

REPÚBLICA BOLIVARIANA DE VENEZUELA  
UNIVERSIDAD DR. RAFAEL BELLOSO CHACÍN  
VICERRECTORADO DE INVESTIGACIÓN Y POSTGRADO  
DECANATO DE INVESTIGACIÓN Y POSTGRADO  
PROGRAMA: DOCTORADO EN CIENCIAS MENCIÓN GERENCIA



## **WEB SEMÁNTICA EN LA CONSULTA DE DATOS GEORREFERENCIADOS**

Trabajo presentado como requisito para optar al grado de  
Doctor en Ciencias, Mención: Gerencia

Autora:  
MgSc. Nelcy Piña  
C.I.10.413.862  
Tutor:  
Dr. Jacinto Dávila

Maracaibo, Febrero de 2013

**WEB SEMÁNTICA EN LA CONSULTA DE DATOS  
GEORREFERENCIADOS**

## VEREDICTO

## DEDICATORIA

A mi familia.

## **AGRADECIMIENTOS**

A mis amigos Jacinto Dávila y Nelson Arapé

A mi equipo GRYN.

## ÍNDICE

VEREDICTO.....	iii
DEDICATORIA.....	iv
AGRADECIMIENTOS.....	v
ÍNDICE.....	vi
RESUMEN.....	viii
ABSTRACT.....	ix
INTRODUCCIÓN.....	1
CAPÍTULO I	
EL PROBLEMA.....	3
1. Planteamiento del Problema.....	3
1.1. Formulación del Problema.....	8
2. Objetivos de la Investigación.....	9
2.1. Objetivo General.....	9
2.2. Objetivos Específicos.....	9
3. Justificación de la Investigación.....	10
4. Delimitación de la Investigación.....	11
CAPÍTULO II	
MARCO TEÓRICO.....	13
1. Antecedentes.....	13
2. Bases Teóricas.....	26
2.1. Web Semántica.....	26
2.1.1. Arquitectura de la Web.....	29
2.1.1.1. Protocolo.....	30
2.1.1.2. Servicio.....	32
2.1.1.3. Integración.....	33
2.1.2. Tipos de Consultas.....	35
2.1.2.1. Estructuradas.....	36
2.1.2.2. No estructuradas.....	37
2.1.2.3. Espaciales.....	38
2.1.3. Servicios Web.....	39
2.1.3.1. Configuración.....	41
2.1.3.2. Especificación.....	42
2.1.3.3. Inteligencia artificial.....	43
2.1.3.4. Servicios asociados.....	44
3. Sistema de Variables.....	45

3.1. Definición nominal.....	45
3.2. Definición conceptual.....	45
3.3. Definición operacional.....	46
3.4. Abreviaturas y Acrónimos.....	48
<b>CAPÍTULO III</b>	
<b>MARCO METODOLÓGICO.....</b>	<b>50</b>
1.Enfoque Paradigmático.....	50
2.Tipo de Investigación.....	52
3.Diseño de investigación.....	52
4.Población y Muestra.....	53
4.1. Población.....	53
4.2. Muestra.....	54
5. Técnicas e Instrumentos de Recolección de Datos.....	56
5.1. Instrumento.....	56
5.2. Validez.....	57
6. Análisis de Datos.....	58
7. Procedimiento de la Investigación.....	59
<b>CAPÍTULO IV</b>	
<b>ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS.....</b>	<b>61</b>
1.Análisis y Discusión de los Resultados.....	61
2.Lineamientos teórico – prácticos de servicios Web semánticos de consulta de datos georreferenciados.....	75
<b>CONCLUSIONES.....</b>	<b>80</b>
<b>RECOMENDACIONES.....</b>	<b>84</b>
<b>REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>86</b>
<b>ANEXOS.....</b>	<b>91</b>
1 FORMATO PARA LA VALIDEZ DE CONTENIDO DE LOS INSTRUMENTOS PERTENECIENTES A LA INVESTIGACIÓN.....	92
2 RESULTADOS DE LA APLICACIÓN DE LA ENCUESTA.....	100

### **LISTA DE CUADROS**

1 Operacionalización de la Variable.....	47
2 Distribución de la Muestra.....	55

Piña Rivera, Nelcy Patricia. "Web Semántica en la consulta de datos georreferenciados". Universidad Dr. Rafael Beloso Chacín. Programa de Doctorado en Ciencias Mención Gerencia. Maracaibo, 2013.

## **RESUMEN**

Esta investigación presenta un estudio del problema de consulta de datos georreferenciados mediante las tecnologías de servicios Web en la Web actual, y su proceso de transformación a una Web Semántica o Web de significados, en la que agentes de software sean capaces de interpretar las solicitudes de los usuarios, asistirlos en sus búsquedas; y la información se almacene en bases de conocimiento llamadas ontologías. Se trazaron como objetivos analizar la Web actual, su arquitectura, sus tipos de consulta, y sus servicios Web, así como formular lineamientos teórico prácticos en la Web semántica de la consulta de datos georreferenciados, de modo que sirva como referencia a otras investigaciones relacionadas. Para lograr estos objetivos se realizó una investigación de tipo descriptiva, con un diseño de investigación ex post-facto o no experimental, transversal o sincrónica, de campo. Para la recolección de datos se utilizó la observación directa simple y documental, en una muestra intencionada o sesgada de 26 servicios web de consulta a datos georreferenciados, a la cual se aplicó una encuesta dicotómica. Se analizaron los resultados obtenidos mediante estadística descriptiva simple, mostrando las tendencias en orden decreciente de la muestra objeto de estudio. Se presentaron conclusiones y recomendaciones.

Palabras Clave: Web semántica, servicios web de consulta, datos georreferenciados, ontologías.



Piña Rivera, Nelcy Patricia. "Semantic Web in georeferenced data query". Universidad Dr. Rafael Beloso Chacín. Programa de Doctorado en Ciencias Mención Gerencia. Maracaibo, 2013.

## **ABSTRACT**

This research presents a study of the problem of georeferenced data query using Web services technologies in the current Web, and its transformation to a Semantic Web or Web of meaning, in which software agents are able to interpret users queries, assisting them in their search; and information would be stored in knowledge bases called ontologies. Objectives were drawn to analyze the current Web, its architecture, its query types, and Web services, moreover to formulate practical and theoretical guidelines on the Web semantics of georeferenced data query, so that it serves as a reference to related research. To achieve these objectives, a descriptive research was made with an ex post facto or nonexperimental research design type, synchronous, of field. To collect data, it was used a simple direct and documentary observation, in a purposive sample of 26 or biased services of georeferenced data query, which were surveyed with dichotomic survey. Results were analyzed using descriptive statistics simple, showing trends in decreasing order of the sample under study. Conclusions and recommendations were presented

Key words: Semantic Web, query web services, georeferenced data, ontologies.

## INTRODUCCIÓN

“La Web es una obra en marcha” Berners-Lee (1999: p.116). Así la define su creador y director del consorcio World Wide Web, Tim Berners-Lee. Una obra en la que se tiene el conocimiento distribuido y a disposición de la gente, al alcance de todos sus usuarios humanos. Es además un reto para quienes contribuyen en su formación y crecimiento descentralizado, habilitar este conocimiento para que pueda estar al alcance de agentes de software que puedan interoperar entre si y asistir a los usuarios en la búsqueda y acceso a la información. Alcanzar un estadio como el descrito, es el que se denomina Web semántica. Es un reto de esta generación, que se construye poco a poco a través del desarrollo de aplicaciones con esta visión de presente y futuro.

La problemática relacionada en el desarrollo de este tipo de aplicaciones, se estudia desde la perspectiva de la Web semántica, en función de los fundamentos técnicos que permiten su construcción y de sus limitaciones. Para ello la presente investigación se ocupa de la revisión, contrastación de teorías y antecedentes, en relación con esta variable, dimensiones, e indicadores, así como de los servicios web que en la

actualidad ofrecen a sus usuarios la posibilidad de consultar datos georreferenciados. Para cumplir con los objetivos planteados, este documento se estructura en cuatro capítulos, constituyendo el Capítulo I El Problema, en el que se presentan las limitaciones de las tecnologías actuales de la Web, utilizadas para proveer a los usuario de servicios Web, la posibilidad de consultar información georreferenciada y el reto que constituye transformar esta Web que conocemos en Web Semántica; en el Capítulo II, las Bases Teóricas, la investigadora realiza una revisión de los autores y sus teorías, así como de los antecedentes más recientes. En el Capítulo III, del Marco Metodológico, se formulan las características, tipo, diseño de la investigación, así como técnicas de recolección y análisis de datos, de la muestra objeto de estudio.

Finalmente en el Capítulo IV, se realiza el Análisis y Revisión de los Resultados obtenidos, formulando lineamientos teórico prácticos de los servicios web semánticos en la consulta de datos georreferenciados. Así mismo se presentan las conclusiones y recomendaciones originadas por la investigación.

## **CAPÍTULO I**

### **EL PROBLEMA**

#### **1. Planteamiento del Problema**

La Web actual es un espacio de información habilitado por computadores y redes que le permiten a la gente comunicarse, luego estos equipos se quitan del medio, según lo afirma Berners-Lee (1999). La ayuda que brindan estas máquinas se ha limitado a poner al alcance de los usuarios gran cantidad de documentos, más que ocuparse de la calidad de información que el usuario necesita. Esta información desplegada se presenta de modo que los seres humanos puedan interpretarla y utilizarla en su beneficio.

En este sentido, esta ingente cantidad de resultados ofrecidos, redundan en pérdida de tiempo del usuario que continúa en su búsqueda extrayendo aquella información útil y desechando el resto. Hoy en días aunque las máquinas han demostrado ser útiles y rápidas en la búsqueda de datos, los documentos que éstas consiguen en la web no están diseñados ni tienen la estructura necesaria para que estos equipos ofrezcan una mejor selección según el tópico de interés del usuario.

Para mejorar tal selección de datos, las estructuras y las interrelaciones de la información son importantes. Estas son propiedades básicas de los Sistemas Manejadores de Bases de Datos Relacionales (SMBDR), pues estos son herramientas que manejan tales estructuras que corresponden a columnas de información que se relacionan unas con otras. Pero de esta manera no está disponible la información de la web actual.

Los documentos de la web apenas comienzan a incorporar *marcas de información* relacionada a la información contenida en ellos, i.e. datos acerca de los datos, en otras palabras, metadatos. Algunos lenguajes como el Hyper Text Markup Language (HTML), les permiten a los ordenadores representar y compartir hipertexto y la posibilidad de navegar de un sitio a otro; sin embargo, es un lenguaje limitado y aún con estas marcas la computadora es incapaz de interpretar los significados de los documentos.

Adicionalmente existen otros metalenguajes de marcado que agregan metadata a documentos como el eXtensible Markup Language (XML). Estas marcas por sí mismas constituyen una forma de metadata mediante la cual se explica (por su nombre) lo que son los elementos que las componen y cómo estos objetos de información están estructurados en unidades coherentes mayores, sin embargo en el XML gobierna formalmente la sintaxis, más no la semántica Cover (1998). Asevera el autor que para un procesador XML, <calle>, <municipio> o <casa> son etiquetas que carecen de significado. De

aquí que las marcas descriptivas de XML tengan una limitada relevancia como mecanismo para habilitar el intercambio de información a nivel de máquina.

El Resource Description Framework (RDF) por su parte, es una tecnología basada en XML que permite expresar significado, el cual lo codifica a través de triplas formadas por una sentencia elemental de sujeto + verbo + objeto, por ejemplo, cosa + tiene propiedades + con ciertos valores; y estos verbos habilitan a cualquier persona a definir nuevos conceptos. Según Berners-Lee (1999), ha habido proyectos para almacenar significados intervinculados en el computador, en un campo que se ha denominado *Representación del Conocimiento*. Sin embargo, el problema consiste en que esos sistemas están diseñados alrededor de una base de datos central que solo tiene espacio para sus definiciones conceptuales, y que no están diseñados para vincularse a otras bases de datos.

Ahora bien, entre instituciones las soluciones computacionales interoperativas, implican la existencia de ontologías compartibles o conjunto común de objetos semánticos. Afirma Levinsonh (2001), que entre los parámetros de interoperabilidad, debe haber interés para interoperar a nivel institucional; se demanda sencillez, transparencia, apertura, efectividad y universalidad; adoptar protocolos estándares de redes bien establecidos; así como estándares de bases de datos; y utilizar herramientas e interfaces estándares para el intercambio de datos; entre otros aspectos.

Estas ontologías son la explicación de algún vocabulario compartido o conceptualización de algún asunto específico, la cual ha nacido para

sobrellevar el problema del conocimiento escondido e implícito, haciendo explícita la conceptualización de un dominio, por ejemplo la Geografía Stuckenschmidt H. et al. (2001). Entonces una ontología es un documento o archivo que define formalmente las relaciones entre términos. Según Guarino (2012), las ontologías desde el punto de vista computacional, son vocabularios de términos que expresan el significado de categorías primitivas, las relaciones que describen su naturaleza y la estructura de un dominio de discurso. Son usadas para desambiguar términos.

El tipo más común de ontología para la Web, es aquel que tiene una taxonomía y un conjunto de reglas de inferencia. Las reglas de inferencia en las ontologías resuelven *confusiones*, pues proveen relaciones de equivalencia Berners-Lee et al. (2001). Ontology Web Language (OWL) es el lenguaje desarrollado para lograr interoperabilidad entre las ontologías. Y con el OWL, unido al RDF Y XML se completan las tres tecnologías hasta ahora disponibles por la web para la representación del conocimiento, de las cuales las dos últimas se han creado precisamente para la interpretación del mismo por humanos y máquinas, y la última para lograr la interoperabilidad.

Básicamente este fue el principio de la Web Semántica, como la define su promotor, es una web o red de datos que pueden directa o indirectamente ser procesados por máquinas y hombres. Un espacio en el que si publicamos los conceptos relacionados de un dominio particular, entonces la Web Semántica en su totalidad conocerá su equivalencia, según Berners-Lee (1999: p.172).

En el caso particular de la data georreferenciada, y específicamente en el dominio de la información territorial, las diferencias de sistemas catastrales no sólo se constituyen en procedimientos legales diferentes; éstas comienzan en un nivel técnico donde la data geográfica acerca de la propiedad de tierras es almacenada en muchos formatos y sistemas diferentes Stuckenschmidt, H. et al. (2001). Las estructuras necesarias para almacenar la información registro-catastral, poseen campos de índices sujetos a las disposiciones legales del país en el cual se inscriben, y a los cambios políticos de estudio.

Según Berners-Lee (1999: p.173), no se pretende crear estándares globales, se busca trabajar con puntos de entendimiento global, regionales y locales. Sin embargo y a pesar de que se logre almacenar y procesar dicha data en repositorios de conocimiento, pueden hallarse pocos servicios web de consultas de datos georreferenciados y específicamente de información territorial, en la plataforma Web Semántica.

Hoy en día, la Web enfrenta un serio reto en la búsqueda, extracción, representación, interpretación y mantenimiento de la información. Según Thuraisingham (2010), para convertir a una Web actual estática en otra dinámica y pueda explotarse todo el potencial de la Web, deben desarrollarse servicios Web, y convertirlos en servicios Web Semánticos. Asegura el autor que las tecnologías actuales permiten el desarrollo de servicios Web, pero que existen deficiencias como información puramente sintáctica en la descripción de la mayoría de los documentos; soporte sintáctico para operaciones de



descubrimiento, descomposición y ejecución; y ausencia de contenidos/servicios con marcado semántico.

De acuerdo con este contexto, donde la Web de contenidos es accesible por usuarios humanos pero carente de semántica, donde los contenidos no tienen una estructura habilitada para la lectura automática de agentes de software, y donde las tecnologías para los servicios web aún se encuentran en un estadio inicial e incipiente, se formuló la siguiente pregunta:

### **1.1. Formulación del Problema**

¿Cómo es la Web Semántica en la consulta de Datos Georreferenciados?

Así mismo, otras preguntas tuvieron lugar para proporcionar en detalle el estado del arte en el que se encuentra la Web, y estas son:

¿Cuál es la arquitectura de la Web semántica en la consulta de datos Georreferenciados?

¿Cómo son los tipos de Consulta de la Web semántica en datos Georreferenciados?

¿Cuáles son los Servicios Web de la Web semántica en la consulta de datos Georreferenciados?

¿Cuáles serían los lineamientos teórico-prácticos de servicios web semánticos de consulta de datos georreferenciados?

## **2. Objetivos de la Investigación**

Para responder a las preguntas antes señaladas, se formularon los propósitos u objetivos, los cuales sirvieron de guía en la búsqueda del conocimiento por parte de la investigadora. Estos son los siguientes:

### **2.1. Objetivo General**

Analizar la web semántica en la consulta de datos georreferenciados.

### **2.2. Objetivos Específicos**

Identificar la arquitectura de la Web semántica en la consulta de datos Georreferenciados.

Describir los tipos de Consulta de la Web semántica en datos Georreferenciados.

Caracterizar los Servicios Web de la Web semántica en la consulta de datos Georreferenciados.

Formular lineamientos teórico-prácticos de servicios web semánticos de consulta de datos georreferenciados.

### **3. Justificación de la Investigación**

Para aquellos quienes construimos o simplemente utilizamos la Web, a quienes nos sirve de medio eficaz de comunicación, de información al instante, de interacción con otros, de repositorio de datos y de conocimiento; que gozamos de la capacidad de visualizar representaciones del lugar en dónde estamos y hacia donde queremos ir; que podemos realizar planificación en función de la información distribuida en diferentes repositorios de distinta índole y, definitivamente de tantos propósitos para la que la Web nos es útil, esta investigación se justifica desde diferentes perspectivas.

Desde el punto de vista teórico, contrasta teorías y se sustenta en principios de la arquitectura de la Web, los tipos de consultas, la inteligencia artificial, las reglas expresadas en lenguajes formales, la semántica de los objetos, en la representación del conocimiento en ontologías, y la definición de Servicios Web.

Así mismo como un valor agregado social, se justifica por el acercamiento de los usuarios beneficiados a la web del significado, pues éstos sin dominio técnico experto previo podrían a través de consultas desde su computador conectado a la internet, informarse de detalles geográficos y tener acceso a información oficial certificada, en algunos casos.

Desde un enfoque práctico, el producto de la investigación constituye un diseño de consulta subyacente, puntual y descentralizado, que sirve de antesala al desarrollo de aplicaciones que contribuyan a la Web Semántica, específicamente en un subconjunto de datos georreferenciados, que se ponen a disposición de los usuarios de la Web desde cualquier punto del planeta.

Se justifica tecnológicamente pues el resultado de la evaluación de servicios Web basados en Web Semántica, permite observar sus tendencias de arquitectura, diseño, tipos de consulta, configuración, especificación, uso de herramientas de inteligencia artificial, estándares, y capacidades ofrecidas a sus usuarios finales.

Finalmente, como justificación metodológica, aporta una guía de referencia en la construcción de Servicios Web basados en Web Semántica, los cuales permiten acceder a datos Georreferenciados de algún dominio de información espacial. Además se convierte en un antecedente en el análisis de unidades de consulta a datos geoespaciales en la Web.

#### **4. Delimitación de la Investigación**

La investigación se encuentra delimitada en el espacio virtual de la Web, también conocida como internet, o red de recursos distribuidos alrededor del planeta. Con esto se señala que no hubo restricción geográfica en los sujetos estudiados.

Según el aspecto temporal, se señala delimitada la investigación dentro del periodo comprendido desde Marzo de 2011 hasta Enero de 2013, tiempo en el cual se revisaron antecedentes, teorías y casos particulares, que contaron con características relevantes, las cuales serán explicadas más adelante.

En cuanto a la delimitación teórica, la base formal del estudio la aportan los autores que sustentan la variable Web Semántica en sus diferentes dimensiones e indicadores son: Arquitectura de la Web Semántica- Pautasso (2008) y Antoniou (2008); para la dimensión Tipos de Consulta, los autores que aportaron teorías fueron Manning (2008) y los estándares de la OGC; y en cuanto a la tercera dimensión correspondiente a los Servicios Web, la autora se apoya en las OGC y Antoniou (2008), así como en una investigación de Piña (2006), la cual se considera fundamental para la presente investigación.

## **CAPÍTULO II**

### **MARCO TEÓRICO**

El marco teórico es una construcción documental de las teorías aportadas por autores, las cuales contribuyen a la formación del estado del arte en la presente investigación. En este capítulo se recopila el conocimiento hasta ahora aceptado por la comunidad científica, se le permite a otros investigadores reconocer la realidad de interés para la que se indaga junto con los caminos recorridos por otros, y vislumbrar lo que falta por recorrer.

#### **1. Antecedentes**

En este apartado se presentarán los aportes teóricos de otras investigaciones en función del cuadro de operacionalización de variables planteado, y se obtendrán referencias en relación a la variable objeto de estudio: la Web Semántica.

Arapé (2012), constituye el antecedente más reciente, en su trabajo de investigación titulado “Arquitectura orientada a servicios basada en software libre para el acceso a datos georreferenciados”, el cual fue presentado como trabajo de ascenso a profesor Asociado de la Universidad del Zulia. En el mismo Arapé propuso un diseño de arquitectura de cinco (5) capas para servicios web los cuales le brindaran soporte a aplicaciones que tuvieran necesidad de acceder a datos georreferenciados. A saber estas cinco capas según el punto de vista lógico, fueron denominadas como capa de datos, capa de acceso a datos, capa de procesamiento, capa de aplicaciones y capa de cliente.

Aunque Arapé (2012) no utiliza un componente razonador, ni una gramática para la interpretación de la consulta del usuario en lenguaje natural, el aporte que hace a la presente investigación, es que presenta cuáles son las características que debería tener una arquitectura orientada a servicios para satisfacer necesidades de aplicaciones que consultan datos georreferenciados. Dicha arquitectura fue validada mediante su aplicación exitosa en dos casos de estudio, uno de ellos correspondiente a un prototipo de datos georreferenciados estructurados según la ontología del registro catastral formulada por Piña (2006), y el otro caso de estudio que corresponde con un servicio de información de centros de salud dedicado a atender solicitudes de usuarios mediante dispositivos inalámbricos realizado por Soto (2010).

Como otro antecedente se cita la investigación de Piña (2012) y otros, titulada “Ontología web semántica del registro catastral venezolano”, artículo publicado en la Revista Ciencia e Ingeniería de la Universidad de Los Andes, en ocasión de las 1eras Jornadas de Modelado y Simulación, en el que la investigadora realiza una aplicación prototipo para acceder a datos georreferenciados, procesando una consulta formulada en lenguaje natural para extraer datos sensibles de propiedad inmobiliaria. Ese Procesamiento de Lenguaje Natural (PLN) lo realiza mediante un razonador que extrae palabras clave de la pregunta del usuario, apoyándose en una gramática y una ontología en la que se estructuran los conceptos del registro catastral venezolano con sus atributos y relaciones.

El aporte del estudio de Piña (2012) y otros, constituye el antecedente más antiguo de aplicación de consulta semántica a datos georreferenciados, mediante el uso de una ontología específica a un contexto de escala nacional, presentado por Piña (2006). Es originalmente el punto de partida de la presente investigación, que utilizó software libre y se ajustó al marco legal venezolano vigente para ese entonces, aunque sin ofrecerle la visualización o representación geográfica al usuario.

El siguiente antecedente presentado por los investigadores Partyka y otros (2011), en su artículo “Comparación de esquemas semánticos geográficamente ensamblados”, de las universidades de Texas en Dallas y de Minnesota, de los Estados Unidos de Norte América, el cual se entrega en la revista especializada de Web Semántica del año 2011.



Referencian en su contenido el esfuerzo computacional realizado mediante la aplicación de algoritmos, para hallar similitud semántica y determinar semejanzas y diferencias entre data geoespacial y sus modelos de datos, de las experiencias previas conocidas como Data Web y LinkedGeoData (Auer, 2009); los cuales representan esfuerzos de transición en el procesamiento de datos hacia una web semántica geoespacial. La propuesta de Partyka (2011) denominada GSim, también es una aplicación para hallar similitud, pero no entre los modelos o estructuras de datos sino entre las instancias de estos o comparación entre tablas a un alto nivel por atributos.

Para lograrlo utilizaron una distribución basada en entropía (EBD), el cual es un valor dentro de un rango de 0 a 1, donde 0 indica ausencia de similitud entre atributos y 1 indica presencia de atributos idénticos. La aplicación GSim primero examina las instancias, determina los tipos geográficos (GT) de las mismas, haciendo uso de fuentes de datos externas denominadas Gazetteer (Zhou, 2004; Newsam, 2008); además resuelve ambigüedades mediante valores de posición de los elementos, como latitud y longitud.

Así mismo señalan que para calcular la diferencia en profundidad entre dos conceptos o determinar un padre común entre los mismos (como se hace en las aplicaciones de lógica descriptiva -DL), es estrictamente necesario el uso de ontologías y su alineación. Partyka (2011) y su equipo (Ibid: p.53), señalan que las aplicaciones de PLN son dependientes de las relaciones entre

las palabras descritas en lenguaje natural de los conceptos geográficos, concluyendo a este respecto que puede haber problemas de ambigüedad.

Concluyen que la comparación del dominio geoespacial presenta retos únicos, debido a las propiedades inherentes en la data geoespacial tales como la geometría, coordenadas georreferenciadas, variaciones en la formación de los sistemas de referencia y muchas otras. Aseguran que la naturaleza geoespacial de la data es compleja, tal que las aplicaciones, incluyendo GSim, manejan un subconjunto limitado de la misma según sus propiedades.

Es éste el aporte fundamental de Partyka (2011) y otros, pues evidencian avances en la web semántica para datos georreferenciados, al tiempo que reconocen el reto que tendrán otras investigaciones relacionadas con otros subconjuntos del dominio de información geoespacial. Tal propuesta señala que los servicios web semánticos, pueden incorporar el dominio de información georreferenciada.

En 2010, Kyong-ho, L. y JongHyun, L. presentaron en el artículo intitulado “Construcción de Servicios Web compuestos desde solicitudes de Lenguaje Natural” en la revista especializada Web Semantic Journal. En el que señalan el incremento en la demanda de servicios web compuestos. Proponen el uso de interfaces de lenguaje natural a los servicios web, los cuales pueden ser usados aun por usuarios que no conozcan las tecnologías de los servicios web. Dada una consulta en lenguaje natural de un usuario a un servicio compuesto, el método propuesto genera un diagrama de flujo de trabajo

abstracto el cual describe las tareas que lo constituyen y sus transiciones en un servicio compuesto.

Los resultados experimentales de esta investigación con una variedad de solicitudes realizadas en lenguaje natural, muestran la extracción exitosa de los diagramas de flujos de trabajo abstractos, resultando en una precisión del 95,2%. Su contribución, radica en no solo la factibilidad de uso del lenguaje natural en servicios web, atendiendo frases complejas de usuarios, sino en su conclusión relacionada al aumento de dicha tendencia en estos servicios web compuestos.

Auer (2009) y otros en su artículo “Agregando una dimensión espacial a la Web de datos mediante LinkedGeoData”, presentado en la Conferencia de Web Semántica Internacional, señalan que con el fin de emplear la Web como un medio para integración de data e información, los conjuntos de datos y los vocabularios son requeridos pues estos permiten tanto la desambigüación como la alineación de otros datos e información. Muchas tareas de agregación e integración de información son imposibles sin un conocimiento de fondo relacionado a elementos espaciales de las vías, estructuras y paisajes que nos rodean.

En este artículo los autores contribuyen a la generación de una dimensión espacial para la data de la Web elaborado así como la data de OpenStreetMap, la cual es recolectada en forma colaborativa y puede ser transformada y representada añadiéndole al modelo de data RDF, según cómo estos datos pueden ser enlazados con otros conjuntos de datos, sobre cómo puede éstos

hacerse accesibles para máquinas de acuerdo al paradigma de data enlazada y para humanos a través del visualizador de geo – data.

Según esta investigación, el paradigma de la Web Semántica combina técnicas de representación de conocimiento sencillas (como RDF, RDF Schema y Ontologías simples) con tecnologías web tradicionales (como HTTP y REST (REpresentational State Transfer, (R.T. Fielding, 2000))) para publicar y entrelazar data e información. Así también presenta un producto que agrega la dimensión espacial al entorno colaborativo web semántico.

En este sentido, el artículo anteriormente citado sirve como base fundamental y teórica para el desarrollo de la variable de estudio, puesto que representa el primer antecedente para la web semántica en el que se creó una aplicación con semántica en datos georreferenciados. También constituye un ejemplo de aplicación colaborativa con un registro semántico de más de 2 billones de triplas que agregan semántica mediante el lenguaje de anotaciones RDF.

Gang (2009) en su artículo intitulado “Una Investigación sobre servicios web semánticos geoespaciales basados en REST”, presentado en el foro internacional en ciencia-tecnología de la computación y aplicaciones, el autor refiere que las tecnologías de servicios web existentes están sólo en el nivel sintáctico. Señala que siguen teniendo gaps semánticos (diferencia entre dos representaciones de un objeto debido a diferentes representaciones lingüísticas) en descubrimiento de recursos con dominio cruzado, consultas de

recursos heterogéneos, traducción de recursos de un dominio a otro en el nivel semántico.

Gang (2009) señala que la web semántica busca exponer los vastos recursos de información de la web como data que las computadoras puedan interpretar automáticamente. Este artículo presenta cómo las tecnologías web semánticas agregan valor en la construcción de servicios web semánticos geoespaciales usando REST, basado en metodologías de diseño. Ilustra aplicaciones de los servicios web semánticos en GIS (Geographic Information Systems), describe las características de la aproximación semántica pregunta/respuesta REST, y mapea las operaciones HTTP REST dentro de la vía semántica.

Gang (2009) señala que las reglas URIs (Uniform Resources Identifiers) semánticos son usadas para identificar recursos geoespaciales y sugiere usar RDF (Resource Description Framework) para describir las relaciones de los recursos geoespaciales y sus conceptos los cuales comparten significado con ontologías y finalmente consultar el conocimiento geoespacial distribuido usando SPARQL actualizado (SPARQL protocolo y lenguaje de consulta RDF).

El resultado de la investigación de Gang (2009), aporta al presente estudio el uso de URIs semánticos para acceder a información georreferenciada bajo la aproximación pregunta/respuesta REST. El autor sugiere esta arquitectura para la construcción de servicios Web semánticos como tecnología vigente.

Otro antecedente a la presente investigación lo aporta Criado (2009) en su tesis doctoral titulada “Procedimiento para transformar la Web en Web Semántica” presentada en la Universidad Nacional de Educación a Distancia (UNED), de España. El autor señala acerca del concepto de Web Semántica, la necesidad de una representación formal de la información de acuerdo a ontologías de referencia que doten a la Web de semántica para los sistemas informáticos. Afirma el autor acerca de la existencia de un acuerdo generalizado para que esto se haga a través de lenguajes de etiquetado, indicando la necesidad de suficientes anotaciones semánticas de este tipo, para que tenga sentido global en la Web.

Criado (2009) asegura sólo cuando se tenga la facilidad de generar suficientes anotaciones semánticas, ya sea de forma automática o semi-automática, se podrá desarrollar aplicaciones que aprovechen o saquen partido de esa semántica. En ese sentido su aporte específico a esta investigación, es la propuesta de un procedimiento en la contribución de la extensión de ontologías, que faciliten a un usuario el etiquetado semántico de la información que gestiona. El proceso propuesto por Criado (2009) para poblar diferentes ontologías desde el mismo contenido, lo denomina Vistas semánticas.

Según Criado (2009), su propuesta de automatización y semi-automatización agregaría valor a los sitios web tradicionales, reconocidos por buscadores semánticos, sin interferir en las labores de búsqueda a sitios carentes de la característica semántica, es decir, carentes de ontologías para

la representación del significado. Así mismo señala que la calidad de las ontologías depende de la afinidad, precisión, estandarización, completitud, entre otros, de los conceptos; así como la claridad del contenido, la capacidad de extracción y el análisis están condicionadas en gran medida al procesamiento del lenguaje natural (PLN).

El aporte del estudio doctoral de Criado (2009) a la presente investigación, está orientado y coincide con la necesidad de migrar la Web actual a una Web semántica, cualquiera que sea el dominio de información, para asistir a los usuarios en el acceso a la misma. Así como en otros antecedentes y en Criado (2009), se invita al uso de ontologías para la representación del conocimiento.

Otra contribución la aporta Klien (2008) en su investigación doctoral intitulada "Anotación semántica de la información Geográfica" presentado en la Universidad de Múnster, Alemania, en el área de Geoinformática. En ella asegura que se está en dirección de responder a la pregunta de cómo explicar el conocimiento de los términos que contienen información geográfica, i.e. que lidian con anotaciones semánticas de entidades geográficas.

Continúa refiriéndose a que las anotaciones semánticas establecen un enlace entre descripciones de tipos de elementos y descripciones de categorías formales. Para evaluar la validez o usar una categoría para una anotación, es necesario saber determinar que los elementos (instancias de un tipo de elemento) representan miembros de esa categoría.

Klien (2008) afirma que las anotaciones manuales tienden a ser erradas si la persona que anota no está familiarizada la conceptualización capturada en

las descripciones de la categoría, o no está con el contenido de la fuente de datos. Además, las anotaciones manuales con frecuencia consideran sólo una perspectiva sobre la fuente de información, así se ignora su potencial de aplicabilidad en contextos diferentes. Los métodos existentes que automáticamente soportan anotaciones son principalmente basados en análisis de texto y sus algoritmos descansan en estadística y heurística.

Sus beneficios en acceso confiable a interoperabilidad semántica son limitados, puesto que ellos no se toman en cuenta las conceptualizaciones subyacentes. Klien (2008) desarrolla en dicha tesis, un método para evaluar la validez y extensibilidad de las anotaciones semánticas que han sido generadas de forma manual o con métodos basados en texto. Los métodos propuestos por Klien (2008) se basan en cualidades físicas y relaciones espaciales para categorizar entidades geográficas.

Así que es posible definir reglas para miembros de categoría que están restringidas a características físicas y relaciones espaciales de entidades geográficas. Estas reglas son formalizadas como parte de ontologías de dominio geoespacial.

Desde que los elementos geográficos tienen representaciones geométricas así como ubicación espacial, es posible computar y analizar relaciones métricas y topológicas de las entidades geográficas representadas. En consecuencia, pueden usar reglas formales para evaluar miembros de categorías para acceder a la validez y extensibilidad de las anotaciones semánticas. Para esto, las reglas formales son traducidas en procedimientos



de análisis espacial que pueden ser ejecutados sobre instancias de elementos con operadores espaciales provistos por un sistema de información geográfica.

El resultado de la investigación de Klien (2008) es triple: (i) define un modelo conceptual para anotaciones semánticas en ambientes para servicios web con información geográfica, (ii) provee una base ontológica formal para el dominio geoespacial, y (iii) desarrolla un método para soportar automáticamente el proceso de anotación que evalúa la validez de anotaciones existentes y sugiere otras posibles.

Las contribuciones de Klien (2008) para la presente investigación son: la propuesta de uso de modelo conceptual para anotaciones semánticas en servicios web y de una base ontológica formal para el dominio geoespacial. Asegura Klien con este precedente, que mediante la especificación de reglas se provee la base para la implementación de algoritmos de análisis espacial.

A diferencia de la postura de la autora de la presente investigación, Klien (2008) en este antecedente sugiere evitar el procesamiento de lenguaje natural. Para Klien las técnicas de extracción de conocimiento basado en análisis de atributos, el cálculo de características espaciales se mantiene independiente desde una descripción textual de los objetos geográficos y sus atributos.

Pautasso (2008) en su artículo intitulado RESTful Web Services vs. "Big" Web Services: Making the Right Architectural Decision, presentado en la 17<sup>ma</sup> Conferencia Internacional del World Wide Web en Beijing, China para abril de

2008, presenta un estudio de las tendencias en tecnologías (hasta ese momento) en el dominio de los Servicios Web (SW) e indica que una solución, eliminando la presunta complejidad de los estándares SW-\* está dada en función de los servicios web que utilizan REpresentational State Transfer (REST).

Pautasso (2008) explica porque los trabajos del World Wide son aplicables a resolver problemas de integración de aplicación empresarial y para simplificar las tareas requeridas para construir arquitecturas orientadas a servicios. En este artículo el investigador objetiva el debate sobre los estándares de los SW-\* Vs. REST, dando una comparación técnica cuantitativa basada en decisiones y principios de arquitectura. Así también muestra que esas dos aproximaciones difieren en el número de decisiones sobre arquitectura que deben ser tomadas y en el número de alternativas disponibles.

Continúa Pautasso (2008) asegurando que hay diferencias significativas en las consecuencias de ciertas decisiones en términos del desarrollo resultante y costos de mantenimiento. Su comparación ayuda a quienes toman decisiones técnicas para evaluar los dos estilos y tecnologías de integración más objetivamente y a seleccionar aquel que mejor encaja en sus necesidades: REST está mejor configurado para lo básico, escenarios de integración AD HOC (desde cualquier lugar del mundo escapando a los contextos locales), SW-\* es más flexible y direccionable, calidad avanzada de

requerimientos de servicio que comúnmente ocurren en computación empresarial.

La contribución de Pautasso (2008) radica en su evaluación técnica cuantitativa basada en decisiones y principios de arquitectura sobre los estándares para la construcción de los servicios web, permite tomar la decisión de cuál estándar se ajusta mejor al desarrollo de un servicio Web semántico basado en un modelo ontológico para la gestión de datos georreferenciados.

## **2. Bases Teóricas**

### **2.1. Web Semántica**

La Web Semántica es un proyecto que busca añadir significado (semántica) procesable por computadoras a la World Wide Web (W3). En febrero de 2004, el Consorcio World Wide Web (W3C) lanzó una tecnología denominada Resource Description Framework (RDF) y un lenguaje para la generación de ontologías a usarse en la Web conocido como Ontology Web Language (OWL); como recomendación para representar información y para intercambiar conocimiento en la Web. OWL es usado para publicar y compartir conjuntos de términos llamados ontologías, las cuales soportan búsquedas avanzadas en la Web, interacción de usuarios con agentes de software y gestión de conocimiento.

El W3C es una comunidad internacional que desarrolla estándares abiertos, protocolos y lineamientos para asegurar el crecimiento de la Web. Según esta organización, el valor social de la Web es que habilita la comunicación entre humanos, el comercio y las oportunidades para compartir conocimiento. Una de sus metas fundamentales es hacer dichos beneficios disponibles para la gente, cualquiera que sea su hardware, software, infraestructura de red, lenguaje nativo, cultura, ubicación geográfica o su habilidad mental o física. Algunos pueden ver la Web como un repositorio gigante de información enlazada mientras otros como un inmenso conjunto de servicios para intercambiar mensajes. Según la W3C ambas visiones se complementan y el uso de la Web dependerá de la aplicación.

Para la International Association for Ontology and its Applications (IAOA) (2012), una ontología aplicada es considerada como un dominio interdisciplinario en la intersección de la filosofía, la lingüística, la lógica, las ciencias cognitivas y las ciencias computacionales, la cual tiene aplicaciones en el modelado conceptual, la ingeniería, la administración del conocimiento, los sistemas de información, las ciencias de bibliotecología y la documentación.

Según Antoniou (2008: p.21) la Web Semántica es una iniciativa que tiene como meta mejorar el estado actual de la World Wide Web, que se concentra en el uso de información web procesable por máquinas y que incluye como tecnologías clave metadata explícita, ontologías, lógica e inferencia y agentes inteligentes.

Kuhn (2005: p.2) antes de referirse a la Web semántica, define la interoperabilidad como un mecanismo de compartir conceptualizaciones, entre agentes de software. Señala además que la Interoperabilidad semántica es la técnica análoga a la comunicación y cooperación humana, y que deben entenderse ampliamente ambas para comprender la interoperabilidad entre seres humanos y sistemas. Su perspectiva de interoperabilidad semántica incluye a seres humanos como parte de sistemas interoperativos en los que las interfaces a usuarios son subsumidas dentro de las interfaces a servicio, de las que es requerido conocer su sintaxis y el vocabulario al que hacen referencia.

Para Kuhn (2005: p.3) el único uso razonable del término “semántica” se refiere al significado de expresiones en un lenguaje. Estas expresiones pueden ser símbolos sencillos como palabras o combinaciones de símbolos. De acuerdo al significado de los términos, ellos son usados para expresar algo. Los lenguajes relevantes en un contexto de sistemas de información son aquellos que permiten a los seres humanos conceptualizar algo con el propósito de representarlo y manipularlo en máquinas.

Por último, según como la define su promotor Berners-Lee (1999), la Web Semántica es una red de datos que pueden ser procesados directa o indirectamente por máquinas. La autora fija posición con esta definición. Así, la variable Web Semántica, objeto de estudio de la presente investigación, se desglosa en tres dimensiones, ampliadas a continuación.

### **2.1.1. Arquitectura de la Web**

Hablar de la arquitectura de la web no es una tarea trivial, debido a que no se trata de definir la arquitectura de una pieza de hardware o de un programa de software. La arquitectura tecnológica representa la estructura física y aspectos del diseño físico de una pieza de tecnología según Erl (2008: p.35) y en el caso de la web, ésta no es una cosa física que existe en un determinado lugar sino un espacio en el que la información puede existir Berners-Lee (2001: p.34).

No existe un ordenador central controlando el web, pero si existen algunas reglas básicas y comunes con las que los ordenadores pueden establecer comunicación entre sí, definidas por el protocolo de transporte de comunicación HTTP (por sus siglas Hypertext Transference Protocol). Sobre este protocolo se construyen los servicios, es decir, éstos se comunican utilizando HTTP. La arquitectura tecnológica limitada al diseño físico de un programa de software diseñado como un servicio, es denominada a la arquitectura del servicio.

Según Erl (2008: p.62), esta forma de arquitectura tecnológica es comparable a la arquitectura de componentes, excepto que esta debe apoyarse en una cantidad mayor de extensiones de infraestructura para soportar sus necesidades de creciente confiabilidad, ejecución, escalabilidad,

comportamiento predecible y especialmente su necesidad de aumentar su autonomía. El alcance de una arquitectura de servicio tiende a ser más grande porque un servicio puede entre otras cosas, abarcar múltiples componentes.

#### **2.1.1.1. Protocolo**

Para Berners-Lee (1999: p.115) los protocolos definen cómo interactúan los ordenadores. El protocolo es el primer principio de la arquitectura según Manning (2008), el cual permite clarificar si el HTTP se considera como un protocolo de aplicación en la Web que es usado como medio para entregar aplicaciones de servicio a clientes o si se considera el HTTP como protocolo de transporte, es decir, como medio de mensajería para integración de aplicaciones.

En el contexto REST, la Web es vista como medio universal para publicar información accesible globalmente. Las aplicaciones forman parte del Web mediante el uso de URIs para identificar los recursos provistos, datos y servicios, y apalancar la completa semántica de los cuatro verbos del HTTP (GET, POST, PUT y DELETE) para exponer operaciones sobre los recursos.

Por su parte, desde la perspectiva de SOAP/SW-\*, el Web es visto como un medio de transporte universal para mensajes, los cuales son intercambiados entre los servicios web finales de las aplicaciones publicadas.

Así las aplicaciones ganan la habilidad de interactuar remotamente a través del Web pero permaneciendo fuera del Web. En otras palabras, el protocolo HTTP se usa como un protocolo túnel para permitir la comunicación a través de la barrera o pared de fuego (firewall), pero no es usado para comunicar la semántica de la interacción del servicio. Esto puede ser visto desde la vía de los SW-\* que utilizan URIs para direccionar mensajería final, la cual típicamente permanecen iguales para todas las operaciones de un servicio, donde los URIs de REST identifican recursos del dominio de aplicación.

Para Thomas Erl (2008: p.150), el protocolo representa las tecnologías requeridas para establecer la comunicación base. Desde la perspectiva de los servicios web esto debe incluir la estandarización del protocolo de transporte HTTP, y un protocolo de mensajes SOAP.

Para la dimensión Arquitectura de la Web indicador protocolo, la autora fija posición a favor de Pautasso (2008), el cual sugiere según una evaluación técnica cuantitativa que existe una tendencia a la tecnología REST sobre los estándares SW.



### **2.1.1.2. Servicio**

En el nivel conceptual, un servicio es un componente de software provisto a través de un terminal con acceso a la red. Un cliente de un servicio y un proveedor utilizan mensajes para intercambiar documentos autocontenidos con la información de la solicitud y la respuesta, que hacen pocas asunciones acerca de las capacidades tecnológicas del receptor Pautasso (2008). En particular, no existe noción de una referencia a un objeto remoto que requiera un gestor de objetos para administrar un espacio de direcciones de memoria distribuida. En el nivel tecnológico, SOAP es un lenguaje XML con el cual se define la arquitectura y formato de un mensaje, por lo tanto provee un rudimentario protocolo de procesamiento.

Un documento SOAP define un elemento de alto nivel denominado envoltorio, el cual contiene una cabecera y un cuerpo. El encabezado SOAP es un contenedor extensible para la infraestructura de la información de la capa de mensajes, que puede ser usada con propósitos de enrutamiento (ej. Direccionamiento) y configuración de servicio de calidad QoS (Quality of Service) (ej. Transacciones, seguridad, confiabilidad). El cuerpo contiene la carga útil del mensaje. El esquema XML es usado para describir la estructura del mensaje SOAP, así que el motor de SOAP puede codificar o decodificar el contenido del mensaje y direccionarlo a la implementación apropiada.

Para Pautasso (2008), los terminales que solicitan y que atienden un servicio son direccionados en la capa de transporte por URIs mediante SOAP/HTTP. El servicio se identifica, se describe, se declara la seguridad/confiabilidad deseada y las políticas de mensajería y transacciones, se reutiliza/compone y se descubre.

### **2.1.1.3. Integración**

El indicador integración según la OGC, siglas para la Open Gis Consortium, puede ser definido como la capacidad de intercambiar componentes de software de diferentes proveedores, con la finalidad de integrar data de incontables fuentes en servicios accesibles por clientes de la web. La OGC es una organización que formula estándares para publicar y tener acceso a la información geoespacial, y es pilar fundamental en el éxito de los nodos de información espacial denominados IDE, siglas de Infraestructura de Datos Espaciales, las cuales se definieron a principios de los años 1990's como una colección de tecnologías, políticas y arreglos institucionales que facilitan la disponibilidad y el acceso a datos espaciales.

Debido al carácter de la información de interés para la presente investigación, a saber, los datos georreferenciados, es relevante esta posibilidad de integrar aplicaciones de acceso a los mismos desde la web, mediante servicios web. Estos recursos de información pueden ser registrados

con metadata en catálogos en línea de manera que puedan ser accesibles y fácilmente descubiertos, por usuarios locales y remotos. Gracias a los estándares hoy es más fácil acoplar o integrar sistemas y aplicaciones.

Equipos como teléfonos inteligentes o dispositivos móviles hacen posibles nuevas aplicaciones. Los datos de posiciones geográficas pueden ser almacenados local o remotamente en servidores, en lo que se denominan servidores de la “nube”. La “nube” o Cloud computing, es un término anglosajón acuñado en países hispanos, que se refiere al resultado de la convergencia de tecnologías tales como comunicación de banda ancha, virtualización o computación ubicua, y servicios web. Esta soporta arquitecturas orientadas a servicio basadas en estándares de la OGC para servicios web.

Kuhn (2005: p.2) antes de referirse a la Web semántica, define la interoperabilidad como un mecanismo de compartir conceptualizaciones, entre agentes de software. Señala además que la Interoperabilidad semántica es la técnica análoga a la comunicación y cooperación humana, y que deben entenderse ampliamente ambas para comprender la interoperabilidad entre seres humanos y sistemas. Su perspectiva de interoperabilidad semántica incluye a seres humanos como parte de sistemas interoperativos en los que las interfaces a usuarios son subsumidas dentro de las interfaces a servicio, de las que es requerido conocer su sintaxis y el vocabulario al que hacen referencia.

### 2.1.2. Tipos de Consultas

Según Manning (2008: p.1), existen tres tipos de consultas que permiten la recuperación de información. Y antes de describir estos tipos de consulta, define y delimita el amplio término recuperación de información, como la búsqueda de material (usualmente documentos) de naturaleza no estructurada (usualmente texto) que satisface una necesidad de información dentro de una gran colección (usualmente almacenada en computadoras).

Tales tipos de consulta son: la estructurada, la no estructurada y la que se realiza sobre bases de datos relacionales. Así mismo este autor define tres escalas según la necesidad de la información y la cantidad de documentos en donde se hará la consulta. La primera de ellas las búsquedas en el web, que se encuentra en uno de los extremos, en el otro la recuperación de información personal, y en medio de éstas, las búsquedas sobre algún dominio específico, institucional o empresarial.

Aunque Manning no incluye en su taxonomía las consultas de tipo espacial, por la naturaleza de la presente investigación la autora agrega este tipo de consulta a datos georreferenciados, a la que otros autores hacen mención, otras aplicaciones demuestran su uso y organizaciones como la OGC y W3C, desarrollan estándares para hacer más accesible y estructurada la construcción de servicios web de este tipo.

### **2.1.2.1. Estructuradas**

Para Manning (2008: p.178), la recuperación estructurada de datos en la web es aquella que realiza búsqueda y recuperación sobre documentos estructurados por el lenguaje de marcado XML, el cual es el estándar más ampliamente utilizado. Dicho autor plantea que existe una diferencia de este tipo de búsqueda con la búsqueda y recuperación de datos semi-estructurada, pues así como otros investigadores la distinguen de las consultas realizadas sobre bases de datos.

La autora de esta investigación hace referencia a Manning en el sentido de otorgarle a una consulta el adjetivo de estructurada, como aquella a la que el usuario tiene acceso a datos bien sea etiquetados con XML o de la extracción de estos de una base de datos. Así mismo, hace notar que denomina consulta estructurada aquella donde el usuario de la web tiene una interfaz a los datos que le ofrecen un menú limitado de opciones para rescatarlos del repositorio donde se encuentren. Es limitado, pues es el diseñador de la interfaz es quien decide a qué información podrá el usuario tener acceso mediante una serie de “clicks” sobre menús desplegables que ejecutarán los comandos necesarios para este fin.

Según Kuhn (2005), existen aplicaciones estándares que introducen vocabularios más o menos controlados (tales como aquellas en catálogos de atributos por elementos o estándares de metadata). La autora hace esta

referencia a Kuhn, pues coincide con este control al que él hace mención, en el que las aplicaciones presentan un dominio de información y términos definidos desde el diseño de las mismas.

Ahora bien, la autora plantea una taxonomía de consultas estructuradas simples y compuestas, no en función de cómo se encuentran estructurados los datos, sino de la forma en la que se realiza la pregunta. El primer tipo de consulta estructurada simple, se refiere a aquel servicio de consulta a datos donde el usuario selecciona de un menú desplegable un término por el que tiene interés y la aplicación se lo proporciona. O bien a aquel en el que el usuario completa el término en un formulario para un dato atómico. Por su parte, una consulta estructurada compuesta puede incluir algún operador lógico como “y” u “o”, y de forma similar el usuario escogerá de un menú desplegable o completará en un formulario dos datos atómicos para componer su pregunta.

#### **2.1.2.2. No estructuradas**

Para Manning (2008), el término data no estructurada se refiere a la data que no tiene una estructura clara, semánticamente abierta, “fácil” de interpretar para una computadora. Es la opuesta a la data estructurada. Pero aún así, ninguna data carece de estructura pues obedece a la estructura lingüística del lenguaje humano.

Según Kuhn (2005), existen también aplicaciones de data y colecciones de metadata con entradas de texto en formato libre (no estructuradas), abren la puerta a un uso casi ilimitado de expresiones en lenguaje natural. Para él hacer frente a la semántica de todas las expresiones en lenguajes que manejen el significado, va más allá de lo que actualmente se tiene en cuanto al significado.

Entonces Manning se refiere al estado estructurado de los datos y Kuhn a la consulta estructurada. La autora, así como Kuhn, hace referencia a la consulta, en la que un usuario formula una pregunta en su lenguaje natural. Para este tipo de consultas, la aplicación debe estar dotada de un componente razonador, que sea capaz de interpretar la pregunta, y es a esto a lo que otros autores como Antoniou (2008), Criado (2009), Kyon-ho (2010), Partyka (2011) denominan PLN, o Procesamiento de Lenguaje Natural. La autora además advierte que no comparte la opinión de Klien (2008), quien señala preferible evitar el uso de PLN.

### **2.1.2.3. Espaciales**

Las consultas de tipo espacial, o aplicaciones que permiten el acceso a data geoespacial, en la actualidad ofrecen similares herramientas que los Sistemas de Información Geográfica (SIG) tradicionales, y añaden la posibilidad a los usuarios de realizar sus búsquedas como antes no era posible

a través del web. Esto gracias a las capacidades de la denominada “computación de nube” y a las colecciones superpobladas de datos a las que los mismos usuarios contribuyen en su formación. Hoy en día, los servicios web son los encargados de proveer, muchas veces sin costo, varios de los servicios que previamente habrían sido accesibles solo para expertos de SIG.

Hoy esta tecnología ha cambiado y continúa haciéndolo, ajustándose a las necesidades de sus usuarios. Según la OGC, datos y servicios de innumerables fuentes pueden ser accedidos por clientes desde la Web, y es por esto que cada año nuevos estándares son formulados por la organización para extender las posibilidades de los IDEs basados en la Web.

La autora fija posición a favor de estos estándares de acceso a datos que la OGC pone a disposición de los usuarios de las instituciones que desarrollan aplicaciones a datos espaciales. Estos estándares, permiten a través de los servicios web de consulta a datos georreferenciados, la representación o visualización de los espacios geográficos a nivel mundial, nacional y regional, y el desarrollo de aplicaciones que puedan interoperar y colaborar entre sí.

### **2.1.3. Servicios Web**

Para McIlraith (2003: p.90), los servicios web son aplicaciones autocontenidas y modulares que pueden ser descritas, publicadas, localizadas e invocadas sobre una red de trabajo- generalmente, la World Wide Web.



Por su parte Antoniou (2004: p.194), define a los servicios web como sitios Web que no solo proveen información estática, sino que además involucran interacción con usuarios y con frecuencia permite a estos realizar algunas acciones. Según Antoniou, estos pueden ser simples y complejos, siendo los primeros aquellos programas que acceden al Web, como sensores o dispositivos que no se apoyan en otros servicios Web y que no requieren más interacción con el usuario, más allá de una simple respuesta. Según esta taxonomía, señala que los servicios Web complejos están compuestos por servicios simples y con frecuencia requieren continua interacción con el usuario, donde éste puede tomar decisiones o proveer información de acuerdo a ciertas condiciones.

Para Antoniou (2004) el reto de los servicios Web desde la perspectiva de la Web Semántica es que éstos puedan automáticamente descubrir, invocar, componer y monitorear información, proveyendo una descripción del mismo, interpretable por agentes de software u otros servicios. A este efecto, Antoniou refiere un conjunto de conocimiento asociado con el servicio que debe ser publicado como su perfil, su modelo y su configuración. Así mismo lo relaciona con la Inteligencia Artificial.

La autora a continuación desarrolla dichos indicadores, detallando la configuración del estándar de la OGC para los servicios Web y los complementa con su visión de otros servicios asociados.

### 2.1.3.1. Configuración

Algunos de los aspectos que deben estar presentes en los servicios Web, son los descritos en los estándares de la OGC. Estos aspectos comunes son principalmente algunos de los parámetros y estructuras de datos utilizados en operaciones de pregunta-respuesta. Actualmente se incluyen dentro de estos estándares los Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS), Catalogue Service Web (CSW), Web Feature Service-Transactional, Web Map Tile Service (WMTS).

Un WMS es un servicio Web que sirve mapas en forma de imagen, generados dinámicamente según el requerimiento del usuario, los cuales se basan en información georreferenciada de algún SIG.

Un WFS es un servicio Web que provee al usuario información codificada en GML (Geographic Markup Language), la cual le servirá para interactuar o generar un WMS.

Un WCS es un servicio Web de Coberturas de información geográfica, que a diferencia de un WMS, sirve al usuario que las solicita de objetos que pueden ser editados y analizados, y no de imágenes.

Un CSW es un servicio Web de Catálogo de metadatos de información geoespacial, el cual contiene datos acerca de la interfaz para el descubrimiento, despliegue, consulta y otros recursos potenciales.

Un WFS-Transactional es un servicio Web por el cual es posible editar elementos de una base de datos georreferenciados a través de un servicio

WFS, i.e. habilita las operaciones de lectura y escritura a través de transacciones de inserción, borrado y actualización de los datos.

Un WMTS es un servicio Web que habilita a una aplicación de servidor a servir series de mapas espacialmente referenciados usando colección de imágenes con un contenido, extensión y resolución predefinidos.

La información ampliada se encuentra en la referencia de la OGC.

### **2.1.3.2. Especificación**

El indicador especificación está descrito por Antoniou (2008: p.194) y a continuación se amplía.

Antoniou señala que el servicio debe tener tres tipos de conocimiento asociado. El primero el perfil del servicio, el cual es una descripción de lo que ofrece y requiere el mismo. Esto es una información esencial que permite ser descubierto por otros servicios, los cuales a través de agentes de software, pueden determinar si el servicio es apropiado para sus propósitos, basados en este perfil. El siguiente es el modelo del servicio, el cual describe como trabaja el mismo o lo que sucede exactamente cuando el servicio es utilizado. Tal información puede ser importante para agentes buscadores de servicios compuestos que ejecutan tareas complejas y para monitorear la ejecución del servicio.

Finalmente Antoniou (2008) señala la configuración del servicio como aquella que especifica detalles de cómo un agente de software puede acceder al servicio. Es este aspecto el que incluye detalles del protocolo de comunicación y del número de puertos a ser usados en contactar el servicio. Así mismo, sugiere este autor sean incluidas una descripción para ser leída por humanos, del proveedor del servicio, una especificación de las funcionalidades provistas por el servicio e información adicional como tiempo de respuesta esperado y restricciones geográficas.

#### **2.1.3.3. Inteligencia artificial**

McIlraith (2003: p.90) señala que un elemento clave para la realización de la Web Semántica es el desarrollo de un lenguaje rico configurable para la codificación y descripción del contenido del Web. Tal lenguaje debe tener una semántica bien definida, ser suficientemente expresivo para describir las complejas interrelaciones y restricciones entre los objetos Web, y ser capaz de manipular automáticamente y de razonar con un límite aceptable de requerimientos de tiempo y recursos.

Según McIlraith (2003: p.90), los lenguajes ontológicos para Web basados en inteligencia artificial y contruidos sobre XML, que permiten representar conocimiento son: Resource Description Framework, RDF Schema, DAML+OIL y OWL.

Son precisamente estos lenguajes los que pueden ser leídos e interpretados por agentes de software. Según Erl (2008: p.65), un aspecto relacionado con la infraestructura del diseño de un servicio es cualquier dependencia que el servicio pueda tener con agentes de servicio, los cuales son programas dirigidos por eventos capaces de interceptar transparentemente y procesar mensajes enviados a o desde un servicio. Los agentes de servicios pueden ser desarrollados a la medida.

Antoniou (2008: p.196) por su parte, asegura que los servicios Web son un área de aplicación en la que las técnicas de Inteligencia Artificial pueden ser usadas efectivamente.

La autora de la presente investigación señala la absoluta necesidad del uso de estos lenguajes estructurados que contienen descripciones semánticas, los cuales pueden no solo ser leídos e interpretados por seres humanos sino también por agentes de software. Un recurso de este tipo es el almacenado en ontologías, las cuales son referidas en varios de los antecedentes citados en la presente investigación. De esta forma la autora concuerda con la posición Antoniou.

#### **2.1.3.4. Servicios asociados**

Los servicios asociados constituyen un indicador motivado a la naturaleza de la investigación, el cual es planteado como necesario por la autora de la misma. Estos tienen que ver con inquietudes de la investigación previamente

realizada por Piña (2006), en la que la aplicación prototipo desarrollada dejó ciertas inquietudes que podrían ser resueltas en la presente investigación.

Aspectos como la privacidad de la información de los usuarios, son señalados por Berners-Lee (2001: p.133), o sobre el carácter oficial de los datos georreferenciados son citados en las Normas Técnicas de Formación y Conservación del Catastro Nacional (2002) Venezolano y en la Constitución de la República Bolivariana de Venezuela (1999). Otros se incluyen para conocer las frecuencias absolutas y porcentajes en los servicios estudiados, como si estos ofrecen al usuario la capacidad de editar datos tabulares o gráficos, o si los servicios solicitan la posición georreferenciada del usuario o la calcula automáticamente. Estos resultados se presentarán en el Capítulo IV de análisis y discusión de los resultados.

### **3. Sistema de Variables**

#### **3.1. Definición nominal**

Web Semántica

#### **3.2. Definición conceptual**

La Web Semántica es una red de datos que pueden ser procesados directa o indirectamente por máquinas, según Berners-Lee (1999). La Web

Semántica es un proyecto que busca añadir significado (semántica) procesable por computadoras a la World Wide Web. (W3C:2013).

### **3.3. Definición operacional**

La Web Semántica es la misma web tal como existe hoy en día, de la que millones de usuarios disponen como herramienta que ofrece servicios de consulta de información, entre otros usos, con una característica adicional: la participación de agentes de software que asisten las búsquedas de información de seres humanos mediante la interpretación de sus solicitudes y resultados, gracias a que esos agentes tienen acceso al significado de los símbolos y palabras usados por los humanos en la Web. Este tipo de servicios semánticos, básicamente le otorgan a la web actual el adjetivo de web semántica.

Los desarrolladores de servicios de consulta de información en la web actual enfrentan el reto de vigilar la pertinencia, veracidad y actualidad de la información que manejan estas aplicaciones; en función de responder responsablemente al requerimiento del usuario y procurar sea satisfactoria su interacción con tal servicio. Así mismo, los desarrolladores deciden en qué forma los usuarios pueden realizar las consultas, al ofrecer servicios web de consultas estructuradas sobre algún dominio de información específico. Sin embargo, y haciendo hincapié en el asunto, el aspecto de servicio semántico

solo se le otorga a la web si las máquinas tienen la capacidad de efectuar algún tipo de razonamiento sobre los datos, y colaborar en la interpretación de lo que el usuario solicitó.

Para lograr que los servicios en la Web sean semánticos, además de contar con las mencionadas características, los desarrolladores tendrán que superar ciertas dificultades como la cantidad de datos anotados, la obsolescencia de las plataformas de desarrollo, la multiplicidad de estándares que compiten; y verificar a través de mediciones, la utilidad y efectividad de estos servicios (usuario satisfecho), cumplir con el Habeas Data (usuario informado) y proveer la herramienta ideal para mejorar la toma de decisiones mediante la visualización de datos georreferenciados (característica apenas hasta ahora mencionada) con la que la web semántica arribaría a otro estadio de consultas de información.

La Web Semántica es ya un proyecto establecido, que convoca a un gran colectivo científico distribuido por todo el mundo. El objeto de trabajo es la propia World Wide Web y cualquier metodología para su desarrollo tiene que asumir su historia y la importante contribución del W3C en la definición de estándares para compartir las contribuciones y el propio contenido de la red.

**Cuadro 1:  
Operacionalización de la Variable**

<b>Objetivo General</b>	Analizar la web semántica en la consulta de datos georreferenciados.			
<b>Objetivos Específicos</b>	<b>Variable</b>	<b>Dimensiones</b>	<b>Indicadores</b>	<b>Autores</b>
	Web Semántica	Arquitectura de la web	Protocolo Servicio	Pautasso (2008)



Identificar la arquitectura de la web semántica en la consulta de datos georreferenciados.			Integración	Antoniou (2008)
Describir los tipos de consulta de la web semántica en datos georreferenciados.		Tipos de Consultas	Estructuradas	Manning (2008)
			No estructuradas	
Caracterizar los servicios web de la web semántica en la consulta de datos georreferenciados.		Servicios Web	Espaciales	OGC Estándares
			Configuración	Antoniou (2008)
			Especificación	
			Inteligencia Artificial	Piña (2006)
Servicios Asociados				
Formular lineamientos teórico-prácticos de servicios web semánticos de consulta de datos georreferenciados.	Objetivo a alcanzar a través de la revisión teórica así como del análisis sistemático por parte de la investigadora por medio de los resultados			

**Fuente:** Piña (2013)

### 3.4. Abreviaturas y Acrónimos

CSW: Catalogue Service Web

HTTP: Protocolo de transferencia de hipertexto (Berners-Lee, 1999: p.34).

IP: Protocolo de Internet (Berners-Lee, 1999: p.17).

OWL: Lenguaje Ontológico para la Web Semántica.

OWL-S (OWL para servicios): Consiste en tres partes interrelacionadas: un perfil ontológico, usado para describir qué hace el servicio; una ontología de procesos y su sintaxis de presentación correspondiente, usada para describir cómo es usado el servicio; y una ontología base, usada para describir cómo interactuar con el servicio. (Martin, 2007: p.244)

REST: REpresentational State Transfer

SOA: Service Oriented Architecture

SOAP: Simple Object Access Protocol

TCP: Protocolo de Control de Transmisión (Berners-Lee, 1999: p.17).

URI: Identificadores de Recursos Universales (Berners-Lee, 1999: p.34).

W3C: World Wide Web Consortium. Cuerpo de creación de estándares y protocolos (Solá, 1999: XII)

WCS: Web Coverage Service

WFS: Web Feature Service

WFS-G: Web Feature Service - Gazetteer

WFS-Transactional: Web Feature Service Transactional

WMS: Web Map Service

WMTS: Web Map Tile Service

WPS: Web Processing Service

WSDL (Web Service Description Language): describe un servicio web vía su conjunto de operaciones (Hull, 2005: p.87).

XML: eXtensible Markup Language

## **CAPÍTULO III**

### **MARCO METODOLÓGICO**

#### **1. Enfoque Paradigmático**

El presente apartado, esboza los aspectos correspondientes al paradigma que orienta la investigación, en este sentido, es concerniente señalar que el estilo de pensamiento de la investigadora se enmarca en el deductivo abstracto, en el que la realidad según Leal (2009), se concibe de forma representacional.

En correspondencia con lo antes expuesto, el paradigma que orienta la construcción del conocimiento se encuentra asentado bajo el enfoque científico-positivista, el cual busca el objeto de los supuestos teóricos y el empírico que se constituye en un todo sobre el estado actual de un hecho, situación o persona.

Según Comte citado por Vieytes (2004: p.48), establece la idea “fuera de nosotros existe una realidad totalmente hecha, acabada, plenamente externa y objetiva, y los seres humanos son como un espejo que pueden reflejarla, de esta forma, ser objetivo es copiar bien esa realidad sin

deformarla". Desde esta postura epistémica la fidelidad de la verdad consiste en la representación de nuestra imagen interior de la realidad.

Siguiendo las reglas esgrimidas por la postura epistémica antes descrita, el enfoque ontológico del estudio se basa en el abordaje empírico analítico, en palabras de Leal (2009), se busca la objetividad del conocimiento caracterizado por lo medible y determinado, es decir lo objetivo, para ello se utilizan instrumentos, válidos y confiables considerados como la vía idónea para producir conocimientos así como resultados basados en representaciones numéricas.

En la misma tendencia expresan Hernández, Fernández y Baptista (2010), que el investigador observa, mide y manipula variables; acepta los conocimientos que proceden de la experiencia (datos empíricos), es decir, los hechos es lo único que cuenta, se enuncia fundamentalmente el principio de verificación basado en la experiencia así como la observación del fenómeno objeto de indagación, como condición se establece que todo debe ser comprobable.

Asimismo, este paradigma hace énfasis en la verificación, abre la observación, contempla a la ciencia como un intento de codificar y anticipar la experiencia por lo que puede considerarse como el método válido de conocimiento basado en datos observacionales y mediciones. Ahora bien, en relación al estilo de pensamiento, se utiliza el lenguaje cuantitativo con una orientación objetiva y precisa del fenómeno en indagación, la expresión

numérica, es la referencia de validación en correspondencia con la realidad de forma objetiva.

## **2. Tipo de Investigación**

En referencia al tipo de investigación, se consideró de tipo descriptiva, en este sentido, Hernández, Fernández y Baptista (2008: p.102), señalan que “miden, evalúan o recolectan datos sobre diversos conceptos (variables), aspectos, dimensiones o componentes del fenómeno a investigar”. Fontaines Ruiz (2012: p.126) también clasifica según su metódica al tipo de investigación descriptiva, como aquella con la que se busca establecer mediante un análisis detallado de la realidad estudiada, las características, frecuencias, composición, funcionamiento del objeto estudiado, sin realizar ningún tipo de modificación de la realidad inicial.

## **3. Diseño de investigación**

El diseño de la investigación, corresponde al ex post-facto o no experimental, transversal o sincrónica, de campo, por cuanto Vieytes (2004: p.118), afirma que la investigación no experimental, “los datos se recogen en ambientes naturales, o utilizan muestras representativas de la población que pretende estudiar, ofrece resultado de mayor validez respecto de la población

real acerca de la cual hacen afirmaciones”. En este caso, se observa la variable sin manipular deliberadamente, se observa el fenómeno tal y como se dan en su contexto.

Asimismo, la investigación se identificó como sincrónica o transversal. Al respecto, Vieytes (2004: p.119), afirma, que los estudios transversales cristalizan los datos recogidos en un tiempo único, es decir se recolectan los datos en un solo momento, su propósito es la medición y ofrecer un conocimiento exhaustivo del objeto de indagación analizando su incidencia e interrelación en un momento dado.

Igualmente, el estudio se considera de campo, porque se verifican los hechos en un lugar determinado. Al respecto, señalan Hernández, Fernández y Baptista (2006: p.583) que “ocurre en los ambientes naturales y cotidianos de los participantes o unidades se análisis.”

## **4. Población y Muestra**

### **4.1. Población**

En este apartado se constituyen las particularidades que determinan el universo de estudio, definiendo la totalidad de atributos de la población a indagar. En este sentido, Vieytes (2004: p.395) expresa “es un conjunto de elementos, finito o infinito, definido por una o más características”. Es relevante

definir con precisión las particularidades de los elementos a investigar en el contexto del universo.

Desde la perspectiva, de Tamayo y Tamayo (2007: p.176) consiste en la totalidad de entidades que integran el fenómeno e incluye las unidades de análisis las cuales deben cuantificarse para un determinado conjunto, integrando unas premisas particulares del fenómeno objeto de investigación, lo cual, permitirá la obtención de información. En este sentido, la población objeto de estudio quedará conformada por los servicios web semánticos de datos georreferenciados.

#### **4.2. Muestra**

Es la parte de la población la cual se someterá a estudio, para Vieytes (2004: p.396) “la muestra hace referencia a una parte de la población o universo. Es decir, corresponde al conjunto de unidades de análisis extraídas del marco muestral para representar la conducta del universo en su conjunto”, en otras palabras, es un subconjunto y en derivación es representativa.

En este sentido, la investigadora realiza un muestreo intencionado o sesgado, que Según Tamayo y Tamayo (2007: p.178), el investigador selecciona los elementos que a su juicio son representativos, lo cual exige un conocimiento previo del fenómeno que se estudia. En correspondencia a lo anteriormente expuesto, los servicios web semánticos de datos

georreferenciados objeto de estudio contarán con todas o algunas de las siguientes particularidades:

Debe ser un servicio web, consultar el dominio de la información georreferenciada, interactuar con agentes de software inteligentes para consultar ontologías en ese dominio, o prestar ayuda a usuarios en consultas realizadas en lenguaje natural, es decir, permitirle al usuario hacer consultas en formato libre de texto. El número de unidades las cuales integran la muestra se representa de la siguiente manera:

**Cuadro 2:  
Distribución de la Muestra**

Nº	URI	Servicio Web de datos georreferenciados
1	Sitar..aragon.es/servicios-wcs.htm	SITAR Sistema de Información Territorial de Aragón
2	Visor.grafican.es/visorweb	SIT GOBIERNO DE CANARIAS
3	www.geoportal-idec.net/geoportal/cat/	IDEC CATALUNYA
4	Idena.navarra.es	IDENA Nodo de Infraestructura de datos espaciales de Navarra
5	www.cartografia.jcy.l.es	IDECyL CASTILLA Y LEÓN
6	www.ideextremadura.es/geoportal	IDE EXTREMADURA
7	www.ideib.cat	IDEIB ISLAS BALEARS
8	www.iderioja.larioja.org	IDE LA RIOJA
9	www.geo.euskadi.net/s69-15375/es	IDE EUSKADI
10	www.cartografia.princast.es/cartositpa/	SITPA IDEAS
11	www.madrid.org	IDE MADRID
12	terrasit.gva.es	TERRASIT
13	http://www.openstreetmap.org/	Open Street Map
14	http://www.osmgb.org.uk/	OSM-GB Open Street Map Gran Bretaña
15	http://www.geospatialplatform.gov	GEOSPATIAL PLATFORM
16	http://www.chennaiemergency.co.in	Chennaiemergency
17	http://linkedgeodata.org	Linkedgeodata
18	http://www.igac.gov.co/igac	Instituto Geográfico Agustín Codazzi
19	https://maps.google.co.ve/	Google Maps
20	http://www.flickr.com/	Flickr
21	http://instagram.com/	Instagram
22	http://www.bing.com/	Bing Microsoft
23	http://here.com/	NOKIA Maps



CONT... Cuadro 2

24	<a href="http://www.mapquest.com/">http://www.mapquest.com/</a>	MapQuest
25	<a href="http://www.waze.com/">http://www.waze.com/</a>	Waze
26	<a href="http://www.wikitude.com/">http://www.wikitude.com/</a>	Wikitude

**Fuente:** Piña (2013)

## 5. Técnicas e Instrumentos de Recolección de Datos

Según Fontaines (2012: p.146), la técnica hace emerger la información de los sujetos de estudio, y los instrumentos permiten su registro para un posterior análisis. Para tal efecto, como técnica se diseñó una Lista de cotejo, definida por el mismo Fontaines (2012: p.147) como aquella que permite identificar la presencia o ausencia del rasgo que se está estimando. Esta lista de cotejo es una encuesta que utiliza una escala dicotómica, de valores verdadero para la característica presente o falso para la ausente.

### 5.1. Instrumento

Como instrumento se utilizó una matriz de doble entrada en la que se registraron la presencia o ausencia de los indicadores de la variable de estudio seleccionada por el investigador. Por filas se señalaron los Servicios Web objeto de estudio y por columnas, cada uno de los ítems a consultar. En este apartado se resalta la característica fundamental de los 26 servicios Web o unidades de análisis inanimadas, los cuales hacen consultas a datos

georreferenciados. Otros criterios a evaluar fueron la arquitectura, tipos de consulta, configuración, especificación del servicio Web, uso de ontologías, agentes de software consultores y procesamiento de lenguaje natural, entre otros indicadores.

Según Hurtado (2000), la observación estructurada bajo enfoque cuantitativo, se realiza en aquellos casos en donde es preciso estudiar los patrones de comportamiento de la variable que se quieren medir, aumenta la precisión y objetividad. Por lo tanto, se utilizó una hoja de cálculo para registrar los resultados, que luego fueron interpretados. Según Fontaines (2012: p.146), existen instrumentos que no requieren de un estudio técnico de medición, el cual es la expresión de validez y confiabilidad. La lista de cotejo es uno de ellos. Aún así, la investigadora sometió a validación el instrumento.

## **5.2. Validez**

La validación del instrumento, se realizó a través del juicio de siete (7) expertos, quienes determinaron la pertinencia de los ítems dicotómicos con los objetivos de la investigación, redacción entre otros; según Pelekais y Belloso (2007), la investigadora debe seguir el procedimiento en el formato del instrumento para la validación: a) Título de Investigación, objetivos (general y específicos), cuadro de operacionalización de la variable y la disgregación de los ítems. b) Encuadrándole en un formato para escribir sus observaciones así

como las recomendaciones en cuanto a la pertinencia del reactivo, la redacción y formulación teórica con respeto a la variable, dimensión e indicador de la investigación. En este sentido la investigadora buscó la condensación del juicio de los expertos en cuanto al análisis de las discrepancias, coincidencias y recomendaciones, las cuales fueron incorporadas en la versión final del instrumento.

## **6. Análisis de Datos**

Para el procesamiento y tabulación se diseñó una matriz de doble entrada donde se asentaron los datos suministrados por las unidades de análisis en atención a la aplicación de una encuesta o lista de cotejo, empleando la estadística descriptiva en lo que respecta a la distribución frecuencial y percentil. Según Hurtado de B. (2005: p.129), para las investigaciones de tipo descriptivas el análisis de datos se realiza por frecuencias y porcentajes. Por su parte Fernández O. (2007: p.30) quien cita a Hurtado de B. (2000), señala que en una investigación descriptiva el análisis de datos se hace por separado para cada uno de los objetivos. Según Eyssautier (2008: p.246), el tipo de análisis de los datos es descriptivo. Luego de este análisis se obtienen las tendencias utilizando métodos estadísticos. En este caso se utilizaron porcentajes y frecuencias absolutas para cada uno de los ítems indagados.

En este sentido la información se mostró en tablas de distribución por dimensiones identificando el comportamiento de los indicadores para su posterior confrontación con los antecedentes y las teorías. Ulteriormente, la discusión de los resultados se elaboró por indicadores, dimensiones y variable, en base a las respuestas obtenidas con la aplicación del instrumento.

## **7. Procedimiento de la Investigación**

Al inicio de la actividad investigativa, la meta es obtener inestimables resultados, y para ello, es necesaria que la misma sea realizada de manera sistemática, secuencial y organizada, para llevar la estructura establecida con la finalidad de lograr efectos placenteros. A continuación, se exhiben las fases de abordaje del proceso de investigación:

1. Observación de la situación objeto de conocimiento, la cual permitió definir la problemática planteada, los objetivos, justificación y las delimitaciones de la investigación.

2. Posteriormente obtuvo la información a través textos, artículos científicos, entre otros, relacionados con la variable, los cuales soportan los antecedentes del estudio y sirven para la elaboración del marco teórico.

3. Consecutivamente, se construyó la Operacionalización la variable en investigación, se procedió al diseño del instrumento de recolección de datos en su versión preliminar.

4. Se aplicó la validez del instrumento a través del juicio de expertos.

5. Posteriormente, se procedió a la aplicación del instrumento final, donde se estructuró una tabla o matriz de doble entrada para el vaciado de los resultados obtenidos de la aplicación de la encuesta o lista de cotejo.

6. Se realizó un estudio ítem por ítem de los resultados obtenidos, por indicadores y dimensiones de la variable Web Semántica. En función de éste se formularon lineamientos teórico prácticos de los servicios Web semánticos actuales en la consulta de datos georreferenciados. Se utilizó un orden decreciente, mostrando primero las tendencias más observadas.

7. Finalmente se formularon las debidas conclusiones, recomendaciones.

## **CAPÍTULO IV**

### **ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS**

En el presente capítulo se hace una revisión de los resultados obtenidos luego de la revisión de los 26 servicios Web señalados en el Capítulo III. La autora presenta estos resultados contrastándolos con los planteamientos teóricos y hace un análisis en el que fija posición a favor o en contra de las teorías. Así mismo se señalan tendencias reflejadas en estos resultados que coinciden con las tendencias reseñadas en algunos de los antecedentes de esta investigación.

#### **1. Análisis y Discusión de los Resultados**

Los resultados de la primera dimensión de la Web Semántica: Arquitectura de la Web, se muestran en la Tabla 1. En ella se observa el comportamiento de la dimensión Arquitectura de la web, específicamente para el indicador Protocolo se realizaron 4 preguntas dicotómicas de SI o NO. Se obtuvieron los siguientes resultados para este indicador: en el ítem 1 se observa que el total de los servicios fueron modelados con una

Arquitectura Orientada a Servicios (SOA), que consiste en un modelo para construir aplicaciones que implementan soluciones interoperativas a través de servicios. En el ítem 2, se observa una tendencia del 38,5% en la arquitectura de los servicios estudiados, que realizan el acceso a recursos mediante REST (Transferencia de Estado Representacional), los cuales son servicios que han sido construidos como sistemas distribuidos con recursos expresados en URIs. En este ítem se observa que el 62,5% no utilizan dicho enfoque.

**Tabla 1**  
**Dimensión: Arquitectura de la web**

Indicador	Ítem		Sí		No	
			Fa	%	Fa	%
Protocolo	1	Fue modelado como una arquitectura orientada a servicios	26	100	0	0
	2	Realiza el acceso a recursos mediante REST	10	38,5	16	62,5
	3	Realiza el acceso a recursos mediante SOAP	2	7,7	24	92,3
	4	Carece de estado	26	100	0	0
Servicio	5	Se implementa como un Web Service	3	11,5	23	88,5
	6	Se implementa como un REST Service	10	38,5	16	61,5
	7	Se implementa como un Componente Distribuido	0	0	26	100
Integración	8	Está disponible a aplicaciones de terceros	25	96,2	1	3,8
	9	Es interoperativo con otros servicios	25	96,2	1	3,8

**Fuente:** Elaboración Propia (2013)

Ahora bien, en lo correspondiente a la pregunta 3, solo 2 de los servicios consultados, equivalentes al 7,7% realizan acceso a recursos mediante SOAP, aunque históricamente SOAP ha sido la manera usual de implementar los servicios y cuenta con los estándares formulados por la OGC, se evidencia

que ha entrado en desuso. En la pregunta 4, el total de los servicios carecen de estado, entendiéndose por esta característica ausente que estos servicios permiten el acceso concurrente de otras solicitudes de otros usuarios.

Para el indicador Protocolo, la autora observa que el comportamiento de los servicios estudiados, concuerda con los autores Manning (2008) y Gang (2009) quienes señalan la tendencia a construir servicios utilizando acceso a recursos mediante REST.

En lo correspondiente al indicador Servicio, los resultados proporcionados en el ítem 5 señalan que el 11,5% de los servicios indagados se implementan como Web Service, sólo tres del total, con un resultado de 23 que no utilizan esta tecnología, correspondiente al 88,5%. Mientras en el interrogante 6, el 38,5% de servicios son implementados como REST Service, igualmente se observa condensación en el cotejo al chequear el ítem 7, evidenciando que el 100 % de los servicios consultados no se implementan como componentes distribuidos, o por lo menos no lo expresan en su documentación.

La autora observa que de 26 servicios 10 se implementan como REST, 3 como Web Service y 13 como otro tipo de API que no expresan en su documentación. Estos resultados concuerdan con la tendencia expresada por Gang (2009).

En lo concerniente al indicador Integración, en el ítem 8 se observa que el 25 de los servicios están disponibles a aplicaciones de terceros e igualmente



son interoperativos con otros servicios, equivalente al 96,2%. Solo 1 no tenía esta característica o no se indicaba en su documentación.

Este indicador representa la tendencia a realizar aplicaciones interoperativas, pues la mayoría de los servicios estudiados cumplen con esta característica del estándar de la OGC, y con lo señalado por Aüer (2009), autor que señala la necesidad de integrar data e información, conjuntos de datos comprensibles y vocabularios. De esta manera se concluye con la revisión de la dimensión Arquitectura de la Web, la cual tiene un comportamiento que evidencia una tendencia a construir aplicaciones orientadas a servicio, las cuales faciliten los recursos a los usuarios, y bajo esquema de acceso a recursos REST, los que se han hecho más usados debido a la facilidad de solicitudes mediante URIs.

En la dimensión Tipos de Consultas, se observan tres indicadores que tienen que ver con las posibilidades de un usuario de realizar una pregunta, mediante la interfaz ofrecida. Estos indicadores se revisaron en 26 servicios y los resultados fueron los siguientes: Para el indicador de Tipos de Consultas Estructuradas, los ítems fueron 3. El ítem 10 obtuvo una frecuencia absoluta de 26, es decir, el 100% de los servicios permiten realizar consultas estructuradas simples, bien sea llenando un dato atómico en un formulario o seleccionando un dato de un menú desplegable.

**Tabla 2**  
**Dimensión: Tipos de Consultas**

Indicador	Ítem		Si		No	
			Fa	%	Fa	%
Estructuradas	10	Procesa consultas estructuradas simples	26	100	0	0
	11	Procesa consultas estructuradas complejas con uso de operadores lógicos	15	57,7	11	42,3
	12	Procesa consultas estructuradas complejas con uso de operadores geoespaciales	7	26,9	19	73,1
No estructuradas	13	Procesa consultas formuladas en lenguaje natural	0	0	26	100
	14	Procesa consultas formuladas en más de un lenguaje	16	61,5	10	38,5
Espaciales	15	Calcula automáticamente georreferencia del usuario	4	15,4	22	84,6
	16	Ofrece coberturas de datos georreferenciados a escala mundial	9	34,6	17	65,4
	17	Ofrece coberturas de datos georreferenciados a escala nacional	12	46,2	14	53,8
	18	Ofrece coberturas de datos georreferenciados a escala regional	26	100	0	0

**Fuente:** Elaboración Propia (2013)

El ítem 11 referido a consultas estructuradas complejas con uso de operadores lógicos obtuvo una frecuencia absoluta de 15, equivalente al 57.7% de los servicios estudiados. Esta tendencia indica que el 42.3% de los servicios disponibles no le permiten al usuario componer una pregunta más elaborada. En función de los resultados del ítem 12, sólo 7 servicios, el 26,9% le ofrece al usuario la posibilidad de consultar datos georreferenciados utilizando operadores espaciales. Esto evidencia que el 73,1% de los servicios son aplicaciones que no le ofrecen a sus usuarios realizar análisis espacial en sus consultas, por ejemplo, carencia en el cálculo de rutas, o de objetos más cercanos a alguna posición en un radio de, o intercepciones, entre otros operadores espaciales.

Los resultados concuerdan con el teórico Manning (2008), en el sentido de que la mayoría de los servicios de consulta estudiados lo hacen sobre datos que obedecen a alguna estructura. La autora agrega, la forma en la cual se

almacenan los datos así como la forma de consultarlos corresponden a una estructura. Pocos servicios (sólo el 26,9%) le ofrecen al usuario la posibilidad de utilizar operadores espaciales para componer sus consultas.

Para el indicador de Tipo de Consulta No Estructurada, los ítems 13 y 14 arrojaron los siguientes resultados. Ningún servicio procesa consultas formuladas en Lenguaje Natural, a pesar de las opiniones de teóricos como Antoniou (2008), Partyka (2011) y Kyong-ho (2010) lo sugieren, y en el antecedente de Piña (2006) se demuestra que es posible realizar una pregunta en formato libre de texto. Además, 16 de los servicios, el 61,5% le brindan al usuario la alternativa de más de un idioma para realizar su consulta. Eso sugiere el carácter de globalidad de la información disponible en estos servicios. La tendencia observada coincide con la postura de Klien (2008) quien señala evitar el procesamiento de lenguaje natural en función de avanzar al estadio de Web Semántica en los servicios, aunque la autora no la comparte.

Los resultados en el indicador de Tipo de Consultas Espaciales, son para el ítem 15, sólo 4 servicios calculan automáticamente la posición del usuario, esto es el 15,4%, e 84,6% no, aunque podrían solicitarla. Los 3 ítems siguientes se refieren al ámbito mundial, nacional y regional de la información ofrecida. Para el ítem 16, sólo 9, equivalente al 34,6% ofrece coberturas de datos georreferenciados a escala mundial; para el ítem 17, sólo 12 servicios, el 46,2% ofrece coberturas de datos georreferenciados a escala nacional, y en el ítem 18, el 100% de los servicios estudiados ofrecen coberturas de datos

georreferenciados a escala regional o local. Esto muestra que todos los servicios ofrecen en detalle algún territorio de interés.

Se observa entonces en función de los resultados obtenidos para la dimensión Tipos de Consulta, con un particular interés, la tendencia de los servicios a brindar un acceso estructurado simple o complejo a la información y que el 73,1% no hace uso de las aplicaciones ofrecidas por la OGC, referidas a incluir la posibilidad de uso de operadores espaciales en las consultas para realizar análisis espacial. La autora opina a favor de una tendencia a la baja de este porcentaje, mientras que los desarrolladores de los servicios incorporen estas facilidades de cálculo espacial a los servicios ofrecidos a sus usuarios.

Para la tercera dimensión de la Web Semántica: los Servicios Web, se hace entrega de los resultados en la Tabla 3 relacionados a sus dimensiones configuración, especificación, inteligencia artificial, y servicios asociados:

**Tabla 3**  
**Dimensión: Servicios Web**

Indicador	Ítem	Sí		No		
		Fa	%	Fa	%	
Configuración	19	Según su tipo es un WMS	15	57,7	11	42,3
	20	Según su tipo es un WFS	8	30,8	18	69,2
	21	Según su tipo es un WCS	4	15,4	22	84,6
	22	Según su tipo es un CSW	2	7,7	24	92,3
	23	Según su tipo es un WFS-Transaccional	0	0	26	100
	24	Según su tipo es un WFS-Gazetteer	1	3,8	25	96,2
	25	Según su tipo es un WPS	2	7,7	24	92,3
	26	Según su tipo es un WMTS	2	7,7	24	92,3
Especificación	27	Ofrece una especificación legible por máquinas de las funcionalidades del servicio	4	15,4	22	84,6
	28	Ofrece una descripción del proveedor del servicio	26	100	0	0
	29	Provee información adicional como tiempo de respuesta esperado	0	0	26	100
	30	Provee información acerca de restricciones geográficas	15	57,7	11	42,3
	31	Ofrece una especificación legible por humanos de las funcionalidades del servicio	26	100	0	0
Inteligencia Artificial	32	Usa ontologías	4	15,4	22	84,6
	33	Permite la interacción Usuario-Agente de software	3	11,5	23	88,5
	34	Permite la interacción Agente-Agente de software	4	15,4	22	84,6
Servicios Asociados	35	Autentica al usuario	13	50	13	50
	36	Permite consultar a usuarios no autenticados	25	96,2	1	3,8
	37	Presenta información oficial certificada	13	50	13	50
	38	Permite la edición gráfica del usuario	9	34,6	17	65,4
	39	Permite la edición tabular del usuario	9	34,6	17	65,4

**Fuente:** Elaboración Propia (2013)

En los ítems 19 al 26, se registró la configuración del servicio de acuerdo a los estándares de la OGC. La mayoría eran de tipo WMS con un 57,7%. Esta tendencia preferida por los desarrolladores de los servicios ofrece la

información geográfica, al usuario al tiempo que la protege de la edición o copia de la misma. Según el ítem 20, solo 8 eran WFS, equivalente al 30,8%, con este modo el usuario es capaz de tomar la información codificada ofrecida y generar del lado del cliente su imagen del mapa de la zona de interés.

Para el resto de los tipos de configuración se observa una decreciente preferencia, solo el 15,4% de los servicios eran de tipo WCS, en los que se le provee al usuario las capas o coberturas vectoriales para su edición y análisis; el 7,7% (sólo 2) eran de tipo CSW, es decir, que contenían un catálogo de metadatos. Ninguno era WFS-Transactional, sólo 1 (el 3,8%) era WFS-Gazeteer, 2 eran WPS, así como 2 eran WMTS.

Para el indicador Configuración, la autora señala que los tipos de servicios más populares fueron el WMS y WFS, y que ser de uno u otro tipo no es excluyente, i.e. un servicio Web puede ser clasificado como de más de un tipo de configuración mencionados.

La especificación de los Servicios Web es el siguiente indicador estudiado en el grupo de 26 servicios, y los resultados a las preguntas de la 27 a la 31, fueron los siguientes:

Para el ítem 27, sólo 4 de los 26 servicios, es decir el 15,4%, ofrece una especificación legible por máquinas de las funcionalidades del servicio. Esto evidencia que un 84,6% de los servicios aún no ofrecen la posibilidad de interoperar con otros agentes de software en funciones como el descubrimiento de los datos, o de mostrar su funcionalidad automáticamente a otros servicios; el total de los servicios ofrece una descripción del proveedor

del servicio; ningún servicio provee información adicional como tiempo de respuesta esperado; 15 servicios, el 57,7% provee información acerca de restricciones geográficas; y el total de los servicios ofrece una especificación legible por humanos de las funcionalidades del servicio.

Estos ítems constituyen el grado de contribución de dichos servicios a la Web Semántica, en los cuales se evidencia un estado incipiente de la tecnología utilizada, y una primera aproximación en la transición de la Web de documentos a la Web de significado que refiere Antoniou (2008). Se observa en estos resultados que aún los servicios están orientados a seres humanos, donde las descripciones son publicadas para estos en el 100% de los casos, dejando a un lado la interoperabilidad con agentes de software en una mayoría del 84,6% de los mismos.

El siguiente apartado lo constituyen los resultados del indicador Inteligencia Artificial, los cuales reflejan una medición sensible del grado de contribución de los servicios estudiados a la web Semántica. El ítem 32, en el que se consulta por el uso de ontologías, obtuvo una frecuencia absoluta de sólo 4 servicios, concuerda con los resultados del ítem 27. Con un valor similar, el ítem 33 en el que se consulta si el servicio permite la interacción Usuario-Agente de software, el resultado arrojado de la documentación es de sólo 3 servicios, i.e. el 88,5% no la permite; y en lo relacionado con el ítem 34, sólo el 15,4% de los servicios permiten la interacción Agente-Agente de software también.

Estos resultados demuestran una baja tendencia en la contribución de estos servicios a la Web Semántica. En contraposición a lo que dicen los autores McIlright (2003), Piña (2006), Antoniou (2008), Criado (2009), Gang (2009) y Aüer (2009), en la actualidad estos servicios en su minoría (15,4%) aprovechan las ontologías para compartir significado, y ofrecen información estructurada a agentes de software para integrarse con otras aplicaciones.

Para el indicador Servicios Asociados, se obtuvieron los siguientes resultados: en el ítem 35, el 50% de los servicios autentica al usuario el otro 50% no. En aquellos que si le solicitan una firma electrónica que identifique a quien realiza la consulta, es en los casos de desplegar información sensible como valor de propiedades inmobiliarias y nombres de los propietarios. Este valor concuerda con el resultado del ítem 37, pues el 50% de los servicios ofrecen información oficial certificada. El resultado del ítem 36, arroja que la mayoría de los servicios, a saber 25 que representan el 96,2%, le permite a cualquier usuario sin autenticación, consultar información no sensible. La información se muestra a la mayoría de usuarios anónimos, es decir, la mayoría de la información y datos georreferenciados son de carácter público.

Los ítems 38 y 39 se incluyeron sólo por curiosidad de la investigadora, pues si la investigación estaba enfocada en servicios de consulta de información, la edición gráfica y tabular no es sobreentendida. Aún así, el 34,6% de los servicios si le permiten a los usuarios realizar sus contribuciones a la Web, en cuanto a edición: creación de nuevos elementos gráficos,



corrección de los preexistentes, y en general modificación tabular de las características de la información ya almacenada.

En este Indicador se hicieron diferentes hallazgos en los servicios de datos georreferenciados relacionados al catastro español, que se mencionan a continuación.

De los 26 servicios estudiados, 12 fueron servicios autónomos de provincias españolas, representando estas el 23% de un total de 52, de los cuales todos ellos tienen configuración de WMS. Ofrecen información oficial de datos georreferenciados, relacionados al catastro regional de bienes y presentan una codificación para cada unidad vendible, descripción detallada descriptiva y gráfica. Así mismo, se reporta como hallazgo que cierta información proporcionada es restringida, i.e. que existen datos protegidos (titularidad y valor catastral) de los inmuebles y certificados telemáticos de los mismos.

Para el resto de la información, no se autentica al usuario, sin embargo se le ofrece una vía de acceso a una plataforma segura para acceder a los servicios electrónicos mediante la aplicación a través de formularios que la institución que la requiera, completa y envía a la espera de la certificación. Otra forma de acceso a esta información se le brinda a usuarios (personas) que siendo estos los titulares se deben autenticar a través de un certificado

Otro hallazgo, es el referido a que el 100% de las provincias españolas estudiadas, desarrollan y ofrecen sus servicios de forma autónoma al público en general, y a la hora de consultar datos protegidos, todos llevan a un servicio

catastral oficial centralizado. Entonces son servicios oficiales distribuidos que devienen en un servicio oficial centralizado.

Los resultados de este indicador, señalan un avance en dirección hacia la Web Semántica, una importante contribución de parte del estado español, que se apoya en los estándares desarrollados y que es un ejemplo a seguir a nivel mundial. A pesar de que solo 1 de las 12 provincias españolas utilizan repositorios de significados en ontologías, y que ofrece la descripción de su servicio para ser interpretado por máquinas, los resultados manifiestan una clara tendencia de trabajo orientado a fortalecer la Web Semántica y a contribuir con el sueño de Berners-Lee (1999).

## **CAPÍTULO V**

### **GUÍA METODOLÓGICA**

La Web Semántica es un proyecto cuyo objetivo es agregar inteligencia a la Web. El estudio que se presenta en esta tesis, realizado sobre una conceptualización particular de ese proyecto, ha permitido obtener la visión del estado actual que se presenta en el capítulo anterior. A partir de estos resultados se puede afirmar que es un proyecto en las primeras etapas de desarrollo, con algunas tecnologías básicas bien desarrolladas (como OWL, RDF, REST y SOAP) pero con muchas otras aún por consolidarse (como los estándares de OGC y el propio SOA). Se trata de otra de esas situaciones de la historia en la que una tecnología prometedora se perfila, pero aún no se consolida.

Esa situación justifica el presente capítulo, en que se intenta un ejercicio de especulación informada acerca de las tendencias que podrían definir el futuro del proyecto. Un ejercicio que pretende ofrecer algunas líneas metodológicas que pudieran servir a emprendedores tecnológicos para

definir su participación en la gesta tecnológica al tiempo que la aplican a la solución de problemas y a la prestación de servicios.

### **Tendencias**

Con ese propósito en mente, se presentan a continuación algunas tendencias tecnológicas generales que se procuran justificar deductivamente y apelando a casos particulares.

<Acá debes nombrar las tendencias de las que hablarás. Sugiero mencionar 1.- las ontologías como medio para mejorar la inteligencia de los sistemas de información web, 2.- como se resolverá la guerra de estándares es decir SOA, 3.- como afectará la web semántica a los sistemas y usuarios en esta parte del mundo>

#### **1. Lineamientos teórico – prácticos de servicios Web semánticos de consulta de datos georreferenciados**

Luego de revisar las teorías relacionadas con la variable de estudio formuladas por otros autores, la Web Semántica, y de aplicar un instrumento de valoración a 26 servicios Web, se obtuvieron los resultados antes mencionados, de los cuales la autora realiza un análisis. A continuación se presentan las tendencias en la elaboración de tales aplicaciones, y se contrasta con las teorías estudiadas.

- Se observó una tendencia del 100% de los servicios web a desplegar públicamente datos georreferenciados, junto con sus características asociadas. La visualización del espacio geográfico es un atractivo para los usuarios de estos servicios al tiempo que es un ejercicio de abstracción en la representación de la ubicación de las personas, y otros sitios de interés a nivel mundial.
- El 100% de los servicios carecen de estado. Esto significa que dichas aplicaciones pueden estar disponibles a sus usuarios de forma concurrente, i.e. varios usuarios accediendo juntos al mismo servicio de búsqueda de información.
- En 100% de los servicios Web se presentó información especializada así como detallada de algún ámbito local, un 46,2% ofrecieron información a escala nacional y un 34,6% en escala mundial. Tomando en cuenta que los de escala mundial, reciben contribuciones de usuarios editadas por los datos georreferenciados desde todos los rincones del planeta, se evidencia la premisa de Berners-Lee (1999: p.115) donde opina sobre el Web como una creación más social que técnica, la cual fue diseñada por su efecto social “para ayudar a la gente a trabajar junta” y no como un juguete técnico.
- En cuanto a la posibilidad que ofrecen estos servicios Web de interoperar con otros y del uso de sus datos públicos para ser usados por terceros, existe una alta tendencia del 96,2%. Lo que refleja el

compromiso con los estándares de interoperabilidad de la OGC, así como con la visión de Berners-Lee (1999: p.121) de lograr una infraestructura del Web para compartir contenido informativo, el cual pueda comunicarse a través de las capas de software, hardware y medios de transmisión subyacentes.

- En el 61,5% de los casos, los proveedores de los servicios Web de consulta a datos georreferenciados, incluyen la posibilidad de hacer la búsqueda en más de un idioma. Esto refleja que se están desarrollando servicios para proveer información local y remota a sus usuarios.
- Según el indicador de las consultas estructuradas, la tendencia de los servicios fue total a la hora de ofrecerle a sus usuarios una manera simple de hacer búsquedas, i.e. por valores atómicos. En cuanto a las búsquedas complejas sólo el 57,7% y apenas el 26,9% mediante operadores espaciales. Resultados que indican que se puede hacer más con este tipo de servicios, pues las herramientas están al alcance de todos y disponibles en la OGC.
- La configuración de los servicios obedece en un 57,7% de los casos a estándares de la OGC. En el 42,3% de los casos no se observa en la documentación el cumplimiento de los mismos, y esto puede significar que no los utilizan o que usan otros métodos para publicar sus contenidos en el Web.

- Se observó que el 50% de los sitios ofrecen acceso a datos sensibles propiedades como valor de inmuebles y sus dueños. Esto indica el interés de quienes administran el dominio de la información catastral en el desarrollo servicios web para atender a los ciudadanos.
- En relación al protocolo de comunicación de datos, para la construcción del servicio se observa preferencia por el REST (38,5%) sobre el SOAP (7,7%), como lo confirma Gang (2009). Debido a su simplicidad de acceso a recursos mediante URIs. Así como el modo de implementación del REST (38,5%) sobre el (11,5%) de los Web Services.
- Sólo el 15,4% de los servicios Web estudiados que consultan datos georreferenciados, se consideran Web semánticos, pues son los que utilizan ontologías para almacenar significados, y le permiten a agentes de software el descubrimiento de información y la comunicación entre ellos.
- Ninguno de los servicios web para consulta a datos georreferenciados le ofrece la posibilidad a sus usuarios de realizar la consulta en lenguaje natural. Coinciden estos resultados con Klien (2008) que sugiere evitar el Procesamiento del Lenguaje Natural (PLN). Esta ausencia en el PLN, podría cambiar si se implementaran servicios Web razonadores del lado del servidor, para atender las solicitudes del usuario. Logrando esto se

conseguiría una Web de significado, abierta a la interpretación de agentes de software que atienden las solicitudes de los seres humanos.



## CONCLUSIONES

Luego de analizar las teorías, antecedentes, casos de servicios Web semánticos y cumplir con los objetivos propuestos por el estudio, la autora llega a las siguientes conclusiones en el contexto de la Web Semántica en la consulta de datos georreferenciados:

Se identificó la arquitectura de la Web Semántica en la consulta de datos georreferenciados, reconociéndolas según el tipo de protocolo de comunicación de datos subyacente, su tipo de implementación, su capacidad para integrar aplicaciones con terceros, y el nivel de interoperabilidad con otros servicios. Las tendencias coinciden con las opiniones de otros autores como Pautasso (2008), Gang (2009) y Erl (2009); quienes señalan la preferencia (38,5% en la muestra estudiada) en el desarrollo de servicios Web utilizando arquitecturas tipo REST; y con teóricos como Aüer (2009) y Antoniou (2008) quienes plantean la necesidad de interoperatividad e integración de estos servicios Web, para los que se obtuvo una tendencia positiva del 96,2% de estas características.

Se describieron los tipos de consulta en la Web Semántica en datos georreferenciados, señalando aspectos de presencia o ausencia de estructura en las preguntas que realiza el usuario, con un

particular interés en el procesamiento de lenguaje natural. Por su carácter espacial, se indagó acerca del ámbito de la información ofrecida, el proceso de captura de la posición o ubicación del usuario. Como resultados se observaron conjuntos no excluyentes de servicios con estas características. El 100% de los servicios estudiados le ofrece al usuario consultas de tipo estructuradas, un 57,7% del tipo complejas mediante uso de operadores lógicos y un 7% mediante uso de operadores espaciales. Ningún servicio le permite al usuario formular su pregunta en lenguaje natural, coincidiendo con Klien (2008), aunque el 61,5% está habilitado para procesar más de un lenguaje. El total de los servicios tiene acceso a datos georreferenciados.

Se caracterizaron los servicios Web Semánticos en la consulta de datos georreferenciados, enumerándolos según su configuración, especificación, inteligencia artificial y servicios asociados, tomando en cuenta aspectos como la privacidad en el acceso a datos sensibles, el carácter oficial de la información, el uso de ontologías y la descripción de los servicios a disposición de humanos y máquinas. Entre las tendencias observadas el 57,5% de los servicios Web utilizan los estándares de la Open Gis Consortium (OGC), el 100% ofrece la especificación del servicio legible para humanos, sin embargo sólo el 4% utiliza ontologías como bases de conocimiento, así como este mismo porcentaje ofrece su descripción para lectura y descubrimiento por agentes de software.

En cuanto a los servicios asociados se concluyó que el 50% de los servicios estudiados ofrecen privacidad en datos sensibles e información

oficial de usuarios. El 96,2% permite el acceso a usuarios no autenticados a datos georreferenciados, adicionalmente, en un 34% de los servicios de consulta también se le ofrece la posibilidad de edición tabular y gráfica a sus usuarios.

Se formularon lineamientos teórico-prácticos de servicios web semánticos de consulta de datos georreferenciados, los cuales fueron obtenidos de acuerdo al análisis de los resultados, así como de la revisión por parte de la autora de las teorías y antecedentes planteados en relación a la variable objeto de estudio. Entre los cuales se concluye que existe una tendencia del 100% de la muestra a desarrollar servicios Web en el dominio de información georreferenciada, aun cuando en el 73,1% de éstos no se ponen a disposición del usuario herramientas de consulta y análisis espacial. Así mismo se refiere la existencia de una relación directa cruzada entre indicadores, a saber, los servicios (el 15,4%) que ofrecen una especificación legible por máquinas de las propias funcionalidades, también utilizan ontologías; el 50% de los servicios estudiados que autentican usuarios para consultar información sensible de propietarios o valores de inmuebles, son aquellos que ofrecen información oficial certificada.

Otros resultados relevantes de las tendencias fueron que todos los servicios estudiados fueron modelados con una arquitectura orientada a servicios, carentes de estado, y con procesamiento de consultas estructuradas simples. Así mismo se registró una tendencia negativa en todos los servicios en función al procesamiento de lenguaje natural.

Como conclusión al objetivo general, se analizó la Web semántica en servicios de consulta a datos georreferenciados, encontrando una contribución de servicios Web distribuidos a escala mundial, los cuales ponen en práctica los estándares y teorías actuales. Esta contribución ha revolucionado el acceso a datos mediante servicios, agregando información semántica a los contenidos que reposan en la Web actual. Aunque los resultados no evidencian el uso de procesamiento de lenguaje natural, interpretado por agentes de software, este sería un valor agregado, indispensable para completar el sueño de su visionario, Tim Berners-Lee.

## RECOMENDACIONES

La autora recomienda desarrollar aplicaciones de servicios Web semánticos de consulta a datos georreferenciados: a) con arquitectura orientada a servicio, b) bajo los lineamientos de los estándares propuestos por la OGC, c) con tecnologías REST por su sencillez en el acceso a recursos mediante URIs, d) carentes de estado para habilitar la concurrencia de múltiples usuarios, e) interoperativos e integrables a aplicaciones de terceros, f) con procesamiento de consultas estructuradas simples, complejas con uso de operadores lógicos y espaciales, g) que procesen consultas formuladas en lenguaje natural, h) que accedan a bases de conocimientos u ontologías habilitadas para varios idiomas, i) que ofrezcan la especificación de las funcionalidades del servicio tanto para usuarios humanos como para agentes de software, y finalmente j) que le permitan el acceso a datos oficiales sensibles bajo esquemas de autenticación, privacidad y protección a éstos.

Cada vez, las necesidades de información por parte de los usuarios se incrementan y el acceso mediante la Web actual, demanda un acercamiento a estas tecnologías que incorporan significado a los documentos en línea. El uso en pleno de las capacidades de un computador, se pone a prueba para

que éstos hagan más que poner datos al alcance de los humanos, como participar activamente en la interpretación de sus necesidades. Se recomienda a los desarrolladores de servicios Web semánticos, continuar con sus contribuciones descentralizadas, aportadas desde cualquier punto del planeta, pues estas apuntan a la consolidación de la Web Semántica.

Estas recomendaciones se hacen en función de las teorías estudiadas y los resultados obtenidos de la revisión de servicios Web semánticos de consulta a datos georreferenciados. Los antecedentes demostraron, cada uno con su contribución a la Web semántica, la factibilidad de incorporar estas recomendaciones en los desarrollos de aplicaciones de ese dominio de información geoespacial.

## REFERENCIAS BIBLIOGRÁFICAS

- Antoniou, G. & van Harmelen, F. A Semantic Web Primer (2008). The MIT Press. 2<sup>nd</sup>. Edition.
- Arapé, N. (2012). Arquitectura Orientada a Servicios basada en Software Libre para el acceso a Datos Geo-referenciados. Trabajo de Ascenso. Universidad del Zulia.
- Aüer, S.; Lehmann, J. y Hellmann, S. (2009). LinkedGeoData – Adding a Spatial Dimension to the Web of Data. International Semantic Web Conference. P. 731-746.
- Berners-Lee, T. (1999). Tejiendo la Red. Siglo veintiuno de España Editores. España.
- Berners-Lee T., Hendler J. & Lassila O. (2001). The Semantic Web, Scientific American.
- Cover, R. (1998). XML and Semantic Transparency. Technology Reports. Disponible para mayo de 2006 en:  
<http://www.oasis-open.org/cover/xmlAndSemantics.html>
- Criado, L. (2009). Procedimiento para transformar la Web en Web Semántica. Tesis doctoral de la Universidad Nacional de Educación a Distancia (UNED). España.

- Eyssautier, M. (2008). Metodología de la Investigación. 5ta Edición. Cengage Learning. México.
- Erl, T. (2009). *SOA design patterns*. Prentice Hall PTR.
- Fernández, I. (2007). Diccionario de Investigación. Quirón Ediciones. Caracas, Venezuela
- Fontaines, T. (2012). Metodología de la Investigación. Júpiter Editores, C.A. Caracas, Venezuela.
- Gang, Y. (2009). A Research on Semantic Geospatial Web Service Based REST. Foro internacional sobre aplicaciones y tecnología-ciencia en computación.
- Guarino, N. (2012). First Interdisciplinary Summer School on Ontological Analysis. 16-20 Julio 2012. Trento, Italia.
- Hernández, Fernández y Baptista (2008). Metodología de la Investigación. 4ta Edición. McGraw Hill.
- Hurtado de B., J. (2000). Metodología de la investigación holística. Fundación SYPAL. Caracas, Venezuela.
- Hurtado de B., J. (2005). Cómo formular objetivos de investigación. Quirón Ediciones. Caracas, Venezuela.
- International Association for Ontology and its Applications (IAOA) (2012). Disponible en: <http://iaoa.org/isc2012/index.php>
- Klien, E. (2008). Semantic Annotation of Geographic Information. Tesis doctoral de la Universidad de Múnster. Alemania.



Kuhn, W. (2005). Geospatial Semantics: Why, of What, and How? Journal on Data Semantics III Lecture Notes in Computer Science, Volume 3534/2005, 587, DOI: 10.1007/11496168\_1

Kyong-ho, L. y JongHyun, L. (2010). Constructing Composite Web Services from Natural Language Requests. Web Semantic Journal. Vol. 8 No.1.

Disponible en:

<http://www.websemanticsjournal.org/index.php/ps/article/view/211/206>

Leal, J. (2009). La autonomía del sujeto investigador y la metodología de Investigación. Edición Azul Intenso. 2da. Edición. Valencia. Venezuela.

Levinsohn, A. (2001). La Interoperabilidad geoespacial. Metadatos.

Manning, C. (2008). Introduction to the retrieval information. Cambridge

University Press. Disponible en:

<http://cesimo.ing.ula.ve/~jacinto/recuperacion/IR-cap-01.PDF>

Martin, D. et al (2007). Bringing semantic to web services with OWL-S.

Springer. Disponible en:

[http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1179199](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1179199)

McIlraith, S. et al (2003). Bringing semantic to web services. Intelligent Systems. IEEE. Disponible en:

[http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1179199](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1179199)

Newsam, S. et al (2008). Integrating gazetteers and remote sensed Imaginery.

OGC Estándares. Disponibles en: <http://www.opengeospatial.org/standards/is>

Pautasso, C. (2008). RESTful Web Services vs. “Big” Web Services: Making the Right Architectural Decision. 17<sup>th</sup> International World Wide Web Conference. Abril, 21 al 25 de 2008. Beijing, China. Disponible el día 4 de mayo de 2011 en:

<http://www.ra.ethz.ch/CDstore/www2008/www2008.org/papers/xhtml/fp179/fp179-pautasso.html>

Partyka, J., Kahn, L., Parveen, P., Thuraisingham y Shekhar, S. (2011). Enhanced Geographically-Typed Semantic Schema Matching. Web Semantic Journal. Volúmen 9, No.1. Disponible en:

[http://www.utdallas.edu/~bxt043000/Publications/Journal-Papers/DS-DMJ97\\_Enhanced%20Geographically%20typed%20Semantic%20Schema%20Matching.pdf](http://www.utdallas.edu/~bxt043000/Publications/Journal-Papers/DS-DMJ97_Enhanced%20Geographically%20typed%20Semantic%20Schema%20Matching.pdf)

Piña, N., Arapé, N. & Dávila, J. (2012). Ontología web semántica del registro catastral venezolano. *Revista Ciencia e Ingeniería*. Jornadas de Modelado y Simulación, 73-84. ISSN 1316-7081.

Piña, N. (2006). Ontología espacio temporal de registro catastral venezolano como base para la creación de sistemas de información territorial. Maestría, Universidad de los Andes.

Soto, J. (2010). Plataforma de geolocalización de centros de salud con tecnología móvil implementando el protocolo de comunicación HL7. Maestría, Universidad del Zulia. Venezuela.

- Stuckenschmidt, H. et al (2001). Modeling Land Transactions: Legal Ontologies in Contest. Disponible en: <http://www.lr.jur.uva.nl/jurix2001/papers/stuckensmitdt.pdf>
- Tamayo y Tamayo (2007). El proceso de la Investigación Científica, 4ta Edición, editorial Limusa, Mexico.
- Thuraisingham, B. (2010). Trustworthy Semantic Webs. University of Texas at Dallas. Disponible en: <http://www.utdallas.edu/~bxt043000/Teaching/CS-7301/Trustworthy-Sem-Webs-S2010/?C=D;O=A>
- Vieytes, R. (2004). Metodología de la investigación en organizaciones, mercado y sociedad. Editorial de las Ciencias.
- Zhou, C. et al (2004). Discovering personal gazetteers: an interactive clustering approach, GIS. P.266-273. Disponible en: [http://www-users.cs.umn.edu/~czhou/pub/acmgis04\\_final.pdf](http://www-users.cs.umn.edu/~czhou/pub/acmgis04_final.pdf)

## **ANEXOS**

**ANEXO 1:** FORMATO PARA LA VALIDEZ DE CONTENIDO DE LOS INSTRUMENTOS PERTENECIENTES A LA INVESTIGACIÓN

**ANEXO 2:** RESULTADOS DE LA APLICACIÓN DE LA ENCUESTA PARA LA VARIABLE

REPÚBLICA BOLIVARIANA DE VENEZUELA  
UNIVERSIDAD DR. RAFAEL BELLOSO CHACÍN  
VICERRECTORADO DE INVESTIGACION Y POSTGRADO  
DECANATO DE INVESTIGACION  
DOCTORADO EN CIENCIAS MENCION GERENCIA



FORMATO PARA LA VALIDEZ DE CONTENIDO DE LOS  
INSTRUMENTOS PERTENECIENTES A LA INVESTIGACIÓN TITULADA:

**WEB SEMÁNTICA EN LA CONSULTA DE DATOS  
GEORREFERENCIADOS**

**REALIZADO POR:  
NELCY PATRICIA PIÑA RIVERA**

Maracaibo, 14 de noviembre de 2012

Ciudadano

**NELSON ARAPÉ**

Su Despacho.-

La presente tiene por finalidad solicitar su colaboración para determinar la validez de contenido de los instrumentos de recolección de datos a ser aplicados en el estudio denominado **“WEB SEMÁNTICA EN LA CONSULTA DE DATOS GEORREFERENCIADOS”**.

Su valiosa ayuda consistirá en la evaluación de la pertinencia de cada una de las preguntas con los objetivos, variables, dimensiones, indicadores, y la redacción de las mismas.

Agradeciendo de antemano su valiosa colaboración, se despide de Usted,

Atentamente,

Nelcy Patricia Piña Rivera

CI.10.413.862

<b>Información del experto</b>	
Nombre:	
Cargo que desempeña:	
Profesión:	
Lugar de trabajo:	

## **IDENTIFICACIÓN DEL TRABAJO**

### **1. TÍTULO DEL TRABAJO**

Web semántica en la consulta de datos Georreferenciados

### **2. OBJETIVOS DEL TRABAJO**

#### **2.1. OBJETIVO GENERAL**

Analizar la web semántica en la consulta de datos georreferenciados.

#### **2.2. OBJETIVOS ESPECÍFICOS**

- Identificar la arquitectura de la web semántica en la consulta de datos georreferenciados.
- Describir los tipos de consulta de la web semántica en datos georreferenciados.
- Caracterizar los servicios web de la web semántica en la consulta de datos georreferenciados.
- Formular lineamientos teórico-prácticos en base al análisis de servicios web semánticos de consulta de datos georreferenciados.

### **3. SISTEMA DE VARIABLES**

#### **3.1. DEFINICIÓN NOMINAL DE LA VARIABLE**

Web Semántica

### **3.2. DEFINICIÓN CONCEPTUAL DE LA VARIABLE**

La Web Semántica es una red de datos que pueden ser procesados directa o indirectamente por máquinas, según Berners-Lee (1999). La Web Semántica y sus tecnologías habilitan a las personas a crear almacenes de datos en la web, construir vocabularios y escribir reglas para manipular datos (W3C:2013).

### **3.3. DEFINICIÓN OPERACIONAL DE LA VARIABLE**

La Web Semántica es la misma web tal como existe hoy en día, de la que millones de usuarios disponen como herramienta que ofrece servicios de consulta de información, entre otros usos, con una característica adicional: la participación de agentes de software que asisten las búsquedas de información de seres humanos mediante la interpretación de sus solicitudes y resultados. Este tipo de servicios semánticos, básicamente le otorgan a la web actual el adjetivo de web semántica.

Los desarrolladores de servicios de consulta de información en la web actual enfrentan el reto de vigilar la pertinencia, veracidad y actualidad de la información que manejan estas aplicaciones; en función de responder responsablemente al requerimiento del usuario y procurar sea satisfactoria su interacción con tal servicio. Así mismo, los desarrolladores deciden en qué forma los usuarios pueden realizar las consultas, al ofrecer servicios web de consultas estructuradas sobre algún dominio de información específico. Sin embargo, y haciendo hincapié en el asunto, el aspecto de servicio semántico solo se le otorga a la web si las máquinas tienen la capacidad de efectuar algún tipo de razonamiento sobre los datos, y colaborar en la interpretación de lo que el usuario solicitó.

Para lograr que los servicios en la Web sean semánticos, además de contar con las mencionadas características, los desarrolladores tendrán que superar ciertas dificultades como la obsolescencia de las plataformas de desarrollo, la multiplicidad de estándares que compiten; y verificar a través de mediciones, la utilidad y efectividad de estos servicios (usuario satisfecho), cumplir con el Habeas Data (usuario informado) y proveer la herramienta ideal para mejorar la toma de decisiones mediante la visualización de datos georreferenciados (característica apenas hasta ahora mencionada) con la que la web semántica arribaría a otro estadio de consultas de información.

La Web Semántica es ya un proyecto establecido, que convoca a un gran colectivo científico distribuido por todo el mundo. El objeto de trabajo es la propia World Wide Web y cualquier metodología para su desarrollo tiene que asumir su historia y la importante contribución del W3C en la definición de estándares para compartir las contribuciones y el propio contenido de la red.



### 3.4. CUADRO DE OPERACIONALIZACIÓN DE LA VARIABLE

Objetivo General	Analizar la web semántica en la consulta de datos georreferenciados.			
Objetivos Específicos	Variable	Dimensiones	Indicadores	Autores
Identificar la arquitectura de la web semántica en la consulta de datos georreferenciados.	Web Semántica	Arquitectura de la web	Protocolo	Pautasso (2008)
			Servicio	
			Integración	Antoniou (2008)
Describir los tipos de consulta de la web semántica en datos georreferenciados.		Tipos de Consultas	Estructuradas	Manning (2008)
			No estructuradas	
			Espaciales	OGC Estándares
Caracterizar los servicios web de la web semántica en la consulta de datos georreferenciados.	Servicios Web	Configuración	Antoniou (2008)	
		Especificación		
		Inteligencia Artificial	Piña (2006)	
Servicios Asociados				
Formular lineamientos teórico-prácticos de servicios web semánticos de consulta de datos georreferenciados.	Objetivo a alcanzar a través de la revisión teórica así como del análisis sistemático por parte de la investigadora por medio de los resultados			

**Fuente:** Piña (2013)

#### 4. TIPO DE INVESTIGACIÓN

Descriptiva Hernández, Fernández y Baptista (2008: p.102).

#### 5. DISEÑO DE LA INVESTIGACIÓN

El diseño de la investigación corresponde al ex post-facto o no experimental, transversal o sincrónica, de campo, según Vieytes (2004: p.118)

#### 6. POBLACIÓN OBJETO DE ESTUDIO

La población objeto de estudio quedará conformada por los servicios web semánticos de datos georreferenciados.

##### 6.1. MUESTRA

Se realiza un muestreo intencionado o sesgado, para lo cual según Tamayo y Tamayo (2007: p.178), el investigador selecciona los elementos que a su juicio son representativos, requiriendo un conocimiento previo del fenómeno que se estudia. De acuerdo a esto, los servicios web semánticos de datos

georreferenciados objeto de estudio podrían contar con todas o algunas de las siguientes particularidades:

Interacción con agentes de software inteligentes u otros servicios web semánticos.

Consultas dentro del dominio de la información georreferenciada.

Uso de ontologías.

Visualización de información georreferenciada.

Procesamiento de consultas formuladas en lenguaje natural.

Las unidades que integran la muestra se representan en el siguiente cuadro:

**Cuadro 1:  
Distribución de la Muestra**

Nº	URI	Servicio Web Semántico de datos georreferenciados
1	<a href="http://Sitar..aragon.es/servicios-wcs.htm">Sitar..aragon.es/servicios-wcs.htm</a>	SITAR Sistema de Información Territorial de Aragón
2	<a href="http://Visor.grafican.es/visorweb">Visor.grafican.es/visorweb</a>	SIT GOBIERNO DE CANARIAS
3	<a href="http://www.geoportal-idec.net/geoportal/cat/">www.geoportal-idec.net/geoportal/cat/</a>	IDEC CATALUNYA
4	<a href="http://Idena.navarra.es">Idena.navarra.es</a>	IDENA Nodo de Infraestructura de datos espaciales de Navarra
5	<a href="http://www.cartografia.jcyl.es">www.cartografia.jcyl.es</a>	IDECyL CASTILLA Y LEÓN
6	<a href="http://www.ideextremadura.es/geoportal">www.ideextremadura.es/geoportal</a>	IDE EXTREMADURA
7	<a href="http://www.ideib.cat">www.ideib.cat</a>	IDEIB ISLAS BALEARS
8	<a href="http://www.iderioja.larioja.org">www.iderioja.larioja.org</a>	IDE LA RIOJA
9	<a href="http://www.geo.euskadi.net/s69-15375/es">www.geo.euskadi.net/s69-15375/es</a>	IDE EUSKADI
10	<a href="http://www.cartografia.prncast.es/cartositpa/">www.cartografia.prncast.es/cartositpa/</a>	SITPA IDEAS
11	<a href="http://www.madrid.org">www.madrid.org</a>	IDE MADRID
12	<a href="http://terrasit.gva.es">terrasit.gva.es</a>	TERRASIT
13	<a href="http://www.openstreetmap.org/">http://www.openstreetmap.org/</a>	Open Street Map
14	<a href="http://www.osmgb.org.uk/">http://www.osmgb.org.uk/</a>	OSM-GB Open Street Map Gran Bretaña
15	<a href="http://www.geoplatform.gov">http://www.geoplatform.gov</a>	GEOSPATIAL PLATFORM
16	<a href="http://www.chennaiemergency.co.in">http://www.chennaiemergency.co.in</a>	Chennaiemergency
17	<a href="http://linkedgeodata.org">http://linkedgeodata.org</a>	Linkedgeodata
18	<a href="http://www.igac.gov.co/igac">http://www.igac.gov.co/igac</a>	Instituto Geográfico Agustín Codazzi
19	<a href="https://maps.google.co.ve/">https://maps.google.co.ve/</a>	Google Maps
20	<a href="http://www.flickr.com/">http://www.flickr.com/</a>	Flickr
21	<a href="http://instagram.com/">http://instagram.com/</a>	Instagram
22	<a href="http://www.bing.com/">http://www.bing.com/</a>	Bing Microsoft
23	<a href="http://here.com/">http://here.com/</a>	NOKIA Maps
24	<a href="http://www.mapquest.com/">http://www.mapquest.com/</a>	MapQuest
25	<a href="http://www.waze.com/">http://www.waze.com/</a>	Waze
26	<a href="http://www.wikitude.com/">http://www.wikitude.com/</a>	Wikitude

**Fuente:** Piña (2013)

## 7. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

Como técnica de recolección de datos se aplicará una encuesta sobre las unidades de análisis presentadas en el cuadro de Distribución de la Muestra, con el fin de evaluar los servicios web semánticos de consulta de datos georreferenciados.

## **7.1. INSTRUMENTO**

Como instrumento se utilizará una encuesta, de tipo dicotómica. Una lista verificable o check list, en la que se registrará la presencia o ausencia de los indicadores de la variable de estudio seleccionada por la investigadora. Según Hurtado (2000), la observación estructurada bajo enfoque cuantitativo, se realiza en aquellos casos en donde es preciso estudiar los patrones de comportamiento de la variable que se quieren medir; aumenta la precisión y objetividad.

## **7.2. VALIDEZ**

La validación del instrumento, se realizará a través del juicio de siete (7) expertos, quienes determinarán la pertinencia de los ítems dicotómicos con los objetivos de la investigación, redacción entre otros; según Pelekais y Belloso (2007), la investigadora debe seguir el siguiente procedimiento en el formato del instrumento para la validación: a) Título de Investigación, objetivos (general y específicos), cuadro de operacionalización de la variable y la disgregación de los ítems. b) Formato para escribir observaciones así como las recomendaciones en cuanto a la pertinencia de la pregunta, la redacción y formulación teórica con respeto a la variable, dimensión e indicador de la investigación. En este sentido la investigadora busca la condensación del juicio de los expertos en cuanto al análisis de las discrepancias, coincidencias y recomendaciones, las cuales serán incorporadas en la versión final del instrumento.

## **8. TÉCNICAS DE ANÁLISIS**

Para el procesamiento y tabulación se diseñará una matriz de doble entrada donde se asentaran los datos suministrados por las unidades de análisis en atención a la aplicación de la encuesta, lista verificable o check list, empleando la estadística descriptiva en lo que respecta a la distribución frecuencial percentil y al cálculo de las medias aritméticas a través del paquete estadístico computarizado. En este sentido la información se mostrará en tablas de distribución por dimensiones identificando el comportamiento de los indicadores para su posterior confrontación con los antecedentes y las teorías. Ulteriormente, la discusión de los resultados se elaborará a través de la distribución aritmética de las medias por indicadores, dimensiones y variable, siendo categorizadas según baremo diseñado previamente en base a

intervalos para análisis de las respuestas obtenidas con la aplicación del instrumento.











Resultados de la aplicación de la encuesta para la variable:																			
WEB SEMÁNTICA EN LA CONSULTA DE DATOS GEORREFERENCIADOS MARZO 2011 - ENERO 2013																			
Dimensión	ARQUITECTURA								TIPOS DE CONSULTA										
Integración	Protocolo				Servicio				Integración		Estruct			NoE		Espaciales			
N	SERVICIOS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	SITAR	1	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0	1
2	SIT GOBIERNO DE CANARIAS	1	0	0	1	0	0	0	1	1	1	0	1	0	1	0	0	0	1
3	IDEC CATALUNYA	1	0	1	1	1	0	0	1	1	1	1	0	0	1	0	0	0	1
4	IDENA	1	0	0	1	0	0	0	1	1	1	1	1	0	1	0	0	0	1
5	IDECyL	1	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0	1
6	IDE EXTREMADURA	1	0	0	1	0	0	0	1	1	1	1	1	0	0	0	0	0	1
7	IDEIB ISLAS BALEARS	1	0	0	1	0	0	0	1	1	1	0	0	0	1	0	0	0	1
8	IDE LA RIOJA	1	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0	1
9	IDE EUSKADI	1	0	0	1	0	0	0	1	1	1	1	0	0	1	0	0	0	1
10	SITPA IDEAS	1	0	0	1	0	0	0	1	1	1	0	0	0	1	0	0	0	1
11	IDE MADRID	1	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0	1
12	TERRASIT	1	0	0	1	0	0	0	1	1	1	0	0	0	1	0	0	0	1
13	Open Street Map	1	1	0	1	0	1	0	1	1	1	0	0	0	0	0	1	1	1
14	Open Street Map GB	1	0	0	1	0	0	0	1	1	1	0	0	0	0	0	0	1	1
15	GEOSPATIAL PLATFORM	1	0	0	1	0	0	0	1	1	1	0	0	0	0	0	0	0	1
16	CHENNAIEMERGENCY	1	0	0	1	0	0	0	0	0	1	1	1	0	0	0	0	0	1
17	Linkedgeodata	1	1	0	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1
18	IGAC	1	0	0	1	0	0	0	1	1	1	1	0	0	1	0	0	1	1
19	Google Maps	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1
20	FLICKR	1	1	1	1	1	1	0	1	1	1	1	0	0	1	0	1	1	1
21	Instagram	1	1	0	1	0	1	0	1	1	1	0	0	0	1	1	1	1	1
22	Bing Microsoft	1	1	0	1	1	1	0	1	1	1	1	0	0	1	0	1	1	1
23	NOKIA Maps	1	1	0	1	0	1	0	1	1	1	0	0	0	1	0	1	1	1
24	MapQuest	1	1	0	1	0	1	0	1	1	1	0	0	0	0	0	0	1	1
25	Waze	1	1	0	1	0	1	0	1	1	1	0	0	0	1	1	1	1	1
26	Wikitude	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1
	Frecuencias Absolutas	26	10	2	26	3	10	0	25	25	26	15	7	0	16	4	9	12	26
	porcentajes %	100	38,5	7,7	100	11,5	38,5	0,0	96,2	96,2	100	57,7	26,9	0,0	61,5	15,4	34,6	46,2	100
	Frecuencias Absolutas negativas	0	16	24	0	23	16	26	1	1	0	11	19	26	10	22	17	14	0
	porcentajes negativos %	0,0	61,5	92,3	0,0	88,5	61,5	100	3,8	3,8	0,0	42,3	73,1	100	38,5	84,6	65,4	53,8	0,0

Resultados de la aplicación de la encuesta para la variable:																						
WEB SEMÁNTICA EN LA CONSULTA DE DATOS GEORREFERENCIADOS MARZO 2011 - ENERO 2013																						
Dimensión	SERVICIOS WEB																					
	Integración	Configuración										Especificación					IA		Serv. Asoc.			
N	SERVICIOS	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
1	SITAR	1	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	1	0	0
2	SIT GOBIERNO DE CANARIAS	1	1	0	0	0	0	1	0	0	1	0	1	1	0	0	0	0	1	1	0	0
3	IDEC CATALUNYA	1	1	1	1	0	0	1	1	0	1	0	1	1	0	0	0	1	1	1	0	0
4	IDENA	1	1	0	1	0	0	0	0	1	1	0	1	1	1	0	1	1	1	1	0	0
5	IDECyL	1	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	1	1	1	0	0
6	IDE EXTREMADURA	1	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	1	0	0
7	IDEIB ISLAS BALEARS	1	1	0	0	0	1	0	0	0	1	0	1	1	0	0	0	1	1	1	0	0
8	IDE LA RIOJA	1	1	1	0	0	0	0	0	0	1	0	1	1	0	0	0	1	1	1	0	0
9	IDE EUSKADI	1	0	1	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	1	0	0
10	SITPA IDEAS	1	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	1	1	1	0	0
11	IDE MADRID	1	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	1	1	1	0	0
12	TERRASIT	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	0
13	Open Street Map	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1	1	1	0	1	1
14	Open Street Map GB	1	1	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	0	1	1
15	GEOSPATIAL PLATFORM	1	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	0	0	0
16	CHENNAIEMERGENCY	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	0	0	0
17	Linkedgeodata	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1	0	1	0	1	1
18	IGAC	1	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	1	1	1	0	0
19	Google Maps	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0
20	FLICKR	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	0	1	1
21	Instagram	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	0	1	1
22	Bing Microsoft	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0
23	NOKIA Maps	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	1
24	MapQuest	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	1
25	Waze	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	1
26	Wikitude	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	1
	Frecuencias Absolutas	15	8	4	2	0	1	2	2	4	26	0	15	26	4	3	4	13	25	13	9	9
	porcentajes %	57,7	30,8	15,4	7,7	0,0	3,8	7,7	7,7	15,4	100	0,0	57,7	100	15,4	11,5	15,4	50,0	96,2	50,0	34,6	34,6
	Frecuencias Absolutas negativas	11	18	22	24	26	25	24	24	22	0	26	11	0	22	23	22	13	1	13	17	17
	porcentajes negativos %	42,3	69,2	84,6	92,3	100	96,2	92,3	92,3	84,6	0,0	100	42,3	0,0	84,6	88,5	84,6	50,0	3,8	50,0	65,4	65,4