instrumento de recolección de información
6. Validar el instrumento
7. Seleccionar participantes
8. Administrar y evaluar el instrumento
9. Analizar los datos
10. Mostrar los resultados

El primer paso para comenzar una encuesta es establecer los objetivos. Cada objetivo es simplemente una declaración de los resultados esperados de la encuesta. Es esencial dedicar tiempo suficiente a la definición de los objetivos para lograr que los mismos sean claros y medibles dado que son esenciales para las actividades posteriores de la encuesta.

El diseño de una encuesta es muy similar al diseño de un experimento en cuanto a que el diseño debe satisfacer los objetivos de la misma y para ello debe proporcionar los medios más efectivos de obtener la información necesaria. Se entiende por medios “efectivos”, que sean resistentes a sesgos, pertinentes y no demasiado costosos.

Una consideración a tener en cuenta en la construcción del cuestionario es el impacto del sesgo propio del investigador. A menudo se tiene una idea de lo que se está buscando, y la forma en que se construye el instrumento de encuesta puede revelar inadvertidamente un sesgo. Por ejemplo, se pueden influenciar las respuestas mediante la forma en que se realizan las preguntas, la cantidad de preguntas realizadas, la gama y tipo de categorías de respuestas o las instrucciones a los encuestados. Para evitar sesgos, es necesario realizar preguntas neutrales, realizar la cantidad suficiente de preguntas para cubrir adecuadamente el tópico en cuestión, tener en cuenta que el orden de las preguntas no inflencie las respuestas de las siguientes, escribir instrucciones claras sin sesgos.

Luego de crear el instrumento, es necesario evaluarlo para verificar que las preguntas son entendibles, estimar la tasa de respuestas y la confiabilidad y validez del instrumento. Esto se realiza mediante grupos de discusión o estudios piloto.

### **5.3.2. Diseño de la Encuesta**

Una encuesta tiene uno de los siguientes objetivos: describir un fenómeno de interés o evaluar el impacto de una intervención [36]. En el primer caso, la encuesta está orientada a esclarecer o conocer algún aspecto confuso o desconocido sobre un producto, proceso o población, requiriéndose un diseño de encuesta descriptivo. Mientras que en el segundo caso, la encuesta pretende evaluar la incidencia de una intervención, para lo cual será necesario un diseño experimental que soporte pruebas de hipótesis.

El diseño de la encuesta puede ser simple o complejo, debiéndose seleccionar el que provea los medios más efectivos para obtener la información requerida. Se entiende por “medios efectivos” a aquellos que presenten las siguientes características:

- Resistencia al sesgo: un diseño no debe estar excesivamente influenciado por un aspecto u opinión particular. Es decir, los resultados de la encuesta deben reflejar representativamente la realidad de la situación.
- Pertinencia: un diseño que tenga sentido en el contexto de la población, con la complejidad suficiente para resolver las cuestiones planteadas por los objetivos de estudio.
- Rentabilidad: un diseño cuya administración y análisis se encuentren dentro del presupuesto asignado.

La encuesta que se elaboró para este trabajo está orientada a desarrolladores de aplicaciones geográficas que utilizan tanto DSBC como desarrollo tradicional. El objetivo consiste en conocer la categorización propuesta por un grupo de desarrolladores de aplicaciones SIG expertos en el dominio, para un conjunto de casos de uso determinado. Así, posteriormente se podrá contrastar los resultados obtenidos por la herramienta ReqGIS contra los resultados de la encuesta.

En consecuencia, conforme al objetivo de nuestra encuesta, la clase de diseño a aplicar será un diseño descriptivo, el cual podría corresponder a alguno de los siguientes tipos:

- Transversal: los participantes responden respecto a un punto fijo en el tiempo.
- Cohorte: provee información sobre cambios en la población a futuro.
- Caso de control: los participantes responden sobre circunstancias previas, es un estudio retrospectivo.

Así, en nuestro caso el diseño será de tipo Transversal, debiendo los participantes responder sobre la clasificación que según su conocimiento del dominio, corresponde a cada caso de uso.

El cuestionario confeccionado para realizar la encuesta se muestra en la Figura 5.2. El mismo está compuesto por dos preguntas cerradas, es decir que el encuestado debe elegir su respuesta, seleccionando una o más opciones de una lista predefinida.

La primera pregunta indaga sobre el conocimiento previo del entrevistado acerca de la categorización de servicios geográficos propuesta por el estándar ISO 19119. Esta pregunta tiene por objetivo establecer el grado de dificultad con que el encuestado responde al cuestionario.

La segunda pregunta es la parte fundamental del cuestionario y es donde el encuestado debe clasificar cada caso de uso con cero, una o más categorías. Podría ser que según su criterio ninguna categoría de servicio satisfaga un caso de uso particular, o por el contrario que una o más categorías brinden los servicios requeridos por el caso de uso. La lista de respuestas que el encuestado tiene a disposición incluye además de las categorías de servicios geográficos, la respuesta “Ninguna”, para contemplar los casos en los que el encuestado entienda que el caso de uso no puede clasificarse con alguna de las categorías propuestas.

Además, se debe considerar la granularidad del caso de uso. Si el caso de uso analizado es muy específico, seguramente va a requerir un servicio de una única

categoría, pero si el caso de uso es más genérico o de alto nivel, podría ser que necesite servicios de más de una categoría.

Cuando se administra una encuesta, no es rentable (y a veces imposible) encuestar a toda la población. Es por ello que se debe seleccionar para la encuesta un subconjunto de la población al que se denomina muestra, y cuyas respuestas se espera sean representativas del grupo completo. Si la muestra no es representativa, entonces los resultados no se pueden generalizar a la población objetivo [37].

Existen dos clases de métodos de muestreo, muestreo probabilístico y muestreo no-probabilístico. Una muestra probabilística es aquella en la cual todo miembro de la población tiene una probabilidad conocida distinta de cero de ser incluida en la muestra, así en una muestra probabilística el objetivo es eliminar la subjetividad y obtener una muestra sin sesgo y representativa de la población objetivo. Las muestras no-probabilísticas se crean cuando los encuestados son elegidos por su facilidad de acceso o porque los investigadores tienen alguna justificación para creer que son representativos de la población. Este tipo de muestreo tiene el riesgo de tener sesgo y en consecuencia no es posible obtener inferencias estadísticas a partir de este tipo de muestras.

No obstante, existen razones para usar muestras no-probabilísticas:

- La población objetivo es difícil de identificar
- La población objetivo es muy específica y de disponibilidad limitada
- La muestra es un estudio piloto, no la encuesta final, y un grupo no aleatorio está disponible.

En nuestro caso de estudio utilizamos una muestra no-probabilística, debido a que la población disponible en la zona es muy limitada. Por lo tanto la encuesta se realizó a un conjunto de desarrolladores de aplicaciones SIG conocidos a través de contactos personales del grupo de investigación.

A continuación, en la Figura 5.2, se muestra el cuestionario elaborado para la encuesta:



### Cuestionario sobre Categorización de Requerimientos

1. ¿Conocía Ud. la clasificación de servicios geográficos propuesta por el estándar ISO 19119?  
 SI  NO
2. ¿Qué categoría/s considera que satisfacen los requerimientos especificados en el caso de uso? *(Anotar en la línea punteada todos los números de categoría que correspondan)*

**CATEGORÍAS DE SERVICIOS GEOGRÁFICOS**

<b>1. Interacción Humana</b>	<b>5. Procesamiento Temporal</b>
<b>2. Gestión de Modelo/Información</b>	<b>6. Procesamiento Metadatos</b>
<b>3. Procesamiento Espacial</b>	<b>7. Comunicación</b>
<b>4. Procesamiento Temático</b>	<b>8. Ninguna</b>

CASO DE USO	CATEGORÍA DE SERVICIO CORRESPONDIENTE
Locate Property	
Assess Property	
Find School District	
Find City Facilities	
Create Landmarks	
Find Schools	
Measure Distance	
Create Formatted Maps and Reports	
Turn On/Off Layers	
Run Statistics	

**Figura 5.2:** Cuestionario de la encuesta

A partir de la información recabada por el cuestionario, se analizará para cada caso de uso, la concordancia entre la categoría seleccionada por la herramienta ReqGIS y la/las categorías elegidas por los encuestados.

### 5.3.3. Resultados de la Encuesta

El cuestionario presentado en la Sección 5.3.2 fue respondido por diez desarrolladores de aplicaciones SIG, a quienes se les suministró junto con el cuestionario, la especificación completa de casos de uso y el listado de categorías de servicio geográfico propuesto en nuestra Taxonomía junto a la descripción de cada una de las categorías. La descripción de categorías se proporcionó como material de referencia, para informar sobre el alcance de cada categoría de servicio geográfico.

En la Tabla 5.2 se detalla el resultado de la pregunta 1 del cuestionario. A través del mismo se puede discernir que un 80 por ciento de los desarrolladores entrevistados tomaron contacto con la clasificación de servicios geográficos recién en el momento de la encuesta. Esta cuestión no se considera una limitación para la encuesta, dado que los desarrolladores que respondieron la encuesta son expertos en el dominio de sistemas geográficos, y aunque no conocían previamente la clasificación, conocen la funcionalidad de los sistemas de información geográficos.

	SI	NO
¿Conocía el Estándar ISO 19119?	2	8

**Tabla 5.2:** Conocimiento previo sobre Categorización de Servicios

A continuación, en la Tabla 5.3 se presentan los resultados tabulados de la pregunta 2 del cuestionario. En el Apéndice D se incluyen los diez cuestionarios individuales.

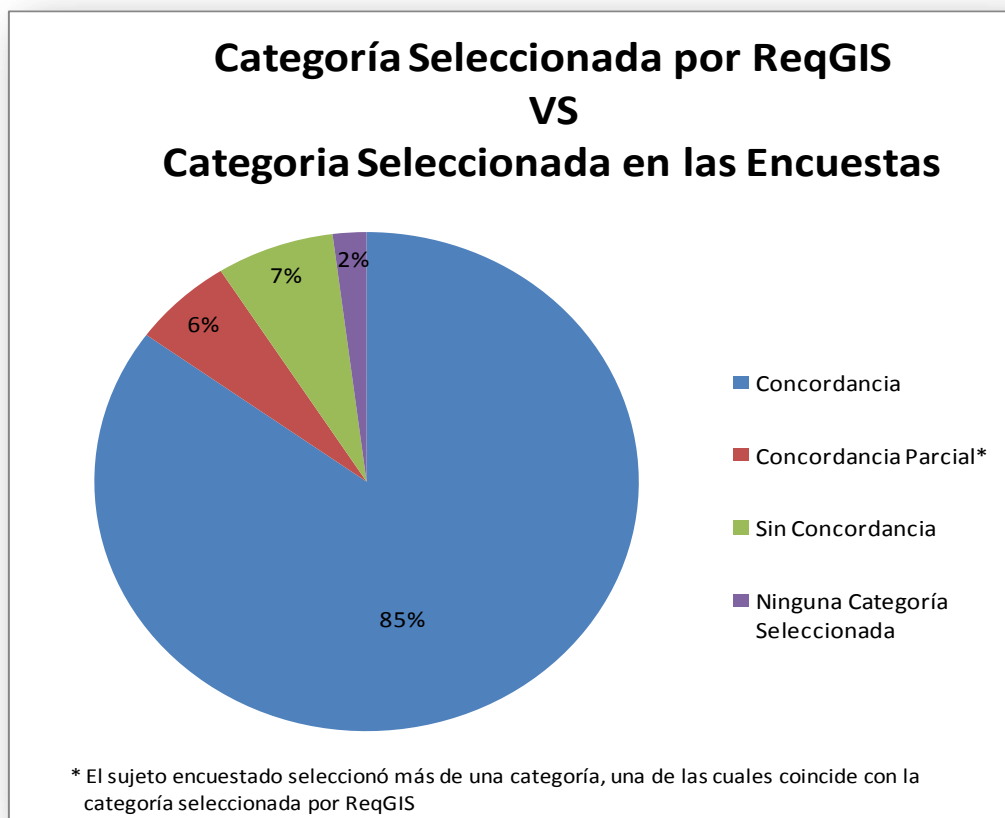
La tabulación de la Tabla 5.3, se compone de las columnas *Caso de Uso*, en la cual se nombra cada uno de los casos de uso analizados en el cuestionario; la columna *Coincidencia Total*, en la cual se registra la cantidad de sujetos encuestados que para un caso de uso determinado coincidieron en la elección de la categoría con la categoría propuesta por ReqGIS. En la columna *Sin Coincidencia*, se anota la cantidad de sujetos encuestados que para un caso de uso en particular eligieron una categoría distinta que la categoría propuesta por ReqGIS. En la columna *Coincidencia Parcial*, se registra la

cantidad de cuestionarios en que el sujeto encuestado seleccionó más de una categoría de servicio, y que entre esas categorías seleccionadas una de ellas coincide con la propuesta por ReqGIS. Por último, la columna *Ninguna Categoría*, incluye la cantidad de encuestados que para un caso de uso en particular seleccionaron como categoría, la opción del cuestionario: “8. Ninguna”.

En este gráfico podemos observar que el 85 por ciento de las categorizaciones realizadas por los sujetos encuestados coincidieron con la categorización de ReqGIS. Del 15 por ciento que no coincidieron totalmente, el 6 por ciento eligió la misma categoría que la propuesta por ReqGIS pero consideró que no satisfacía completamente los requerimientos del caso de uso, por lo cual seleccionó una o más categorías adicionales, y el 2 por ciento no encontró ninguna categoría que cumpliera con los requerimientos del caso de uso, por lo cual seleccionó la opción “Ninguna Categoría”. Solamente un 7 por ciento de las categorías elegidas por los sujetos encuestados no coincidieron con la categorización de ReqGIS.

CASO DE USO	Coincidencia Total	Sin Coincidencia	Coincidencia Parcial	Ninguna Categoría
Locate Property	10			
Assess Property	10			
Find School District	6	2	2	
Find City Facilities	10			
Create Landmarks	5	3	2	
Find Schools	10			
Measure Distance	10			
Create Formatted Maps and Reports	4	2	2	2
Turn On/Off Layers	10			
Run Statistics	10			
<b>TOTALES</b>	<b>85</b>	<b>7</b>	<b>6</b>	<b>2</b>

**Tabla 5.3:** Tabulación de Cuestionarios



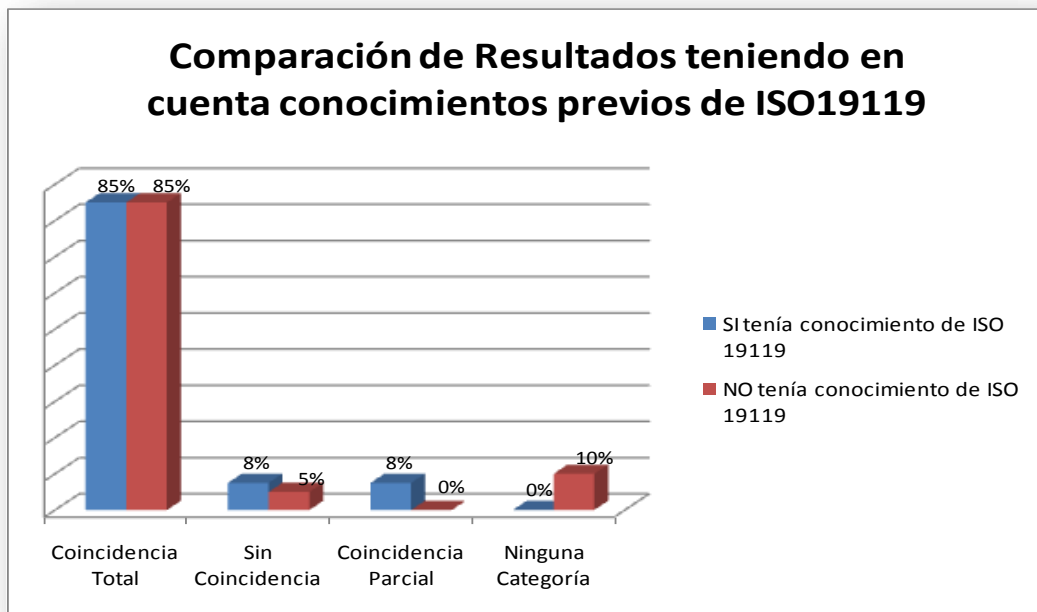
**Figura 5.3:** Gráfico circular “Categoría Seleccionada por ReqGIS vs Categoría Seleccionada en las Encuestas”

La Figura 5.3 muestra un gráfico de los resultados descriptos en la Tabla 5.3. En la Tabla 5.4, se visualiza la misma información presentada en la Tabla 5.3, pero discriminada según el conocimiento previo que el encuestado tiene sobre la clasificación de servicios geográficos del estándar ISO 19119.

En la Figura 5.4, se muestra un gráfico en el que se comparan los valores porcentuales de los distintos resultados según si el encuestado tiene o no conocimiento previo de la categorización de servicios geográficos del estándar ISO 19119

CASO DE USO	SI tenía conocimiento de ISO 19119				NO tenía conocimiento de ISO19119			
	Coincidencia Total	Sin Coincidencia	Coincidencia Parcial	Ninguna Categoría	Coincidencia Total	Sin Coincidencia	Coincidencia Parcial	Ninguna Categoría
Locate Property	8				2			
Assess Property	8				2			
Find School District	5	1	2		1	1		
Find City Facilities	8				2			
Create Landmarks	3	3	2		2			
Find Schools	8				2			
Measure Distance	8				2			
Create Formatted Maps and Reports	4	2	2					2
Turn On/Off Layers	8				2			
Run Statistics	8				2			
<b>TOTALES</b>	<b>68</b>	<b>6</b>	<b>6</b>	<b>0</b>	<b>17</b>	<b>1</b>	<b>0</b>	<b>2</b>

**Tabla 5.4:** Tabulación de cuestionarios teniendo en cuenta el conocimiento previo de ISO 19119



**Figura 5.4:** Gráfico de columnas comparativo según conocimiento previo de ISO 19119.

Así, se puede observar que el porcentaje alcanzado por “*coincidencia total*” es el mismo para ambas situaciones de los encuestados. Es decir, sea que los entrevistados conocieran o no previamente el estándar, el porcentual es equivalente para “*coincidencia total*”, cuando la categoría de servicio geográfico seleccionada por el encuestado coincide con la categoría de servicio geográfico propuesta por ReqGIS.

En cuanto al porcentaje correspondiente a “*sin coincidencia*” es similar en ambos casos, siendo del 8 por ciento cuando el entrevistado tenía conocimiento previo del estándar y 5 por ciento en caso contrario. Con referencia a “*coincidencia parcial*”, se observa que sólo se han presentado casos cuando el encuestado tenía conocimiento previo del estándar, y para el caso de “*ninguna categoría*” sucede a la inversa, sólo se presentan casos cuando el encuestado no tenía conocimiento del estándar.

En consecuencia, el conocimiento previo de la categorización de servicios geográficos no es una variable que modifique sustancialmente el resultado de la encuesta, siendo esto favorecido por el conocimiento del dominio geográfico con que cuentan todos los encuestados.

Estos resultados fueron obtenidos a partir de una muestra no-probabilística de desarrolladores, por lo tanto no se puede generalizar el análisis a la totalidad de la población de desarrolladores SIG. Sin embargo, este resultado es alentador, dado que para este caso de estudio real, la comparación de resultados obtenidos con la herramienta ReqGIS y los resultados de las encuestas muestra un alto grado de coincidencias.

## **5.4. Resumen**

En este capítulo se muestra el resultado de una experiencia realizada con un equipo de desarrolladores de aplicaciones SIG. Se partió de una especificación de requerimientos para un Sistema de Información Geográfico para información pública, basado en el Proyecto SIT Santa Cruz, la cual está expresada mediante casos de uso textuales.

Esta especificación fue procesada utilizando la herramienta ReqGIS, y se obtuvieron las categorías de servicios geográficas necesarias para satisfacer esos requerimientos.

A continuación se efectuó una validación preliminar de los resultados obtenidos mediante la realización de una encuesta.

Esta encuesta fue respondida por desarrolladores SIG, sobre qué categoría de servicio geográfico es la que soluciona cada caso de uso, según su criterio y experiencia.

Luego se realizó la comparación de los resultados obtenidos por ReqGIS y las respuestas brindadas por los expertos encuestados, y se constató que hubo un 85 por ciento de coincidencia en la selección de categorías de servicios geográficos. Entre los casos en que no hubo coincidencia, en parte se debió al concepto de granularidad de los casos de uso. Los desarrolladores consideraron que para satisfacer un caso de uso era necesaria más de una categoría, concepto que no se contempla en nuestra herramienta. Otro de los factores por los que no hubo coincidencia, aunque no fue muy relevante, es el desconocimiento de la categorización de servicios geográficos, que provocó que algunos desarrolladores respondieran que ninguna categoría satisfacía un caso de uso particular. Así, sólo un 7 por ciento del total de respuestas son las que específicamente no coincidieron con la clasificación de nuestra herramienta ReqGIS.

Dado que los resultados obtenidos no pueden generalizarse a la población objetivo, resta realizar a futuro un muestreo probabilístico que nos permita realizar inferencias estadísticas.





# Capítulo 6

## Conclusión y Trabajo Futuro

En este capítulo se analizan los resultados del trabajo llevado a cabo en esta tesis. De esta manera, los apartados incluidos son los siguientes: análisis de la consecución de objetivos, principales aportes de esta tesis, contraste de los mismos en publicaciones científicas y líneas de trabajo abiertas.

### 6.1. Análisis de la Consecución de Objetivos

En el primer capítulo de esta tesis se han presentado los objetivos parciales que se pretendían cumplir para satisfacer el objetivo principal de nuestra investigación que consiste en:

*Mejorar el modelo de demanda al definir herramientas que estandaricen los servicios requeridos.*

A continuación, se analiza cada sub-objetivo propuesto, en términos de los logros actuales.

a) *Clasificar Servicios SIG.*

En el Capítulo 3, se analizaron distintas categorizaciones propuestas para servicios geográficos así como también estándares existentes para definición y categorización de servicios SIG. Tomando como base la clasificación de servicios geográficos propuesta en el estándar ISO 19119, en el Capítulo 4 se definió una Taxonomía de Servicios Geográficos. Esta clasificación consta de 8 categorías, las cuales están descritas por un vocabulario asociado que permite discernir a qué categoría corresponde un servicio geográfico particular.

b) *Construir el Modelo de Demanda.*

En el Capítulo 2, se examinaron diversos enfoques existentes para efectuar la selección de componentes y se analizó la técnica de casos de uso como herramienta para la especificación de requerimientos. A su vez, se estudiaron distintas herramientas lingüísticas que permiten extraer información de los casos de uso. Finalmente, en el Capítulo 4, se presenta la propuesta de un Método para la Extracción de Conocimiento en Casos de Uso, y se describe la herramienta ReqGIS, que automatiza el Método propuesto, clasificando los requerimientos expresados mediante casos de uso textuales según las categorías definidas en la Taxonomía de Servicios Geográficos

c) *Validar el Modelo*

En el Capítulo 5, se describe un caso de estudio real al que se aplica la herramienta ReqGIS, y se muestran los resultados de una encuesta realizada a desarrolladores expertos en el dominio de sistemas SIG en la que se les consulta sobre cómo categorizarían los requerimientos especificados para el caso de estudio real. Del análisis comparativo entre las respuestas a las encuestas y los resultados brindados por ReqGIS surge una alta coincidencia en la selección de categorías; pero dado que los resultados obtenidos no pueden generalizarse a la población objetivo, es necesario realizar un muestreo probabilístico que nos permita realizar inferencias estadísticas.

## **6.2. Principales Aportes**

Como principales aportes de esta Tesis se pueden mencionar los siguientes:

- La taxonomía de servicios geográficos, integrada por categorías de servicios geográficos y vocabulario que facilita la identificación de la categoría geográfica correspondiente a un servicio geográfico particular.
- Un proceso para extracción de conocimiento en casos de uso y
- Una herramienta que permite clasificar un requerimiento según las categorías de servicios geográficas propuestas en la taxonomía.

### 6.3. Publicaciones derivadas de esta Tesis

A continuación se detallan las publicaciones realizadas hasta el momento de resultados parciales de esta tesis. Las mismas se listan clasificadas según su tipo:

#### Artículos en Publicaciones y Journals

- Saldaño V.; Buccella A.; Cechich A. *ReqGIS Classifier: A Tool for Geographic Requirements Normalization*. XVII Argentine Congress of Computer Science Selected Papers, Computer Science & Technology Series. Armando De Giusti / Javier Diaz (Eds). VIII Software Engineering Workshop, 2011.
- Saldaño V.; Buccella A.; Cechich A. *Discovering Geographic Services from Textual Use Cases*, Journal of Computer Science & Technology, Vol. 10 - No. 2 – June 2010 - ISSN 1666-6038.

#### Conferencias Internacionales

- Gaetán G, Saldaño V, Buccella A, Cechich A. *A Domain-Oriented Approach for GIS Component Selection*. En Proceedings de ICSEA 2010, 5th International Conference on Software Engineering Advances, en Nice, Francia, Agosto 2010. Editado por: Jon Hall, H. Kaindl, L. Lavazza, G. Buchgeher. O. Takaki. ISBN 978-0-7695-4144-0. Editorial: IEEE Computer Society. Páginas: 94-100.

#### Conferencias Nacionales

- Saldaño, V.; Buccella, A.; Cechich, A. *A Tool for Geographic Requirements Normalization*, Actas del XVII Congreso Argentino en Ciencias de la Computación. ISBN 978-950-34-0756-1. (877-886), 2011.

- Gaetán, G.; Saldaño, V.; Buccella, A.; Cechich, A.; Martin, A.; Molina, S. *Selección de Componentes para SIG a partir de Información en Servicios y Catálogos Web Normalizados*. Actas del XIII Workshop de Investigadores en Ciencias de la Computación, Rosario, Argentina, 2011.
- Gaetán, G.; Saldaño, V.; Buccella, A.; Cechich, A.; Martin, A.; Molina, S. *Publicación y Selección de Componentes para SIG*. En Anales del XII Workshop de Investigadores en Ciencias de la Computación, El Calafate, Argentina, 2010.
- Gaetán, G.; Saldaño, V.; Buccella, A.; Cechich, A.; Martin, A.; Molina, S. *Recuperación de Información de Componentes SIG Enriquecido con Tecnologías Semánticas*, Encuentro de Investigadores de Ciencias Básicas, Aplicadas y Experimentales de Patagonia Austral EIBAE, 2010.
- Saldaño V.; Buccella A.; Cechich A. *Descubrimiento de Servicios Geográficos a partir de Casos de Uso Textuales*, Actas del XV Congreso Argentino en Ciencias de la Computación, ISBN 978-897-24068-4-1, (921—930), 2009.
- Gaetán, G.; Saldaño, V.; Buccella, A.; Cechich, A. Clasificación de Componentes SIG. En Anales del XI Workshop de Investigadores en Ciencias de la Computación, San Juan, Argentina, 2009.
- Gaetan, G.; Saldaño, V.; Buccella, A.; Cechich, A.; Mayorga, C.; Gelman, B.; Pejcich, G.; Diaz, M. Clasificación de componentes OTS (Off-The-Shell) para sistemas de información geográfica. En Anales del X Workshop de Investigadores en Ciencias de la Computación, WICC08. Gral. Pico, La Pampa, 2008.
- Saldaño, V.; Buccella, A.; Cechich, A. *Una Taxonomía de Servicios*

*Geográficos para Facilitar la Identificación de Componentes.* En Actas del XIV Congreso Argentino en Ciencias de la Computación, ISBN 987-24611-0-2, 2008.

## 6.4. Trabajos Futuros

El trabajo presentado en esta tesis solo trata una faceta de la selección de componentes geográficos. Futuras investigaciones podrían extender los resultados obtenidos:

- Desarrollar el Modelo de Mediación, para integrar los Modelos de Oferta y de Demanda. Esto permitirá introducir el uso del método propuesto en un proceso integral de selección de componentes. Este proceso integral ha sido publicado en la conferencia internacional anteriormente citada. El Modelo de Oferta es objeto de otro trabajo de Tesis [20].
- Ampliar la experimentación de la herramienta ReqGIS con nuevos casos de uso, para ajustar el vocabulario asociado a la Taxonomía, y mejorar la confiabilidad de la clasificación obtenida con la herramienta.
- Realizar un muestreo probabilístico que nos permita realizar inferencias estadísticas, generalizando los resultados a una población objetivo.



# APÉNDICE A

## Categorización ISO 19119

En este apéndice se muestra la categorización de servicios geográficos propuesta por el estándar ISO 19119. Se detallan las categorías y los ejemplos de servicios comprendidos en cada categoría, junto a una descripción de cada uno de ellos.

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
GEOGRAPHIC HUMAN INTERACTION SERVICES	CATALOGUE VIEWER	Client service that allows a user to interact with a catalogue to locate, browse, and manage metadata about geographic data or geographic services.
	GEOGRAPHIC VIEWER	Client service that allows a user to view one or more feature collections or coverages. This viewer allows a user to interact with map data, e.g., displaying, overlaying and querying.
	GEOGRAPHIC SPREADSHEET VIEWER	Client service that allows a user to interact with multiple data objects and to request calculations similar to an arithmetic spreadsheet but extended to geographic data.
	SERVICE EDITOR	Client service that allows a user to control geographic processing services. Views include understanding a service, composing/scripting service chains, invoking a service, statusing a service, scheduling

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		services for peak performance times, and invoking a service chain
	CHAIN DEFINITION EDITOR	Provides user interaction with a chain definition service
	WORKFLOW ENACTMENT MANAGER	Provides user interaction with a workflow enactment service
	GEOGRAPHIC FEATURE EDITOR	Geographic viewer that allows a user to interact with feature data, e.g., displaying, querying. Supports feature annotation. The user controls view orientation, perspective, depth cueing, hiddenline/surface, light-sources, transparency, and texture mapping onto the objects. Objects in view can be picked or drawn onto to generate new objects in the model.
	GEOGRAPHIC SYMBOL EDITOR	Client service that allows a human to select and manage symbol libraries
	FEATURE GENERALIZATION EDITOR	Client service that allows a user to modify the cartographic characteristics of a feature or feature collection by simplifying its visualization, while maintaining its salient elements – the spatial equivalent of simplification
	GEOGRAPHIC DATA-STRUCTURE VIEWER	Client service that allows a user to access part of dataset to see



CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		its internal structure. Allows user to request creation of new objects from parts of an object being browsed. Allow user to request a check of an object, e.g., type checking
<p style="text-align: center;"><b>GEOGRAPHIC MODEL/INFORMATION MANAGEMENT SERVICES</b></p>	FEATURE ACCESS SERVICE	Service that provides a client access to and management of a feature store. An access service may include a query that filters the data returned to the client
	MAP ACCESS SERVICE	Service that provides a client access to a geographic graphics, i.e., pictures of geographic data
	COVERAGE ACCESS SERVICE	Service that provides a client access to and management of a coverage store. An access service may include a query that filters the data returned to the client
	SENSOR DESCRIPTION SERVICE	Service that provides the description of a coverage sensor, including sensor location and orientation, as well as the sensor's geometric, dynamic, and radiometric characteristics for geoprocessing purposes
	PRODUCT ACCESS SERVICE	Service that provides access to and management of a geographic product store. A product can be a predefined feature collection and metadata

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		with known boundaries and content, corresponding to a paper map or report. A product can alternately be a previously defined set of coverages with associated metadata
	FEATURE TYPE SERVICE	Service that provides a client to access to and management of a store of feature type definitions
	CATALOGUE SERVICE	Service that provides discovery and management services on a store of metadata about instances. The metadata may be for dataset instances, e.g., dataset catalogue, or may contain service metadata, e.g., service catalogue
	REGISTRY SERVICE	Service that provides access to store of metadata about types. Types are vocabularies that can be organized and related to each other. Example registries are information community registries, type dictionaries, service registries and schema registries
	GAZETTEER SERVICE	Service that provides access to a directory of instances of a class or classes of real-world phenomena containing some information regarding position
	ORDER HANDLING SERVICE	Service that provides a client with the ability to order

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		products from a provider including: formulation of quotes on orders, selection of geographic processing options, submission of an order, statusing of orders, and billing and accounting of users' orders
	STANDING ORDER SERVICE	Order handling service that allows a user to request that a product over a geographic area be disseminated when it becomes available. Such dissemination includes receive, prepare (i.e., reformat, compress, decompress, etc.), prioritize, and transmit geographic information requested through standing queries or profiles
WORKFLOW / TASK MANAGEMENT SERVICES	CHAIN DEFINITION SERVICE	Service to define a chain and to enable it to be executed by the workflow enactment service. This includes information about its starting and completion conditions, constituent activities and rules for navigating between them, user tasks to be undertaken, references to applications which may be invoked, definition of any workflow relevant data which may need to be referenced, etc. Chain definition service may

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		also provide a chain validation service
	WORKFLOW ENACTMENT SERVICE	The workflow enactment service interprets a chain and controls the instantiation of services and sequencing of activities. This is done through one or more co-operating workflow management engines, which manages the execution of individual instances of the various services. A workflow enactment service maintains control data either centralised or distributed across a set of workflow engines. Workflow control data includes the internal state information associated with the various services under execution and may also include check-pointing and recovery/restart information used by the workflow engines to co-ordinate and recover from failure conditions
	SUBSCRIPTION SERVICE	Service to allow clients to register for notification about events. Events are defined by a service that performs an activity resulting in the event. Events are catalogued by the subscription service. Clients

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		<p>identify events of interest, e.g. receipt of data with a specific geographic extent. When an event occurs, the subscription service sends notification to all clients who have registered an interest in the event. Once an event occurs a subscription service may cause an activity to occur, e.g., delivery of a product</p>
<p>SPATIAL PROCESSING SERVICES</p>	<p>COORDINATE CONVERSION SERVICE</p>	<p>Service to change coordinates from one coordinate system to another coordinate system that is related to the same datum. In a coordinate conversion the parameters' values are exact. Coordinate conversion services include map projection services. ISO 19111 is relevant to coordinate conversion</p>
	<p>COORDINATE TRANSFORMATION SERVICE</p>	<p>Service to change coordinates from a coordinate reference system based on one datum to a coordinate reference system based on a second datum. A coordinate transformation differs from a coordinate conversion in that the coordinate transformation parameter values are derived empirically: therefore there may be several different estimations</p>

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		(or realizations)
	COVERAGE/VECTOR CONVERSION SERVICE	Service to change the spatial representation from a coverage schema to a vector schema, or vice versa
	IMAGE COORDINATE CONVERSION SERVICE	A coordinate transformation or coordinate conversion service to change the coordinate reference system for an image
	RECTIFICATION SERVICE	Service for transforming an image into a perpendicular parallel projection and therefore a constant scale
	ORTHOECTIFICATION SERVICE	A rectification service that removes image tilt and displacement due to terrain elevation. Orthorectification requires use of digital elevation data, usually in grid form
	SENSOR GEOMETRY MODEL ADJUSTMENT SERVICE	Service that adjusts sensor geometry models to improve the match of the image with other images and/or known ground positions
	IMAGE GEOMETRY MODEL CONVERSION SERVICE	Service that converts sensor geometry models into a different but equivalent sensor geometry model
	SUBSETTING SERVICE	Service that extracts data from an input in a continuous spatial region either by geographic location or by grid coordinates
	SAMPLING SERVICE	Service that extracts data from

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		an input using a consistent sampling scheme either by geographic location or by grid coordinates
	TILING CHANGE SERVICE	Service that changes the tiling of geographic data
	DIMENSION MEASUREMENT SERVICE	Service to compute dimensions of objects visible in an image or other geodata. An alternative name for this service is “image mensuration services.”
	FEATURE MANIPULATION SERVICES	Register one feature to another, an image, or another data set or coordinate set; correcting for relative translation shifts, rotational differences, scale differences, and perspective differences. Verify that all features in the Feature Collection are topologically consistent according to the topology rules of the Feature Collection, and identifies and/or corrects any inconsistencies that are discovered
	FEATURE MATCHING SERVICE	Service that determines which features and portions of features represent the same real world entity from multiple data sources, e.g., edge matching and limited conflation
	FEATURE GENERALIZATION SERVICE	Service that reduces spatial variation in a feature collection

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		to increase the effectiveness of communication by counteracting the undesirable effects of data reduction
	ROUTE DETERMINATION SERVICE	Service to determine the optimal path between two specified points based on the input parameters and properties contained in the Feature Collection. May also determine the measured distance between two points along a specified path based on the properties supported in the Feature Collection. Further may determine the length of time it takes to follow a route through the geographic data in the Feature Collection
	POSITIONING SERVICE	Service provided by a position-providing device to use, obtain and unambiguously interpret position information, and determine whether the results meet the requirements of the use
	PROXIMITY ANALYSIS SERVICE	Given a position or geographic feature, finds all objects with a given set of attributes that are located within an user-specified distance of the position or feature
THEMATIC PROCESSING	GEOPARAMETER	Service to derive application-



CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
SERVICES	CALCULATION SERVICE	oriented quantitative results that are not available from the raw data itself
	THEMATIC CLASSIFICATION SERVICE	Service to classify regions of geographic data based on thematic attributes
	FEATURE GENERALIZATION SERVICE	Service that generalizes feature types in a feature collection to increase the effectiveness of communication by counteracting the undesirable effects of data reduction.
	SUBSETTING SERVICE	Service that extracts data from an input based on parameter values
	SPATIAL COUNTING SERVICE	Service that counts geographic features of a given type within a specified area
	CHANGE DETECTION SERVICES	Service to find differences between two data sets that represent the same geographical area at different times
	GEOGRAPHIC INFORMATION EXTRACTION SERVICES	Services supporting the extraction of feature and terrain information from remotely sensed and scanned images
	IMAGE PROCESSING SERVICE	Service to change the values of thematic attributes of an image using a mathematical function. Example functions include: convolution, data compression, feature extraction, frequency filters, geometric operations,

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		non-linear filters, and spatial filters
	REDUCED RESOLUTION GENERATION SERVICE	Service that reduces the resolution of an image
	IMAGE MANIPULATION SERVICES	Services for manipulating data values in images: changing colour and contrast values, applying various filters, manipulating image resolution, noise removal, "striping", systematic-radiometric corrections, atmospheric attenuation, changes in scene illumination, etc
	IMAGE UNDERSTANDING SERVICES	Services that provide automated image change detection, registered image differencing, significance-of-difference analysis and display, and area-based and model-based differencing
	IMAGE SYNTHESIS SERVICES	Services for creating or transforming images using computer-based spatial models, perspective transformations, and manipulations of image characteristics to improve visibility, sharpen resolution, and/or reduce the effects of cloud cover or haze
	MULTI-BAND IMAGE MANIPULATION	Services that modify an image using the multiple bands of the image. Examples include:

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		ratioing; principal components transformation, Intensity-Hue-Saturation colour space transformation, de-correlation-stretching
	OBJECT DETECTION SERVICE	Service to detect real-world objects in an image
	GEOPARSING SERVICE	Service to scan text documents for location-based references, such as a place names, addresses, postal codes, etc., in preparation for passage to a geocoding service
	GEOCODING SERVICE	Service to augment location-based text references with geographic coordinates (or some other spatial reference)
TEMPORAL PROCESSING SERVICES	TEMPORAL REFERENCE SYSTEM TRANSFORMATION SERVICE	Service to change the values of temporal instances from one temporal reference system to another temporal reference system
	SUBSETTING SERVICE	Service that extracts data from an input in a continuous interval based on temporal position values
	SAMPLING SERVICE	Service that extracts data from an input using a consistent sampling scheme based on temporal position values
	TEMPORAL PROXIMITY ANALYSIS SERVICE	Given a temporal interval or event, find all objects with a given set of attributes that are

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		located within an user-specified interval from the interval or event
METADATA PROCESSING SERVICES	STATISTICAL CALCULATION SERVICE	Service to calculate the statistics of a data set, e.g., mean, median, mode, and standard deviation; histogram statistics and histogram calculation; minimum and maximum of an image; multiband cross correlation matrix; spectral statistics; spatial statistics; other statistical calculations
	GEOGRAPHIC ANNOTATION SERVICES	Services to add ancillary information to an image or a feature in a Feature Collection (e.g., by way of a label, a hot link, or an entry of a property for a feature into a database) that augments or provides a more complete description
COMMUNICATION SERVICES	ENCODING SERVICE	Service that provides implementation of an encoding rule and provides an interface to encoding and decoding functionality
	TRANSFER SERVICE	Service that provides implementation of one or more transfer protocols, which allows data transfer between distributed information systems over off-line or on-line

CATEGORÍA	EJEMPLO DE SERVICIO	DESCRIPCIÓN
		communication media. To successfully transfer data between two systems the sender and receiver need to agree on the transfer protocol to be used
	GEOGRAPHIC COMPRESSION SERVICE	Service that converts spatial portions of a feature collection to and from compressed form
	GEOGRAPHIC FORMAT CONVERSION SERVICE.	Service that converts from one geographic data format to another
	MESSAGING SERVICE	Service that allows multiple users to simultaneously view, comment about, and request edits of feature collections. This service allows collaboration involving geographic data
	REMOTE FILE AND EXECUTABLE MANAGEMENT	Service that provides access to secondary storage of geographic features as if it were local to the client

**Tabla A.1.:** Categorización de servicios geográficos del estándar ISO 19119



# APÉNDICE B

## Casos de Uso

En este Apéndice detallamos los casos de uso especificados para satisfacer los requerimientos del Caso de Estudio presentado en el Capítulo 5.

Los casos de uso presentados para esta experiencia, representan un subconjunto de la totalidad de casos de uso especificados para el sistema SIG, los cuales fueron seleccionados teniendo en cuenta que la funcionalidad que proporcionan sea representativa de la totalidad de casos uso especificados.

En la Figura B.1, se muestra el caso de uso denominado “Locate Property”. Su objetivo consiste en proveer un medio al ciudadano para ubicar su propiedad en un mapa mediante la dirección o número de catastro.

<b>Nombre Caso de Uso</b>	<b>LOCATE PROPERTY</b>
<b>Objetivo</b>	Citizen wants to find his property on a map by address or tax account number
<b>Escenario Principal</b>	1 Citizen inserts address
	2 System looks for location on the map
	3 System shows property's location on the map

**Figura B.1:** Caso de Uso “Locate property”

En la Figura B.2, se muestra el caso de uso denominado “Assess Property”. Su finalidad es permitir al ciudadano visualizar el valor de valuación fiscal de su propiedad.

<b>Nombre Caso de Uso</b>	<b>ASSESS PROPERTY</b>	
<b>Objetivo</b>	Citizen wants to see the property's assessment value	
<b>Escenario Principal</b>	1	Citizen selects property on map
	2	System looks for assessment value
	3	System shows assessment value

**Figura B.2:** Caso de Uso “Assess property”

En la Figura B.3, se detalla el caso de uso “Find School District”. Su objetivo es brindar al ciudadano información sobre el distrito escolar en que reside.

<b>Nombre Caso de Uso</b>	<b>FIND SCHOOL DISTRICT</b>	
<b>Objetivo</b>	Citizen wants to know what school district he lives in	
<b>Escenario Principal</b>	1	Citizen selects a location on map
	2	Citizen chooses option “school district” on menu
	3	System shows school's district boundaries
	4	Citizen pans map to see all boundaries

**Figura B.3:** Caso de Uso “Find school district”



En la Figura B.4, se muestra el caso de uso “Find City Facilities”. Su finalidad es que el ciudadano pueda localizar las dependencias de servicios: policía, bomberos, hospitales, así como también plazas, zonas de recreación, etc.

<b>Nombre Caso de Uso</b>	<b>FIND CITY FACILITIES</b>
<b>Objetivo</b>	Citizen wants to know where all the parks, recreation facilities and trails are located around his neighborhood
<b>Escenario Principal</b>	1   Citizen chooses kind of facility from menu
	2   System shows selected facilities on map
	3   Citizen zooms map
	4   Citizen pans map to see facilities out of screen

**Figura B.4:** Caso de Uso “Find city facilities”

En la Figura B.5, se puede ver el caso de uso denominado “Create Landmarks”. El mismo detalla la forma en que un ciudadano puede marcar en un mapa sus puntos de interés.

<b>Nombre Caso de Uso</b>	<b>CREATE LANDMARKS</b>
<b>Objetivo</b>	Citizen wants to mark points of interest on the map, by dropping a pin or selecting a parcel on the map
<b>Escenario Principal</b>	1   Citizen pans map to locate point of interest
	2   Citizen selects “Add a landmark” option
	3   Citizen selects point of interest
	4   System shows map with new landmark added

**Figura B.5:** Caso de Uso “Create landmarks”

En la Figura B.6, se muestra el caso de uso “Find Schools”. El objetivo del mismo es posibilitar al ciudadano la búsqueda de instituciones educativas más cercanas a su lugar de residencia.

<b>Nombre Caso de Uso</b>	<b>FIND SCHOOLS</b>
<b>Objetivo</b>	Citizen wants to know where the nearest elementary/middle/high school is located
<b>Escenario Principal</b>	1 Citizen selects level of school
	2 Citizen selects location on map
	3 System shows schools in neighborhood

**Figura B.6:** Caso de Uso “Find schools”

En la Figura B.7, se describe el caso de uso “Measure Distance”. La finalidad del mismo es brindar al ciudadano una herramienta de medición de distancia entre dos puntos seleccionados.

<b>Nombre Caso de Uso</b>	<b>MEASURE DISTANCE</b>
<b>Objetivo</b>	Citizen wants to measure physical distance between two points
<b>Escenario Principal</b>	1 Citizen selects “how to get” option
	2 Citizen selects first point of interest on map
	3 Citizen selects second point of interest on map
	4 System shows route between both points with a colored line
	5 System shows distance between both points

**Figura B.7:** Caso de Uso “Measure distance”

En la Figura B.8, se muestra el caso de uso “Create Formatted Maps and Reports”. El objetivo del mismo consiste en imprimir mapas o reportes generados por un usuario.

<b>Nombre Caso de Uso</b>	<b>CREATE FORMATTED MAPS AND REPORTS</b>	
<b>Objetivo</b>	Citizen wants to print or email formatted maps and reports	
<b>Escenario Principal</b>	1	Citizen customizes map
	2	Citizen saves updated map
	3	Citizen copies map’s address
	4	Citizen selects “print” option
	5	System prints formatted map

**Figura B.8:** Caso de Uso “Create formatted maps and reports”

En la Figura B.9, se detalla el caso de uso “Turn On/Off Layers”. En el mismo, se especifica la manera en que el usuario puede agregar o quitar información que se visualiza en un mapa.

<b>Nombre Caso de Uso</b>	<b>TURN ON/OFF LAYERS</b>	
<b>Objetivo</b>	User wants to be able to turn on/off layers, to see only what he wants to see	
<b>Escenario Principal</b>	1	User selects a layer on the menu
	2	User toggles layer on/off
	3	System shows map

**Figura B.9:** Caso de Uso “Turn on/off layers”

En la Figura B.10, se muestra el caso de uso “Run Statistics”. Su objetivo consiste en proveer al personal de la administración pública funcionalidad para la obtención de mediciones estadísticas.

<b>Nombre Caso de Uso</b>	<b>RUN STATISTICS</b>	
<b>Objetivo</b>	Government employee needs to run web statistics reports, to see what part of the application is most frequently used	
<b>Escenario Principal</b>	1	Government employee selects statistic report from menu
	2	Government employee runs statistics
	3	System generates statistics

**Figura B.10:** “Run statistics”

En estos casos de uso se puede observar que la funcionalidad especificada consiste principalmente en consultas de información que cualquier ciudadano puede realizar sobre sus propiedades, tal como la valuación fiscal de las mismas o su localización en un mapa, como en los casos de uso “Locate Property” y “Assess Property” o también sobre distintos servicios públicos disponibles en una zona en particular, como en el caso de uso “Find City Facilities”. Se presentan además requerimientos de empleados que actualizan la información del sistema y evalúan la utilización del mismo, como por ejemplo el caso de uso “Run Statistics”.

# APÉNDICE C

## Resultados del Procesamiento de ReqGIS

En este apéndice se detallan los resultados obtenidos de realizar el procesamiento de los casos de uso especificados en el Apéndice B mediante la herramienta ReqGIS.

En la Tabla C.1 se enumera cada caso de uso y las categorías de servicios geográficos correspondientes a cada caso de uso. En dicha Tabla se observa que cada fila corresponde a uno de los casos de uso de la especificación de requerimientos presentados anteriormente en el Apéndice B. Cada fila está compuesta por los pasos que componen el Escenario Principal de ese caso de uso. En la columna *Categoría de Servicios Geográficos*, se detalla la categoría de servicios más relacionada semánticamente con cada paso del caso de uso especificándose el valor de relación correspondiente en la columna *Valor de Relación Semántica*. Las columnas *Valor Seleccionado* y *Categoría Seleccionada* contienen el nombre y valor de relación semántica de aquella categoría de servicio geográfico que haya obtenido el valor mayor en ese caso de uso. Si una categoría tiene más de una ocurrencia por caso de uso, entonces el valor que la representa es el valor promedio de los valores de relación semántica que obtuvo esa categoría en ese caso de uso.

Así, al procesar este conjunto de casos de uso, la herramienta ReqGIS nos propone realizar la búsqueda de componentes que brinden servicios geográficos categorizados como: *Human Interaction*, *Spatial* y *Metadata*.

NOMBRE CASO DE USO	CATEGORÍA DE SERVICIOS GEOGRÁFICOS	VALOR DE RELACIÓN SEMÁNTICA	VALOR SELECCIONADO	CATEGORÍA SELECCIONADA
LOCATE PROPERTY	Communication	53	<b>60</b>	<b><i>Human Interaction</i></b>
	Human Interaction	60		
	Spatial	58		
ASSESS PROPERTY	Spatial	76	<b>76</b>	<b><i>Spatial</i></b>
	Human Interaction	53		
	Human Interaction	46		
FIND SCHOOL DISTRICT	Spatial	69	<b>173,5</b>	<b><i>Human Interaction</i></b>
	Metadata	71		
	Human Interaction	55		
	Human Interaction	292		
FIND CITY FACILITIES	Metadata	69	<b>297</b>	<b><i>Human Interaction</i></b>
	Human Interaction	308		
	Human Interaction	291		
	Human Interaction	292		
CREATE LANDMARKS	Human Interaction	308	<b>308</b>	<b><i>Human Interaction</i></b>
	Metadata	65		
	Spatial	75		
	Human Interaction	308		
FIND SCHOOLS	Spatial	54	<b>69</b>	<b><i>Spatial</i></b>
	Spatial	69		
	Human Interaction	49		
MEASURE DISTANCE	Metadata	65	<b>157,33</b>	<b><i>Spatial</i></b>
	Spatial	75		
	Metadata	65		
	Spatial	217		
	Spatial	180		
CREATE FORMATTED MAPS AND REPORTS	Human Interaction	296	<b>296</b>	<b><i>Human Interaction</i></b>
	Human Interaction	295		
	Communication	53		
	Metadata	65		
	Human Interaction	297		
TURN ON/OFF LAYERS	Metadata	66	<b>308</b>	<b><i>Human Interaction</i></b>
	Thematic	40		
	Human Interaction	308		
RUN STATISTICS	Metadata	71	<b>212,33</b>	<b><i>Metadata</i></b>
	Metadata	298		
	Metadata	268		

**Tabla C.1:** Resultados de la estandarización de los casos de uso

# APÉNDICE D

## Respuestas a Encuesta

En este apéndice se muestran los resultados obtenidos en la encuesta (descrita en Capítulo 5) respondida por un grupo de desarrolladores de sistemas SIG de la zona.

***Cuestionario sobre Categorización de Requerimientos***

1. ¿Conocía Ud. la clasificación de servicios geográficos propuesta por el estándar ISO 19119?  
SI  NO

2. ¿Qué categoría/s considera que satisfacen los requerimientos especificados en el caso de uso? (Anotar en la línea punteada todos los números de categoría que correspondan)

<u>CATEGORÍAS DE SERVICIOS GEOGRÁFICOS</u>	
1. Interacción Humana	5. Procesamiento Temporal
2. Gestión de Modelo/Información	6. Procesamiento Metadatos
3. Procesamiento Espacial	7. Comunicación
4. Procesamiento Temático	8. Ninguna

CASO DE USO	CATEGORÍA DE SERVICIO CORRESPONDIENTE
Locate Property	1
Assess Property	1
Find School District	1
Find City Facilities	1
Create Landmarks	2
Find Schools	3
Measure Distance	3
Create Formatted Maps and Reports	2
Turn On/Off Layers	1
Run Statistics	6

Figura D.1.: Cuestionario 1

## Cuestionario sobre Categorización de Requerimientos

1. ¿Conocía Ud. la clasificación de servicios geográficos propuesta por el estándar ISO 19119?

SI

NO

2. ¿Qué categoría/s considera que satisfacen los requerimientos especificados en el caso de uso? (Anotar en la línea punteada todos los números de categoría que correspondan)

### CATEGORÍAS DE SERVICIOS GEOGRÁFICOS

**1. Interacción Humana**

**5. Procesamiento Temporal**

**2. Gestión de Modelo/Información**

**6. Procesamiento Metadatos**

**3. Procesamiento Espacial**

**7. Comunicación**

**4. Procesamiento Temático**

**8. Ninguna**

CASO DE USO	CATEGORÍA DE SERVICIO CORRESPONDIENTE
Locate Property	<b>1</b>
Assess Property	<b>1</b>
Find School District	<b>4</b>
Find City Facilities	<b>1</b>
Create Landmarks	<b>1</b>
Find Schools	<b>3</b>
Measure Distance	<b>3</b>
Create Formatted Maps and Reports	<b>1</b>
Turn On/Off Layers	<b>1</b>
Run Statistics	<b>6</b>

**Figura D.2.:** Cuestionario 2



## Cuestionario sobre Categorización de Requerimientos

1. ¿Conocía Ud. la clasificación de servicios geográficos propuesta por el estándar ISO 19119?

SI

NO

2. ¿Qué categoría/s considera que satisfacen los requerimientos especificados en el caso de uso? *(Anotar en la línea punteada todos los números de categoría que correspondan)*

**CATEGORÍAS DE SERVICIOS GEOGRÁFICOS**

- |   |                                   |
|---|-----------------------------------|
| <b>1. Interacción Humana</b>            | <b>5. Procesamiento Temporal</b>  |
| <b>2. Gestión de Modelo/Información</b> | <b>6. Procesamiento Metadatos</b> |
| <b>3. Procesamiento Espacial</b>        | <b>7. Comunicación</b>            |
| <b>4. Procesamiento Temático</b>        | <b>8. Ninguna</b>                 |

CASO DE USO	CATEGORÍA DE SERVICIO CORRESPONDIENTE
Locate Property	<b>1</b>
Assess Property	<b>1</b>
Find School District	<b>1</b>
Find City Facilities	<b>1</b>
Create Landmarks	<b>1</b>
Find Schools	<b>3</b>
Measure Distance	<b>3</b>
Create Formatted Maps and Reports	<b>8</b>
Turn On/Off Layers	<b>1</b>
Run Statistics	<b>6</b>

**Figura D.3.:** Cuestionario 3

## Cuestionario sobre Categorización de Requerimientos

1. ¿Conocía Ud. la clasificación de servicios geográficos propuesta por el estándar ISO 19119?  
 SI  NO

2. ¿Qué categoría/s considera que satisfacen los requerimientos especificados en el caso de uso? (Anotar en la línea punteada todos los números de categoría que correspondan)

**CATEGORÍAS DE SERVICIOS GEOGRÁFICOS**

<b>1. Interacción Humana</b>	<b>5. Procesamiento Temporal</b>
<b>2. Gestión de Modelo/Información</b>	<b>6. Procesamiento Metadatos</b>
<b>3. Procesamiento Espacial</b>	<b>7. Comunicación</b>
<b>4. Procesamiento Temático</b>	<b>8. Ninguna</b>

CASO DE USO	CATEGORÍA DE SERVICIO CORRESPONDIENTE
Locate Property	<b>1</b>
Assess Property	<b>1</b>
Find School District	<b>4</b>
Find City Facilities	<b>1</b>
Create Landmarks	<b>1</b>
Find Schools	<b>3</b>
Measure Distance	<b>3</b>
Create Formatted Maps and Reports	<b>8</b>
Turn On/Off Layers	<b>1</b>
Run Statistics	<b>6</b>

**Figura D.4.:** Cuestionario 4

## Cuestionario sobre Categorización de Requerimientos

1. ¿Conocía Ud. la clasificación de servicios geográficos propuesta por el estándar ISO 19119?  
 SI  NO

2. ¿Qué categoría/s considera que satisfacen los requerimientos especificados en el caso de uso? (Anotar en la línea punteada todos los números de categoría que correspondan)

**CATEGORÍAS DE SERVICIOS GEOGRÁFICOS**

<b>1. Interacción Humana</b>	<b>5. Procesamiento Temporal</b>
<b>2. Gestión de Modelo/Información</b>	<b>6. Procesamiento Metadatos</b>
<b>3. Procesamiento Espacial</b>	<b>7. Comunicación</b>
<b>4. Procesamiento Temático</b>	<b>8. Ninguna</b>

CASO DE USO	CATEGORÍA DE SERVICIO CORRESPONDIENTE
Locate Property	1
Assess Property	1
Find School District	1
Find City Facilities	1
Create Landmarks	2
Find Schools	3
Measure Distance	3
Create Formatted Maps and Reports	1
Turn On/Off Layers	1
Run Statistics	6

**Figura D.5.:** Cuestionario 5

## Cuestionario sobre Categorización de Requerimientos

1. ¿Conocía Ud. la clasificación de servicios geográficos propuesta por el estándar ISO 19119?

SI

NO

2. ¿Qué categoría/s considera que satisfacen los requerimientos especificados en el caso de uso? (Anotar en la línea punteada todos los números de categoría que correspondan)

### CATEGORÍAS DE SERVICIOS GEOGRÁFICOS

**1. Interacción Humana**

**5. Procesamiento Temporal**

**2. Gestión de Modelo/Información**

**6. Procesamiento Metadatos**

**3. Procesamiento Espacial**

**7. Comunicación**

**4. Procesamiento Temático**

**8. Ninguna**

CASO DE USO	CATEGORÍA DE SERVICIO CORRESPONDIENTE
Locate Property	<b>1</b>
Assess Property	<b>1</b>
Find School District	<b>1 – 4</b>
Find City Facilities	<b>1</b>
Create Landmarks	<b>1 – 2</b>
Find Schools	<b>3</b>
Measure Distance	<b>3</b>
Create Formatted Maps and Reports	<b>8 – 2 – 1</b>
Turn On/Off Layers	<b>1</b>
Run Statistics	<b>6</b>

**Figura D.6.:** Cuestionario 6

## Cuestionario sobre Categorización de Requerimientos

1. ¿Conocía Ud. la clasificación de servicios geográficos propuesta por el estándar ISO 19119?  
 SI  NO

2. ¿Qué categoría/s considera que satisfacen los requerimientos especificados en el caso de uso? (Anotar en la línea punteada todos los números de categoría que correspondan)

**CATEGORÍAS DE SERVICIOS GEOGRÁFICOS**

<b>1. Interacción Humana</b>	<b>5. Procesamiento Temporal</b>
<b>2. Gestión de Modelo/Información</b>	<b>6. Procesamiento Metadatos</b>
<b>3. Procesamiento Espacial</b>	<b>7. Comunicación</b>
<b>4. Procesamiento Temático</b>	<b>8. Ninguna</b>

CASO DE USO	CATEGORÍA DE SERVICIO CORRESPONDIENTE
Locate Property	1
Assess Property	1
Find School District	1
Find City Facilities	1
Create Landmarks	1
Find Schools	3
Measure Distance	3
Create Formatted Maps and Reports	2
Turn On/Off Layers	1
Run Statistics	6

**Figura D.7.:** Cuestionario 7

## Cuestionario sobre Categorización de Requerimientos

1. ¿Conocía Ud. la clasificación de servicios geográficos propuesta por el estándar ISO 19119?

SI

NO

2. ¿Qué categoría/s considera que satisfacen los requerimientos especificados en el caso de uso? (Anotar en la línea punteada todos los números de categoría que correspondan)

### CATEGORÍAS DE SERVICIOS GEOGRÁFICOS

- |   |                                   |
|---|-----------------------------------|
| <b>1. Interacción Humana</b>            | <b>5. Procesamiento Temporal</b>  |
| <b>2. Gestión de Modelo/Información</b> | <b>6. Procesamiento Metadatos</b> |
| <b>3. Procesamiento Espacial</b>        | <b>7. Comunicación</b>            |
| <b>4. Procesamiento Temático</b>        | <b>8. Ninguna</b>                 |

CASO DE USO	CATEGORÍA DE SERVICIO CORRESPONDIENTE
Locate Property	<b>1</b>
Assess Property	<b>1</b>
Find School District	<b>1 – 4</b>
Find City Facilities	<b>1</b>
Create Landmarks	<b>1 – 2</b>
Find Schools	<b>3</b>
Measure Distance	<b>3</b>
Create Formatted Maps and Reports	<b>1 – 2</b>
Turn On/Off Layers	<b>1</b>
Run Statistics	<b>6</b>

**Figura D.8.:** Cuestionario 8

## Cuestionario sobre Categorización de Requerimientos

1. ¿Conocía Ud. la clasificación de servicios geográficos propuesta por el estándar ISO 19119?  
 SI  NO

2. ¿Qué categoría/s considera que satisfacen los requerimientos especificados en el caso de uso? *(Anotar en la línea punteada todos los números de categoría que correspondan)*

**CATEGORÍAS DE SERVICIOS GEOGRÁFICOS**

<b>1. Interacción Humana</b>	<b>5. Procesamiento Temporal</b>
<b>2. Gestión de Modelo/Información</b>	<b>6. Procesamiento Metadatos</b>
<b>3. Procesamiento Espacial</b>	<b>7. Comunicación</b>
<b>4. Procesamiento Temático</b>	<b>8. Ninguna</b>

CASO DE USO	CATEGORÍA DE SERVICIO CORRESPONDIENTE
Locate Property	1
Assess Property	1
Find School District	4 – 1
Find City Facilities	1
Create Landmarks	2 – 1
Find Schools	3
Measure Distance	3
Create Formatted Maps and Reports	1 – 2
Turn On/Off Layers	1
Run Statistics	6

**Figura D.9.:** Cuestionario 9

## *Cuestionario sobre Categorización de Requerimientos*

1. ¿Conocía Ud. la clasificación de servicios geográficos propuesta por el estándar ISO 19119?

SI  NO

2. ¿Qué categoría/s considera que satisfacen los requerimientos especificados en el caso de uso? *(Anotar en la línea punteada todos los números de categoría que correspondan)*

**CATEGORÍAS DE SERVICIOS GEOGRÁFICOS**

<b>1. Interacción Humana</b>	<b>5. Procesamiento Temporal</b>
<b>2. Gestión de Modelo/Información</b>	<b>6. Procesamiento Metadatos</b>
<b>3. Procesamiento Espacial</b>	<b>7. Comunicación</b>
<b>4. Procesamiento Temático</b>	<b>8. Ninguna</b>

CASO DE USO	CATEGORÍA DE SERVICIO CORRESPONDIENTE
Locate Property	<b>1</b>
Assess Property	<b>1</b>
Find School District	<b>1</b>
Find City Facilities	<b>1</b>
Create Landmarks	<b>2</b>
Find Schools	<b>3</b>
Measure Distance	<b>3</b>
Create Formatted Maps and Reports	<b>1</b>
Turn On/Off Layers	<b>1</b>
Run Statistics	<b>6</b>

**Figura D.10.:** Cuestionario 10



# Bibliografía

- [1] ALBRECHT, J.: Universal GIS Operations for Environmental Modeling, *Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, Santa Fe, NM. Santa Barbara: National Center for Geographical Information and Analysis, 1996.
- [2] ALVES, C.; CASTRO, J.: CRE: A Systematic Method for COTS Components Selection. *XV Brazilian Symposium on Software Engineering (SBES)* Rio de Janeiro, Brazil, October 2001.
- [3] ALVES, C.; FINKELSTEIN, A.: Challenges in COTS decision-making: a goal-driven requirements engineering perspective. *In Proceeding SEKE '02 Proceedings of the 14th international conference on Software engineering and knowledge engineering* ACM New York, NY, USA ©2002
- [4] ARMOUR, F., MILLER, G.: *Advanced Use Case Modeling Volume One, Software Systems*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.
- [5] AVISON D.: Information systems development: A broader perspective. *In Proceedings of the IFIP TC8 Working Conference on Method Engineering: Principles of method construction and tool support*, R. W. S. Brinkkemper, K. Lyytinen, Ed.:(Chapman-Hall, London, 1996).
- [6] AYALA, C.: *Systematic Construction Of Goal-Oriented COTS Taxonomies – Doctoral Thesis - Universitat Politècnica de Catalunya – 2008*
- [7] BANERJEE, S., PEDERSEN, T.: An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet. In: Gelbukh, A. (ed.) *Computational Linguistics and Intelligent Text Processing*. LNCS, vol. 2276, pp. 117--171. Springer, Heidelberg, 2002.
- [8] BASKERVILLE R., WOOD-HARPER T.: A critical perspective on action research as a method for information systems research. *Journal of Information Technology* 11, 3, 1996.
- [9] BITTNER, K.; SPENCE, I.: *Use case modeling*. Addison-Wesley Professional, ISBN: 9780201709131, 2003

- [10] BURROUGH, P.; MCDONNELL, R.: Principles of Geographical Information Systems. Oxford University Press, 1998
- [11] BROWN, A.: Building Systems from Pieces: Principles and Practice of Component-Based Software Engineering, In Constructing Superior Software, Macmillan Technical Publishing. ISBN: 15-787-01473, 1999.
- [12] CECHICH A., RÉQUILÉ A., AGUIRRE J., LUZURIAGA J.: Trends on COTS Component Identification. *5th International Conference on COTS-Based Software Systems*. Orlando, USA. IEEE Computer Science Press, 2006.
- [13] CHUNG, L., COOPER, K.: Defining Goals in a COTS-Aware Requirements Engineering Approach. *System Engineering*, Vol. 7, No.1, 2004.
- [14] COCKBURN, A.: Writing Effective Use Cases, Addison-Wesley Pub Co, ISBN: 0201702258, 1ra ed., Enero 2000.
- [15] DEAN M., SCHREIBER G., BECHHOFFER S., VAN HARMELEN F., HENDLER J., HORROCKS I., MCGUINNESS D., PATEL-SCHNEIDER P., AND STEIN L.: OWL Web Ontology Language Reference - W3C Recommendation 10 February 2004. <http://www.w3.org/TR/owl-ref/>.
- [16] DIAZ B.G. Y CALVIÑO P. (Compiladores): El Proyecto SIT Santa Cruz, *Jornadas Regionales de Información Geográfica y Ordenamiento Territorial*, © Ministerio Secretaría General de la Gobernación, Proyecto SIT SantaCruz / ISBN 978-987-25302-0-4, pp. 9-19, 2009.
- [17] DRAZAN, J., MENCL, V.: Improved Processing of Textual Use Cases: Deriving Behavior Specifications. *In Proceedings of SOFSEM 2007*. LNCS, vol. 4362. pp. 856-886. Harrachov, Czech Republic, Enero 20-26, 2007.
- [18] ESPRIT/ESSI Project no 21580.: Guidelines for Best Practice in User Interface for GIS, Section 6 "List of key GIS operations". European Commission.
- [19] FELLBAUM, C., editor.: WordNet: An electronic lexical database. MIT Press, 1998.
- [20] GAETÁN, G.: Normalización de Información en Catálogos de Componentes OTS para Sistemas de Información Geográficos. Tesis Maestría, Universidad Nacional de La Plata, 2011.

- [21] GAETÁN, G., BUCCELLA, A., CECHICH A.: Clasificación de Componentes OTS (OFF-THE-SELF) Para Sistemas de Información Geográficos, *Actas del X Workshop de Investigadores en Ciencias de la Computación*, ISBN 978-950-863-101-5, 2008.
- [22] GAETÁN, G., BUCCELLA, A., CECHICH A.: Un Esquema de Clasificación Facetado para Publicación de Catálogos de Componentes SIG, *Actas del XIV Congreso Argentino en Ciencias de la Computación*, ISBN 987-24611-0-2, 2008.
- [23] GAETÁN, G., BUCCELLA, A., CECHICH A., SALDAÑO V.: Clasificación de Componentes SIG, *Actas del XI Workshop de Investigadores en Ciencias de la Computación*, ISBN 978-950-605-570-7, 2009.
- [24] GAETÁN G., CECHICH A., BUCCELLA A.: Aplicación de Técnicas de Procesamiento de Lenguaje Natural y Web Semántica en la Publicación de Componentes para SIG, *Actas del ASSE (Simposio Argentino de Ingeniería de Software)*, JAIIO, ISSN 1850-2792 (110—121), 2009.
- [25] GAETÁN G., CECHICH A., BUCCELLA A.: Extracción de Información a partir de Catálogos Web de Componentes para SIG, *Actas del XV Congreso Argentino en Ciencias de la Computación*, ISBN 978-897-24068-4-1, (891—900), 2009.
- [26] GAETÁN G., SALDAÑO, V., BUCCELLA, A., CECHICH, A.: A Domain-Oriented Approach for GIS Component Selection, *The Fifth International Conference on Software Engineering Advances, ICSEA 2010*, August 22-27, 2010 - Nice, France
- [27] GRAHAM, I.: *Object-Oriented Methods: Principles and Practice*, Addison-Wesley, 2000.
- [28] GRAU, G., CARVALLO, J.P., FRANCH, X., QUER, C.: DesCOTS: A Software System for Selecting COTS Components. *In Proceedings of the 30th EUROMICRO Conference*, IEEE Computer Society, 2004.
- [29] HARMON, J.E.; ANDERSON, S. J.: *The Design and Implementation of Geographic Information Systems*. John Wiley & Sons, 2003.
- [30] HOFFMANN, V., LICHTER, H., NYBEN, A., WALTER, A.: Towards the Integration of UML and Textual Use Case Modeling. *Journal of Object Technology*, Vol. 8, No. 3, 2009.

- [31] IRIBARNE L., TROYA J., VALLECILLO A.: Trading for COTS Components in Open Environments, *In 27th Euromicro Conference 2001: A Net Odyssey (euromicro'01)*, 2001.
- [32] IRIBARNE L.: Un Modelo de Mediación para el Desarrollo de Software Basado en componentes COTS – Tesis Doctoral – Universidad de Almería – 2003
- [33] ISO/IEC 10746-1:1998 Information technology -- Open Distributed Processing -- Reference model: Overview
- [34] JURAFSKY, D., MARTIN, J.: *Speech and Language Processing*, Prentice-Hall, 1999.
- [35] KHOLKAR, D., KRISHNA, G., SHROTRI, U., AND VENKATESH, R.: Visual Specification and Analysis of Use Cases. *In SoftVis '05: Proceedings of the 2005 ACM symposium on Software visualization*, pp 77-85, New York, NY, USA, ACM, 2005.
- [36] KITCHENHAM, B., PFLEEGER, S.: Principles of Survey Research Part 3: Constructing a Survey Instrument. *SIGSOFT Softw. Eng. Notes 27, 2 (March 2002), 20-24.*
- [37] KITCHENHAM, B., PFLEEGER, S.: Principles of Survey Research Part 5: Populations and Samples. *SIGSOFT Softw. Eng. Notes 27, 5 (September 2002), 17-20.*
- [38] KONTIO, J.: OTSO: A Systematic Process for Reusable Software Component Selection – University of Maryland, Maryland, CS-TR-3478, 1995
- [39] KULAK, D., GUINEY, E.: *Use Cases: Requirements in Context*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, ISBN: 0321154983, 2da ed., Agosto 2003.
- [40] LAURINI, R.; THOMPSON, D.: *Fundamentals of Spatial Informations Systems*. The APIC Series N 37, Academic Press, 1992.
- [41] LESK, M.: Automatic Sense Disambiguation Using Machine Readable Dictionaries: How to Tell a Pine Cone From a Ice Cream Cone. *In: 5th ACM International Conference on System Documentation*, pp. 24-26. ACM, Toronto (1986)

- [42] LI W., ZHAO S., SUN H., ZHANG X.: Ontology-Based QoS Driven GIS Grid Service Discovery. *IEEE Proceedings of the Second International Conference on Semantics, Knowledge, and Grid (SKG'06)*, 2006
- [43] LONGLEY, P.; GOODCHILD, M.; MAGUIRE, D.; RHIND, D.: *Geographic Information Systems and Science*. John Wiley & Sons, 2001.
- [44] LUCENA, V.: Flexible Web-based Management of Components for Industrial Automation, Phd thesis. Stuttgart University, 2002.
- [45] LUTZ, M.: Ontology-Based Service Discovery in Spatial Data Infrastructures. *ACM GIR'05*, Bremen, Germany, 2005.
- [46] MAIDEN, NCUBE, C.: PORE: Procurement-Oriented Requirements Engineering Method for the Component-Based Systems Engineering Development Paradigm. *International Workshop on Component-Based Software Engineering*, May 1999.
- [47] MAIDEN, NCUBE, C.: COTS Software Selection: The Need to make Trade-offs Between Systems Requirements, Architectures and COTS/Components. *Proceedings of ICSE-2000*, Limerick, Ireland, 2000.
- [48] MANNING, C.; SHÜTZE, H.: *Foundations of Statistical Natural Language Processing*. MIT Press, 1999. ISBN 0-262-13360-1
- [49] MENCL, V.: Deriving Behavior Specifications from Textual Use Cases. *In: Proceedings of Workshop on Intelligent Technologies for Software Engineering (WITSE04, part of ASE 2004)*, Linz, Austria, Oesterreichische Computer Gesellschaft, Sept. 2004.
- [50] MEYERS, C.; OBERNDORF, P.: *Managing Software Acquisition*, Addison-Wesley, 2001.
- [51] MOHAMED, A., RUHE, G., EBERLEIN, A.: Cots Selection: Past, Present, and Future. *Proceedings of the 14th Annual IEEE International Conference and Workshops on the Engineering of Computer-Based Systems (ECBS'07)*, Arizona, United States, 2007.
- [52] OBERNDORF, P.: Facilitating Component-Based Software Engineering: COTS and Open Systems. *Proceedings of the Fifth International Symposium on Assessment of Software Tools- SAST.97*, 1997.

- [53] OBERNDORF, P.; BROWNSWORD, L.: Are You Ready for COTS?, Software Institute Engineering. August 1997.
- [54] OBERNDORF, P.; BROWNSWORD, L.; MORRIS, E.; SLEDGE, C.: Special Report CMU/SEI-97-SR-019, *Workshop on COTS-Based Systems – SEI Institute*, CMU, Nov. 1997.
- [55] O'DEA D., GEOGHEGAN S., EKINS C.: Dealing with Geospatial Information in the Semantic Web. Australasian Ontology Workshop (AOW 2005), Sydney, Australia, 2005.
- [56] OGC. OpenGIS Web Services Architecture Description. Best Practices Paper. 2005.
- [57] OGC. Topic 12: OpenGIS Service Architecture. Open GIS Consortium, 2002.
- [58] OMG. UML Superstructure Specification, v2.1.2. OMG Formal Document 2007-11-02, Noviembre 2007.
- [59] OWL-S Coalition. OWL-S: Semantic Markup for Web Services, 2004. <http://www.w3.org/Submission/OWL-S/>.
- [60] PADRÓ, L.; COLLADO, M.; REESE, S.; LLOBERES, M.; CASTELLÓN, I.: FreeLing 2.1: Five Years of Open-Source Language Processing Tools. *In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, 2010.
- [61] PEDERSEN, T., PATWARDHAN, S., MICHELIZZI, J.: WordNet::Similarity – Measuring the Relatedness of Concepts. Demonstration Papers at HLT-NAACL 2004, pp. 38—41. Boston, Massachusetts, 2004.
- [62] PERCIVALL G.: ISO 19119 and OGC Service Architecture. FIG XXII International Congress, Washington, D.C. USA, 2002.
- [63] PETRITSCH, H.: Service-Oriented Architecture (SOA) vs. Component Based Architecture. Vienna University of Technology white paper, available at <http://whitepapers.techrepublic.com/abstract.aspx>, 2006..
- [64] PLASIL, F., MENCL, V.: Getting “Whole Picture” Behavior in a Use Case Model. *Transactions of SDPS: Journal of Integrated Design and Process Science* 7(4), pag.63-79, Dic. 2003.

- [65] RIGAUX, P.; SCHOLL, M.; Voisard, A.: Spatial Databases With Application To GIS. Academic Press, 2001.
- [66] RODRÍGUEZ LUACES, M.A.: A Generic Architecture for Geographic Information Systems. PhD thesis, Univerdade da Coruña, 2004.
- [67] SAATY, T.L.: The Analytic Hierarchy Process. McGraw-Hill. New York. 287 pp., 1980.
- [68] SAATY, T.L.: Decision making for Leaders. The Analitic Herarchy Process for decision in a complex World. University of Pittsburgh. RWS Publications, Pittsburgh, USA. 292 pp., 1990.
- [69] SALDAÑO, V., BUCCELLA, A., CECHICH, A.: Una Taxonomía de Servicios Geográficos para Facilitar la Identificación de Componentes, *Actas del XIV Congreso Argentino en Ciencias de la Computación*, ISBN 987-24611-0-2, 2008
- [70] SALDAÑO, V., BUCCELLA, A., CECHICH, A.: Descubrimiento de Servicios Geográficos a partir de Casos de Uso Textuales. *In: XV Congreso Argentino en Ciencias de la Computación. Jujuy, 2009.*
- [71] SALDAÑO, V., BUCCELLA, A., CECHICH, A.: Discovering Geographic Services From Textual Use Cases, *Journal of Computer Science & Technology*, Vol. 10 - No. 2 – June 2010 - ISSN 1666-6038
- [72] SPIVEY, J.: The Z Notation: A Reference Manual. Prentice Hall, 1992.
- [73] TORCHIANO M., MORISIO M.: Overlooked Facts on COTS-Based Development, *IEEE SOFTWARE*, Vol. 21(2), pp. 88-93, ISSN: 0740-7459, 2004
- [74] WHITTLE, J AND JAYARAMAN, P.: Generating Hierarchical State Machines from Use Case Charts. *In RE '06: Proceedings of the 14th IEEE International Requirements Engineering Conference (RE'06)*, pp 16-25. Washington, DC, USA, IEEE Computer Society, 2006.
- [75] WOOD-HARPER T.: Research methods in information systems: Using action research. *In Research Methods in Information Systems (North-Holland, 1985)*, E. Mumford, R. Hirschheim, G. Fitzgerald, and A. Wood-Harper, Eds., Elsevier Science Publishers B.V.

- [76] WORBOYS, M. F.: GIS: A Computing Perspective. Taylor & Francis, 1995.
- [77] YUE P., DI L., ZHAO P., YANG W., YU G., WEI Y.: Semantic Augmentations for Geospatial Catalogue Service. *IEEE International Conference on Geoscience and Remote Sensing Symposium, IGARSS 2006*, 2006.