

2012



UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS

Facultad 6

**Trabajo de diploma para optar por el título de Ingeniero en
Ciencias Informáticas**

Título: “Diseño de una geo-ontología para el sistema de información geográfica en la Universidad de las Ciencias Informáticas.”

Autor:

Yelenis Díaz Larquin

Tutor:

Ing. Adrian Gracia Aguila

La Habana, junio del 2012

“Año 54 de la Revolución”



“La idea es convertir la informática en una de las ramas más productivas y aportadoras de recursos para la nación. Es el empleo a fondo de la inteligencia y del capital humano que tenemos y principalmente del que podemos crear casi como espina dorsal de la economía.”

A handwritten signature in black ink, which appears to be 'Fidel Castro', written in a cursive style.

Declaración de Autoría

Declaro que soy la única autora de este trabajo y autorizo a la Universidad de las Ciencias Informáticas a hacer uso del mismo, con carácter exclusivo.

Para que así conste firmo la presente a los ____ días del mes de _____ del año _____.

Yelenis Díaz Larquin

Adrian Gracia Aguila

Autora

Tutor

Datos de Contactos

Ing. Adrian Gracia Aguila

Graduado en la Universidad de las Ciencias informáticas. Se ha desempeñado como profesor y actualmente ocupa el cargo de Jefe del proyecto Aplicativos SIG. Tiene experiencia laboral en la línea de desarrollo de los SIG.

Correo electrónico: agracia@uci.cu

Agradecimientos

Agradezco infinitamente a mi mamá que ha sido mi todo desde que vine al mundo, me ha brindado todo el amor, el cariño, las fuerzas y el ánimo que he necesitado para seguir adelante y llegar a la meta. Y como ella siempre me ha dicho "Todo lo que uno se proponga en la vida lo logra" y con esa mente positiva que siempre me alumbró en cada examen en los cuales ella compartía sentimentalmente conmigo, pude salir adelante y hacer nuestro sueño realidad. Mil gracias por todo el sacrificio y la dedicación mami.

A papi y a villaya (mis abuelos) que conjuntamente con mi mamá me han apoyado en todo, sufriendo cada día la distancia que nos separa, pero siempre deseándome los mayores éxitos del mundo. A ellos y a mi mamá les debo la persona que soy hoy. Por eso este triunfo es de ellos también.

A mi papá y hermanos, tíos y primos que siempre me tuvieron presentes y tan solo con una llamada telefónica me alegraban el corazón. A mi tía Oxana que en estos años de la carrera ha sido como una madre para mí y conté siempre con su apoyo.

A mi novio Carlos que siempre estuvo a mi lado en los momentos que más lo necesité, brindándome todo su amor, comprensión y ayuda.

Ha sido mi ángel guardián todos estos años

A mi equipo VIP (Jane y Nana) que más que mis amigas de siempre, se han convertido en mis hermanas. En los momentos que más las necesité estuvieron ahí, tratando de alegrarme o buscando una solución. Por eso solo me queda decirles que nunca las olvidaré y siempre estarán en mi corazón.

A mis hijos postizos (El dany, Berna, Andy, y Tati (Rober)) que igual siempre me brindaron la mano amiga en todo lo que necesité, un consejo, un favor. Compartimos varios cumpleaños donde me hacían olvidar todas las preocupaciones. Formamos un gran equipo.

Mis amistades del IPVE (Liset y Jose Carlos) que aunque no estábamos muy cerca siempre estuvieron presentes, y dispuestos a brindarme su apoyo, confianza y sobre todo buenos consejos para la vida.

Mis amigos Aylin, Roe y Yanio que desde que entré a la universidad el 1er día, me encaminaron y ayudaron en todo. Siendo muy serviciales y cariñosos conmigo y con mi familia.

A Nurka y demás familiares de Carlos que siempre estuvieron allí apoyándonos.

A Miguel y a Mirna que siempre pudimos contar con su apoyo, nos dieron siempre muy buenos, útiles y oportunos consejos para la vida.

Gracias a todos los que de una forma u otro posibilitaron que hiciera este sueño realidad.

Dedicatoria

Este triunfo de mi vida profesional se lo dedico especialmente a mi mamá y a mis abuelos que han sido los protagonistas principales y las personas que más amor me han brindado. Gracias por todo.

Resumen

Hoy día el enorme volumen de datos contenidos en la web dificulta la navegación. A partir de la introducción y utilización de las geo-ontologías en los sistemas de información geográfica, el usuario podrá obtener, de una manera más fácil y automática lo que necesita. Las búsquedas geo-ontológicas basadas en esquemas conceptuales generan resultados coherentes con el criterio del usuario sobre uno o varios dominios de información. Todo ello depende de que la información se describa mediante un mecanismo de representación del conocimiento donde los datos se relacionen entre sí, y se establezca un lenguaje de definición común. Por lo antes mencionado se tiene como objetivo de este trabajo crear una geo-ontología para el SIGUCI. Se especifica el lenguaje y herramienta utilizada en su diseño, la metodología de desarrollo y la ontología de dominio que se empleó como base. Con el apoyo de los razonadores que ofrece el *Protégé* y el validador de la W3C se valida la consistencia de la geo-ontología.

Palabras Claves: geo-ontología, interoperabilidad semántica, ontología, sistema de información geográfica y web semántica.

Índice

Introducción	14
Capítulo 1: Fundamentación Teórica	19
1.1 Introducción	19
1.2 Definiciones y características de los términos usados en la investigación	19
1.2.1 Ontologías	19
1.2.2 Geo-ontología	24
1.2.3 Sistema de información geográfica	25
1.2.4 Web semántica	27
1.2.5 Interoperabilidad semántica	29
1.3. Herramientas para construir la geo-ontología	31
1.3.1 Ontolingua	31
1.3.2 Chimaera	31
1.3.3 Protégé	31
1.4 Metodologías usadas para el desarrollo de geo-ontologías	33
1.4.1 Methontology	34
1.4.2 On-to-Knowledge	34
1.4.3 GEONTO – MET	35
1.5 Lenguaje de representación de geo-ontologías	35
1.6. Conclusiones parciales	38
Capítulo 2: Diseño de la Geo-ontología	39
2.1 Introducción	39
2.2 Ontología de dominio para diseñar la geo-ontología para el SIGUCI	39
2.3 Pasos a seguir para la construcción de la geo-ontología	40
2.3.1 Construcción del glosario de término	40
2.3.2 Construcción de la taxonomía de conceptos	42

2.3.3 Describir relaciones binarias.....	43
2.3.4 Propiedades de datos (Data properties).....	46
2.3.5 Describir axiomas formales.....	47
2.3.6 Describir instancias.....	48
2.4 Conclusiones parciales	48
Capítulo 3: Análisis de criterios para la evaluación de la geo-ontología.....	49
3.1 Introducción.....	49
3.2 Validación mediante el Servicio de Validación Web de la W3C.....	49
3.3 Razonadores semánticos.....	50
3.3.1 Razonador FaCT++.....	50
3.3.2 Razonador Hermit 1.3.5	51
3.4 Pruebas de Razonamiento	51
3.5 Comparación de los resultados por ambos razonadores.....	58
3.6 Conclusiones parciales	58
Conclusiones.....	59
Recomendaciones	60
Bibliografía Citada.....	61
Bibliografía Consultada	¡Error! Marcador no definido.
Anexos	65

Índice de Tablas

Tabla 1: Glosario de términos.....	42
Tabla 2: Taxonomía de conceptos.	43
Tabla 3: Propiedades de objetos (object properties).	45

Índice de figuras

Figura 1: Tipos de ontologías (Vera, 2009).....	23
Figura 2: Niveles del modelo de interoperabilidad conceptual (2011).	30
Figura 3: Taxonomía de conceptos.....	43
Figura 4: Relaciones binarias.	45
Figura 5: Relaciones entre conceptos (Data properties).....	46
Figura 6: Axiomas formales.	47
Figura 7: Instancias.	48
Figura 8: Servicio de Validación de Ontologías.	49
Figura 9: Resultado de la geo-ontología por el Validador W3C.	50
Figura 10: Inferencia de datos con razonadores semánticos (1).	52
Figura 11: Inferencia de datos con razonadores semánticos (2).	53
Figura 12: Inferencia de datos con razonadores semánticos (3).	54
Figura 13: Inferencia de datos con razonadores semánticos (4).	55
Figura 14: Inferencia de datos con razonadores semánticos (5).	56
Figura 15: Inferencia de datos con razonadores semánticos (6).	57

Introducción

Hoy día los contenidos que se manejan en Internet son inmensurables, necesitando integrarle diferentes tecnologías y formatos existentes para lograr que los datos sean fácilmente accesibles y enlazables como las páginas webs actuales. La web actual, describe la evolución del uso y la interacción en Internet, esta viene evolucionando y en busca de mejoras paulatinamente, alcanzando gran interés por parte de sus usuarios, aún cuando el conjunto enorme de páginas contienen texto no estructurado, es decir, texto cuyo contenido no está caracterizado. A su vez afronta dificultades a la hora de manejar y recuperar la gran cantidad de información que procesa. En los resultados obtenidos se detecta la falta de precisión que presenta un buscador cuando se encuentran páginas que no tienen relación alguna con la necesidad informativa, esto sucede cuando existen palabras que poseen varios significados. La falta de exhaustividad puede venir provocada por la utilización de un sinónimo en una página en lugar de la palabra empleada en la consulta.

Como forma de solucionar estos problemas Tim Berners-Lee propone la creación de la web semántica, donde las aplicaciones son capaces de efectuar un procesamiento de la información mucho más profundo. Esta se caracteriza por incorporar programas capaces de comprender y relacionar el contenido de las páginas web, de procesarla, de discriminar la más fiable en un momento dado e incluso de deducir o inferir información no registrada previamente, tomando decisiones con un cierto grado de autonomía. Para lograr aplicaciones y servicios más inteligentes es necesario que la información esté perfectamente descrita y clasificada de manera que su significado exacto esté al alcance de las máquinas. De esta manera los ordenadores podrían manipular y procesar la información adecuadamente y los usuarios quedarán satisfechos con los resultados obtenidos en la búsqueda.

A medida que la web se convierte día a día en un canal de comunicación importante y ofrece mayores posibilidades para transmitir y recibir todo tipo de información, los sistemas de información geográfica (SIG) se complementan con este desarrollo, y en consecuencia, reemplazando sus medios de trabajo por medios de comunicación interactivos a través de la red. La representación geoespacial y la información geográfica se viene desarrollando desde años atrás. La cartografía y la elaboración de mapas han constituido una herramienta de precisión donde las matemáticas y las proyecciones de una esfera (o geoide) sobre un plano han guiado con éxito la confección de mapas básicos y temáticos, por lo que es posible afirmar que la confección de mapas es una ciencia

madura. Esta unión semántica es la base para la obtención de una verdadera interoperabilidad e intercambio de los datos geoespaciales entre diferentes usuarios (Garea).

En la actualidad, existe un enorme volumen de datos recopilados acerca de la Tierra, no solamente por los nuevos sistemas de información espaciales, sino también por las tecnologías de recolección de datos cada vez más sofisticadas. Este escenario ha suscitado en los últimos años una gran preocupación en la comunidad científica por el desarrollo de tecnologías para su integración y de herramientas para su manejo y análisis. Los SIG recogen gran parte de estos datos, desde su creación han sido utilizados para la resolución de problemas de planificación territorial, medioambientales, geográficos y de manipulación (Garea).

Desde los años 90, los SIG han incrementado su uso en diversos campos y aplicaciones, apreciándose también su potencial para la evaluación de actividades humanas y relaciones sociales. Estos son sistemas que facilitan la gestión y la representación de datos espaciales, modelando comportamientos de actividades en un contexto espacial. Los SIG actualmente son de uso intensivo de datos, multidisciplinarios, dinámicos y complejos que basan su aproximación en un modelo ambiental de la realidad (García, 2000).

Las expectativas creadas sobre los SIG están presentes en Cuba con sus correspondientes limitaciones e incongruencias. Los usuarios de los SIG en Cuba no son los cibernéticos o especialistas informáticos, la mayoría son geólogos, cartógrafos, geógrafos, desarrolladores, arquitectos e ingenieros, quienes conocen y operan sobre ellos en sus investigaciones y proyectos. Muchos de estos especialistas han aprendido a manejar los SIG de forma autodidacta, sin un adiestramiento previo y trabajan en organizaciones oficiales del gobierno, tales como GEOCUBA, Ministerio de Ciencia, Tecnología y Medio Ambiente, Instituto de Planificación Física, Ministerio de las Fuerzas Armadas Revolucionarias y universidades.

El trabajo con los SIG se dificulta a raíz de los términos que se manejan en ellos y que los usuarios en ocasiones no son capaces de comprender ni relacionar, como vía de solución y para facilitar el trabajo con los SIG surgen las geo-ontologías. Estas son un intento por definir el mundo de tal forma que todos los subsistemas que trabajan sobre él tengan una visión común del mismo. La razón de ser de una geo-ontología es poder comunicarse, por tanto es necesario emplear lenguajes que permitan representar el conocimiento bajo una base de definición común.

Los términos antes expuestos se pueden centralizar en los Sistemas de Información Geográfica Gobernados por Ontología (SIGGO) en los cuales las geo-ontologías son un componente más, como lo es la base de datos temáticos o espaciales que interviene y coopera de la misma manera para alcanzar los objetivos para los cuales fueron creados los SIG. Los SIGGO pueden constituir la base tecnológica para lograr una verdadera interoperabilidad semántica entre las diferentes comunidades que utilizan los datos espaciales en el proceso de toma de decisiones (Garea).

En Cuba se ha venido trabajando en función de desarrollar geo-ontologías con el fin de concentrar información manipulada de forma tal que sea una motivación para los usuarios. Ejemplo de ello y como solución para la construcción del modelo conceptual ontológico del Dominio Web de la Prensa Cubana para el Caso de los Cinco Prisioneros Políticos en Estados Unidos es la ontología de Dominio Web para el caso “Los Cinco” (Barceló).

La Universidad de las Ciencias Informáticas (UCI) creada al calor de la Batalla de Ideas por el Comandante en Jefe, con el objetivo de crear profesionales altamente comprometidos con la revolución y la informatización del país, se mantiene en constante desarrollo de diferentes productos. Aplicativos SIG perteneciente al Centro de Desarrollo GEYSED de la Facultad 6, tiene la función principal de desarrollar SIG dedicados a las diferentes esferas dentro o fuera del país tales como: la salud, el transporte, recursos hidráulicos, meteorología y otras. Uno de sus productos es el SIGUCI, encaminado a realizar la representación y análisis geoespacial de la información geográfica referente a objetivos socioeconómicos de la UCI, como edificios docentes, de residencia, plazas, manzanas, consultorios, cajeros, y otros. Los estudiantes, trabajadores y profesores de la universidad podrán contar con una aplicación de características web, sostenida sobre una plataforma de desarrollo libre.

Generalmente en los productos realizados en la UCI como en otras empresas de Cuba y del mundo en las representaciones geoespaciales entre mapas y bases geográficas, no existe un común acuerdo en la representación de la semántica espacial. Diferentes organizaciones dibujan con exactitud determinadas líneas, puntos o polígonos sobre un plano para representar ciudades, pozos de abasto de agua, puntos de la red altimétrica, líneas de transmisión eléctrica y viales. Pero no se ponen de acuerdo sobre el significado semántico u ontológico de estos trazos (Garea).

Las personas, organizaciones y sistemas de software que forman parte del trabajo con los SIG deben comunicarse entre sí para alcanzar los objetivos que se plantean en sus áreas de trabajo. Sin embargo, debido a diferentes necesidades, pueden variar ampliamente los puntos de vista y la

interrelación entre ellos. Por lo general, cada uno de estos actores posee un lenguaje propio que da lugar a diferencias de conceptos, estructuras y métodos.

Los SIG son de gran ayuda para el manejo de información geográfica en el mundo, pero distan de ser completamente precisos por no obedecer a una jerarquía única de conocimiento, mas poseen características que proveen una clasificación útil para determinar requerimientos (Medina, 2007). El SIGUCI como una representación espacial no está aislado de este problema, no posee en su integración de datos geográficos una consistencia en el significado terminológico que maneja, la información semántica existente que se encuentra embebida en los datos almacenados no se explota y no logra minimizar la complejidad de navegación. Además se ha podido apreciar que existen datos repetidos con un mismo significado, lo que provoca que los usuarios se sientan un poco desorientados al tener la información de los datos algo dispersa. La situación no es cómo representar con exactitud un elemento geográfico, sino que dos mapas o bases de datos geográficas representen lo mismo o tengan una unidad semántica común (Larin).

Por lo antes expuesto se define como **problema a resolver** ¿Cómo lograr la interoperabilidad semántica del sistema de información geográfica de la UCI? Para darle solución al problema se tiene como **objetivo general** diseñar una geo-ontología para el sistema de información geográfica de la UCI (SIGUCI) en el proyecto Aplicativos SIG, teniendo como **objeto de estudio** interoperabilidad semántica de la información geográfica y para centralizar la investigación se tiene como **campo de acción** interoperabilidad semántica de la información en el sistema de información geográfica de la UCI.

Como **idea a defender** se plantea:

El diseño de una geo-ontología para el tratamiento y razonamiento de los conceptos relacionados en el SIGUCI, favorecerá el manejo del conocimiento y la interoperabilidad semántica de la información.

Las tareas de investigación que se proponen son las siguientes:

- Analizar el estado del arte relacionado con las geo-ontologías como un sub-conjunto de las ontologías.
- Definir el lenguaje de representación formal de una geo-ontología de software para el SIGUCI.
- Definir la metodología y las herramientas a utilizar para el desarrollo de la geo-ontología.

- Desarrollar la geo-ontología para palabras claves a partir de conceptos reutilizados de geo-ontologías existentes y nuevos conceptos identificados.
- Evaluar la geo-ontología desarrollada.

Lo que se espera como **resultado de la investigación** es una Geo-Ontología para el SIGUCI.

Para el desarrollo de las tareas científicas se utilizan diferentes métodos de investigación en la búsqueda y procesamiento de la información, permitiendo su rapidez y eficiencia, éstos son: la **Modelación** para el diseño y creación de la geo-ontología, el **Hipotético–deductivo** para conocer el funcionamiento del SIGUCI y las posibilidades existentes de modelarle una geo-ontología y el **Histórico-lógico** para hacer un análisis retrospectivo sobre el desarrollo, utilización y alcance de las geo-ontologías en el mundo de la informática.

La estructura del presente documento se compone de tres capítulos en los que se recoge la información obtenida desde el punto de vista investigativo. Además de lo referente al análisis y diseño de la geo-ontología para el SIGUCI. **En el CAPÍTULO 1:** Fundamentación Teórica, se aborda todo lo relacionado con la fundamentación teórica que sustenta la presente investigación. Se define el concepto de ontología y se indican los distintos tipos de ontologías existentes, así como otros conceptos relacionados con la misma. También se exponen las tendencias, técnicas, tecnologías, métodos y herramientas usadas durante la investigación y construcción de la solución propuesta. Se define además la metodología y el lenguaje de representación a usar. **En el CAPÍTULO 2:** Diseño de la geo-ontología, se abordan los temas relacionados al proceso de su construcción, definiendo la taxonomía de conceptos, las relaciones entre ellos, axiomas e instancias. **En el CAPÍTULO 3:** Análisis de criterios para la evaluación de la geo-ontología, se aplican pruebas de razonamiento a la geo-ontología para comprobar su consistencia y nivel de inferencia de conocimiento.

Capítulo 1: Fundamentación Teórica

1.1 Introducción

En este capítulo se abordan los principales conceptos relacionados con las geo-ontologías, así como los tipos de ontologías existentes, sus componentes y ventajas. Se define la herramienta, la metodología y el lenguaje a utilizar en el diseño de la misma. También se tratan elementos de los SIG y de otros conceptos relacionados con la investigación.

1.2 Definiciones y características de los términos usados en la investigación

1.2.1 Ontologías

El término ontología lo introdujo Thomasius y Wolf, en el siglo XVII, para designar precisamente, en la filosofía primera de Aristóteles, la parte de la Metafísica que estudia el ser en cuanto tal. Sin embargo, el origen epistemológico del término se encuentra mucho antes, en la disciplina filosófica surgida con Parménides (h. 540-h. 450 a.C.), cuyo objetivo era también, el estudio del ser (Ramírez).

“La ontología se ocupa de las categorías generales del ser, entendidas de forma abstracta, de las que participa el ser concreto.” (Ramírez).

“Una Ontología es una especificación formal y explícita de una conceptualización compartida” (Ramírez).

En el entorno de la hipertextualidad, la ontología ha sido definida como: Una representación explícita y formal de una conceptualización compartida (Ramírez).

Ramírez la define como la conceptualización correspondiente a una parte del mundo o universo que es objeto de tratamiento, como una forma de entender y describir un dominio, por lo que constituye un modelo abstracto de algún fenómeno en el mundo. Modelo que se construye a partir de identificar los conceptos que componen un dominio del conocimiento, y las relaciones relevantes establecidas entre dichos conceptos, por lo que la base de toda ontología es una taxonomía (o clasificación) de conceptos.

En las ciencias de la computación y de la información, una ontología es una representación formal de un conjunto de conceptos dentro de un dominio y de relaciones entre dichos conceptos. Es usada para razonar acerca de propiedades de un dominio, y quizás usada para definir el dominio.

Capítulo 1: Fundamentación Teórica

Una ontología brinda un vocabulario compartido, el cual puede ser usado para modelar un dominio que es, el tipo de los objetos y/o conceptos que existe, y sus propiedades y relaciones (Cruz).

Una ontología estará formada por una taxonomía relacional de conceptos, es decir, un conjunto de conceptos organizados jerárquicamente y por un conjunto de axiomas o reglas de inferencia mediante los cuales se podrá inferir nuevo conocimiento, es una forma de modelar sentencias que son siempre ciertas (García , 2004).

Según (García, 2004) se delimita la ontología como un instrumento que define los términos básicos y relaciones a partir del vocabulario de un área, así como las reglas de combinación de estos términos y relaciones para definir extensiones a un vocabulario. Las ontologías son construcciones que estructuran contenidos explícitos y que son capaces de codificar las reglas implícitas de una parte de la realidad, pese a trabajar con declaraciones explícitas independientes del fin y del dominio de la aplicación.

Una ontología define un vocabulario común para investigadores que necesitan compartir la información en un dominio. Ella contiene definiciones de conceptos básicos y sus relaciones que no pueden ser interpretadas por una máquina (Noy, 2005).

De acuerdo con las definiciones realizadas por los diferentes autores citados anteriormente, los cuales hacen referencia al término ontología desde que si introdujo por primera vez hasta hoy, demostrando como ha ido evolucionando y tomando protagonismo en varios escenarios de la ciencia. Según los objetivos de la investigación y en lo que se basará su estudio, se hace necesario puntualizar los que proponen la ontología como una representación del dominio de un modelo abstracto de algún fenómeno en el mundo, basado en conceptos y las relaciones entre ellos, siendo la base de toda ontología.

Según los criterios antes analizados el autor de este trabajo lo define como la forma de modelar términos que se manejan en un dominio, lo que hace posible una comprensión común de la estructura de la información entre personas o agentes de software, es decir, los términos con los que se trabajan tengan el mismo significado y se hable en un lenguaje común. Esta conceptualización se representa mediante clases o conceptos, propiedades o atributos de las clases y sus relaciones.

Teniendo en cuenta que para un mismo dominio del conocimiento, pueden existir varios dominios ontológicos, ya que pueden existir diversas formas de entender dicho dominio y por lo tanto,

diversas formas de representarlo, el uso de las ontologías de dominio en el ámbito organizacional puede ser de mucha utilidad para: Representar de manera explícita y a través del lenguaje natural, el conocimiento tácito y el operacional de los distintos grupos que integran la organización, permitiendo la comprensión de estructuras de información entre personas o agentes de software.

- Posibilitar el aprendizaje organizacional a través de tales representaciones.
- Reutilizar el conocimiento representado para el desarrollo de servicios y productos de información en la organización, especialmente en procesos de búsqueda y recuperación de información.
- Posibilitar una cultura de colaboración a escala grupal y/o organizacional a través del proceso de consenso en la construcción de la ontología (Ramírez).
- Permite hacer explícitos los supuestos de un dominio. Esta aseveración puede conducir a conclusiones muy interesantes para la representación del conocimiento más allá de consideraciones técnicas, operativas e informáticas.
- Hace posible analizar el conocimiento de un campo, por ejemplo en lo que se refiere al estudio de los términos y relaciones que lo configuran ya sea formalmente o no (Urquijo, 2009).
- Permite representar diferentes estructuras o modelos de datos.
- Permite mayor flexibilidad y reutilización.

En resumen el uso de ontologías en diversas esferas de la vida moderna facilita un protocolo de comunicación entre varios sistemas interrelacionados, proporcionando una forma para representar y compartir el conocimiento utilizando un vocabulario común.

Componentes de una ontología

Los componentes de una ontología varían de acuerdo al dominio de interés y a las necesidades de los desarrolladores. Por lo general entre los componentes se encuentran los siguientes (Ramos, 2007):

- **Clases:** Las clases son la base de la descripción del conocimiento en las ontologías ya que describen los conceptos (ideas básicas que se intentan formalizar) del dominio. Las clases

usualmente se organizan en taxonomías a las que por lo general se les aplican mecanismos de herencia.

- **Relaciones:** Representan las interacciones entre los conceptos del dominio. Las ontologías por lo general contienen relaciones binarias, el primer argumento de la relación se conoce como el dominio y el segundo como el rango.
- **Funciones:** Son un tipo concreto de relación donde se identifica un elemento mediante el cálculo de una función que considera varios elementos de una ontología.
- **Instancias:** Representan objetos determinados de un concepto.
- **Taxonomía:** Conjunto de conceptos organizados jerárquicamente. Las taxonomías definen las relaciones entre los conceptos pero no los atributos de éstos.
- **Axiomas:** Se usan para modelar sentencias que son siempre ciertas. Los axiomas permiten, junto con la herencia de conceptos, inferir conocimiento que no esté indicado explícitamente en la taxonomía de conceptos. Los axiomas definidos en una ontología pueden ser estructurales o no estructurales: un axioma estructural establece condiciones relacionadas con la jerarquía de la ontología, conceptos y atributos definidos; un axioma no estructural establece relaciones entre atributos de un concepto y son específicos de un dominio. Los axiomas se utilizan también para verificar la consistencia de la ontología.
- **Propiedades (Slots):** Son las características o atributos que describen a los conceptos. Las especificaciones, rangos y restricciones sobre los valores de las propiedades se denominan *facets*. Para un concepto dado, las propiedades y las restricciones sobre éstos son heredadas por las subclases y las instancias de la clase.

Tipos de datos de los slots:

- **String:** es el tipo de valor más simple, el cual es usado por slots tales como nombre: el valor es una simple cadena de caracteres
- **Number:** describe slots con valores numéricos (algunas veces los tipos de valores float e integer son usados por ser más específicos).
- **Boolean:** son simples banderas si/no.

- **Enumerated:** especifican una lista determinada de valores admitidos para el slot.
- **Instance:** admiten la definición de relaciones entre individuos. Los slots con tipo de valor Instance deben también definir una lista de clases admitidas de las cuales las instancias pueden provenir.

En el diseño de la geo-ontología se hará uso del tipo string debido a que las propiedades de los datos que se van a representar son nombres de objetos y de personas.

Tipos de ontologías:

De acuerdo al nivel de generalización, (Vera, 2009) distingue cuatro tipos de ontologías:

- Las ontologías de alto nivel describen los conceptos generales como el espacio, tiempo, materia, objeto, evento, acción, los cuales son independientes de un problema o dominio en particular.
- Las ontologías de dominio y de tareas describen, respectivamente, el vocabulario relacionado a un dominio genérico, por ejemplo, medicina, o una tarea o actividad genérica, como diagnosis, especializando los términos introducidos en la ontología de alto nivel.
- Las ontologías de aplicación describen conceptos dependiendo de un dominio y de una tarea en particular, la cual es una especialización de ambas ontologías relacionadas (ontología de dominio y ontología de tarea).

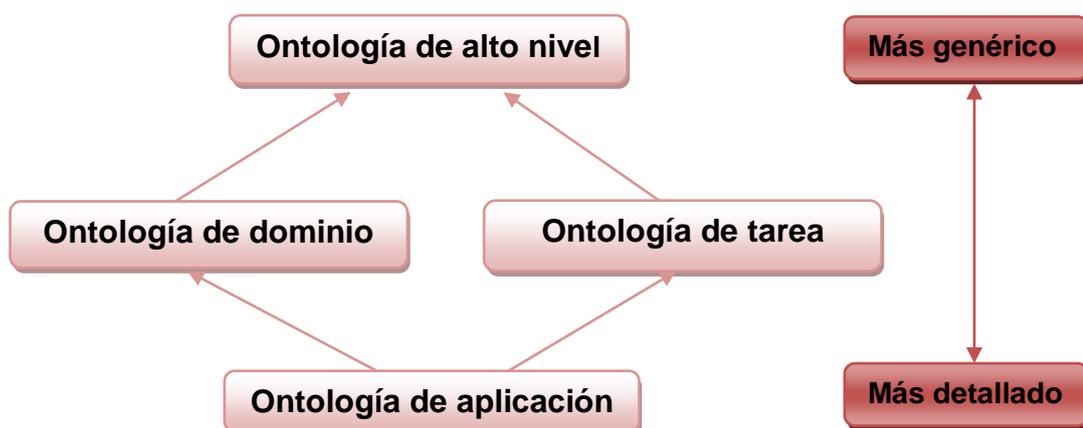


Figura 1: Tipos de ontologías (Vera, 2009).

Las Ontologías de Dominios son reutilizables en un dominio específico (médico, farmacéutico, negocios). Estas ontologías proveen un vocabulario sobre los conceptos dentro de un dominio y sus relaciones, sobre las actividades que tienen lugar y sobre las teorías y principales elementos que lo gobiernan. Los conceptos en ontologías de dominio son usualmente especializaciones de conceptos ya definidos en ontologías de alto nivel, y lo mismo ocurre con las relaciones. Por ejemplo, la instancia de *Ciudad* en una ontología de dominio es una especialización de un concepto más genérico *Localización* definido en una ontología de nivel superior. Las ontologías de dominio contienen los términos básicos de un dominio que se combinan y extienden en ontologías aplicación con el fin de describir la semántica más compleja (Fructuoso).

Se desarrollará una ontología de dominio donde se describirán los términos del dominio definiendo sus conceptos y relaciones, de esta forma se logra una interoperabilidad semántica de la información que se maneja. Las ontologías de dominio establecen la base conceptual que hace posible la anotación semántica y constituyen un común acuerdo sobre un dominio. Se escoge este tipo de ontología para el diseño de la geo-ontología porque el dominio que se representará es específico, relacionando solo términos del SIGUCI, que reflejan la actividad socioeconómica de la UCI.

1.2.2 Geo-ontología

Las ontologías que tratan la temática geoespacial, presentan características particulares de este campo de estudio. Debido a su nivel de especialización, a estas ontologías geográficas se les han denominado geo-ontologías. Estas cumplen con todas las características de una ontología convencional, pero tienen además propiedades propias de este dominio, por ejemplo, un par de coordenadas (x,y) que representa la posición geográfica de un objeto (Vera, 2009).

Las geo-ontologías son estructuras que posibilitan la integración de datos, metadatos y conocimiento. En la geo-ontología como estructura integradora debe estar recogida la conceptualización de los datos y metadatos que representa, así como la formalización del conocimiento en torno a ellos y sus relaciones, unido al enlace físico de donde pueden ser recuperados en caso de que no estén codificados los mismos en la propia estructura (Garea).

Una geo-ontología es una ontología que ofrece una descripción de entidades geográficas y difiere de otras ontologías por la presencia predominante de relaciones semánticas del contenido de la información geográfica de servicios web, con el fin de mejorar el descubrimiento y recuperación en

la web. Es una estructura que relaciona, modifica y asocia información de datos sobre otros datos (Oliva).

Los criterios antes expresados corresponden con el significado del término geo-ontología que se le quiere dar en la presente investigación, ya que en ella el dominio que se representa es geográfico y por tanto es la característica principal que la difiere de las ontologías y la que conllevó al autor hacer uso de ese tipo en específico para darle solución a la problemática planteada.

Las geo-ontologías son estructuras suficientemente adaptables donde la integración de distintos tipos de información es posible. Su uso proporciona un salto cualitativo en funcionalidad y posibilidad, permitiendo la comunicación entre personas, organizaciones y aplicaciones. Son estructuras que pueden crecer, integrarse y reutilizarse en otros dominios, estas características constituyen beneficios de las geo-ontologías como subconjunto de las ontologías.

Una geo-ontología se define como una ontología que se delimita a relacionar información, datos y metadatos geográficos, con el objetivo fundamental de lograr la interoperabilidad semántica entre ellos.

1.2.3 Sistema de información geográfica

Los sistemas de información geográfica (SIG), han sido definidos de diversas maneras por muchos autores, algunas de estas definiciones se incluyen a continuación:

“Conjunto de herramientas para seleccionar, almacenar, recuperar, transformar y exhibir datos espaciales del mundo real para un sistema particular con propósito definido” (Almazán, 2009).

“Un sistema computarizado que proporciona cuatro sistemas de capacidades para manejar datos georreferenciados: entrada de datos, administración de datos (almacenaje y recuperación de datos) manipulación, análisis y salida de datos” (Almazán, 2009).

“Conjunto de hardware, software y procedimientos elaborados para facilitar la obtención, gestión, manipulación, análisis, modelado, representación y salida de datos espacialmente referenciados, para resolver problemas complejos de planificación y gestión” (National Center for Geographic Information & Analysis, 1990).

Técnicamente se puede definir un SIG como una tecnología de manejo de información geográfica formada por equipos electrónicos (hardware) programados adecuadamente (software) que permiten

Capítulo 1: Fundamentación Teórica

manejar una serie de datos espaciales (información geográfica) y realizar análisis complejos con éstos siguiendo los criterios impuestos por el equipo científico (personal) (Mi Tecnológico, 2005).

Los SIG pueden definirse como el conjunto de herramientas informáticas que captura, almacena, transforma, analiza, gestiona y edita datos geográficos (referenciados espacialmente a la superficie de la Tierra) con el fin de obtener información territorial para resolver problemas complejos de planificación, gestión y toma de decisiones apoyándose en la cartografía (Mi Tecnológico, 2005).

Un SIG es un sistema geográfico porque permite la creación de mapas y el análisis espacial, es decir, la modelización espacial; es un sistema de información porque orienta en la gestión, procesa datos almacenados previamente y permite eficaces consultas espaciales repetitivas y estandarizadas que permiten añadir valor a la información gestionada; y es un sistema informático con hardware y software especializados que tratan los datos obtenidos (bases de datos espaciales) y son manejados por personas expertas (Mi Tecnológico, 2005).

Los sistemas de información geográfica (SIG), permiten presentar los datos en forma consolidada y ubicados geográficamente. Los SIG son una herramienta que fortalece la gestión de planificación, administración y prestación de servicios de información. Su capacidad para almacenar, recuperar, analizar, modelar y representar grandes extensiones de terrenos con grandes volúmenes de datos espaciales ha determinado su éxito frente a una gran gama de aplicaciones similares.

De acuerdo con las definiciones expresadas por los diferentes autores acerca de los SIG, el autor de la investigación lo define como: un sistema geográfico que almacena, analiza y manipula datos geográficos, a partir de métodos, software, hardware y personas.

Los SIG posibilitan representar geográficamente objetos y situaciones sociales. En Cuba para brindar información sobre los datos representados acerca de

Según (Larin) “Dicha herramienta ofrece la capacidad de almacenar, acceder, analizar, manipular, desplegar e integrar información ambiental, económica y social en un solo sistema” facilitando las siguientes operaciones:

- Superposición de datos para fines comparativos.
- Actualización de información para ilustrar cambios en el tiempo.
- Cambios de escala para microanálisis.

- Derivación de datos no disponibles mediante manipulación de factores conocidos.
- Integración de paquetes de datos de ciencias sociales y físicas.
- Incorporación de datos adquiridos mediante sensores remotos tales como imágenes de satélite con fines de monitoreo ambiental continuo.
- Modelado de procesos sociales y físicos con propósitos de simulación y predicción.

Aplicaciones de los SIG

En la mayoría de los sectores los SIG pueden ser utilizados como una herramienta de ayuda a la gestión y toma de decisiones, siendo muy diversos sus campos de aplicación. Algunos de ellos son (Marín, 2010):

- Infraestructuras: vías de comunicación, redes eléctricas y de teléfono, canalizaciones de gas.
- Protección Civil: riesgos, desastres y catástrofes.
- Gestión territorial.
- Marketing.
- Demografía.
- Recursos mineros.
- Análisis de mercados.
- Medio ambiente y Recursos Naturales.

Ventajas de las geo-ontologías en los SIG

El desarrollo de geo-ontologías se presenta como el instrumento adecuado para alcanzar la integración semántica en el entorno de la infraestructura de los datos espaciales, representado específicamente en este marco de la investigación por los SIG, de una manera mucho más abstracta en la que el conocimiento juega un papel fundamental. A continuación algunas de las ventajas de las geo-ontologías en los SIG:

- Esta es una herramienta capaz de realizar complejos análisis de la información combinando datos y conocimientos dentro de los SIG.

- Facilita la integración de manera flexible y se basa en el valor semántico de la información, sin tener en cuenta su representación. Se evita la clasificación de los datos sobre la base de sus representaciones solamente.
- Infiere conocimientos que no están explícitamente en el sistema, lo que posibilita que la información que se devuelva sea más consistente y por tanto los usuarios queden satisfechos.
- Constituyen el elemento básico integrador en estos sistemas, las mismas establecen conceptualizaciones de los objetos geográficos compartidas por diversas comunidades lo que permite que los datos integrados al SIG, así como los resultados de las búsquedas y análisis puedan ser compartidos y procesados por los mismos.
- Permiten reutilizar, manipular y compartir el conocimiento que tiene implícito en sus estructuras de datos y bases de datos espaciales.

1.2.4 Web semántica

La Web Semántica fue creada por Tim Berners-Lee, inventor de la WWW, URIs, http y HTML. Existe un equipo en el World Wide Web Consortium (W3C) que se dedica a mejorar, extender y estandarizar el sistema; muchos lenguajes, publicaciones y herramientas han sido ya desarrollados (Van, 1997).

La web semántica se puede definir como un área en la que las máquinas puedan entender y utilizar los datos de la web actual, permitiendo a los usuarios interactuar entre sí, buscar, añadir información sin mucho trabajo, delegando funciones a herramientas y software para lograr una navegación rápida en un ambiente amigable, flexible y económico.

La Web Semántica es una extensión de la World Wide Web en la que los contenidos de la web pueden ser expresados mucho más que en un lenguaje natural y también en un formato que pueda ser entendido, interpretado y usado por diferentes software, permitiéndoles buscar, compartir e integrar información más fácil (Van, 1997).

La Web Semántica tiene como objetivo crear un medio universal para el intercambio de información basado en representaciones del significado de los recursos de la web, de una manera entendible para las máquinas (Pérez, 2007). Con ello se pretende ampliar la interoperabilidad entre los sistemas informáticos y reducir la mediación de operadores humanos en los procesos inteligentes

de flujo de información. La Web Semántica ayuda a resolver estos dos importantes problemas permitiendo a los usuarios delegar tareas en el software. Gracias a la semántica en la web, el software es capaz de procesar su contenido, razonar con este, combinarlo y realizar deducciones lógicas para resolver problemas cotidianos automáticamente.

La visión de la Web Semántica es ampliar los principios de la web desde los documentos a los datos. La misma brindará mayor potencial a la web, permitiendo que los datos sean compartidos con eficiencia por grandes comunidades y sea procesada automática o manualmente (Pérez, 2007).

En el mundo, la web es la base principal de la comunicación y la interrelación entre las personas y las empresas. Muchas son las peticiones y necesidades de los usuarios, pero eso se hace necesario enfocar y convertirla en una web semántica, de manera tal que al ser las herramientas y los software los que le den respuestas a estas peticiones y no expresamente el hombre, estas serán más rápida y certera.

1.2.5 Interoperabilidad semántica

La interoperabilidad semántica es la capacidad derivada de la aplicación de tecnologías que permiten inferir, relacionar, interpretar y clasificar el conocimiento implícito en el contenido digital que dirige los procesos de negocio (Borrego).

El Instituto de Ingenieros Eléctricos y Electrónicos (IEEE por sus siglas en inglés) define interoperabilidad como la habilidad de dos o más sistemas o componentes para intercambiar información y utilizar la información intercambiada.

Se define la interoperabilidad semántica como la capacidad que tienen los sistemas de información y la aplicación de tecnologías de compartir datos, posibilitar el intercambio de información y conocimiento entre varios sistemas. Es la forma de manejar los datos de forma tal que se vayan modificando a partir de la utilización de la información intercambiada, logrando un vocabulario común entre los datos que se manejan.

Niveles del modelo de interoperabilidad conceptual (Metamodel, 2011).

El modelo conceptual de interoperabilidad propone que existen tres fases en la interoperabilidad conceptual: la integrabilidad (el nivel físico), la interoperabilidad (el nivel lógico de implementación)

y la componibilidad (el nivel conceptual, de modelo y abstracción). Además, divide cada fase en dos niveles:



Figura 2: Niveles del modelo de interoperabilidad conceptual (Metamodel, 2011).

- **Nivel 0, No interoperabilidad:** El más simple de explicar. No existe interoperabilidad alguna, lo cual suele significar que es un único sistema o varios homogéneos.
- **Nivel 1, Interoperabilidad técnica:** Existe una definición del sistema y protocolo de comunicaciones.
- **Nivel 2, Interoperabilidad sintáctica:** El protocolo incluye una estructura común de intercambio de información, es decir, un formato común de datos. A nivel práctico: hay una forma de intercambiar datos de la base de datos.
- **Nivel 3, Interoperabilidad semántica:** Además de intercambiar datos, los sistemas implicados conocen el significado de los mismos, y se aseguran los medios de que así sea. A nivel práctico: no sólo se intercambian datos, sino que son semánticamente completos, y poseen un significado.
- **Nivel 4, Interoperabilidad pragmática:** Los sistemas participantes conocen el uso que se le da a los datos y el contexto en el que se utilizan. A nivel práctico: la información está contextualizada y se conoce su uso y propósito.
- **Nivel 5, Interoperabilidad dinámica:** Cada sistema es consciente de los cambios de estado y de información, y de los cambios de limitaciones, a lo largo de la línea temporal, y es capaz de interpretar estos cambios y su significado sobre la información. Este nivel de

interoperabilidad es el que ofrece el concepto de *Awareness* (Conciencia) de la situación y de la información.

- **Nivel 6, Interoperabilidad conceptual:** El nivel conceptual de la información, en el que cada sistema y cada operador de sistema, tienen conciencia de la información, entienden su significado y la interpretan de la misma manera.

Según el modelo antes planteado y las características de la investigación, se identificó que se estará trabajando sobre la fase dos interoperabilidad que se encuentra en el nivel lógico de implementación y dentro de este en el nivel tres que corresponde a la interoperabilidad semántica.

1.3. Herramientas para construir la geo-ontología.

Los editores de geo-ontologías son herramientas especializadas que apoyan la construcción de estas. Las facilidades que proporcionan van desde la definición y modificación de conceptos, propiedades, relaciones, axiomas y restricciones, hasta la inspección y navegación (Ramos, 2007). El uso de editores facilita la tarea de definición de geo-ontologías, la unión y el desarrollo distribuido de ellas.

1.3.1 Ontolingua

Ontolingua es una herramienta de desarrollo para navegar, crear, editar, modificar, verificar, evaluar y usar ontologías. Contiene una librería de ontologías cuyas definiciones, axiomas y términos no lógicos, pueden ser reutilizadas en la construcción de nuevas ontologías (Ramos, 2007).

1.3.2 Chimaera

Chimaera es una herramienta que permite crear y mantener geo-ontologías en la web, proporciona un ambiente distribuido para navegar, crear, editar, modificar y usar geo-ontologías. Entre las facilidades que ofrece la herramienta están: cargar bases de conocimiento en diferentes formatos, reorganizar taxonomías, resolver conflictos de nombres y editar términos. Destaca la capacidad para cargar datos de entrada en quince diferentes formatos, tales como, KIF, Ontolingua, OKBC, Protégé y otros (Ramos, 2007).

1.3.3 Protégé

Protégé es un software libre de código abierto implementado en el lenguaje de programación Java, desarrollado en la Universidad de Stanford, que permite la construcción de geo-ontologías de

dominio. Es capaz de operar como una plataforma para acceder a otros sistemas basados en conocimiento, aplicaciones integradas o como una librería que puede ser usada por otras aplicaciones para acceder y visualizar bases de conocimiento. La herramienta ofrece una interfaz gráfica que permite al desarrollador de geo-ontologías enfocarse en la modelación conceptual sin que requiera de conocimientos de la sintaxis de los lenguajes de salida (Ramos, 2007).

El modelo de conocimiento de *Protégé* está basado en marcos (frames). Las primitivas de representación internas en *Protégé* pueden ser redefinidas declarativamente, permitiendo tener representaciones apropiadas para una variedad de lenguajes de ontologías. *Protégé* puede correr como una aplicación local o a través de un cliente en una comunicación con un servidor remoto. El navegador web de *Protégé* permite a los usuarios compartir, navegar y editar sus ontologías utilizando un navegador web estándar, lo que proporciona un ambiente de colaboración que puede ayudar a las comunidades en el desarrollo de ontologías (Ramos, 2007).

Características generales de Protégé (Sánchez, 2010):

- Cuenta con una arquitectura de tres capas.
- La extensibilidad de la herramienta se basa en extensiones o plug-ins.
- Permite el almacenamiento de geo-ontologías por medio de archivos en formato OWL.
- Permite importar a lenguajes como RDF y XML.
- Permite exportar a lenguajes como XML, RDF, Flogic y CLIPS.
- La representación del conocimiento se basa en marcos y metaclasses.
- Proporciona un lenguaje de axiomas basado en PAL (Programmable Array Logic, por sus siglas en inglés).
- Cuenta con un motor de inferencia nativo.
- Soporte para el manejo de restricciones y revisión de inconsistencias.
- Mantiene la representación gráfica basada en una taxonomía.
- Presenta un soporte para la generación de bibliotecas de ontologías.

- Cuenta con un mecanismo de poda gráfica, basado en restricciones y relaciones.

Ventajas de la herramienta Protégé (Sánchez, 2010):

- Permite exportar geo-ontologías a diversos lenguajes, tales como RDF (Resource Description Framework, por sus siglas en inglés) y XML (Extensible Markup Language, por sus siglas en inglés).
- Está basado en la plataforma de desarrollo de Java, lo cual permite un acceso transparente, vía código nativo o mediante una interfaz programable de aplicación. Además de ser un sistema multiplataforma y flexible en cuanto a su implementación.
- *Protégé* trabaja bajo la teoría de marcos, donde cada concepto se define como una clase que puede ser especializada en subclases. Las relaciones entre las clases son binarias y no existe una posibilidad abierta, en la cual puedan implantarse restricciones n-arias en una forma simple.
- Cuenta con el manejo de instancias sobre las clases; así como restricciones para generarlas. En conjunto, los módulos o extensiones de la herramienta están abiertos para ser programados por cualquier usuario que se interese en el modelado e ingeniería de ontologías.

Se escogió como herramienta para el diseño el *Protégé* en su versión 4.2 porque fue diseñado bajo licencia de software libre, permite el manejo de ontologías desde varios formatos, cuenta con un mecanismo de poda gráfica, lo que permite restringir y agrupar la información. Sus módulos están abiertos para ser programados por cualquier usuario que lo desee. Es un sistema muy usado en el mundo por ser multiplataforma, flexible en cuanto a su implementación, es dinámico y de fácil manejo.

1.4 Metodologías para el desarrollo de geo-ontologías

En general, las metodologías proporcionan un conjunto de directrices que indican cómo deben realizarse las actividades identificadas en el proceso de desarrollo, qué técnicas son las más apropiadas en cada actividad y qué produce cada una de ellas. Su objetivo es planificar, guiar y controlar los pasos a seguir para el diseño de la ontología, exponiendo los procedimientos y las herramientas que pueden usarse para el desarrollo y validación de las mismas (Ramos, 2007).

A pesar de la variedad de metodologías existentes, un denominador común entre ellas es el conjunto de pasos esenciales que pueden ser identificados de la siguiente manera (Ramírez):

- Determinar el propósito de la geo-ontología, sus objetivos y su alcance. Esto permite saber qué tipo de ontología se va a crear.
- Identificar el dominio, donde no solo se incluye el área de conocimientos sino los integrantes, quienes determinarán los conceptos y las relaciones entre los conceptos.
- Capturar el conocimiento que será representado. Es la parte más compleja, debido a que hay que aplicar una serie de técnicas de investigación grupal, que permitan obtener los conocimientos que serán representados y sus relaciones, los cuales deben estar consensuados por los integrantes de ese dominio del conocimiento.
- Formalizar el modelo obtenido a través de un lenguaje de representación y determinar posibles funciones y axiomas para realizar inferencias.

1.4.1 Methontology

Methontology fue desarrollada por el Grupo de Ingeniería Ontológica de la Universidad Politécnica de Madrid. Esta metodología permite construir ontologías en el nivel de conocimientos, tiene sus raíces en las actividades identificadas por el proceso de desarrollo de software propuesto por la organización IEEE y en otras metodologías de ingeniería de conocimientos. Methontology proporciona guías sobre cómo llevar a cabo el desarrollo de la geo-ontología a través de las actividades de especificación, conceptualización, formalización, implementación y mantenimiento (Corcho).

1.4.2 On-to-Knowledge

La metodología On-To-Knowledge, es uno de los resultados más importantes del proyecto europeo (EU-IST-1999-10132) del mismo nombre, que proporciona una guía específica para el desarrollo y mantenimiento de geo-ontologías conjuntamente con el desarrollo de aplicaciones de gerencia de conocimiento. On-To-Knowledge se basa en el análisis de escenarios de uso y considera la identificación de los objetivos que deberían alcanzar las aplicaciones. Esta metodología tiene como objetivo aplicar ontologías a la información disponible electrónicamente para mejorar la calidad de la gestión de conocimiento en organizaciones grandes y distribuidas, fuertemente relacionadas con el WWW (Ramos, 2007).

1.4.3 GEONTO – MET

La metodología GEONTO-MET está compuesta por cuatro etapas de desarrollo: La etapa de análisis proporciona un modelo de abstracción sobre los posibles objetos del dominio geográfico. La etapa de síntesis obtiene la conceptualización del dominio geográfico. La etapa de procesamiento proporciona un conjunto de ontologías de aplicación (contexto turístico y topográfico); así como una ontología de dominio denominada Kaab-Ontology. Por último, la etapa de descripción da una representación alterna de los objetos geográficos; así como la integración de los mismos en una plantilla semántica de descripción y en formato XML (Torres, 2007).

Esta metodología de conceptualización está basada en un conjunto mínimo de relaciones axiomáticas, lo cual permite traducir las relaciones entre conceptos directamente a la conceptualización, con la ventaja de obtener una resolución semántica mayor o más refinada en la definición de tales relaciones. GEONTO-MET proporciona soluciones viables hacia estos tópicos y otros, como el hecho de compactar datos mediante estructuras alternas a modelos tradicionales y evitar la ambigüedad de términos al utilizar una conceptualización del dominio. GEONTO-MET permite coleccionar y organizar jerárquicamente un conjunto de conceptos geográficos, partiendo de diversas fuentes de información, con el objeto de generar una ontología de dominio (Kaab-Ontology) (Torres, 2007).

La metodología que se escogió para guiar la investigación es la GEONTO-MET porque sus cuatro fases permiten llevar un proceso progresivo de construcción de la geo-ontología, haciendo un estudio de cada uno de los elementos y conceptos relacionados, basándose en un conjunto mínimo de relaciones axiomáticas, lo que permite obtener exactitud en las relaciones entre los conceptos, buscando soluciones viables y óptimas en los términos analizados, evitando así ambigüedades de términos en la ontología que se desea diseñar.

1.5 Lenguaje de representación de geo-ontologías

Las geo-ontologías son representaciones formales de la semántica de los datos y para constituirse deben contar con un mecanismo que le permita personalizar dicha información. Este mecanismo no es más que los lenguajes de representación de las ontologías:

OWL (Web Ontology Language o Lenguaje de Ontologías para la web) es un lenguaje de etiquetado semántico para publicar y compartir ontologías en la web. Se trata de una recomendación del W3C y puede usarse para representar ontologías de forma explícita, permite

definir el significado de términos en vocabularios y las relaciones entre aquellos términos (Ontologías). En realidad, OWL es una extensión del lenguaje RDF y emplea las tripletas de RDF, aunque es un lenguaje con más poder expresivo que este. Se trata de un lenguaje diseñado para usarse cuando la información contenida en los documentos necesita ser procesada por programas o aplicaciones, en oposición a situaciones donde el contenido solamente necesita ser presentado a los seres humanos. El OWL está diseñado para ser usado en aplicaciones que necesitan procesar el contenido de la información en lugar de únicamente representar información para los humanos. OWL tiene tres sublenguajes, con un nivel de expresividad creciente: OWL Lite, OWL DL, y OWL Full (Harmelen, 2004).

- **OWL Lite:** Está diseñado para aquellos usuarios que necesitan principalmente una clasificación jerárquica y restricciones simples. Por ejemplo, a la vez que admite restricciones de cardinalidad, sólo permite establecer valores cardinales de 0 ó 1. Debería ser más sencillo proporcionar herramientas de soporte a OWL Lite que a sus parientes con mayor nivel de expresividad. OWL Lite proporciona una ruta rápida de migración para tesauros y otras taxonomías y tiene también una menor complejidad formal que OWL DL.
- **OWL DL** (Description Logic, por sus siglas en inglés): Está diseñado para aquellos usuarios que quieren la máxima expresividad conservando completitud computacional (se garantiza que todas las conclusiones sean computables) y resolubilidad (todos los cálculos se resolverán en un tiempo finito). OWL DL incluye todas las construcciones del lenguaje de OWL, pero sólo pueden ser usados bajo ciertas restricciones (por ejemplo, mientras una clase puede ser una subclase de otras muchas clases, una clase no puede ser una instancia de otra).

Esta versión ya tiene todo el vocabulario OWL completo. Las limitaciones son que las clases no son instancias ni tipos y los tipos no son ni instancias ni clases. No permite restricciones de cardinalidad en propiedades transitivas.

- **OWL Full:** Está dirigido a usuarios que quieren máxima expresividad y libertad sintáctica de RDF sin garantías computacionales. Por ejemplo, en OWL Full una clase puede ser considerada simultáneamente como una colección de clases individuales y como una clase individual propiamente dicha. OWL Full permite una ontología para aumentar el significado del vocabulario preestablecido (RDF u OWL). Es poco probable que cualquier software de

razonamiento sea capaz de obtener un razonamiento completo para cada característica de OWL Full.

OWL es el más completo de los lenguajes utilizados para la representación del conocimiento en la web. OWL abarca las potencialidades ofrecidas por el lenguaje DAML+OIL y al igual que este es una extensión del modelo RDF-S. También incorpora otro gran número de primitivas para representar el significado de los elementos y sus relaciones dentro de la ontología (Urquijo, 2009).

La elección entre OWL Lite y OWL DL depende de las necesidades de los usuarios sobre la expresividad de las construcciones, proporcionando OWL DL las más expresivas. La elección entre OWL DL y OWL Full depende principalmente de las necesidades de los usuarios sobre los recursos de metamodelado del esquema RDF (por ejemplo, definir clases de clases, o definir propiedades de clases). Cuando se usa OWL Full en comparación con OWL DL, el soporte en el razonamiento es menos predecible, ya que no existen en este momento implementaciones completas de OWL Full (Harmelen, 2004).

OWL Full puede ser considerada como una extensión de RDF, mientras que OWL Lite y OWL DL pueden ser consideradas como extensiones de una visión restringida de RDF. Por este motivo, se ha de tener cuidado cuando un usuario quiera migrar un documento de RDF a OWL (Harmelen, 2004).

Características significativas que incorpora OWL (Urquijo, 2009):

- Definición de clases mediante restricciones sobre propiedades, valores o cardinalidad.
- Definición de clases mediante operaciones booleanas sobre otras clases: intersección, unión y complemento.
- Relaciones entre clases (ej. Inclusión, disyunción, equivalencia).
- Cardinalidad (ej. “únicamente una”).
- Igualdad y desigualdad de clases.
- Igualdad y desigualdad de instancias.
- Clases enumeradas.

Ventajas del lenguaje OWL (Gómez):

- Capacidad de representar ontologías distribuidas a través de varios sistemas.
- Escalable a las necesidades de la Web.
- Compatible con los estándares Web de accesibilidad e internacionalización o Abierto y extensible.
- Cada uno de estos sub-lenguajes es una extensión de su predecesor más simple, respecto a lo que puede ser expresado legalmente y a la validación de sus conclusiones. Para ello se establecen estas relaciones, pero las relaciones inversas no se permiten.

Se escogió como lenguaje el OWL en su sublenguaje OWL DL, este ofrece la máxima expresividad conservando completitud computacional (se garantiza que todas las conclusiones sean computables) y resolubilidad (todos los cálculos se resolverán en un tiempo finito). De forma general, OWL DL requiere separación de tipos (una clase no puede ser un individuo o una propiedad y esta a su vez no puede ser tampoco un individuo o una clase). Estas restricciones no afectan el trabajo con la geo-ontología que se diseña en la investigación, pues ciertas clases, propiedades o instancias se definen según la semántica que se necesite representar.

1.6. Conclusiones parciales

En este capítulo se analizaron conceptos, características, componentes de las geo-ontologías y de otros términos manejados en la investigación como los SIG, la interoperabilidad semántica, la web y otros, los cuales son la base para el diseño de la geo-ontología. Teniendo como criterio de selección lo estudiado en el capítulo, se escogió como metodología a utilizar la GEONTO-MET, como herramienta el *Protégé* y como lenguaje de representación el OWL DL.

Capítulo 2: Diseño de la Geo-ontología

2.1 Introducción

En este capítulo se define la geo-ontología de dominio que se usará como base para el diseño de la geo-ontología del SIGUCI. Concluida esta selección, se precisan los pasos a seguir para la construcción de la misma, concretando el glosario de términos, la taxonomía de conceptos, y las relaciones entre ellos. Con esta secuencia de pasos se obtendrá la geo-ontología que se persigue con la investigación, la cual tiene como objetivo minimizar la complejidad de búsqueda de los usuarios en este sistema y que los términos que se manejen tengan un mismo significado.

2.2 Ontología de dominio para diseñar la geo-ontología para el SIGUCI

Para guiar la investigación se escogió como metodología la GEONTO-MET, la cual proporciona una ontología de dominio, denominada Kaab-Ontology, elaborada por el profesor Miguel Jesús Torres-Ruiz (Torres, 2007). Esta ontología permite conceptualizar formalmente el dominio geográfico caracterizando semánticamente el comportamiento de los objetos geográficos a través de sus propiedades y relaciones. La Kaab-Ontology elaborada con este objetivo, propone un mecanismo de recuperación a través de conceptos donde se asocian los representados en la ontología con alguna fuente de información geográfica.

A pesar que las consultas utilizadas en esta ontología son estáticas, sirven para ilustrar cómo acceder a las ontologías. Utilizando la API Jena de Java es posible extraer las instancias de conceptos clases de Kaab-Ontology, ya que dicha ontología fue elaborada en *Protégé* utilizando OWL. De aquí se obtiene un modelo persistente que funciona como enlace entre Jena y *Protégé* de donde se recuperan instancias de conceptos geográficos en OWL. Por otro lado las consultas implementadas en SPARQL¹ posibilitan el acceso a datos de recursos OWL, dado que previamente las clases de entidades abstractas y los conceptos que heredan de estas clases fueron llevados a una forma tabular. Así SPARQL a través de Jena puede acceder a todos los recursos de la ontología.

Al ser el modelo de consultas orientado a datos, una vez definido el significado de un determinado elemento en un glosario dado y representado en un lenguaje, es posible entonces obtener

¹ SPARQL únicamente incorpora funciones para la recuperación de sentencias RDF. Sin embargo, algunas propuestas también incluyen operaciones para el mantenimiento (creación, modificación y borrado) de datos.

instancias de conceptos de forma semántica que estén relacionados dentro del contexto de la ontología (Martínez, 2010).

2.3 Pasos a seguir para la construcción de la geo-ontología

Elementos para desarrollar una ontología (Noy, 2005):

- Construir el glosario de términos.
- Definir clases en la ontología.
- Organizar las clases en una jerarquía taxonómica (subclase-superclase).
- Definir slots y describir valores permitidos para esos slots.
- Llenar los valores de los slots para las instancias.

2.3.1 Construcción del glosario de término

Para la construcción del glosario de términos, se debe comenzar realizando una lista con todos los términos que se manejan en el SIG sin preocuparse del recubrimiento entre los conceptos que representan, relaciones entre los términos, cualquier propiedad que los conceptos puedan tener o si los conceptos son clases o slots. Después de realizar una búsqueda en las funcionalidades del sistema se pudo identificar los siguientes términos:

Cafeterías (Casa del Perro, El framboyán, El puente, Los guanos, Los pinos, Piscina, Manticore, Heladería, Pizzería-Soda).

Complejos Comedores (uno, dos y tres).

Edificios docentes (Aulas de Gastex, Docente 1, Docente 2, Docente 3, Docente 4, Docente 5, Docente 6).

Edificios de Producción (Dirección Authoring DVD, Dirección de Producción, Dirección de televisión universitaria, Docente Viejo, Laboratorio IP, Laboratorios MININT).

Lugares de Interés (Biblioteca, Centro Cultural, Cátedra de Ajedrez, Librería, Maqueta).

Otras construcciones (Rectorado).

Parques y Plazas (Explanada del Nodo, Parque de las Esculturas, Plaza Cultural “Wilfredo Lam”, Plaza Niemeyer, Plaza de las Banderas, Plaza Mella).

Puntos de ventas (Carnicería, Mercadito (Edificio 120), Mercado Agropecuario Estatal “La arboleda”, Panadería, Pescadería, Tienda Mixta, Mercado Agropecuario Particular).

Restaurantes (El Framboyán, El Ranchón, La Casona, La Pizzería).

Seguridad Protección (Garita Entrada Hospital, Garita Entrada IP, Garita Entrada Principal, Garita Entrada Servicio, PNR, Puesto de mando SEPCOM, Puesto de mando UCI).

Servicios Bancarios (Banco, Cadena Principal, Cajero Automático Principal, Cajero Automático de la Residencia, Cajero Automático Docente 4).

Servicios Médicos (Consultorio Médico (Edificio 125), (Edificio 49), (Edificio 72), (Edificio 87), Policlínico-Hospital Ernesto Che Guevara).

Servicios Técnicos (División Territorial de Copextel, Nodo Principal, Nodo Secundario).

Servicios Varios (Atelier-Sastrería, Barbería, Correo, Lavandería (Edificio 122), (Edificio 92), Locutorio, Manicure, Peluquería, Zapatería).

Áreas Deportivas (Campo de Beisbol, Campo de Fútbol, Canchas deportivas facultad 8, Canchas deportivas facultad 9, Canchas de Fútbol Sala, Canchas de Tenis, Gimnasio).

Áreas recreativas (Discoteca, Piscina, Pistas de baile, Salón de juegos).

Una vez seleccionado todos los términos relacionados en el SIGUCI se hizo un análisis de cada uno y se determinó reestructurar e integrar dichos concepto en dependencia de los ya definidos en la Kaap-ontology. Con la reestructuración de los conceptos, el apoyo de sitios especializados (por ejemplo la Real Academia de la Lengua Española) para conocer el significado de los conceptos con los que se trabajan y haciendo referencia a sus sinónimos se concluyó el glosario de términos. Una parte del glosario se muestra a continuación en la **Tabla 1** y el resto se presenta en el Anexo I: Glosario de términos.

Nombre	Sinónimo	Descripción	Tipo
puntos_de_venta		Un punto de venta viene a automatizar el proceso de salida y cobro de la mercancía en las tiendas departamentales, restaurantes y otras instituciones.	Clase
instalacion_diversa			Clase

entidad_geografica		Colectividad considerada como unidad. Especialmente, cualquier corporación, compañía, institución. Pertenece o relativo a la geografía.	Clase
cafeteria	Merenderos	Una cafetería, cafetín o simplemente un café, es un despacho de café y otras bebidas, donde se sirven aperitivos y comidas.	Clase

Tabla 1: Glosario de términos.

2.3.2 Construcción de la taxonomía de conceptos

Después de haber elaborado el glosario de términos el siguiente paso es la construcción de la taxonomía relacionando dichos conceptos. Hay varios enfoques para desarrollar una jerarquía (taxonomía) de clases (conceptos) (Noy, 2005).

Un proceso de desarrollo **top-down** comienza con la definición de los conceptos más generales en el dominio, la subsecuente especialización de los conceptos.

Un proceso de desarrollo **bottom-up** comienza con la definición de las clases más específicas, las hojas de la jerarquía, con el subsecuente agrupamiento de esas clases en conceptos más generales.

Un proceso de desarrollo combinado es el resultado de una combinación de los enfoques **top-down** y **bottom-up**: primero se definen los conceptos más sobresalientes, luego se generalizan y especializan apropiadamente.

Para desarrollar la taxonomía de conceptos del SIGUCI se utilizó el proceso **top-down**, se definieron las clases generales y luego las subclases, de esta forma se abarcaron todos los términos declarados. En la **Tabla 2** se observan algunas de las superclases seleccionadas con sus subclases descendientes formando la taxonomía del SIGUCI y las demás se pueden apreciar en el Anexo II: Taxonomía de conceptos del SIGUCI.

Superclases	Subclases
dominio_fijo	tipo de edificacion, tipo de instalacion
entidad_geografica	instalacion, edificacion
instalacion	instalacion_de_comunicacion, instalacion_deportiva_o_recreativa, instalacion_diversa , instalacion_industrial.
edificacion	edificio_docente, edificio_produccion, edificio_residencia

Tabla 2: Taxonomía de conceptos.

A continuación en la **Figura 3** se muestra como queda en el *Protégé* la taxonomía de conceptos definida en la geo-ontología.



Figura 3: Taxonomía de conceptos.

2.3.3 Describir relaciones binarias

Las relaciones binarias son propiedades que representan un tipo de asociación entre los conceptos del dominio. Una propiedad específica es la forma en que se va a relacionar un concepto con otro. Las mismas poseen diferentes características (Pérez, 2011):

Simétrica: Si se tiene una propiedad simétrica y el par (x, y) es una instancia de esa propiedad, entonces el par (y, x) también es una instancia de esa propiedad. Ejemplo: Si tenemos una

propiedad simétrica llamada *hermano* y dos instancias de esa propiedad (Pepe, Juan) se está especificando que Pepe es hermano de Juan, por lo tanto Juan es hermano de Pepe.

Inversa: Una propiedad puede ser la inversa de otra propiedad. Si se establece una *propiedad 1* como inversa de la *propiedad 2*, y se relaciona *X* con *Y* mediante la *propiedad 2*, entonces *Y* estaría relacionado con *X* mediante la *propiedad 1*. Si la propiedad *tieneHijo* es la propiedad opuesta de *tienePadres* y Juan *tieneHijo* Pepe, entonces Pepe *tienePadres* Juan.

Transitiva: Una propiedad es transitiva, si el par (*x*, *y*) es una instancia de la propiedad *transitiva 1*, y el par (*y*, *z*) es otra instancia de la propiedad *transitiva 1*, entonces el par (*x*, *z*) también es una instancia de la propiedad. Por ejemplo, se indica que la propiedad *Antepasado* es transitiva, si Juan es un antepasado de Pepe y Pepe es un antepasado de Alberto, entonces Juan es un antepasado de Alberto.

Funcional: Es posible definir propiedades para que tengan un valor único. Si una propiedad es funcional no tendrá más de un valor para cada individuo o ninguno. Esta característica se denomina propiedad única, es una forma de indicar que la cardinalidad mínima de la propiedad es 0 y la cardinalidad máxima es 1. Por ejemplo, *tieneMadre* puede establecerse como una propiedad funcional, cada individuo tiene una sola madre.

Funcional inversa: Si una propiedad es funcional inversa, entonces la inversa de la propiedad será funcional. Por tanto, la inversa de la propiedad tiene como máximo de un valor para cada individuo.

Las propiedades también traen relacionado un rango y un dominio, las clases admitidas para los Slots de tipo Instance son a menudo llamadas rango de un Slot. Las clases a las cuales un Slot está adosado o las clases cuyas propiedades son descritas por un Slot son llamadas dominio del Slot (Noy, 2005). Las reglas básicas para determinar un dominio y un rango de un Slot son similares.

Cuando se define un dominio o rango de un Slot, se debe encontrar las clases o clase más generales que puedan ser respectivamente el dominio o rango de los mismos. Por otro lado, no definir un dominio ni rango que sea demasiado general: todas las clases en el dominio de un Slot deben ser descritas por él y las instancias de todas las clases en su rango deben poder ser rellenos potenciales del Slot. No se debe elegir una clase demasiado general para el rango (i.e, es inútil de crear un rango COSA (THING)) pero es posible elegir una clase que cubre todos los valores de relleno (Noy, 2005).

Se establecieron las propiedades a los objetos de las clases, las cuales permiten inferir nuevos conocimientos a partir de las relaciones que existen entre ellos. Se determina el dominio y el rango a cada una de las propiedades establecidas, haciendo referencia a la clase más general del dominio en la que se enmarca. Ejemplos de algunas propiedades se muestran a continuación y las demás en el Anexo III: Descripción de las relaciones binarias (Object properties).

Nombre de la relación	Concepto de Origen	Cardinalidad Máxima	Concepto de destino
_tiene	entidad geografica	1	dominio_fijo
alLadoDerechoDe	entidad geografica	N	entidad geografica
alLadoIzquierdoDe	entidad geografica	N	entidad geografica
frenteA	entidad geografica	N	entidad geografica
situadoEn	instalacion	N	edificacion

Tabla 3: Propiedades de objetos (object properties).

En la **Figura 4** se representan las relaciones binarias establecidas en la geo-ontología, mediante gráficas que aporta el *Protégé*:

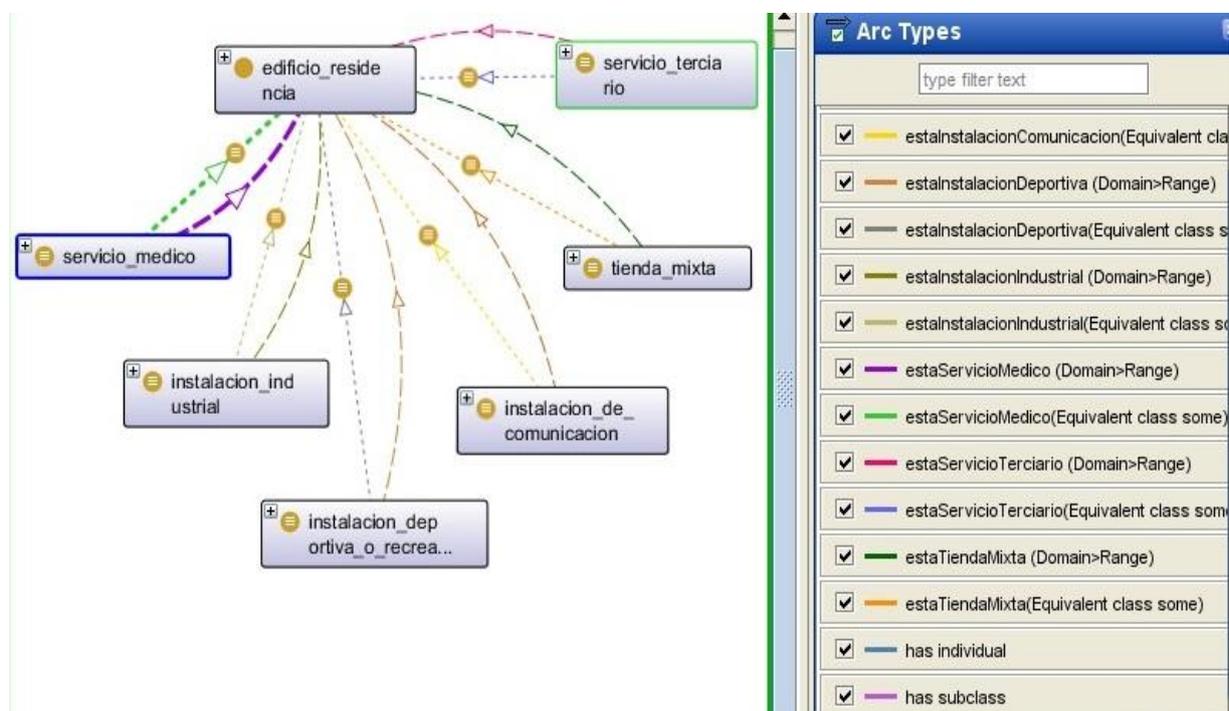


Figura 4: Relaciones binarias.

Como se muestra en la figura, el color de la línea discontinua es la relación que se establece entre las clases, las líneas más largas simbolizan el dominio y el rango y las líneas con puntos representan que ambas clases son equivalentes. Por ejemplo la línea naranja y la verde están sujetas por la relación *estaTiendaMixta* que se establece entre la clase edificio de residencia y Tienda Mixta, especificando con la línea naranja que son clases equivalentes y con la verde que Tienda Mixta es el dominio y edificio de residencia el rango.

2.3.4 Propiedades de datos (Data properties)

Las propiedades de los datos son aquellas relaciones que definen valores para las clases, describiendo los atributos de los objetos. En el Anexo IV: Descripción de las propiedades de los datos, se encuentran las relaciones que se definieron. Ejemplo de ellas es *edificaciónNombre*, esta relación muestra que la edificación que se representa va a tener un nombre propio. En el dominio de la relación se define la clase *edificación* y en el rango el tipo de dato de ese atributo, en este caso el nombre de la *edificación* es de tipo string. En la **Figura 5** se representan estas propiedades de los datos en el *Protégé*.

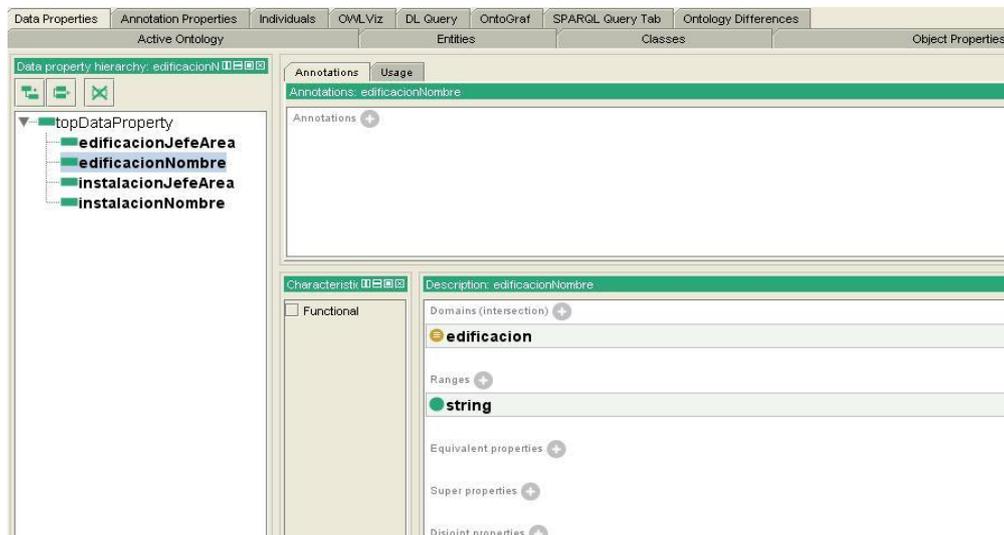


Figura 5: Relaciones entre conceptos (Data properties).

Las relaciones y propiedades que se establecen entre los conceptos proporcionan una representación semántica a tal nivel que junto a los axiomas facilita la interpretación y búsquedas de los usuarios a la hora de la navegación.

En el siguiente epígrafe se hace referencia a los axiomas definidos para la geo-ontología.

2.3.5 Describir axiomas formales

Los axiomas formales se usan para modelar sentencias que son siempre ciertas, describiendo el comportamiento de las clases. Permiten junto con la herencia de conceptos, inferir conocimiento no codificado explícitamente en la geo-ontología. Suele utilizarse para definir restricciones y se establecen a partir de valores específicos de las propiedades. A continuación se hace referencia a algunos de los axiomas definidos y en la **Figura 6** se muestra como queda representado en el *Protégé*.

edificación \longrightarrow **_tiene only** tipo_de_edificación.

edificación \longrightarrow **poseeUn some** instalación.

instalación_comunicación \longrightarrow **estaEdificacionResidencia some** edificio_residencia.

instalación_deportiva_o_recreativa \longrightarrow **estaInstalacionDeportiva some** edificio_residencia.

Instalación_industrial \longrightarrow **estaInstalacionIndustrial some** edificio_residencia.

mercado_ideal \longrightarrow **estaMercadoIdeal only** edificio_residencia.

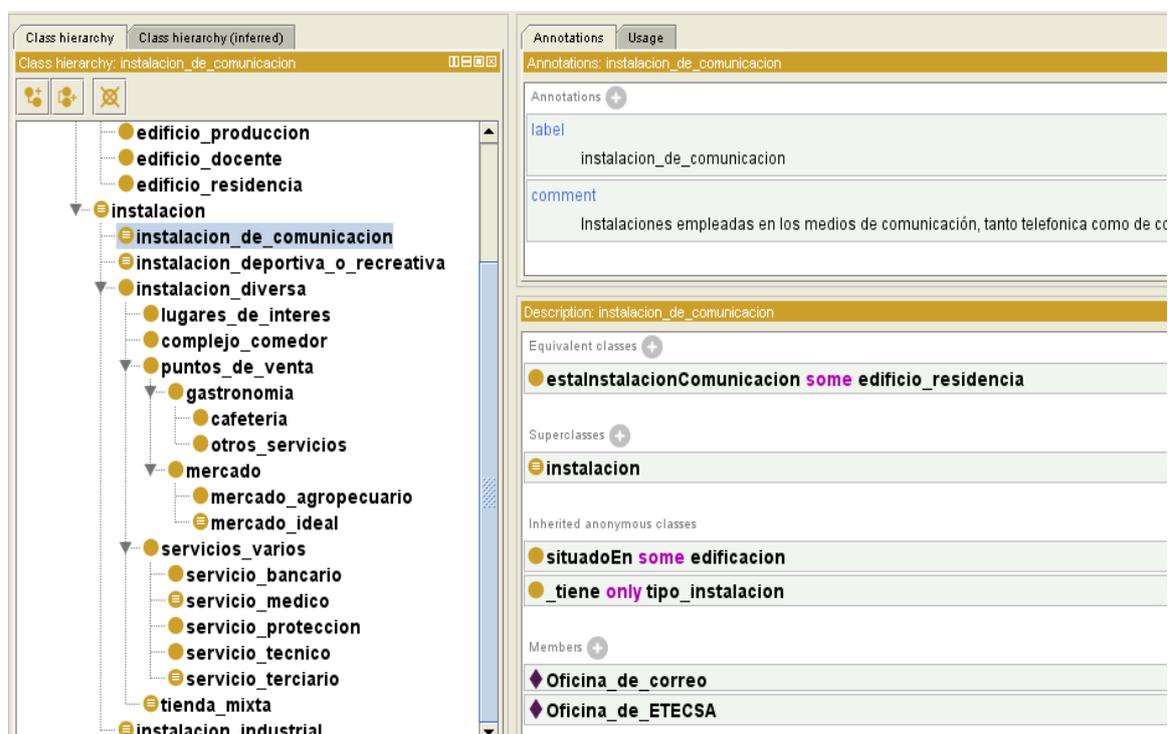


Figura 6: Axiomas formales.

2.3.6 Describir instancias

Las instancias permiten hacer referencia al objeto concreto que se modela en las clases, esta constituye la base de conocimientos de la geo-ontología.

En la **Figura 7** se muestra una de las instancias definidas, **El Framboyán** que es un individuo de la clase Cafetería, tiene como relación binaria **alLadoDe** la piscina, **cercaDe** restaurante_La Casona y **detrasDe** Complejo_Comedor1 y como propiedad **institucionNombre** e **institucionJefeÁrea**. De esta forma se especifica que es un punto de venta con un nombre y un jefe de área, además las relaciones binarias brindan las posibles opciones de localización de cada objeto representado.

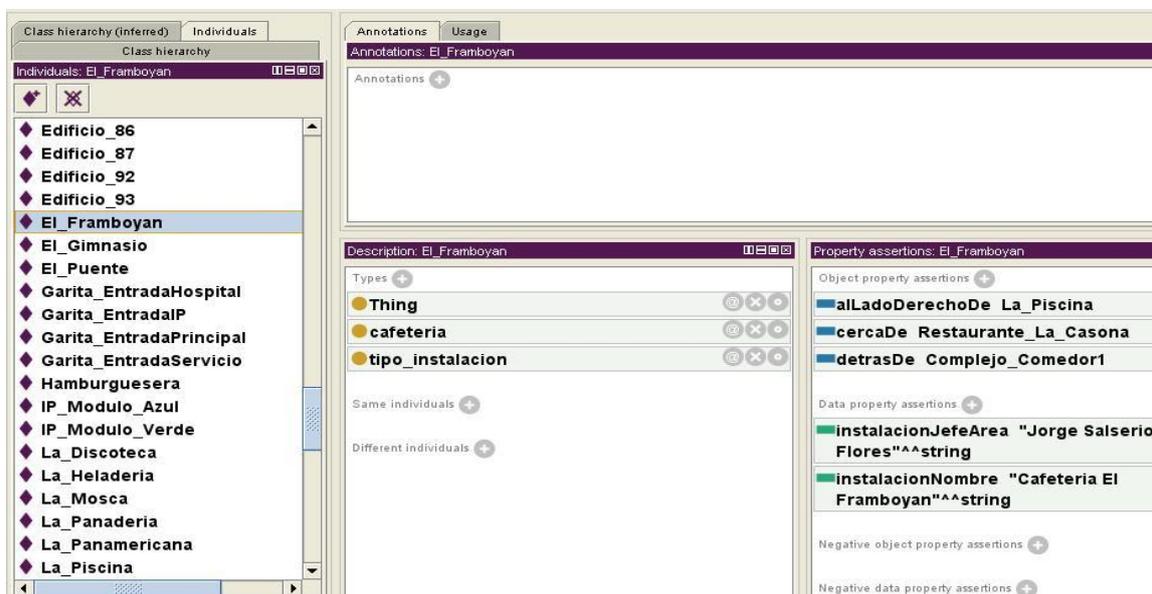


Figura 7: Instancias.

2.4 Conclusiones parciales

En este capítulo de la investigación se elaboró el glosario de términos, la taxonomía de conceptos, las relaciones entre ellos, las propiedades de los objetos y de los datos. Se definió la ontología de dominio con la que se trabajó para el diseño de la geo-ontología, la cual permitió reutilizar conceptos ya definidos y asociarlos con fuentes geográficas del SIGUCI. Con el apoyo de la metodología escogida y utilizando como base la Kaab-Ontology como ontología de dominio, se concretaron los conocimientos y se llegó a crear la versión alpha de la geo-ontología.

Capítulo 3: Análisis de criterios para la evaluación de la geo-ontología

3.1 Introducción

En este capítulo de la investigación se aplicarán mecanismos para validar tanto semántica como funcionalmente la geo-ontología. Con el apoyo de la World Wide Web Consortium (W3C), el usuario conocerá si la geo-ontología posee inconsistencias. Así mismo, mediante el uso de razonadores semánticos se realizarán pruebas para verificar el funcionamiento y cumplimiento de la geo-ontología. Según las necesidades requeridas por el sistema, se podrá comprobar la interrelación entre los conceptos, relaciones y propiedades establecidas. Los razonadores con que cuenta la versión 4.2 del *Protégé* son el FaCT++ y el Hermit 1.3.5, ambos permiten obtener el conocimiento inferido por la geo-ontología.

3.2 Validación mediante el Servicio de Validación Web de la W3C

El World Wide Web Consortium (W3C) es una comunidad internacional que desarrolla estándares que aseguran el crecimiento de la web a largo plazo. La W3C cuenta con un servicio de validación de archivos RDF, que permite validar semánticamente la geo-ontología, posibilitándole al usuario conocer si esta posee inconsistencias y confirmar la estructura de la taxonomía de conceptos.

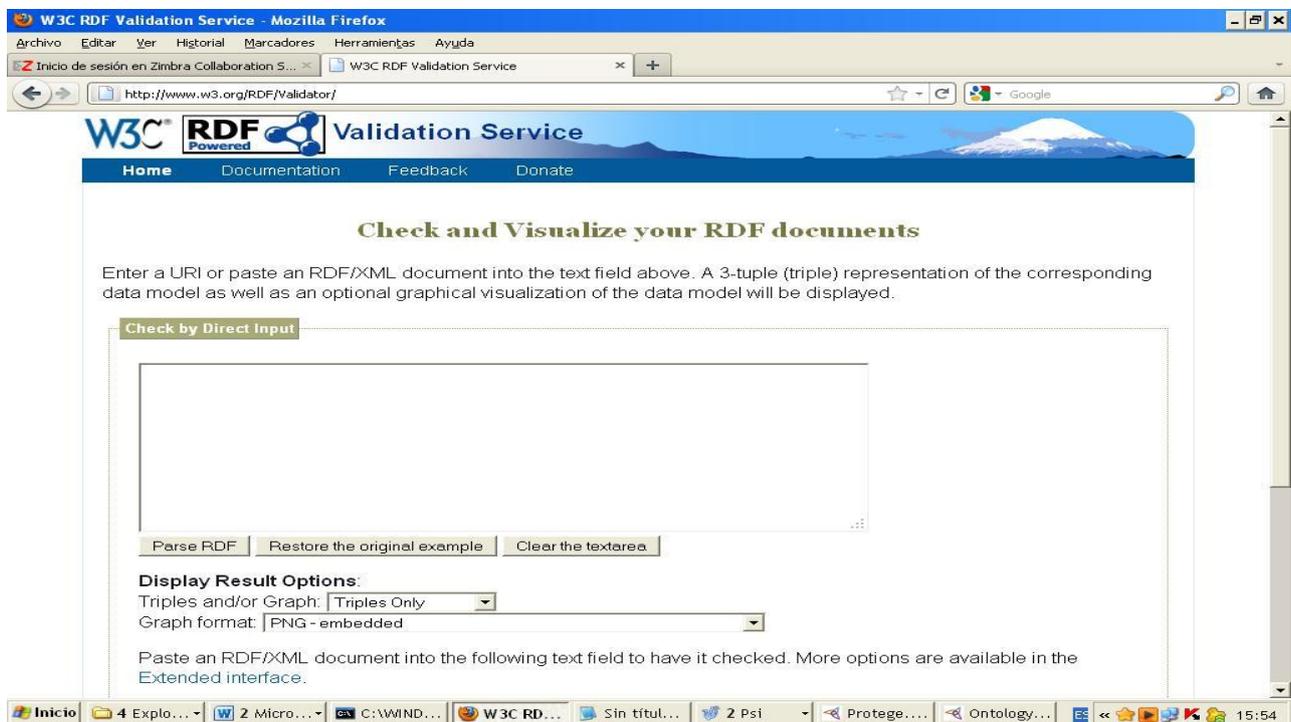


Figura 8: Servicio de Validación de Ontologías.

En la página web mostrada en la **Figura 8**, se inserta todo el código RDF de la geo-ontología y el sistema publicará, luego de parsear el código, los datos de la misma. Luego muestra el resultado de la revisión, con un mensaje si la validación fue satisfactoria o con un error que no se pudo validar el archivo.

Realizando la validación del código de la geo-ontología propuesta, se obtuvo el resultado mostrado en la **Figura 9**, demostrando que está semánticamente correcta.



The screenshot shows the W3C RDF Validation Service interface. At the top, there is a navigation bar with links for Home, Documentation, and Feedback. The main heading is "Validation Results". Below this, a message states "Your RDF document validated successfully." To the right, there is a "Jump To:" menu with links for Source, Triples, Messages, Graph, Feedback, and Back to Validator Input. Below the message, there is a section titled "Triples of the Data Model" which contains a table with three columns: Number, Subject, and Predicate.

Number	Subject	Predicate
1	http://www.semanticweb.org/ontologies/2012/4/Ontology1335976535895.owl	http://www.w3.org/1999/02/22-rdf-syntax-ns#type
2	http://www.semanticweb.org/ontologies/2012/4/Ontology1335976535895.owl#_tiene	http://www.w3.org/1999/02/22-rdf-syntax-ns#type
3	http://www.semanticweb.org/ontologies/2012/4	http://www.w3.org/1999/02/22-rdf-syntax-ns#type

Figura 9: Resultado de la geo-ontología por el Validador W3C.

3.3 Razonadores semánticos

Los razonadores semánticos son aplicaciones computacionales que permiten realizar inferencias y generar conocimiento a partir de un conjunto de axiomas y hechos. Utilizan un motor de inferencias y un conjunto de reglas expresadas en un lenguaje semántico como OWL. La complejidad del razonamiento en ontologías expresadas en OWL ha sido optimizada con Description Logic (DL) y actualmente existen razonadores como FACT++, Racer, Pellet, Flora-2, Hermit y otros, que pueden ser integrados a herramientas para el desarrollo de geo-ontologías como el *Protégé* (Milena,2010).

3.3.1 Razonador FaCT++

FaCT++ es un razonador DL que para realizar inferencia implementa nuevas características y optimizaciones, las cuales permiten adicionar nuevas tácticas de razonamiento y la capacidad de razonar lógicas descriptivas más potentes y cercanas a la expresividad de OWL-DL.

FaCT++ es un razonador eficiente para la TBox² de una geo-ontología. Sin embargo, tiene dos desventajas importantes; por un lado, no tiene soporte para otros tipos de dato que no sean string o integer, como sí ocurre por ejemplo, con Pellet.

Entre las ventajas que ofrece el FaCT++ están: Licencia GPL, todavía se encuentra en mantenimiento, sigue en desarrollo (actualmente se está investigando cómo aumentar la capacidad de razonamiento del motor de inferencia y tiene soporte de OWL 2) y trabaja de forma eficiente con la T-Box de geo-ontologías de tamaño grande y mediano (Milena,2010).

3.3.2 Razonador Hermit 1.3.5

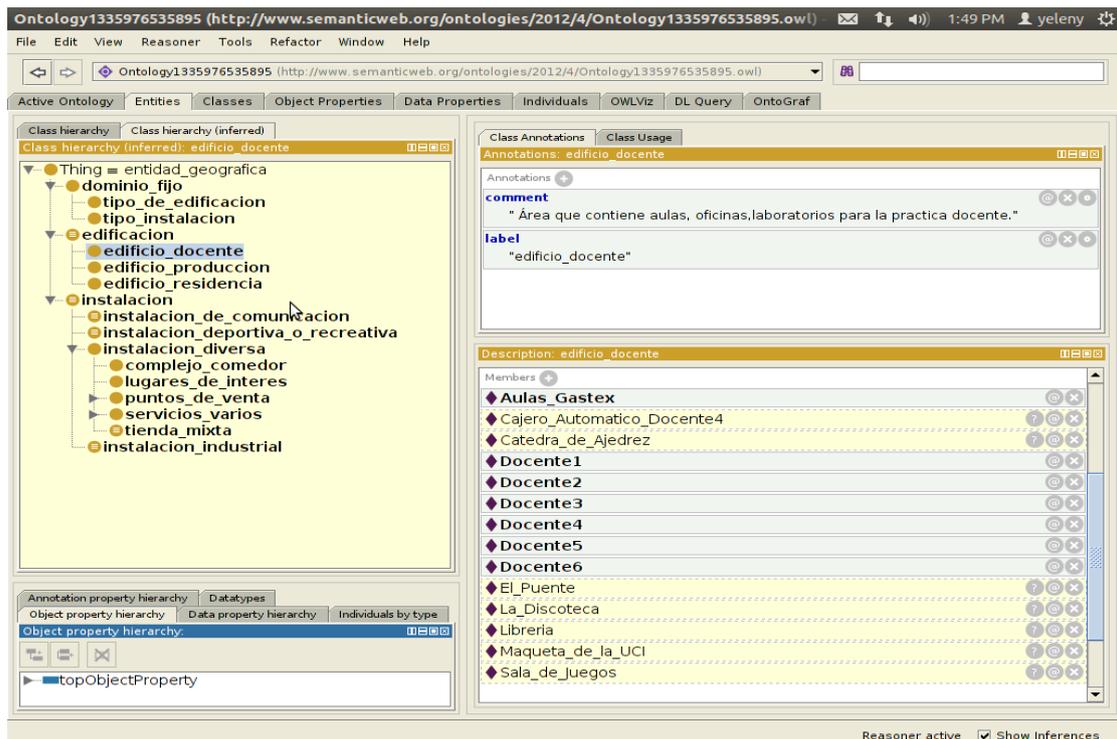
Hermit es un razonador de geo-ontologías que utiliza OWL. Dado un fichero OWL, este puede determinar si la geo-ontología es consistente, identificar las relaciones entre las clases, lo que permite comprobar el buen funcionamiento de la misma. Esta aplicación utiliza la semántica directa y pasa todas las pruebas de conformidad de OWL 2. Desde la versión 1.1, Hermit puede manejar las reglas DL y añadirlas directamente a la geo-ontología de entrada en el estilo funcional u otras sintaxis OWL soportados por la API de OWL. Hermit es de código abierto y liberado bajo licencia LGPL (Motik, 2010) .

3.4 Pruebas de Razonamiento

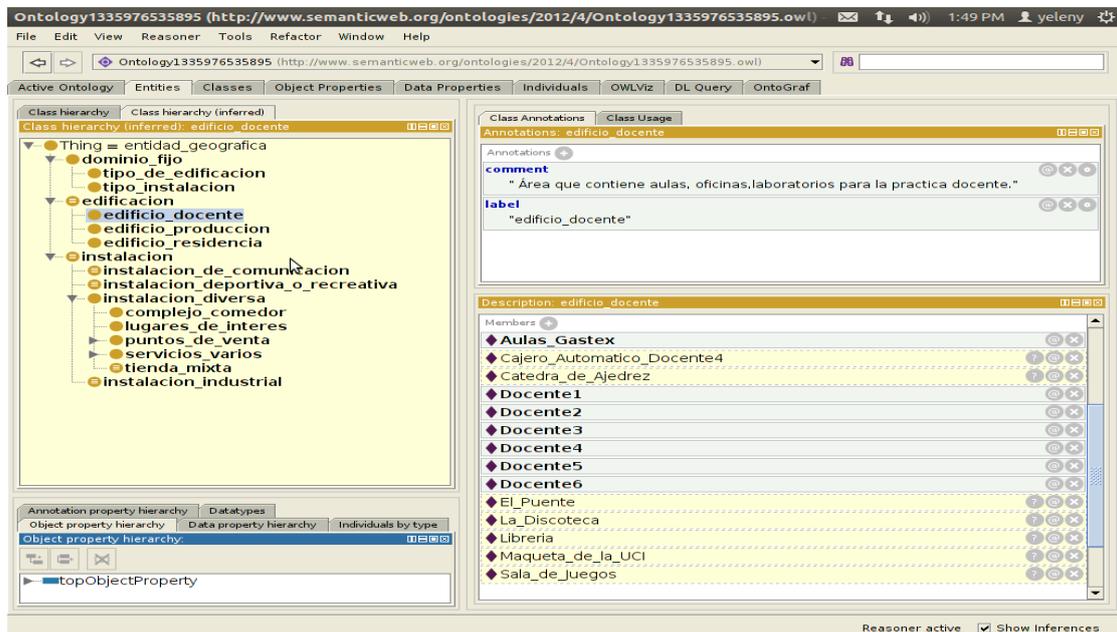
Los razonadores semánticos son plug-ins que brinda el *Protégé* para hacer inferencia sobre los datos representados en la geo-ontología. Al aplicarle un razonador a una geo-ontología los datos que se infieren se muestran sombreados en amarillo, de esta forma se comprueba la consistencia de la misma en cuanto a los conceptos y las relaciones definidas entre ellos. Realizando una corrida de ambos razonadores se obtuvieron los siguientes resultados de las clases y relaciones inferidas.

En la **Figura 10** se representa que la clase edificio docente tiene un *cajero automático*, una *cátedra de ajedrez*, la *cafetería El puente*, una *discoteca*, una *librería*, la *maqueta de la UCI* y una *sala de juegos*, que son datos inferidos a partir de las relaciones establecidas entre las clases.

² TBox: contiene sentencias describiendo conceptos jerárquicos (i.e. interrelaciones entre conceptos).



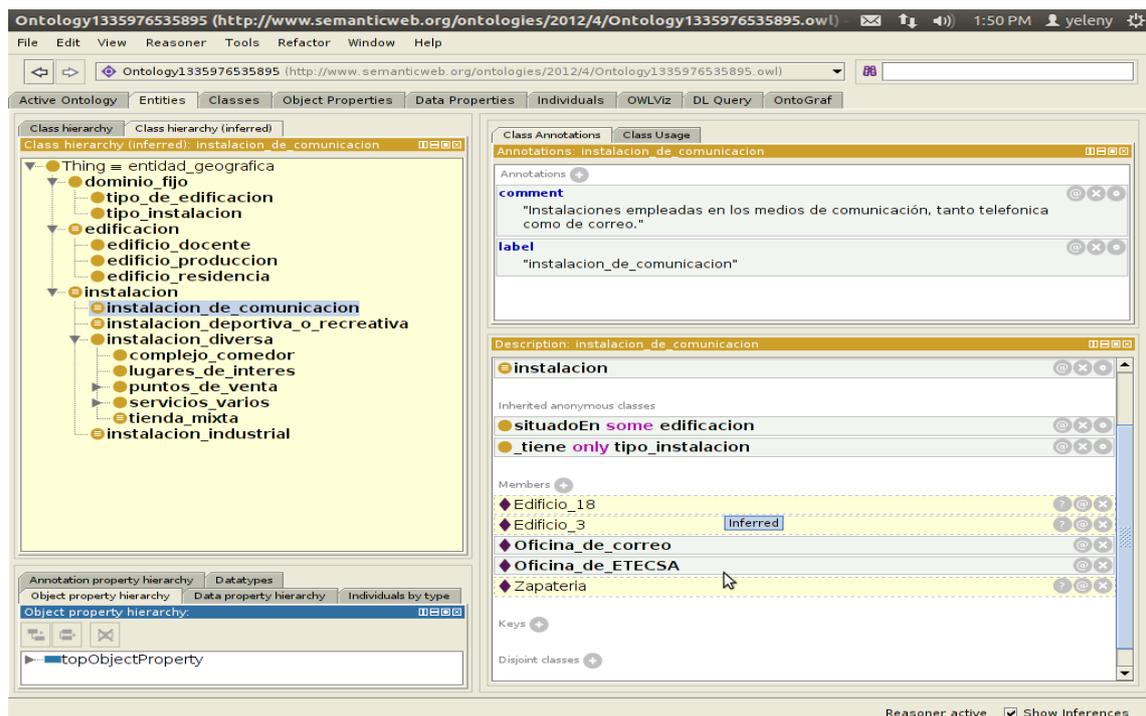
Razonador FaCT++



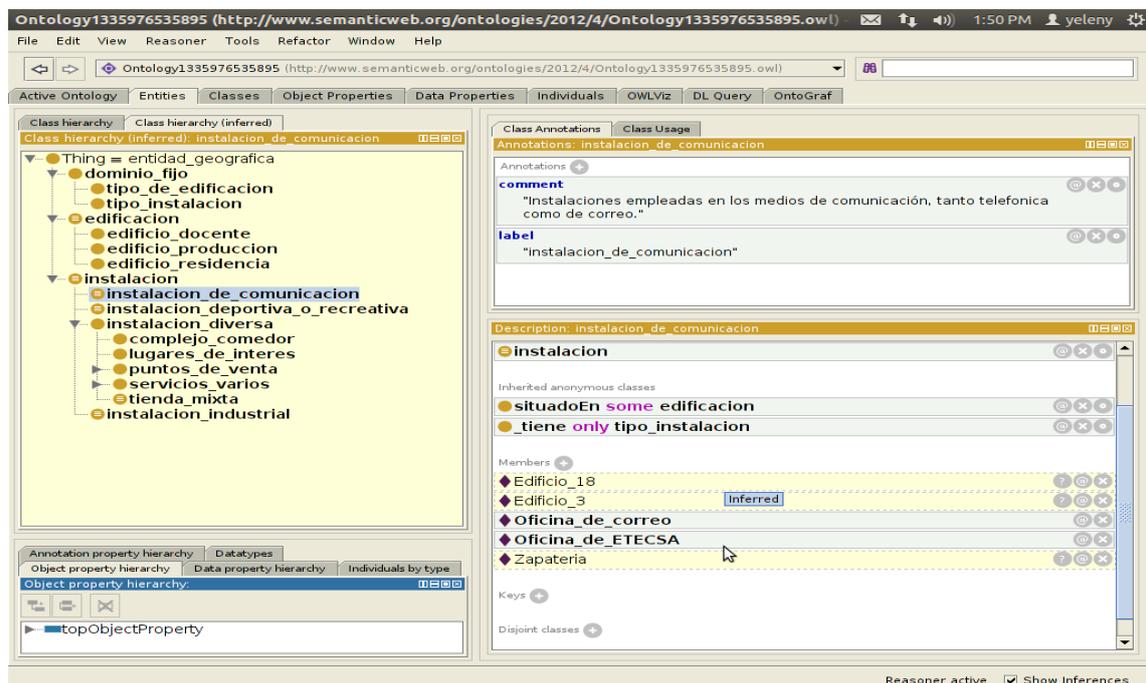
Razonador Hermit

Figura 10: Inferencia de datos con razonadores semánticos (1).

En la Figura 11 se muestra que la clase instalación de comunicación infiere el edificio 18 porque el individuo oficina de correo se localiza allí, así mismo sucede con la oficina de etecsa que está situada en el edificio 3 y la zapatería en el edificio 18



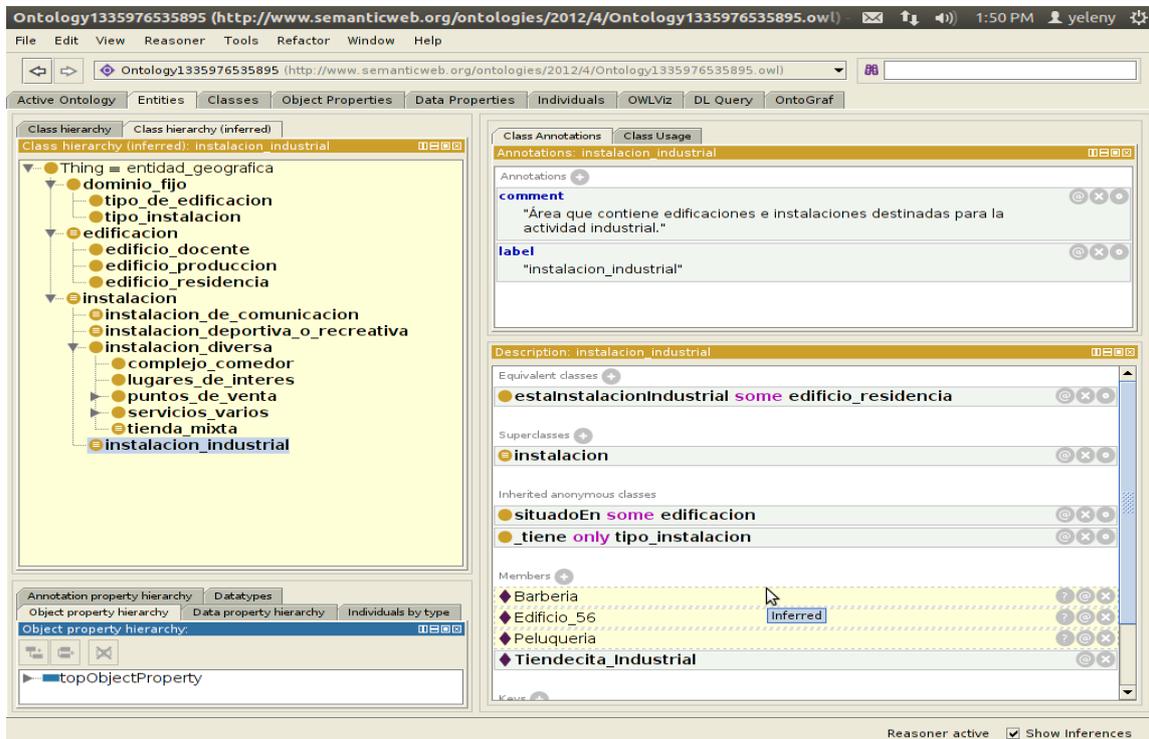
Razonador FaCT++



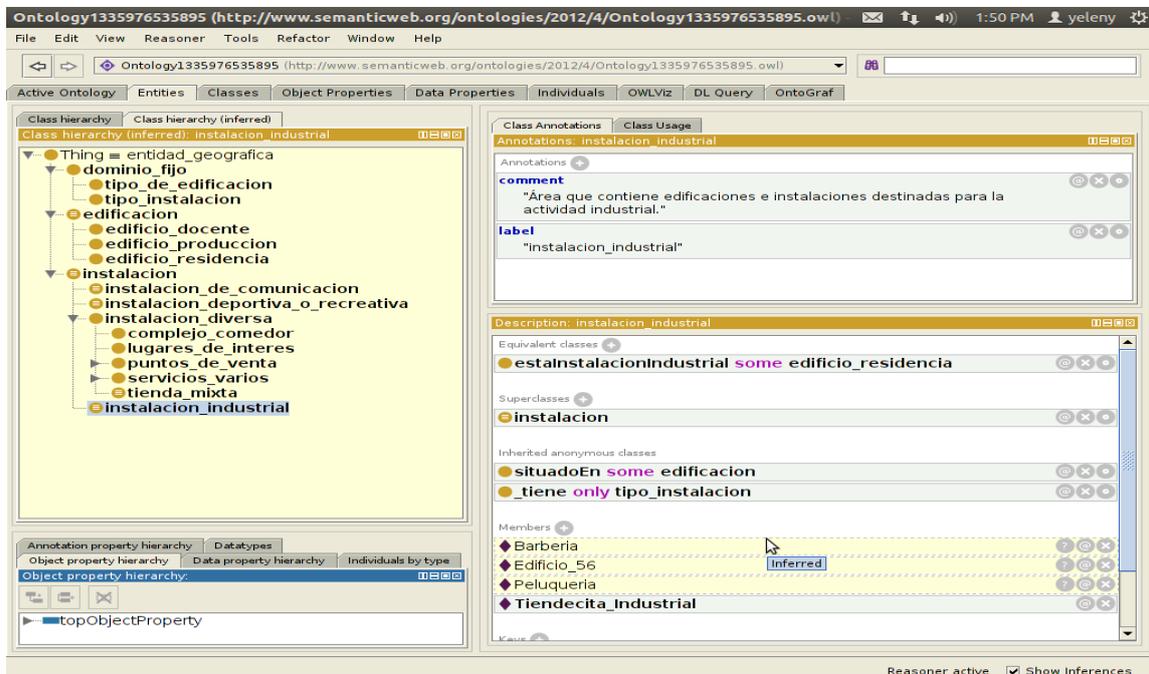
Razonador Hermit

Figura 11: Inferencia de datos con razonadores semánticos (2).

En la **Figura 12** se hace referencia que la instalación industrial tiene como individuo la *tiendecita industrial*, infiere el *edificio 56* porque es donde se localiza, al igual que la *barbería* y la *peluquería*.



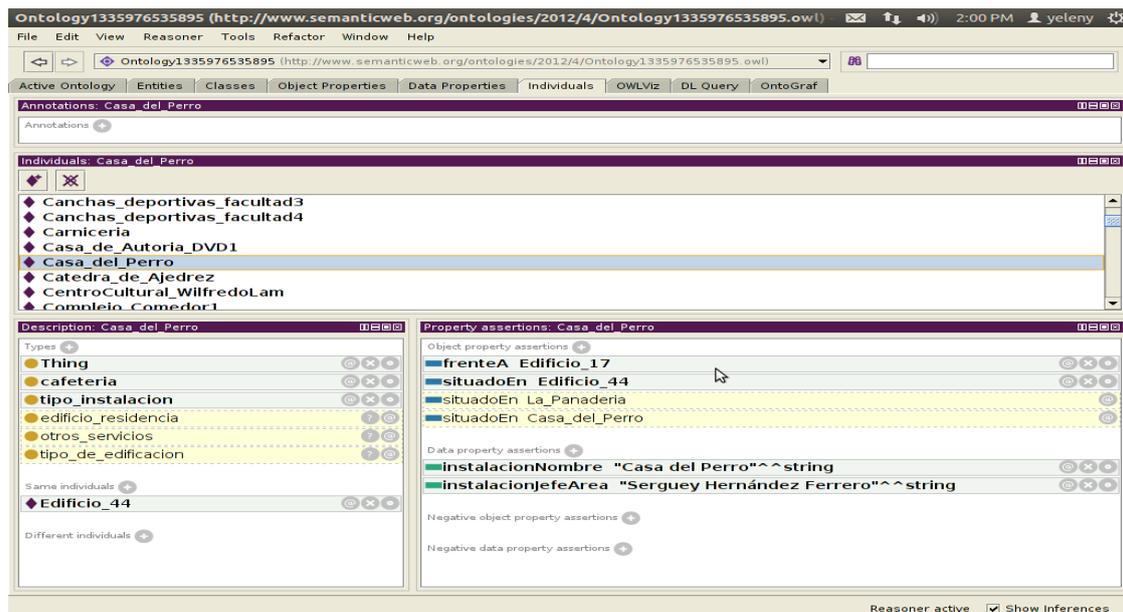
Razonador FaCT++



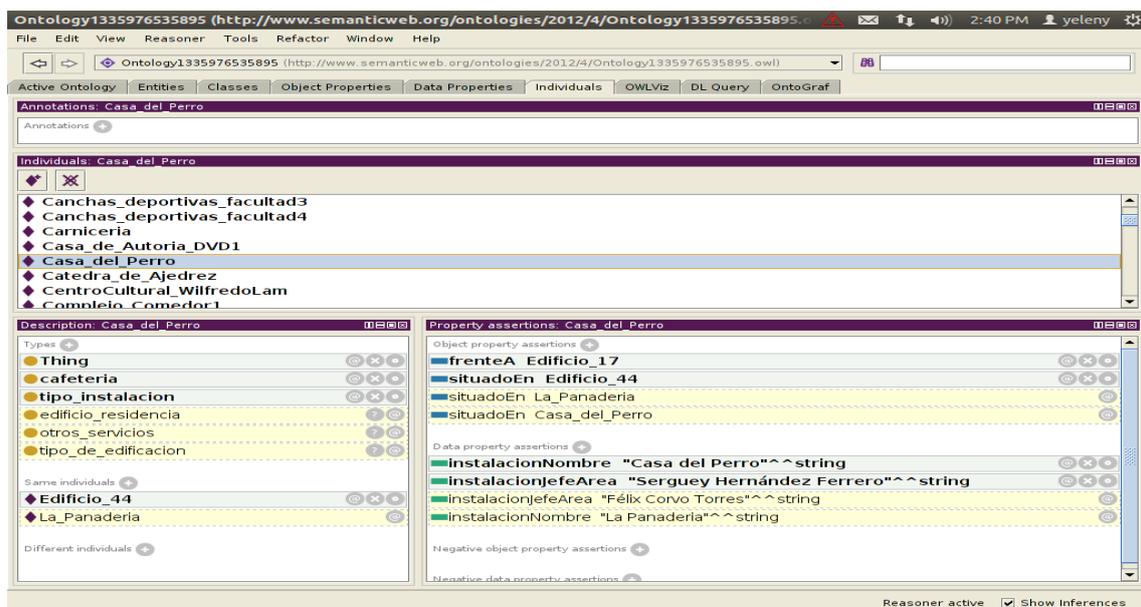
Razonador Hermit

Figura 12: Inferencia de datos con razonadores semánticos (3).

En la **Figura 13** se representa el individuo *Casa del Perro* que infiere que se encuentra en un *edificio de residencia*, que tiene relación con la clase *otros servicios*, además por estar en un *edificio de residencia*, es un *tipo de edificación*. Infiere como dato individual *la panadería*, por encontrarse situada en el mismo edificio que la *Casa del perro*. En el caso del razonador Hermit añade las propiedades de los datos del individuo *la panadería*.



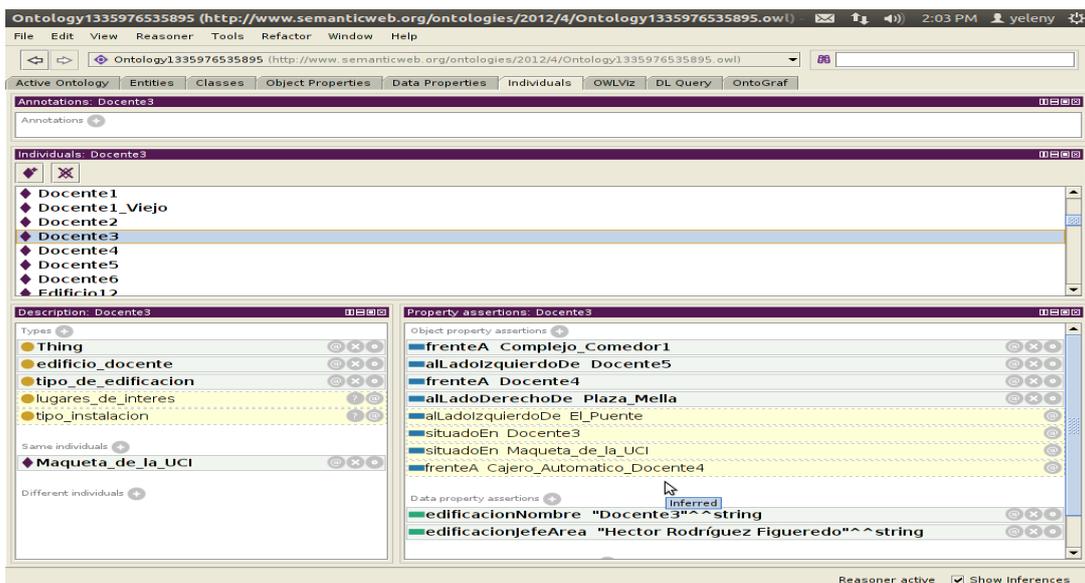
Razonador FaCT++



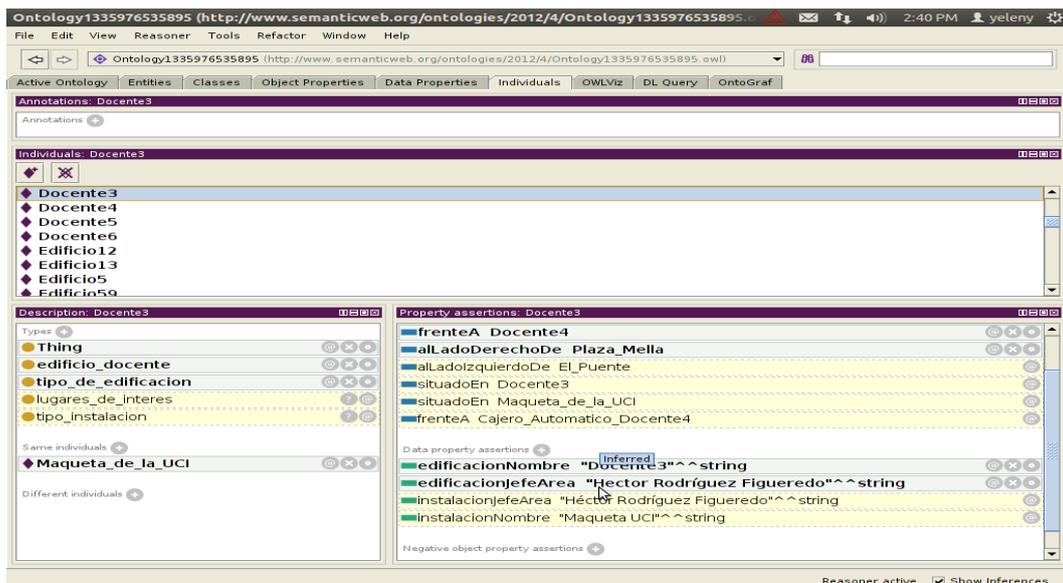
Razonador Hermit

Figura 13: Inferencia de datos con razonadores semánticos (4).

En la **Figura 14** se muestra el individuo *Docente3*, donde se infiere que es un *tipo de instalación que tiene una relación con la clase Lugares de interés* al tener en su estructura *La maqueta de la UCI*. Además se añade que *estaAlLadoIzquierdo* del Puente, el cual está situado en el docente5 y al frente del *cajero del docente 4* que se encuentra en el docente 4, esto se debe a las relaciones entre *Docente3*, *Docente4* y *Docente5*. Y en el caso del Hermit añade las propiedades de los datos de la clase instalación.



Razonador FaCT++

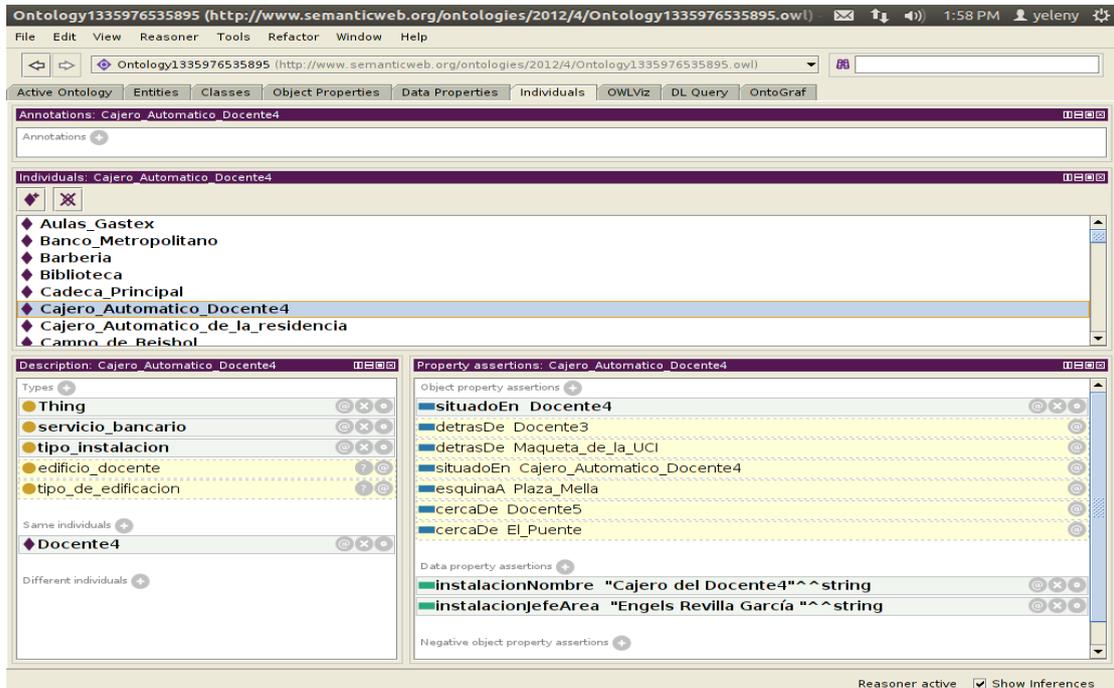


Razonador Hermit

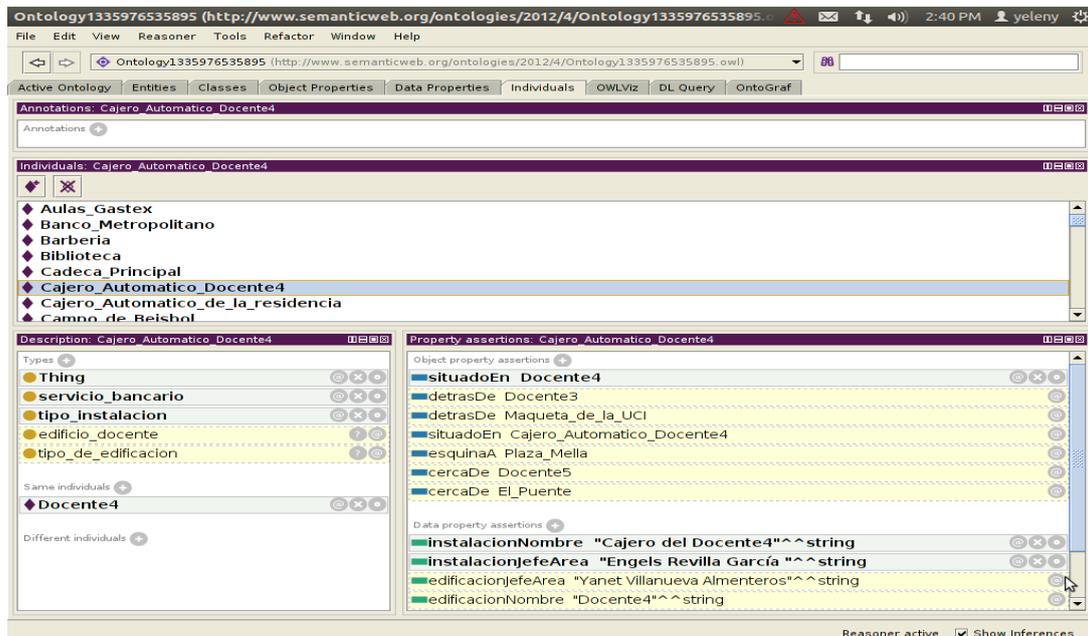
Figura 14: Inferencia de datos con razonadores semánticos (5).

Capítulo 3: Análisis de criterios para la evaluación de la geo-ontología

En la **Figura 15** se muestra el individuo *Cajero Automático*, que al estar situado en el *docente 4*, se infiere que tiene relación con la clase *edificio docente* y por tanto es un *tipo de edificación*. Además infiere que se encuentra *detrasDe* el *docente 3*, *detrásDe* *La maqueta de la UCI*, *esquina a* la *plaza Mella*, *cercaDe* el *docente 5*, y *cercaDe* la cafetería del docente 5 *El puente*. En el resultado mostrado por el razonador Hermit se muestran las propiedades de los datos de la clase edificación.



Razonador FaCT++



Razonador Hermit

Figura 15: Inferencia de datos con razonadores semánticos (6).

3.5 Comparación de los resultados por ambos razonadores

Después de haberle realizado las pruebas de razonamiento a la geo-ontología con los razonadores FaCT++ y Hermit, se comprueba que la geo-ontología posee consistencia en cuanto a la representación de los datos. La información inferida por ambos razonadores es la misma, solo que el Hermit infiere más conocimiento, principalmente en las propiedades de los objetos. Por tanto, se puede afirmar que la geo-ontología cumple con los objetivos que se plantearon al principio de la investigación, solucionando los problemas de navegación y manejo del SIGUCI. Las pruebas realizadas demuestran que la geo-ontología es capaz de inferir todo el conocimiento posible para que los usuarios que trabajen con el sistema, obtengan los datos requeridos según la necesidad informativa, debido a las relaciones establecidas entre clases e individuos y sus propiedades.

3.6 Conclusiones parciales

Con la ayuda de los razonadores semánticos escogidos, el FaCT++ y el Hermit se pudo comprobar la consistencia de los datos en la geo-ontología, sus clases, relaciones y propiedades. Se ha podido evaluar los resultados obtenidos evidenciándose la igualdad de los conocimientos inferidos por ambos razonadores. Por tanto se afirma que la geo-ontología cumple con los objetivos trazados estando sintáctica y funcionalmente correcta.

Conclusiones

Una vez obtenido el diseño de la geo-ontología y cumplido los objetivos trazados a lo largo de la investigación se arriban a las siguientes conclusiones:

Se diseñó la geo-ontología para el dominio que comprende los términos manejados en el SIGUCI, utilizando como base la ontología de dominio Kaap-ontology, en la cual se relacionan conceptos que pueden asociarse con el dominio en cuestión.

La geo-ontología diseñada permite un lenguaje de definición común, además de permitir organizar jerárquicamente los conceptos utilizados, evitando que se definan diferentes términos para un mismo tipo de información.

Con el uso de la geo-ontología se mejora el proceso de navegación en el sistema de información geográfica, debido a que se realiza con exactitud a la necesidad informativa de los usuarios.

Las pruebas realizadas mediante los razonadores semánticos integrados al *Protégé* y el validador de la W3C, demuestran que la geo-ontología presenta consistencia y un alto nivel de razonamiento para inferir la información manejada en el SIGUCI.

Recomendaciones

La geo-ontología diseñada está en su primera versión por lo que se recomienda:

- Continuar enriqueciendo semánticamente la geo-ontología, añadiendo nuevas clases, relaciones y propiedades para inferir más conocimiento.
- Realizarle pruebas de razonamiento a la geo-ontología utilizando otros mecanismos más potentes para comprobar con más exactitud su funcionamiento.
- Integrar la geo-ontología al SIGUCI.

Bibliografía

Cardona, Luis Fernando Medina. 2007. Desarrollo de una metodología para ingeniería de requerimientos en aplicaciones de sistema de información geográfica. Bogotá : s.n., 2007.

Castells, Pablo. La Web Semántica. Madrid : s.n.

Castillo, Jorge Mederos López Del. 2010. Sistema Basado en Conocimientos para apoyar la toma de decisiones en la Subdirección de Investigación y Postgrado del CDAE. La Habana : s.n., 2010.

Céspedes, Zulia Ramírez. Las ontologías como herramienta en la Gestión del Conocimiento. Ciudad de La Habana, Cuba : s.n.

Cibulsky., Indalecio Fructuoso Bezos. Lenguajes y arquitecturas de geoservicios. Tecnologías asociadas al W3C-Web Semántica. Argentina : s.n.

Concepthub.org. [En línea] [Citado el: 29 de 11 de 2011.] http://wikitel.info/wiki/Interoperabilidad_sem%C3%A1ntica.

Damián Pérez Valdés. 2007. [En línea] 25 de 06 de 2007. [Citado el: 24 de 11 de 2011.] <http://www.maestrosdelweb.com/editorial/web-semantica-y-sus-principales-caracteristicas>.

Damián Pérez Valdés. Maestros del Web. [En línea] [Citado el: 24 de 11 de 2011.] <http://www.maestrosdelweb.com/editorial/web-semantica-y-sus-principales-caracteristicas>.

Descripción de la Ontología. **López, Alberto Gómez.** España : s.n.

Díaz, Joaquín Borrego. Complementos del Tema 1. Sevilla : s.n.

Eduardo Garea Llano, Rafael Oliva Santos, Rainer Larin Fonseca, Francisco Vera Voronisky. Sistemas de información geográfica como herramienta para la integración e interpretación semántica de la información espacial.

Eduardo Garea Llano, Rafael Oliva Santos, Cynthia Costales-Llerandi. Integración Ontológica de Datos Metadatos y Conocimiento en Sistemas de Información Geográfica como Herramienta para la Interpretación Semántica de la Información Espacial.

Jose Luis Almazán Gárate, M.Carmen Palomino, Hilda Araceli Caba. 2009. Sistemas de Información Geográfica en la Gestión Integral del Litoral. Madrid : s.n., 2009.

L.McGuinness, Natalya F. Noy y Deborah. 2005. Desarrollo de Ontologías -101:Guía Para Crear Tu Primera Ontología. Stanford : s.n., 2005.

Lic. Lenay Barceló Soto. . Ontología de Dominio para el Caso "Los cinco". [En línea] [Citado el: 25 de 11 de 2011.] <http://es.scribd.com/doc/54494408/ontologias-articulo-cientifico>.

Lic.Francisco Vera Voronosky, Aduardo Garea Llano. 2009. Alineamiento de Ontologías en el dominio geoespacial. Reconocimiento de Patrones. 2009.

Llo. [En línea] [Citado el: 12 de 11 de 2011.] http://www.mappinginteractivo.com/plantilla-ante.asp?id_articulo=1051.

M.J García Rodríguez, A Urrutia zambrana y M.A Bernabé Poveda. 2000. de herramientas de análisis espacio-temporales para el estudio de bases históricas. Madrid : s.n., 2000.

Méndez,Jose A. Senso. 2004. [En línea] 2004. [Citado el: 23 de 11 de 2011.] <http://www.sedic.es/autoformacion/metadatos/tema2.htm>.

Morales, Yanssel Urquijo. 2009. Análisis y diseño de un agente semántico basado en ontologías para el dominio de la salud. La Habana : s.n., 2009.

Nuñez, Aurelio Díaz. 2011. Desarrollo de un Visor de Ontologías en OWL para la plataforma GeneSIG. La Habana. : s.n., 2011.

Oscar Corcho, Mariano Fernández López, Asunción Gómez Pérez, Angel López Cima. Construcción de ontologías legales con la metodología METHONTOLOGY y la herramienta WebODE. Madrid,España : s.n.

Rafael Oliva Santos, Eduardo Garea Llano,Francisco Maciá Pérez. Ideas para una arquitectura de persistencia basada en geo-ontologías para anotaciones semánticas de datos geográficos. La Habana : s.n.

Raimil Cruz Concepción, liz Mary Cruz,Liannet Lucia Reyes Hernández. Utilización de la Representación Formal de Conocimiento en la Toma de decisiones.

Rainer Larin Fonseca, Eduardo Garea Llano. Integración Semántica de datos Espaciales con Sistemas de Información Geográfica. La Habana : s.n.

Ramos Esmeralda, Haydemar Nuñez. 2007. Ontologías:componentes,metodologías, lenguajes, herramientas y aplicaciones. Caracas : s.n., 2007.

Ruiz, Miguel Jesús Torres. 2007. Representación Ontológica Basada en Descriptores Semánticos Aplicada A Objetos Geográficos. México : s.n., 2007.

Sosa, Alianis Pérez. 2011. Búsqueda basada en ontologías sobre Catálogo Temático de Medias Audiovisuales. La Habana : s.n., 2011.

Métodos de Recuperación semántica de datos espaciales en formato vectorial. **Alexis Martínez Guerrero, Eduardo Garea LLano. 2010.** 2141, La Habana : s.n., 2010.

Instrumentos de Representación del conocimiento:Tesauros Versus Ontologías. **Jiménez, Antonio García. 2004.** 007, Espinardo,España : s.n., 2004.

Razonadores Semánticos: un estado del Arte. **Claudia Milena Rodríguez A, Willian Cano Montaña, José Mauricio Martínez. 2010.** 21, 2010.

Belém, Priego Sánchez Angeles. 2010. Desarrollo de una Ontología de los lugares turísticos de México. México : s.n., 2010.

Henst, Christian Van Der. 1997. Maestros del web. [En línea] 1997. [Citado el: 06 de 02 de 2012.] <http://www.maestrosdelweb.com/editorial/web-semantica-y-sus-principales-caracteristicas/>.

Marín, Isabel. 2010. Ambientum.com " El portal profesional del Medio Ambiente". [En línea] 2 de 2010. [Citado el: 20 de 04 de 2012.] <http://www.ambientum.com/revista/2010/febrero/aplicaciones-medioambientales-SIG.asp>.

McGuinness, Frank van Harmelen y Deborah L. 2004. Recomendaciones del W3C. [En línea] 10 de 2 de 2004. [Citado el: 25 de 11 de 2011.] <http://www.w3.org/2007/09/OWL-Overview-es.html>.

Metamodel. 2011. [En línea] 03 de 2011. [Citado el: 20 de 11 de 2011.] <http://metamodel.wordpress.com/2011/03/04/concepto-de-interoperabilidad>.

Mi Tecnológico. 2005. Sistema De Información Geográfica. [En línea] 2005. [Citado el: 20 de 11 de 2011.] <http://www.mitecnologico.com/Main/SistemasDeInformacionGeografica>.

Motik, Boris. 2010. Information Systems Group. [En línea] 07 de 11 de 2010. [Citado el: 15 de 03

Jose Luis Almazán Gárate, M.Carmen Palomino, Hilda Araceli Caba. 2009. Sistemas de Información Geográfica en la Gestión Integral del Litoral. Madrid : s.n., 2009.

Castillo, Jorge Mederos López Del. 2010. Sistema Basado en Conocimientos para apoyar la toma de decisiones en la Subdirección de Investigación y Postgrado del CDAE. La Habana : s.n., 2010.

Nuñez, Aurelio Díaz. 2011. Desarrollo de un Visor de Ontologías en OWL para la plataforma GeneSIG. La Habana. : s.n., 2011.

Castells, Pablo. La Web Semántica. Madrid : s.n.

Ciencias de la Información . **Lic.Raúl Entralgo Amaro, Lic.Jorge Dayán Aguilar Cedeño. 2009.** 3, 2009, Vol. 40.

Ontología para el control y recuperación de información onomástica en televisión. **Jorge Caldera Serrano, Rodrigo Sánchez Jiménez. 2008.** 1, 2008, Vol. 17.

Stefan Gramann. [En línea] [Citado el: 15 de 11 de 2011.] [http://www.digitalpreservationeurope.eu/publications/briefs/es_interoperabilidad.pdf\(stefan.gradmann@rrz.uni-hamburg.de\)Stefan Gradmann.](http://www.digitalpreservationeurope.eu/publications/briefs/es_interoperabilidad.pdf(stefan.gradmann@rrz.uni-hamburg.de)Stefan Gradmann)

Ontologías Aplicación de lógicas descriptivas. [En línea] [Citado el: 05 de 03 de 2012.] [http://www.dsi.fceia.unr.edu.ar/downloads/IIA/presentaciones/Ontologias&LD.pdf.](http://www.dsi.fceia.unr.edu.ar/downloads/IIA/presentaciones/Ontologias&LD.pdf)

Peñalvo, Francisco José García. Web Semántica y ontologías. Departamento de Informática y Automática – Facultad de Ciencias Universidad de Salamanca : s.n

Anexos

Anexo I: Glosario de términos de la geo-ontología del SIGUCI.

Nombre	Sinónimo	Descripción	Tipo
dominio_fijo			Clase
tipo_edificacion		Contiene todos los tipos de edificación que existen en la universidad.	Clase
tipo_instalacion		Contiene todas las instalaciones de la universidad.	Clase
gastronomia		Arte de preparar una buena comida. Brindar servicios alimenticios a la población.	Clase
instalacion_comunicacion		Acción y efecto de comunicar o comunicarse. Correos, telégrafos, teléfonos.	Clase
lugares_de_interes		Instalaciones de disfrute según los gustos de cada individuo.	Clase
edificacion		Edificio o conjunto de edificios.	Clase
instalacion		Recinto provisto de los medios necesarios para llevar a cabo una actividad profesional o de ocio.	Clase
mercado		Conjunto de operaciones comerciales que afectan a un determinado sector de bienes.	Clase
mercado_agropecuario		Que tiene relación con la agricultura y la ganadería. Los productos que se ofertan en un mercado agropecuario son obtenidos del trabajo en la agricultura.	Clase
mercado_ideal	Mercadito	Es un tipo de mercado que oferta diversos tipos de productos.	Clase
complejo_comedor		Un área en la que se encuentran un grupo de comedores que se comunican entre sí.	Clase

edificio		Construcción fija, hecha con materiales resistentes, para habitación humana o para otros usos.	Clase
edificio_docente		Perteneciente o relativo a la enseñanza. Instalación que se utiliza para impartir clases.	Clase
edificio_produccion		Acción de producir.Instalación para producir software.	Clase
edificio_residencia		Casa donde conviven y residen, sujetándose a determinada reglamentación, personas afines por la ocupación, el sexo, el estado, la edad, etc.	Clase
servicios_varios		Organización y personal destinados a cuidar intereses o satisfacer necesidades del público o de alguna entidad oficial o privada.	Clase
servicio_proteccion		Acción y efecto de proteger.	Clase
servicio_bancario		Perteneciente o relativo a la banca mercantil.	Clase
servicio_tecnico		Habilidad para ejecutar cualquier cosa, o para conseguir algo.	Clase
servicio_Medico		Perteneciente o relativo a la medicina.	Clase
servicio_terciario		Conjunto de las actividades humanas intermedias.	Clase
instalacion_deportiva_o_recreativa		Es un área en la que se va a relajar el cuerpo y la mente en el tiempo libre. Ya sea mediante la práctica de deporte o de actividades culturales.	Clase
instalacion_industrial	Tiendecita	Es un lugar donde se venden productos industriales por moneda nacional.	Clase
tienda_mixta	TRD	Es un lugar donde se venden varios	Clase

		productos de primera categoría en este caso por moneda convertible (CUC).	
situadoEn			Object properties
perteneceA			Object properties
poseeUn			Object properties
alLadoIzquierdoDe			Object properties
alLadoDerechoDe			Object properties
cercaDe			Object properties
detrasDe			Object properties
esquinaA			Object properties
frenteA			Object properties
estaInstalacionComunicacion			Object properties
estaInstalacionDeportiva			Object properties
estaInstalacionIndustrial			Object properties
estaMercadoIdeal			Object properties
estaServicioMedico			Object properties
estaServicioTerciario			Object properties

estaTiendaMixta			Object properties
edificacionJefeArea			Data properties
edificacionNombre			Data properties
institucionJefeArea			Data properties
institucionNombre			Data properties

Anexo II: Taxonomía de conceptos de la Geo-ontología del SIGUCI.

Superclases	Subclases
instalacion_diversa	complejo_comedores, lugares_de_interes, puntos_de_venta, servicios_varios , tienda_mixta.
puntos_de_venta	gastronomia, mercado.
mercado	agropecuario, Ideal.
gastronomia	cafetería, otros_servicios.
servicio_varios	servicio_proteccion, servicio_bancario, servicio_tecnico, servicio_medico, servicio_terciario.

Anexo III: Descripción de las relaciones binarias (Object properties)

Nombre de la relación	Concepto de Origen	Cardinalidad Máxima	Concepto de destino
cercaDe	entidad_geografica	N	entidad_geografica
esquinaA	entidad_geografica	N	entidad_geografica
poseeUn	edificacion	N	instalacion
detrasDe	entidad_geografica	N	entidad_geografica

estaInstalacionComunicacion	instalacion_comunicacion	N	entidad_geografica
estaInstalacionDeportiva	instalacion_deportiva_o_recreativa	N	entidad_geografica
estaInstalacionIndustrial	instalacion_industrial	N	entidad_geografica
estaMercadoIdeal	mercado_ideal	N	entidad_geografica
estaServicioMedico	servicio_medico	N	entidad_geografica
estaServicioTerciario	servicio_terciario	N	entidad_geografica
estaTiendaMixta	tienda_mixta	N	entidad_geografica

Anexo IV: Descripción de las propiedades de los datos (Data properties).

Nombre de la relación	Concepto de Origen	Rango
edificacionJefeArea	edificacion	string
edificacionNombre	edificacion	string
instalacionJefeArea	instalacion	string
instalacionNombre	instalacion	string