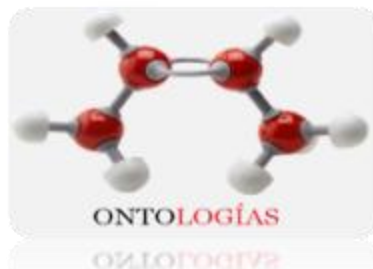




UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS
2012



*Ontología para el observatorio tecnológico
gestionado por el grupo Gestión de Información y el
Conocimiento (GIC) del Centro de Tecnologías para
la Formación (FORTES)*

Autor: Milena Pavón Acosta

Tutores: Ing. Maikel Aparicio Reytor

Lic. Yenieris Moyares Norchales

Frases



Cuando crezcas, descubrirás que ya defendiste mentiras, te engañaste a tí mismo o sufriste por tonterías. Si eres un buen guerrero, no te culparás por ello, pero tampoco dejarás que tus errores se repitan.

Pablo Neruda



(...) Detrás de cada logro, hay otro desafío. Mientras estés vivo, siéntete vivo. Si extrañas lo que hacías vuelve a hacerlo. No vivas de fotos amarillas... Sigue aunque todos esperen que abandones. No dejes que se oxide el hierro que hay en tí. Haz que en vez de lástima, te tengan respeto.

Cuando por los años no puedas correr, trota. Cuando no puedas trotar, camina. Cuando no puedas caminar, usa el bastón. ¡Pero nunca te detengas!

Madre Teresa de Calcuta

Declaración de autoría

Declaramos que somos los autores de este trabajo y autorizamos a la Facultad 4 de la Universidad de las Ciencias Informáticas, así como a dicho centro para que hagan el uso que estimen pertinente con este trabajo.

Para que así conste firmamos la presente a los _____ días del mes de _____ del año 2012.

Autora

Milena Pavón Acosta

Tutora

Lic. Yenieris Movares Norchales

Tutor

Ing. Maikel Aparicio Reyor

Dedicatoria

En especial a la principal promotora de que yo exista, la que me ha dado las ganas, las fuerzas, el apoyo que se necesita, la que ha dado toda su vida para que no me falte nada. La que me ha enseñado todo lo que sé, la que me ha hecho ser la mujer que hoy por hoy soy: MI MAMA. Eres el mejor ejemplo de madre, la que cualquier persona quisiera tener. Quiero que sepas que estoy súper orgullosa de la educación que me has dado, y quisiera que cuando me toque a mí ser madre, ser aunque sea la mitad de madre que has sido tú. Te amo.

A mi profe de los años, Felá, quisiera que sepas que nunca vas a dejar de ser mi profe, porque cada año me enseñas algo nuevo que me sirve de mucho para salir adelante en la vida.

A dos amigos que fallecieron en el transcurso de mi tiempo en la universidad, no me dio tiempo de despedirme pero quiero que sepan que los quise mucho: Paquito y David. Este diploma para ustedes que por las adversidades de la vida no se pudieron graduar. Que dios los salve en su seno.

Agradecimientos

A toda mi familia por confiar en todo momento en mí, por darme las garras para salir adelante ante cualquier obstáculo que impidiera que terminara mi carrera: Mis dos abuelas, mis tías, mis primas que son como hermanas para mí, y mi papá que aunque no se haya portado muy bien en estos 5 años, siempre me apoyó y confió en que yo sí podía.

Agradecer ante todo a las personitas más importantes para mí en esta escuela: mis verdaderos amigos: Darlenís, Danir, Yisel, Danay, Víctor Manuel, Adrián y Alberto. Gracias a ustedes tuve las fuerzas en el momento que las necesité. Gracias por soportar todas mis malcriadeces en estos 5 años. Sólo le pido a Dios que no sea este el último vínculo social en que podamos compartir. Los quiero mucho.

A mi gran amiga Yailén que tanta falta me hizo, cuando pensé que me caía.

A mis tutores que tanto me ayudaron y me guiaron por el camino correcto para realizar mi tesis.

A Nersa por desempeñar el papel de cotutora y consejera, gracias por soportar todas mis malcriadeces, también fuiste cómplice en el buen sentido de la palabra de que este sueño se me realizara. Muchas Gracias.

A Arcadio, jejejejeje...gracias por revisarme el documento y hacerme formar parte del mundo del espectáculo en la facultad, el año que viene vengo, así que guárdame el puesto que Nersa me lo quiere robar...jejejejeje.

A todos los amigos del baile: Yasmany, mi compañero de Equipo UCI, sé que me ponía pesá cuando estábamos montando las coreo, pero sabes que era en buena onda.

A Mícha y Lisbet, mis dos quantanameras, fue un placer conocer a gente más menos parecidas a mí (más menos, no se vallan a creer que son iguales a mí, ustedes saben que estoy muy dura, jejejejeje), gracias por apoyarme cuando más lo necesitaba, es una lástima que me tenga que ir de la escuela, hubiéramos sido muy buenas amigas. Y al resto de mi gente del baile, que sí me pongo a mencionarlos se me llenan los agradecimientos.

A todos los profesores que han hecho de mí hoy en día una Ingeniera Informática. Por último y no menos importante, la persona que le debo la vida, a la que más tengo que agradecerle, por la vida que me ha dado, por acogerme en su seno todos

Frases

estos años. Por darme amor, cariño, por darme las fuerzas cuando las necesité, la que más confió en mí, la que depositó toda una vida, toda las garras para que ambas saliéramos adelante. La persona que ha sido mi madre, mi padre, mi hermana, mi amiga, mi bastón. Mi mamá. Gracias mamá, y como dije en los agradecimientos y vale la pena repetir, en un futuro cuando sea madre, quisiera ser aunque sea, la mitad de madre que has sido tú. Estoy súper orgullosa de la educación que me has dado y por hacerme la mujer que soy hoy, una mujer con garras que no se rinde ante cualquier adversidad que se presente en la vida, es una de las mejores cosas que he aprendido de ti. Te amo.

Resumen

Con el desarrollo de las tecnologías de la comunicación se ha mejorado significativamente el acceso a Internet, convirtiéndose en un instrumento de uso cotidiano en la sociedad actual. La web actual, a pesar de haber alcanzado un gran éxito, presenta como principales problemas la sobrecarga y heterogeneidad de las fuentes de información existentes en ella. Una posible solución a estos problemas es la Web Semántica, permitiendo el procesamiento automático y organización de la información en la web.

La Universidad de las Ciencias Informáticas (UCI) dentro de su estructura cuenta con el grupo de Gestión de Información y el Conocimiento (GIC). La misión del Grupo es brindar servicios y productos de información con valor añadido sobre tecnología educativa para contribuir a la toma de decisiones, en función de estos son generados un gran cúmulo de información. El grupo GIC desarrolla además, un observatorio tecnológico que utiliza herramientas que gestionan un tipo específico de información cada una.

La información gestionada por estas herramientas no está organizada, integrada e interrelacionada. Trayendo como consecuencias lentitud en la recuperación de información. Por lo que para mejorar el sistema de búsquedas y el acceso a la información, se utilizará una de las tecnologías necesitadas para la implementación de la Web Semántica: la ontología para la organización de la información generada por el grupo GIC.

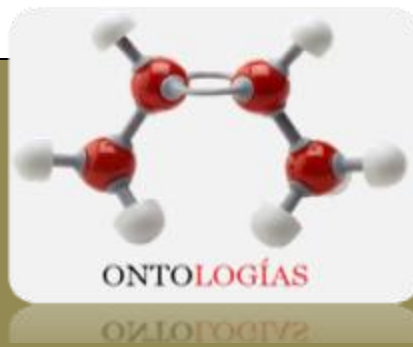
Como resultado de la investigación se diseñó una ontología para la organización, interrelación de la información que es gestionada por el grupo GIC, permitiendo mejoras en el proceso de búsqueda.

Palabras claves: Grupo de Gestión de Información y el Conocimiento, Ontología, Observatorio Tecnológico.

INTRODUCCIÓN	1
CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA	5
INTRODUCCIÓN	5
1.1: INVESTIGACIONES REALIZADAS SOBRE LA WS Y ONTOLOGÍAS	5
1.1.1 Investigaciones realizadas a nivel nacional	7
1.1.2 Investigaciones en la UCI	8
1.2: ONTOLOGÍAS: OBJETIVOS, USOS Y TIPOLOGÍA	8
1.2.1: Objetivos de las ontologías.....	10
1.2.2: Usos de las ontologías.....	10
1.2.3: Clasificación de ontologías	12
1.2.4: Componentes	14
1.2.5: Metodologías para el desarrollo de ontologías	14
1.2.6: Lenguajes para describir ontologías.....	17
1.2.7: Herramientas para el desarrollo de ontologías	19
CONCLUSIONES DEL CAPÍTULO.....	21
CAPÍTULO 2: PROPUESTA DE LA ONTOLOGÍA	22
INTRODUCCIÓN	22
2.1: CONSTRUCCIÓN DE LA ONTOLOGÍA	22
Pasos:.....	22
1. Determinar el domino y alcance de la ontología.....	22
2. Considerar la reutilización de ontologías existentes.....	23
3. Enumerar los términos importantes en la ontología	23
4. Definir las clases y la jerarquía de clases	28
5. Definir las propiedades de las clases: properties	29
6. Definir las facetas de las propiedades	30
7. Crear instancias.....	31
CONCLUSIONES DEL CAPÍTULO.....	32
CAPÍTULO 3: PRUEBA	33
INTRODUCCIÓN	33
3.1: EVALUACIÓN DE LA ONTOLOGÍA	33
3.1.1 Fases:.....	35
1. Uso correcto del lenguaje	35

Índice

2. Exactitud de la estructura taxonómica	36
3. Validez del vocabulario	37
4. Adecuación a requerimientos	38
CONCLUSIONES DEL CAPÍTULO.....	38
CONCLUSIONES GENERALES	39
RECOMENDACIONES.....	40



Introducción

La web ha cambiado profundamente la forma de comunicación, negociación y realización del trabajo. Hoy en día se pueden realizar transacciones económicas mediante Internet, además de tener acceso a millones de recursos, sin importar la localidad ni el idioma. Todos estos factores han contribuido al éxito de la Web. Sin embargo, al mismo tiempo han propiciado sus principales problemas: sobrecarga de información y heterogeneidad de fuentes de información con el consiguiente problema de interoperabilidad.

*“La **Web Semántica (WS)** ayuda a resolver estos dos importantes problemas permitiendo a los usuarios delegar tareas en 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”. Constituye “una Web extendida, dotada de mayor significado en la que cualquier usuario en Internet podrá encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida. Esta Web extendida y basada en el significado, se apoya en lenguajes universales que resuelven los problemas ocasionados por una WS casi inexistente en la que, en ocasiones, el acceso a la información se convierte en una tarea difícil y frustrante.” (W3C, 2010)*

Se conoce además que la WS resolvería problemas aún existentes en la web como el que a continuación se plantea: diferentes sistemas web pueden utilizar diferentes identificadores para un mismo concepto. Así, un programa que quiera comparar o combinar información entre dichos sistemas tiene que conocer qué términos significan lo mismo. Idealmente, el programa debería tener una forma de descubrir los significados comunes de cualquier base de datos que encuentre. Es por eso que una solución a este problema es la WS, incluyéndole un nuevo elemento a colecciones de información denominadas **ontologías**.

*“Una **ontología** es una herramienta conceptual que define un vocabulario común para quien necesita compartir información dentro de un determinado dominio. Esto incluye definiciones de los*

conceptos básicos del dominio, así como sus relaciones, que tienen que ser interpretables por máquinas. (Peñalvo, 2004)”

Las ontologías son usadas para el favorecimiento de la comunicación entre distintas entidades, desde organizaciones, personas, hasta aplicaciones, permitiendo la interoperabilidad entre sistemas informáticos con el objetivo final de llegar al razonamiento automático. Dicho razonamiento permitiría acceder y gestionar la inmensa cantidad de información existente de una forma más eficaz.

El desarrollo de ontologías, que se ha visto activamente impulsado con la WS, tendría también repercusión directa en el ámbito del procesamiento del lenguaje natural, ya que las ontologías como representaciones de un dominio e independientes de la lengua pueden ser el punto de encuentro entre dos o más lenguas. Seguidamente están los sistemas de búsqueda de respuestas, lo que permitiría ir un paso más allá en la búsqueda efectiva de información al responder directamente a las preguntas formuladas por los usuarios. En lugar de devolver el documento completo, daría nada más la zona del texto donde se encuentra la información requerida. Estos sistemas abrirían la puerta a la comunicación en lenguaje natural, que es la forma más ágil y sencilla de expresarse para las personas.

Las ontologías tienen una gran ventaja para las bases de datos y para la recuperación de información. Mediante las ontologías se pueden usar bases de datos que contienen información con un contenido completamente dispar de manera simultánea, a las que se le pueden hacer preguntas de cualquier índole para la recuperación de información. Esto puede suceder sin tener que renunciar a la fuente original de la información, permitiendo que ambas puedan convivir y reutilizarse al mismo tiempo. Nada más se debe desarrollar un sistema informático que actualice automáticamente la ontología con los datos de la base de datos (manualmente se puede hacer si lo desea). El sistema informático transformará el formato de estos a las características propias de la ontología. (Contreras & Martínez Comeche, 2002)

Un ejemplo de instituciones que están investigando e implantando el uso de ontologías es la Universidad de las Ciencias Informáticas (UCI), dedicada al desarrollo de software además de la actividad investigativa, está dividida en centros productivos. El Centro de Tecnologías para la Formación (FORTES), cuenta con un Grupo de Gestión de Información y el Conocimiento (GIC). La misión del grupo es brindar servicios y productos de información con valor añadido sobre tecnología educativa para contribuir a la toma de decisiones, en función de estos servicios y productos se genera un gran cúmulo de información. El grupo GIC desarrolla además, un observatorio tecnológico que utiliza herramientas como: Open Journal System (OJS¹), Open

¹ <http://www.pkp.sfu.ca/ojs>

Harvester System², Alfresco³, Drupal⁴; cada una de estas herramientas gestiona un tipo específico de información.

Open Journal System gestiona información referente a la creación de boletines, permite la búsqueda y recuperación de tesis, artículos y las ediciones anteriores de los boletines por año y mes.

Open Harvester System gestiona los metadatos de los artículos añadidos al Open Journal System.

Alfresco gestiona los documentos administrativos.

Drupal es un CMS⁵ (Content Management System) en el cual se está desarrollando el portal del observatorio que tendrá integrado la información de todas las herramientas mencionadas anteriormente. Además permite la búsqueda y recuperación según las necesidades de los usuarios del centro.

Independientemente de que la información gestionada por estas herramientas sea completamente diferente, no está organizada, integrada e interrelacionada. Todos estos problemas traen como consecuencias lentitud en la recuperación de información. Además, cabe la posibilidad de que alguna herramienta necesitada esté fuera de servicio en ese momento, o que los resultados de la búsqueda realizada por parte del usuario no estén en correspondencia con la necesidad informativa, provocando insatisfacción por parte del mismo.

A partir de la necesidad anteriormente expresada se plantea el siguiente **problema a resolver**:

¿Cómo contribuir a interrelacionar la información gestionada por el grupo GIC?

Se define como **objeto de estudio** la Ingeniería Ontológica, enmarcado en el **campo de acción** ontología para la gestión de la información almacenada por el grupo GIC del centro FORTES.

Como **objetivo general** se propone elaborar una ontología para mejorar la gestión e interrelación de la información del grupo GIC del centro FORTES.

El objetivo general se desglosa en los **objetivos específicos** siguientes:

- Identificar elementos principales de las ontologías empleados para describirlas, abordados por autores en el tema.
- Identificar los conceptos claves que gestiona el grupo GIC para el desarrollo de la ontología.
- Representar en una ontología los conceptos claves gestionados por el grupo.
- Implementar la ontología.
- Realizar pruebas a la ontología diseñada.

La **idea a defender** se plantea de la siguiente manera:

² pkp.sfu.ca/harvester/

³ www.alfresco.com/

⁴ www.drupal.org/

⁵ Sistema para gestionar contenido

Con la creación de una ontología se contribuirá a interrelacionar la información gestionada por el grupo GIC.

Para dar cumplimiento a los objetivos propuestos se plantearon las siguientes **tareas investigativas**:

- Sistematización de los referentes teóricos metodológicos que sustentan: tecnologías referentes a las ontologías.
- Análisis del estado del arte de las ontologías y WS.
- Elaboración del diseño de la ontología.
- Implementación de la ontología.
- Valoración de la ontología.

Métodos utilizados:

-Métodos teóricos:

Se empleó el **analítico-sintético** con el objetivo de analizar los conceptos, metodologías, lenguajes y herramientas sobre ontologías, abordados por los autores en el tema con el fin de sintetizar los aspectos necesarios para el estudio. El **histórico-lógico** para realizar el estudio de algunos elementos de las ontologías, que servirán de guía para un buen diseño de la propuesta de solución. Por último el método de **modelación** para el diseño de la ontología del grupo GIC.

-Método empírico:

La **Entrevista** para el levantamiento de la información con el objetivo de recopilar los datos que son gestionados por el grupo GIC con la finalidad de realizar el modelamiento ontológico.

El presente trabajo consta de introducción, tres capítulos, conclusiones, bibliografías utilizadas en el desarrollo de la investigación y los anexos que complementan el cuerpo del trabajo.

Capítulo 1. Fundamentación teórica

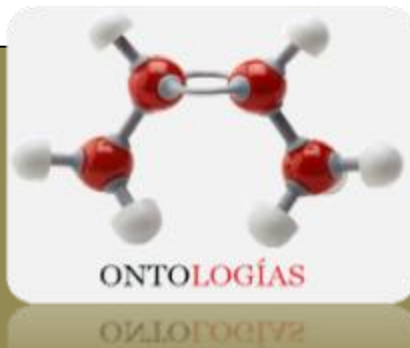
En este capítulo se realizará la Ingeniería Ontológica, donde se analizarán los principales aspectos teóricos: conceptualizaciones, herramientas, metodologías y lenguajes para describir ontologías.

Capítulo 2: Propuesta de la ontología

Se realizará un diagnóstico de los resultados de una entrevista realizada al grupo GIC para el levantamiento de la información gestionada por el mismo y de esta manera desarrollar la ontología en el lenguaje y herramienta propuesta siguiendo la metodología ontológica seleccionada.

Capítulo 3: Prueba

Se llevarán a cabo pruebas a la ontología realizada. En este capítulo se aclara cuál es el método y técnica utilizados y se recogen los resultados de la ejecución de las pruebas.



Capítulo 1: Fundamentación teórica

Introducción

El presente capítulo tratará brevemente elementos fundamentales de las ontologías: conceptos, objetivos, usos, tipologías y los distintos elementos a tener en cuenta para su creación, dígame metodologías, lenguajes y herramientas. Así como también, un estado del arte de las ontologías, el cual abordará investigaciones realizadas de las ontologías y la WS por distintos autores.

1.1: Investigaciones realizadas sobre la WS y ontologías

Son muchos los autores e instituciones que han realizados investigaciones sobre la WS y las ontologías:

“En mayo del 2001, Tim Berners Lee, James Hendler⁶ y Ora Lassila⁷ popularizan la idea de la WS al publicar un artículo en la revista Scientific American titulado "The Semantic Web: a new form of Web content that is meaningful to Computers will unleash a revolution of new possibilities", donde explican de forma sencilla su idea de la WS y los primeros pasos que hay que dar para llevarla a cabo.” (Báez, 2011)

También Mari Vallez en su artículo: “La Web Semántica y las Tecnologías del Lenguaje Humano” motiva y explica las propuestas de la WS. Este artículo expresa además las tecnologías más importantes desarrolladas, así como el punto en el que se encuentra el proyecto promovido por el propio inventor de la web. (Vallez, 2009)

El artículo de María Jesús Lapuente “Hacia la Web Semántica”, describe conceptualizaciones, funciones de las distintas capas de la WS, así como los lenguajes que la misma precisa, y proyectos más importantes que aplican la semántica. (Lapuente, 2011)

⁶ investigador de la inteligencia artificial en el Instituto Politécnico Rensselaer , EE.UU. , y uno de los creadores de la Web Semántica .

⁷ finlandés informático que vive y trabaja como investigador en el Nokia Research Center, en Cambridge, Massachusetts , EE.UU. Él ha estado llevando a cabo investigaciones sobre la Web Semántica desde 1996, y fue co-autor, junto con Tim Berners-Lee y James Hendler , del artículo "La Web Semántica"

La Oficina Española del W3C⁸ posee documentos y recursos sobre la WS. La misma contiene las siguientes guías introductorias: “Guía breve de la Web Semántica⁹” y “Guía breve de Servicios Web¹⁰”.

Francisco José García expone la problemática de la web actual y plantean como posible solución la WS, con la introducción de un nuevo elemento: ontología, donde plantea su concepto, componentes, tipos, aplicaciones y metodologías para su construcción. (Peñalvo, 2004)

Miguel Ángel en su investigación, refiere que el desarrollo de la Web influirá en nuestra forma de usarla, y de cómo se admite que la web del futuro será "inteligente", ya que estará llena de información que las máquinas podrán comprender y sacar conclusiones para los usuarios. (Abián, 2005)

Hoy en día la semántica ya está siendo utilizada por distintos proyectos. Ejemplo de estos son: Haystack del MIT¹¹ (Lapuente, 2011), buscadores semánticos, y el proyecto RETSINA Calendar Agent (RCAL¹²) (Soltero, Navarro Hernández, Sánchez Schmitz, & Barceló Valenzuela, 2007).

Al incluirle la semántica a estos proyectos, se ha logrado facilitar la búsqueda, recuperación y manipulación de la información gestionada. Gracias a la WS se puede compartir los datos y lograr el intercambio de la información, con el objetivo final de conseguir proyectos más eficientes y competitivos con ayuda de las tecnologías de la información. Lo que se quiere con la WS es que la información contenida en ella esté definida de modo tal que pueda ser reutilizada y comprendida por las máquinas automáticamente. De esta manera, la web se convertiría en un espacio capaz de realizar deducciones lógicas además de pensar como lo haría el hombre, todo esto gracias a las ontologías. En los últimos años, varios proyectos han creado sistemas de software basados en ontologías.

Son varios los autores que han estudiado esta tecnología:

Los autores Torcoroma Velásquez, Andrés Mauricio y Jaime A. Guzmán, los dos primeros profesores de la Universidad Francisco de Paula en Colombia, y el último profesor de la Universidad Nacional de Colombia, publicaron el artículo: "Ontologías: una técnica de representación de conocimiento" en la Revista Avances en Sistemas e Informática¹³. El artículo presenta la Ingeniería Ontológica como buena práctica para la representación de conocimientos; dando las principales conceptualizaciones de la WS y el papel que representa la ontología dentro

⁸<http://www.w3c.es/>

⁹<http://www.w3c.es/divulgacion/guiasbreves/WebSemantica>

¹⁰<http://www.w3c.es/Divulgacion/Guiasbreves/ServiciosWeb>

¹¹<http://groups.csail.mit.edu/haystack/>

¹² creado por tres personas del Instituto de Robótica de la universidad estadounidense de Carnegie Mellon: Terry R. Payne, Rahul Singh y Katia Sycara (directora de tesis). Se empezó a desarrollar a finales de los años 90.

¹³ publicación seriada, perteneciente a la Escuela de Ingeniería de Sistemas de la Facultad de Minas de la Universidad Nacional de Colombia Sede Medellín, la cual está dedicada a la difusión de los resultados de la investigación básica y aplicada sobre asuntos relacionados con desarrollos en el área de los sistemas, informática, computación y afines. avances_med@unal.edu.co

de su estructura. Se exponen además las distintas metodologías y tecnologías para describir ontologías. (Velásquez Pérez, Puentes Velásquez, & Guzmán Luna, 2011)

El autor Francisco Javier Honrubia de la Escuela Universitaria Politécnica de Albacete, realiza un artículo titulado "Introducción a las ontologías". En el mismo, el autor plasma diferentes conceptualizaciones de las ontologías dadas por autores como Gruber¹⁴, además de identificar los distintos usos de las ontologías y sus beneficios. (López, 2002)

Muchos son los proyectos en la actualidad que están llevando a cabo el uso de las ontologías. Uno de estos proyectos es el proyecto TELOS¹⁵(Representing Knowledge About Information Systems), es la integración de los subproyectos del e-learning LORNET¹⁶ PanCanadian, explica cómo las ontologías se utilizan en el proceso de desarrollo de la arquitectura conceptual de TELOS, para la implementación del sistema y en el uso real del mismo. Presenta la transición de una ontología conceptual a la técnica ontología que realmente acciona el sistema. Este proyecto va a liderar el camino para futuras investigaciones, desarrollos y aplicaciones fructíferas de aprendizaje basado en Web y los sistemas de gestión del conocimiento. (Paquette & Magnan, 2008)

Otro proyecto que incluye ontología en su estructura es EU4ALL¹⁷(European Unified Approach for Assisted Life Long Learning). El objetivo de este proyecto es desarrollar un mecanismo que implemente la ontología del usuario definida en EU4ALL y que pueda ser utilizado en dotLRN¹⁸. Para ello, dicho mecanismo deberá ofrecer los servicios web necesarios que permitan su integración en dicha plataforma. El proyecto no sólo tendrá resultados técnicos sino también apoyo y buenas prácticas desde el punto de vista de los proveedores de educación a distancia e-learning como desde el punto de vista del estudiante. (CALVO, GHIGLIONE, & ELLIS, 2006)

1.1.1 Investigaciones realizadas a nivel nacional

En Cuba se han realizado investigaciones sobre el tema con resultados que han aportado al desarrollo del país. Perojo en su investigación explora componentes y tecnologías de la WS, explica brevemente el modelo constructivo de capas que garantizará el funcionamiento de toda su infraestructura. Ofrece algunas consideraciones en cuanto a los beneficios brindados por la Web Semántica en el desarrollo de sistemas de gestión de contenidos. (PEROJO, 2006)

Labrada ilustra cuáles serían las nuevas posibilidades de desarrollo del profesional con la Web Semántica. Señala un nuevo ambiente de trabajo para el profesional de la información con las

¹⁴ científico americano de la computación, inventor y empresario con un enfoque de sistemas para el intercambio de conocimientos y la inteligencia colectiva. Es bien conocido por su definición de ontologías en el contexto de la Inteligencia Artificial

¹⁵ sistema operativo especializado en la parte superior de una red informática diseñada para apoyar educación a distancia. Es la producción principal de la integración producida dentro del proyecto LORNET.

¹⁶ es una red de agrupación experiencia de más de cien investigadores, socios de investigación y profesionales, y estudiantes de posgrado de Canadá. <http://www.lornet.ca/>

¹⁷ ambicioso proyecto que tiene como objetivo influenciar la forma en que universidades e instituciones educativas de dentro y fuera de Europa distribuyen sus servicios de aprendizaje a lo largo de la vida

¹⁸ completo sistema de gestión de cursos, una herramienta de código abierto o software libre, enfocada a la gestión de comunidades de prácticas colaborativas en línea.

nuevas tecnologías y aborda conceptos que la sociedad de la información de hoy desconoce y no se puede dejar pasar por alto. (LABRADA, 2003)

1.1.2 Investigaciones en la UCI

Suárez y Paz en su tesis que lleva por título “Mecanismos Semiautomáticos y Automáticos para la Generación de Ontologías de Dominio” exponen una propuesta teórica para la creación de ontologías de forma automática y semiautomática, que podría ser aplicada en un futuro, específicamente a los repositorios de objetos de aprendizaje para que de esta forma sean convertidos en repositorios semánticos. (SUÁREZ & PAZ, 2008.)

Bravo y Reytor emprendieron una investigación para la creación de un módulo de inferencia para la gestión del conocimiento contenido en los objetos de aprendizaje almacenados en el Repositorio de Objetos de Aprendizaje de la UCI, dicha investigación tenía como objetivo garantizar una recuperación efectiva del conocimiento existente. (BRAVO & M., 2009)

La presente investigación aportará conocimiento de autores que tratan el tema de la WS, así como temas tratados por ellos a partir de la investigación científica y relaciones de colaboración con otros investigadores. Se conocerán además los principales aportes de investigaciones sobre el tema realizados en la UCI por estudiantes ya graduados, y las investigaciones realizadas a nivel nacional.

La WS pretende superar la actual orientación de Internet, mediante la WS las personas encuentran la información de manera clara y precisa, pretende tener además un modelo donde la información está preparada para su procesamiento automatizado por parte de programas de computadora. De esta manera una ontología se convierte en la vía para representar el conocimiento de la Web de forma que se haga legible y reutilizable por los ordenadores. En el presente trabajo se llevará a cabo un estudio de la misma, desde los conceptos que han dado diferentes autores, usos, tipologías, así como las metodologías, lenguajes y herramientas para la descripción de la misma, dichos componentes serán descritos en el siguiente subepígrafe.

1.2: Ontologías: objetivos, usos y tipología

Uno de los elementos fundamentales en la WS es el uso de las ontologías para la integración semántica, porque *“intentan formular un esquema conceptual exhaustivo y riguroso de un dominio particular, con la finalidad de facilitar la comunicación, reusar y compartir información entre organizaciones, computadores y humanos.”* (Ramos & Nuñez, 2007)

Existen diferentes denominaciones de ontología, Swartout plantea:

- ✓ *Ontología es la teoría o estudio del ser, más específicamente, de las características básicas de toda realidad. Durante el siglo XVII el término ontología se usaba como sinónimo de metafísica.*

- ✓ *Ontología es una rama de la metafísica que investiga y explica la naturaleza y las propiedades esenciales y relaciones de todo ser, o los principios y causas del ser.*
- ✓ *Ontología es el estudio metafísico de la naturaleza del ser y la existencia.*
- ✓ *Ontología es la rama de la metafísica que trata con la naturaleza del ser.* (Swartout, 1999)

La ontología tiene otras conceptualizaciones dadas por autores que han facilitado su aporte en varias áreas como la Inteligencia Artificial y la Informática, ejemplo de ello se muestra a continuación:

En la Inteligencia Artificial:

- ✓ Ceccaroni la describe como *“la especificación explícita de una conceptualización”*. (Ceccaroni, 2001)
- ✓ Según Gómez es la *“librería de definiciones que pueden ser utilizadas para diferentes propósitos en diferentes dominios, que permiten compartir y reutilizar conocimiento y métodos de razonamiento entre agentes”*. (Gómez Pérez, 1995)
- ✓ Otros como Devedzic explican que la ontología es un *“sistema de conceptos definidos e interpretados de manera declarativa. El sistema define el vocabulario de un dominio y las restricciones que aplican para que éstos puedan ser combinados en el modelo del dominio”*. (Devedzic, 2002)
- ✓ Tramullas la define como el *“resultado de seleccionar un dominio y aplicar sobre el mismo un método para obtener una representación formal de los conceptos que contiene y las relaciones que existen entre estos”*. (Tramullas, 1999)
- ✓ Gruber dijo *“una ontología es una especificación formal y explícita de una conceptualización compartida”*. (Gruber, 1993)
- ✓ Por último Arano planteó *“las Ontologías son para la Inteligencia Artificial recursos contruidos que permiten representar el conocimiento compartido y común sobre algo”*. (Arano, 2005)

En la Informática:

- ✓ *“Una Ontología define los términos a utilizar para describir y representar un área de conocimiento. Las Ontologías son utilizadas por las personas, las bases de datos, y las aplicaciones que necesitan compartir un dominio de información (un dominio es simplemente un área de temática específica o un área de conocimiento, tales como medicina, fabricación de herramientas [...]). Las Ontologías incluyen definiciones de conceptos básicos del dominio, y las relaciones entre ellos, que son útiles para los ordenadores”*. (W3C, 2010)

La presente investigación se basará en el concepto dado por Peñalvo en su artículo “Web Semántica y Ontologías”, que define la ontología como: *“Especificación formal y explícita de una conceptualización compartida”*. (Peñalvo, 2004) Donde el término conceptualización es basado en el uso de conceptos, atributos, relaciones. La frase “especificación específica”, refiere a que cada

término o conocimiento por muy básico que sea deberá ser representado. “Formal” define que todas estas representaciones deberán ser exactamente iguales, permitiendo ser reutilizados y leídos por cualquier sistema sin importar el idioma, la plataforma y la localidad. Por último, y no menos importante, el término “compartida”, que no es más que se podrá tener acceso a cada recurso, términos, conceptos de la ontología, de esta manera se podrá modificar la información según la necesidad que se tenga, la misma tiene varios objetivos.

1.2.1: Objetivos de las ontologías

Los autores Noy y McGuinness presentan como objetivo de las ontologías (Noy & McGuinness, 2001):

-Compartir entendimiento común de la estructura del conocimiento, entre personas o agentes de software: pone a disposición términos y conceptos del dominio de interés, lo cual permitirá agregar y extraer información según la necesidad que se tenga.

-Permitir reutilizar el dominio de conocimiento: muchos dominios hacen uso de un conocimiento específico, si este conocimiento está constituido en una ontología podrá ser reutilizado por aquellos individuos que la necesiten sin necesidad de desarrollar una ontología propia.

-Permitir separar el conocimiento de dominio del conocimiento operacional: una ontología expresa el conocimiento del dominio de manera general de forma tal que pueda ser utilizado y manipulado por diversas técnicas o algoritmos.

-Analizar el conocimiento del dominio: específicamente en lo que se refiere al estudio de los términos y relaciones que lo configuran, ya sea formalmente o no.

Las ontologías tienen como principal objetivo especificar una conceptualización, un determinado punto de vista, además de compartir un entendimiento común de un dominio que servirá de base para resolver los problemas conectados a ese dominio. De aquí que se deriven los usos de ellas que explicaremos por consiguiente.

1.2.2: Usos de las ontologías

La **Fig.2** muestra los distintos usos de las ontologías (Ramos & Nuñez, 2007):

-La interacción entre las distintas entidades con diferentes necesidades y puntos de vista.

-“*La interoperabilidad entre sistemas alcanzada por una traducción entre diferentes modelos, métodos, paradigmas, lenguajes y herramientas de software*”.

-Sirve además de soporte para la ingeniería de sistemas a la hora de hacer el chequeo de confiabilidad del software, de la especificación de requisitos, y a la hora de hacer el levantamiento de los requerimientos.



Fig.2 Usos de las ontologías

Fuente: “ONTOLOGÍAS: componentes, metodologías, lenguajes, herramientas y aplicaciones.”

(Ramos & Nuñez, 2007)

“La principal utilidad de una ontología es ayudar a la compartición del conocimiento entre diferentes actores de un dominio determinado, como pueden ser personas, organizaciones y sistemas de software”. (Ramos & Nuñez, 2007)

De todas las actividades realizadas por las ontologías, una de las más importantes es la comunicación, tanto bidireccional como unidireccional y en cualquier dominio, además la ambigüedad del lenguaje implica errores y malentendidos. Estos problemas son causados por la diferencia de conocimientos del dominio por parte de las entidades que utilizan el proyecto, así como la diferencia de lenguajes a la hora de la comunicación. Gracias a la ontología estos problemas son aplacados al congregarse en un solo lugar los conocimientos y sus respectivas relaciones, evitando de esta manera errores en el manejo de la información con la combinación de los diferentes lenguajes.

“En cuanto a los sistemas de software, el uso de ontologías en su desarrollo permite establecer correspondencia y relaciones entre los diferentes dominios de entidades de información. Pueden existir dos o más ontologías conectadas compartiendo información de dominios distintos, o se puede dar el caso que agentes de software requieran usar la información contenida en una ontología específica. Todo esto trae consigo la interoperabilidad entre diferentes sistemas de computación ayudando a un entendimiento compartido más amplio y reutilizable”.

María Jesús Lamarca Lapuente expone que el uso de las ontologías proporciona una serie de beneficios entre ellos se pueden enunciar los siguientes (Lapuente, 2011):

- ✓ “Proporcionan una forma de representar y compartir el conocimiento utilizando un vocabulario común.
- ✓ Permiten usar un formato de intercambio de conocimiento.

- ✓ Proporcionan un protocolo específico de comunicación.
- ✓ Permiten una reutilización del conocimiento.”

Al la ontología permitir un formato de intercambio del conocimiento gestionado por el grupo GIC, permitiría entonces que mejore la forma de búsqueda del usuario, y que la misma sea más segura. Las ontologías capturan un conocimiento consensuado de modo genérico, de forma que pueda ser compartido y reutilizado por distintos grupos de personas, en el caso del presente trabajo, los que trabajan en el grupo GIC (Cantera Fonseca, Hierro Sureda, & Romo Zabala, 2011). Así como también la interacción de todas las herramientas donde se encuentra distribuida la información gestionada por el mismo, y de esta manera, obtener con la ontología, que dicha información se encuentre contenida en un solo lugar. Logrado esto, se resuelven los problemas de lentitud en la búsqueda y que la información esté disponible e interrelacionada.

En dependencia de los usos que brinden las ontologías, es el tipo de ontología que se debería usar en el proyecto de investigación.

1.2.3: Clasificación de ontologías

Las ontologías se clasifican en 4 tipos de criterios en dependencia del contexto, cantidad y tipo de conceptualización (granularidad), por su propósito de uso y por su nivel de formalidad. Cada una presenta diferentes objetivos. Los autores Gómez, Uschold y Gruninger mencionan algunas de ellas (Gómez y otros, 2004), (Uschold & Gruninger, 1996) & (FIPA, 2000):

Por su dependencia del contexto:

- ✓ **Ontologías de nivel superior:** describen conceptos de manera general y no específico independientemente del dominio. Ejemplo: ontologías sobre el tiempo, el espacio, la materia, entre otros.
- ✓ **Ontologías de dominio:** proveen la terminología específica para un determinado dominio. Estas presentan términos afines con los objetivos de este dominio y sus componentes.
- ✓ **Ontologías de tareas:** proporcionan tareas susceptibles de realizar en un determinado dominio.
- ✓ **Ontologías de aplicación:** sujetan definiciones necesarias para componer el conocimiento necesitado para una específica aplicación en un determinado dominio.

En la siguiente figura se representa la relación entre los distintos tipos de ontologías antes mencionados.

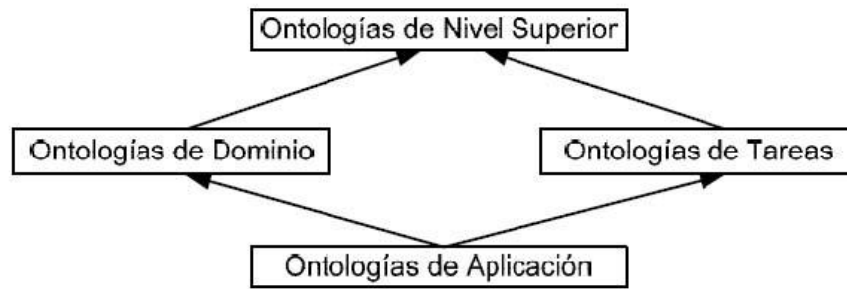


Fig.3 Relaciones entre los tipos de ontologías por dependencia del contexto.

Fuente: "FIPA Ontology Service Specification." (FIPA, 2000)

Cantidad y tipo de conceptualización (granularidad):

- ✓ **Ontologías terminológicas:** utilizadas para agrupar vocabulario en un dominio determinado (contenido léxico y no semántico).
- ✓ **Ontologías de información:** brindan una base para el almacenamiento normalizado de información (estructura de los registros de una base de datos).
- ✓ **Ontologías de representación de conocimiento (Meta-ontologías):** facilitan el vocabulario que se necesita para formar otras ontologías mediante un determinado paradigma.

Por su propósito de uso:

- ✓ **Ontologías para la comunicación entre personas:** sirven para reconocer términos necesarios para referir y representar un área de conocimiento.
- ✓ **Ontologías para la interoperabilidad entre sistemas:** realizan interpretaciones entre diferentes métodos, paradigmas, lenguajes y herramientas de software. Es usada para el intercambio de conocimiento.
- ✓ **Ontologías para beneficiar la ingeniería de sistemas:** es utilizada para la reutilización de componentes, adquisición de conocimiento e identificación de requerimientos, así como para la confiabilidad de los sistemas, al ofrecer consistencia al conocimiento del mismo.

Por su nivel de formalidad (grado de formalismo del lenguaje usado para expresar la conceptualización):

- ✓ **Ontología altamente informal:** expresada en lenguaje natural.
- ✓ **Ontología semi-formal:** usa un lenguaje de definición formal
- ✓ **Ontología rigurosamente formal:** se usan semántica y teoremas para rigurosamente llevar a cabo la definición de los términos.
- ✓ **Ontología informal estructurada:** permite incrementar la claridad y reducir la ambigüedad.

La principal ontología que necesita el grupo GIC es la **ontología de dominio** para proporcionar el vocabulario necesario para describir un determinado dominio. De esta manera se creará una ontología que cumplirá con los requisitos que se precisan para resolver los problemas presentados

en el presente trabajo. Las ontologías están compuestas por 4 componentes fundamentales para su diseño.

1.2.4: Componentes

Los componentes de las ontologías varían en dependencia de las necesidades del desarrollador. Tallarico en su investigación menciona los componentes principales de una ontología (Tallarico, 2003):

- ✓ **Conceptos:** Son las ideas básicas que se intentan formalizar. Los conceptos pueden ser clases de objetos, métodos, planes, estrategias, procesos de razonamiento, etc.
- ✓ **Relaciones:** Representan la interacción y enlace entre los conceptos del dominio. Ejemplos de relaciones pueden ser subclase-de, conectado-a, parte-de.
- ✓ **Axiomas:** Son teoremas sobre relaciones que deben cumplir los elementos de la ontología.
- ✓ **Instancias:** Son usadas para representar objetos determinados de un concepto.

Como tal, estos son los principales componentes utilizados por cualquier ontología que serán utilizados a su vez, en el presente trabajo. Una vez identificados dichos componentes, se describirán los principales elementos a tener en cuenta para la construcción de la ontología.

1.2.5: Metodologías para el desarrollo de ontologías

“El proceso de construir una ontología no difiere mucho, en líneas generales, del usado para construir software. Según la definición dada por la IEEE, un software es un “programa de computación, procedimientos y documentación asociada, además de los datos para que se ejecute”. Las ontologías son productos de software y por lo tanto su desarrollo deberá seguir los estándares establecidos, por supuesto, adaptados a las características de las ontologías.” (Ramos & Nuñez, 2007)

En la siguiente tabla se muestran varias metodologías que han sido desarrolladas a medida que se ha extendido el uso de las ontologías. (Velásquez Pérez, Puentes Velásquez, & Guzmán Luna, 2011)

Metodologías para el desarrollo de ontologías

Metodología	Uso	Desarrollador	Fases
Metodología de Uschold y Grüninger	Desarrollo de Enterprise Ontology	Universidad de Edimburgo junto a IBM Lloyds y Unilever	Identificación de Propósito, Construcción de la Ontología, Evaluación y Documentación
Metodología Gruninger y Fox	Desarrollo de la Ontología TOVE	(<i>Toronto Virtual Enterprise</i>)	Captura de Escenarios, Formulación de Preguntas de Competencia Informales, Especificar la Terminología en un Lenguaje Formal y, Establecer las condiciones que determinan cuando la Ontología esta Completa.
Metodología Amaya	Investigaba la posibilidad de rehuso de conocimiento en procesos técnicos complejos y el beneficio de uso de Ontologías	proyecto ESPRIT KACTUS	Especificación de la Aplicación, Diseño Preliminar Basado en Categorías del Nivel de Abstracción más General y, Refinamiento y Estructuración de la Ontología
Metodología basada en SENSUS	Es usada para construir Ontologías tomando como base los conceptos abstractos identificados en la Ontología SENSUS		Descartar los términos no relevantes de la Ontología SENSUS; para ello, se parte de unos términos relevantes del dominio que se toman como semilla y se excluyen todos los que no se encuentran en el camino hasta la raíz, y Se incluyen en este grafo otros términos relevantes del dominio.
Methontology	Permite la construcción de ontologías en un nivel de conocimiento	<i>Ontology Group</i> de la Universidad Politecnica de Madrid	Especificación de requisitos, adquisición de conocimiento, conceptualización, formalización, integración, implementación, evaluación, documentación

Tabla.1 Metodologías para el desarrollo de ontologías

Fuente: “Ontologías: una técnica de representación de conocimiento.”

(Velásquez Pérez, Puentes Velásquez, & Guzmán Luna, 2011)

Además estos autores afirman que las semejanzas de estas metodologías es que todas comienzan con la identificación del propósito de la ontología y la necesidad de adquirir conocimiento del dominio en cuestión. La metodología Uschold propone codificar en un lenguaje formal mientras que Methontology propone expresar la idea como un conjunto de representaciones intermedias, las cuales establecen un puente entre la manera en que las personas ven un dominio y los lenguajes en que se formalizan las ontologías.

Todas estas metodologías antes expresadas identifican la necesidad de evaluar las ontologías. La metodología de Uschold incluye esta actividad, pero no establece como podría llevarse a cabo. Gruninger y Fox proponen la identificación de un conjunto de interrogantes, por su parte Methontology propone que deben llevarse a cabo actividades de evaluación durante todo el tiempo de vida del proceso de desarrollo de la ontología.

Existen otras metodologías para la creación de ontologías que van surgiendo con el devenir de las ontologías como son On-To-Knowledge¹⁹, Kactus²⁰. Lo que no se puede asegurar es qué tan buena es una, o qué tan mala es otra, ya que cada una de ellas puede ser bien usada en el proceso para el que fue creada. Es por eso que ese elige la metodología creada por Noy y McGuinness²¹, la misma enumera de manera sencilla los pasos a seguir para la construcción de ontologías, son pasos de desarrollo iterativo además de ser generales para todas las metodologías para la creación de ontologías (Noy & McGuinness, 2005).

Pasos:

- 1. Determinar el dominio y alcance de la ontología.** Según la bibliografía antes expuesta para comenzar se define un dominio y alcance que no es más que darle respuesta a las siguientes preguntas:
 - ✓ ¿Cuál es el dominio que la ontología cubrirá?
 - ✓ ¿Para qué usar la ontología?
 - ✓ ¿Para qué tipos de preguntas la información en la ontología debería proveer respuestas?
 - ✓ ¿Quién usará y mantendrá la ontología?
- 2. Considerar la reutilización de ontologías existentes.** En este paso se considerará la reutilización de ontologías existentes con el fin de depurar y extender los recursos que estas poseen.
- 3. Enumerar los términos importantes en la ontología.** Es útil escribir una lista con todos los términos con los que quisiéramos hacer enunciados o dar explicación a un usuario.
- 4. Definir las clases y la jerarquía de clases.** *“De la lista creada en el Paso 3, seleccionamos los términos que describen objetos que tienen existencia independiente en lugar de términos que describen esos objetos. Esos términos serán las clases.”*
- 5. Definir las propiedades de las clases: properties.** Las clases aisladas no ofrecen la información necesaria como para dar respuesta a las preguntas de competencia. A causa de esto después de definir las clases se pasa a crear la estructura interna de los conceptos.
- 6. Definir las facetas de las properties.** Las propiedades tienen diferentes facetas que describen el tipo de valor, valores admitidos, la cardinalidad entre otras.

¹⁹ 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 ontologías conjuntamente con el desarrollo de aplicaciones de gerencia de conocimiento.

²⁰ metodología de reutilización de ontologías, que se basa en reutilizar el conocimiento técnico. A partir de una base de conocimiento determina y conceptualiza los términos y relaciones más importantes en la ontología.

²¹ Dra. Deborah McGuinness: directora en funciones y científica principal de investigación en los sistemas de conocimiento del Laboratorio de Inteligencia Artificial de la Universidad de Stanford. Líder en la representación del conocimiento y los idiomas y sistemas de razonamiento y ha trabajado en la creación de ontologías y evolución de los entornos por más de 20 años.

Natasha Noy: estudió matemáticas aplicadas en la Universidad Estatal de Moscú, tiene una Maestría en Ciencias de la Computación de la Universidad de Boston, se especializa en arquitectura de computadores, y tiene un doctorado en la Universidad de Northeastern en el cual estudió los temas de diseño de la ontología de las ciencias experimentales y el uso de ontologías en la recuperación de información basada en el conocimiento. Ahora trabaja en el Centro de Investigación de Informática Biomédica de la Universidad de Stanford (BMIR).

7. Crear instancias. Este último paso radica en crear instancias de las clases de la jerarquía propuesta.

No existe una única metodología para la creación de ontologías pero si varios principios fundamentales que ayudarán en este proceso (Ramos & Nuñez, 2007):

- ✓ **Claridad y objetividad:** una ontología debe proveer al usuario con el significado del término de forma objetiva, y también, en lenguaje natural para su comprensión.
- ✓ **Compleitud:** las definiciones han de ser expresadas en términos necesarios y suficientes.
- ✓ **Coherencia:** Para asegurar que las inferencias realizadas de ésta, sean consistentes con las definiciones.
- ✓ **Máxima extensibilidad monótona:** las especializaciones o generalizaciones deben ser incluidas en la ontología de tal forma que no se requiera una revisión de las definiciones preexistentes.
- ✓ **Principio de distinción ontológica:** las clases en una ontología deben ser disjuntas, es decir, en conjuntos diferentes.
- ✓ **Diversificación:** se han de diversificar las jerarquías incluidas para aumentar la potencia de los mecanismos de herencia múltiple²².
- ✓ **Modularidad:** para disminuir el acoplamiento entre los módulos.
- ✓ **Estandarización:** los nombres deberán seguir reglas y estándares según sea posible.
- ✓ **Minimización:** *“conceptos similares estarán agrupados y representados utilizando las mismas primitivas”.*

Estos principios, servirán de gran importancia en el presente trabajo de investigación, ya que serán estándares o bases por los que se regirá el equipo de trabajo para desarrollar la ontología de manera organizada.

Una vez seleccionada la metodología para la descripción de la ontología, así como definidos los componentes, la ontología puede ser representada en varios lenguajes que serán mencionados en el siguiente apartado.

1.2.6: Lenguajes para describir ontologías

La utilización de un lenguaje de ontologías para definir la información de gestión trae consigo varios beneficios, el primero de ellos es la posibilidad de utilizar herramientas ya existentes para trabajar completamente con ontologías.

El autor Lapuente dentro de los principales lenguajes de ontologías destaca los siguientes (Lapuente, 2011):

- ✓ **SHOE:** Simple HTML Ontology Extensions. Fue el primer lenguaje de etiquetado para diseñar ontologías en la Web. Este lenguaje nació antes de que se ideara la WS. Las

²² característica de la programación orientada en la que una clase puede heredar comportamientos y características de más de una superclase. Esto contrasta con la herencia simple, donde una clase sólo puede heredar de una superclase

ontologías y las etiquetas se incrustaban en archivos HTML. Este lenguaje permite definir clases y reglas de inferencia, pero no negaciones o disyunciones.

- ✓ **OIL:** Ontology Inference Layer. Este lenguaje, derivado en parte de SHOE, fue impulsado también por el proyecto de la Unión Europea On-To- Knowledge. Utiliza ya la sintaxis del lenguaje XML y está definido como una extensión de RDFS. Se basa tanto en la lógica descriptiva (declaración de axiomas) y en los sistemas basados en frames (taxonomías de clases y atributos). Posee varias capas de sublenguajes, entre ellas destaca la capa base que es RDFS, a la que cada una de las capas subsiguientes añade alguna funcionalidad y mayor complejidad. La principal carencia de este lenguaje es la falta de expresividad para declarar axiomas.
- ✓ **DAML y OIL:** Este lenguaje nació fruto de la cooperación entre OIL y DARPA y unifica los lenguajes DAML (DARPA's Agent Markup Language) y OIL (Ontology Inference Layer). Se basa ya en estándares del W3C. Desarrollado como una extensión del lenguaje XML y de Resource Description Framework (RDF) y para extender el nivel de expresividad de RDFS. Hereda muchas de las características de OIL, pero se aleja del modelo basado en clases (frames) y potencia la lógica descriptiva. Es más potente que RDFS para expresar ontologías. En su última revisión ofrece ya un conjunto de elementos con los cuales se pueden crear ontologías y marcar la información para que sea legible y comprensible por la máquina. Funciona como formato de intercambio, sin embargo, este lenguaje presenta algunas carencias debido a su complejidad conceptual y de uso, complejidad que se intentó solventar con el desarrollo de OWL.
- ✓ **OWL:** 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, es decir, 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. OWL surge como una revisión al lenguaje DAML y OIL y es mucho más potente que este. Al igual que OIL, OWL se estructura en capas que difieren en la complejidad y puede ser adaptado a las necesidades de cada usuario, al nivel de expresividad que se precise y a los distintos tipos de aplicaciones existentes (motores de búsqueda, agentes, etc.)
- ✓ **KIF:** Knowledge Interchange Format es un lenguaje para representar ontologías basadas en la lógica de primer orden. KIF está basado en la lógica de predicados con extensiones para definir términos, metaconocimiento, conjuntos, razonamientos no monotónicos, etc.; y

pretende ser un lenguaje capaz de representar la mayoría de los conceptos y distinciones actuales de los lenguajes más recientes de representación del conocimiento. Se trata de un lenguaje diseñado para intercambiar conocimiento entre sistemas de computación distintos, diferentes lenguas, etc.; y no para la interacción entre seres humanos.

Para la representación de la ontología se definió como lenguaje a utilizar OWL. Tiene como objetivos realizar un modelo de marcado sobre RDF, reutilizando las tripletas de RDF. Está codificado en XML permitiendo representar ontologías mediante un vocabulario mucho más expresivo y una sintaxis más fuerte que la permitida por RDF y el resto de los lenguajes. Es una revisión del lenguaje DAML+OIL, del que aprendió de sus deficiencias para convertirse en un lenguaje más expresivo y completo. Por último y no menos importante, facilita que las máquinas puedan comprender e interpretar los contenidos web que son soportados por XML, RDF y RDF Schema.

1.2.7: Herramientas para el desarrollo de ontologías

Los editores de ontologías son herramientas que ayudan a proveer soporte para el desarrollo y uso de las ontologías, al mismo tiempo, permiten las facilidades de definición y modificación de los componentes principales de las ontologías: conceptos, propiedades, relaciones, axiomas y restricciones, hasta la inspección y navegación (Ramos & Nuñez, 2007). En las siguientes tablas se muestra una comparación entre las herramientas más potentes que existen:

Herramienta	Extensibilidad	Lenguajes Compatibles	Descripción	Consistencia	Página
Ontolingua Server	Mediante plug-ins.	KIF, Prolog, LOOM, OKBC, IDL, CML, Clips.	Provee, en un entorno colaborativo, un buscador, generador y modificador de ontologías.	En tiempo real mientras se construye la ontología.	http://www.ksl.stanford.edu/software/ontolingua/
OntoEdit	Mediante plug-ins.	XML, RDFs, Flogic, DamL+OIL.	Para construir ontologías usando significaciones gráficas.	A través del plug-ins Ontobroker.	http://www.ontoknowledge.org/tools/ontoedit.shtml

WedODE	Mediante plug-ins.	XML, RDFs, Oil, DamL+OIL, Flogic, Prolog, HTML.	Permite desarrollar ontologías sobre ingeniería.	Utiliza restricciones, verificadores de consistencia y otros parámetros.	http://webode.dia.fi.upm.es/WebODEWeb/index.html
Protegé	Mediante plug-ins.	RDFS, OWL, XML Schema.	De código abierto para que sea multiplataforma.	Mediante plug-ins.	http://protege.semaniticweb.org
SWOOP	Mediante plug-ins.	RDF, XML, OWL.	Editor de ontologías hipermedia, para la navegación rápida y fácil y el desarrollo de ontologías OWL..	Mediante plug-ins.	http://www.mindswap.org/2004/SWOOP/

Tabla 2 Tabla Comparativa

Fuente: “Hipertexto: El nuevo concepto de documento en la cultura de la imagen.” (Lapuente., 2010)

Se propone Protegé²³ por las siguientes razones:

Es un editor de ontologías de código abierto, se le conoce también como framework para bases de conocimientos.

- ✓ Desarrollado en Java para que fuera multiplataforma.
- ✓ Posee gran cantidad de plug-ins que permite su conexión con base de datos.
- ✓ Cuenta con una comunidad bastante amplia de desarrolladores.
- ✓ Provee un API que permite la exploración, edición, y creación de ontologías.
- ✓ Permite modelado sencillo, y varios modos de visualización y puede ser exportado a una gran variedad de formatos.

²³ Software libre de código abierto implementado en Java, desarrollado en la Universidad de Stanford, que permite la construcción de ontologías de dominio.<http://protege.stanford.edu/>

- ✓ Puede correr como una aplicación local o a través de un cliente en una comunicación con un servidor remoto.
- ✓ Los componentes de su interfaz de usuario para mostrar y adquirir datos pueden ser reemplazados con nuevos componentes que se pueden adaptar mejor a determinadas ontologías.
- ✓ Es completamente compatible con el lenguaje elegido.

Conclusiones del capítulo

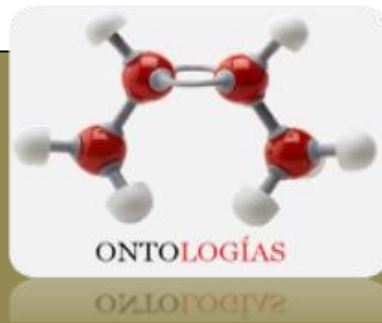
En este capítulo se expusieron los principales conceptos relacionados con esta investigación, así como las posibles soluciones a los problemas planteados en la situación problémica. Fueron estudiados aspectos relacionados con las Ontologías y Web. Se mostraron las principales metodologías, herramientas y lenguajes utilizados para construir y representar Ontologías. Donde se llega a la conclusión que para la realización de este trabajo se ha decidido utilizar:

La metodología propuesta por los autores **Noy y McGuinness** por ser una serie de pasos generales para todas las metodologías que describen ontologías.

El lenguaje **OWL** debido a que es un lenguaje mucho más completo que el resto de lenguajes de especificación de ontologías, por la razón de que añade más vocabulario para describir clases y propiedades, además de otras relaciones entre clases, como cardinalidad en las relaciones. Además de que se trata de una recomendación de la W3C²⁴ junto a RDF.

La herramienta **Protegé** por ser completamente compatible con el lenguaje escogido: OWL; cuenta además con un entorno abierto y fácil de entender, que ha generado toda una comunidad de desarrolladores. Posee un API que permite construir, explorar y modificar ontologías.

²⁴ <http://www.w3c.es/>



Capítulo 2: Propuesta de la ontología

Introducción

Una vez identificados en el capítulo 1 los principales elementos para la construcción de la ontología, dígase metodologías, lenguajes y herramientas, se procede entonces a la implementación de la misma. Por tanto, en el presente capítulo se detallarán dichos elementos así como se llevará a cabo un levantamiento de la información mediante una entrevista realizada a cada miembro del grupo GIC, con el objetivo de realizar el modelado de la ontología para luego proceder a la implementación de la misma.

En el capítulo 1, se describieron 7 pasos fundamentales que ayudarán en el proceso para la creación de ontologías.

2.1: Construcción de la ontología

En el capítulo 1 se toma como propuesta de metodología los pasos realizados por los autores Noy y McGuinness.

Pasos:

1. Determinar el dominio y alcance de la ontología

- ✓ ¿Cuál es el dominio que la ontología cubrirá?
- ✓ ¿Para qué usar la ontología?
- ✓ ¿Para qué tipos de preguntas la información en la ontología debería proveer respuestas?
- ✓ ¿Quién usará y mantendrá la ontología?

-Se define como dominio de la ontología del presente trabajo al grupo GIC, la misma será usada para mejorar la gestión e interrelación de la información de dicho grupo. Para determinar el alcance de la misma se utilizan las llamadas preguntas de competencia que podrán ser utilizadas para futuras pruebas de control de calidad, ya que serán preguntas que la base de conocimiento deberá

ser capaz de responder. Serán preguntas como: ¿Qué servicios y productos son brindados por el grupo? La ontología será usada por todo el personal del grupo GIC, y será mantenida por los encargados de gestionar el conocimiento en el grupo.

2. Considerar la reutilización de ontologías existentes

-Para el presente trabajo no existe una ontología de la que se podrá reutilizar su conocimiento, por tanto se obvia este paso.

3. Enumerar los términos importantes en la ontología

-Con el objetivo de conseguir los principales conceptos o terminologías de la ontología se hace necesario un levantamiento de la información gestionada por el grupo GIC y de esta manera realizar el modelado de la ontología. Por tanto, para el cumplimiento de esta tarea, se llevo a cabo una entrevista a cada miembro del grupo GIC. Después de un análisis de dicha entrevista los resultados fueron:

El grupo está compuesto por 6 integrantes: una coordinadora Licenciada en Bibliotecología y Ciencias de la Información que organiza y controla las actividades llevadas a cabo por cada miembro del grupo; una editora de boletín, Máster en Calidad de Software; dos técnicos de Informática; dos desarrolladores que al mismo tiempo ejercen el rol de montador del Boletín, los mismos están encargados de prestar servicios según las necesidades de cada usuario.

En el grupo se trabaja con buscadores de Web superficial (Google) y metabuscadores, portales web especializados, base de datos dentro de la universidad y herramientas de escritorio como: Office Word, Acrobat Reader, Excel y Power Point. En cuanto a los desarrolladores trabajan con herramientas como Open Journal System, Open Harvester System, Drupal, Open Conference.

El **soporte** realizado en dicho grupo es el siguiente:

Se utiliza Drupal y Open Journal System para la gestión de información y Open Harvester System para la recopilación de metadatos. Se desarrollan modificaciones de código cuando son detectados errores, se les da mantenimiento constantemente a las aplicaciones utilizadas en el grupo, velando que las mismas estén online en todo momento y se realizan también la revisión de los permisos. Todo esto con el objetivo de dar apoyo al grupo y que las herramientas estén en plena disposición para lo que se necesite.

El grupo brinda además tanto servicios como productos según la necesidad informativa que se tenga:

Servicios:

- Inteligencia Empresarial:

Crea tres tipos de servicios diferentes: Estudios Métricos; Estudios de Mercado y los de Competencia que generan el producto de informe; y un producto de Perfiles Estratégicos que a su vez se dividen en cuatro tipos: de Sector, País, Personalidad y Competencia.

- Arquitectura de Información:

Brinda los servicios de Talleres de Formación y los Cursos de Pre y Post-grado, que terminan en el producto de Taller y el Curso respectivamente.

- **Gestión de la Información y el Conocimiento:**

Brinda dos tipos de servicios y crea además dos tipos de productos. Los servicios son cursos de Gestión de Información y el Conocimiento y la gestión de conferencias; y los productos: el curso y la conferencia como tal.

- **Alfabetización Informacional:**

Genera tres tipos de servicios que obtienen un producto cada uno: diagnóstico de competencia crea un informe, y los Talleres de Formación y cursos de Post-grado, crean los productos taller y curso respectivamente.

Productos:

- **Observatorio Tecnológico:**

Encargado de la búsqueda y recuperación de información, que puede ser utilizada con diferentes fines en dependencia de la finalidad. Dicho observatorio es la base tecnológica que va a tributar al servicio Gestión de la Información y el Conocimiento.

- **Boletín TeduScopio:**

Es un producto de información brindado a la Facultad 4 donde se publican áreas temáticas de interés que identifican al centro.

- **Catálogo Referativo:**

Producto que brinda información sobre determinados temas donde son incluidos indicadores que están en correspondencia con los intereses del usuario.

- **Compendio Informativo:**

Producto que brinda información sobre determinados temas donde son incluidos los indicadores: fuentes de información, congresos, eventos, observatorios y aquellos de interés por parte del usuario.

- **Conferencias, Talleres y Cursos:**

Clases magistrales en dependencia del tema que se requiera.

Las **tareas** realizadas de manera general en el grupo son:

- ✓ Asesoramientos y cursos Pre-grado y Pos-grado a Arquitectos de Información.
- ✓ Cursos sobre Alfabetización Informacional para elevar competencias informacionales de los profesionales y conocer necesidades informativas.
- ✓ Edición del boletín TeduScopio.
- ✓ Actualización y divulgación de la información.
- ✓ Búsqueda de información sobre temáticas en función de la necesidad de la comunidad de usuario.
- ✓ Identificación de fuentes informativas a utilizar para el producto de información a desarrollar.
- ✓ Análisis de la información.

- ✓ Obtención de los resultados a través de informes que responden a la necesidad informacional del cliente.
- ✓ Registro de tesis y artículos en el Open Journal System.
- ✓ Administrar y configurar el Open Journal System y el Open Harvester System.
- ✓ Maquetar²⁵ y montar el boletín.

Las **terminologías** comunes entre los miembros del grupo son: Alfabetización Informacional, boletín, necesidades informativas, fuentes de información.

De forma general estas terminologías son palabras claves utilizadas en la búsqueda de informaciones afines con las distintas temáticas, de esta manera se generará y socializará un producto que brinde lo acontecido en el ámbito científico en estas áreas temáticas pertenecientes al centro en correspondencia con la necesidad informativa. Por tanto se toman como principales terminologías las siguientes:

No.	Conceptos	Descripción
1.	Necesidades informativas	Necesidad de información sobre determinado tema que tiene un usuario dado del centro.
2.	Fuentes de información	Tipo de fuentes de información.
3.	Artículos	Fuentes de información que se utilizan para generar un producto y socializar información sobre un tema en específico.
4.	Tesis	Documento utilizado para la investigación por parte de los usuarios del centro FORTES y la Facultad 4.
5.	Talleres	Actividades prácticas enfocadas en una temática determinada.
6.	Cursos	Cursos de diferentes temáticas impartidos por especialistas del grupo.
7.	Informes	Informes que son generados producto de una necesidad informativa, clasificándose en un tipo de producto de información.
8.	Revistas	Fuente de información que se utiliza para la búsqueda y recuperación de información en la toma de decisiones.

²⁵ Organizar imágenes y textos en la aplicación, que tengan el formato correcto.

9.	Eventos	Suceso realizado en un país que se utiliza para la búsqueda y recuperación de información en la toma de decisiones.
10.	Congresos	Eventos realizados que abordan áreas temáticas de interés para los usuarios del centro y la Facultad 4.
11.	Conferencias	Charlas expuestas sobre determinada temática por un experto en el tema.
12.	Herramientas	Empleadas en la gestión y organización de la información y la toma de decisiones del grupo GIC.
13.	Metadatos	Datos que describen datos.
14.	Productos	Productos de información generados de una necesidad informativa.
15.	Boletín TeduScopio	Producto informativo realizado por el grupo GIC para socializar noticias recientes que se han publicado sobre las temáticas del centro. Divulga resultados investigativos sobre otras áreas que pueden ser de interés para los usuarios.
16.	Observatorio tecnológico	Conjunto de herramientas informáticas que permite la búsqueda y recuperación de información.
17.	Servicios	Servicios que brinda el grupo al centro
18.	Asesoría y Capacitación de Arquitectura Informacional	Servicio específico que brinda el grupo al centro.
19.	Talleres	Servicio brindado por los servicios de Alfabetización Informacional y Arquitectura Informacional con el objetivo de capacitar a un grupo de usuarios interesados sobre determinada temática o herramienta.
20.	Cursos	Servicio brindado por los servicios de Alfabetización y Arquitectura Informacional con el

		objetivo de capacitar a un grupo de usuarios interesados sobre determinada temática o herramienta.
21.	Inteligencia empresarial	Servicio específico que brinda el grupo al centro.
22.	Perfiles estratégicos	Producto de Inteligencia Empresarial que se clasifica en cuatro tipos: de sector, país, personalidad y corporativos.
23.	Perfiles de país	Tipo de perfil estratégico, describe información sobre determinado país que contribuye a la toma de decisiones.
24.	Perfiles de sector	Tipo de perfil estratégico, describe información sobre un determinado sector.
25.	Perfiles corporativos	Tipo de perfil estratégico, caracteriza una empresa o compañía.
26.	Perfiles de personalidad	Tipo de perfil estratégico, describe información sobre determinada personalidad.
27.	Estudios métricos	Análisis de la producción científica sobre determinado tema en un período de tiempo determinado.
28.	Estudios de competencia	Identifican las empresas que desarrollan productos afines al centro.
29.	Estudios de mercado	Analizan comportamiento del mercado acerca de determinados productos y servicios sobre empresas que desarrollan productos afines al centro FORTES.
30.	Gestión de la información y el conocimiento	Servicio específico que brinda el grupo al centro.
31.	Catálogos referativos	Producto que brinda información sobre determinados temas donde son incluidos indicadores que están en correspondencia con los intereses del usuario.
32.	Compendios informativos	Producto que brinda información sobre

		determinados temas donde son incluidos los indicadores: fuentes de información, congresos, eventos, observatorios y aquellos de interés por parte del usuario.
33.	Alfabetización Informacional	Servicio específico que brinda el grupo al centro.
34.	Autor del servicio y producto	Responsable de llevar a cabo el servicio y controlar la culminación del producto final.
Total de Conceptos: 31		

Tabla 4: Principales terminologías del grupo GIC

4. Definir las clases y la jerarquía de clases

La **Fig.9** muestra un parte de la jerarquía de clases de la ontología propuesta para el dominio del grupo GIC. Dicha jerarquía se realizó iniciando por los conceptos más generales como son los productos y servicios brindados por el grupo GIC en dependencia de la necesidad de información que se tenga, terminando por los conceptos más específicos que serían los distintos productos y servicios brindado. **Ver la vista completa en Fig.10.**



Fig.9 Jerarquía de clases de la ontología

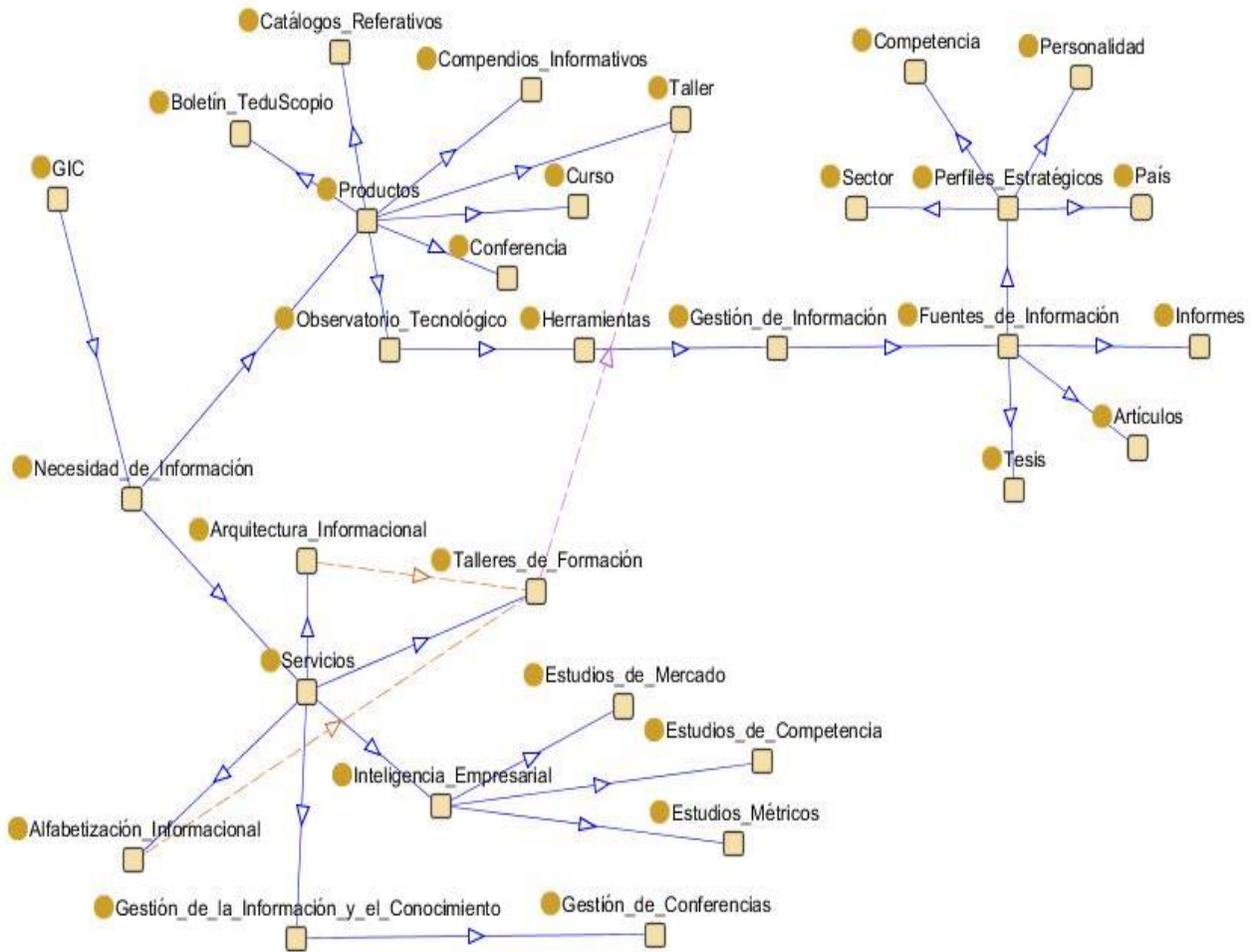


Fig.10 Jerarquía de clases de la ontología

Para definir las clases se seleccionaron, de los conceptos extraídos en el paso 3, los que tienen existencia independiente. Luego se organizaron las clases jerárquicamente, determinando si una clase es una superclase de otra, por lo que cada instancia de la clase padre lo es también de la clase hija. Para el dominio del grupo GIC, las necesidades de información (Productos o Servicios), fueron colocadas en un mismo nivel, y luego se desglosaron en niveles inferiores con las clases que definen los tipos de productos y servicios.

5. Definir las propiedades de las clases: properties

Una vez definidas las clases se procede a definir las propiedades de las mismas, un ejemplo de esto lo constituye la clase **Necesidad de Información** de la cual se necesita conocer la fecha de inicio y fin en la que se hace la solicitud (**fecha_Inicio** y **fecha_Fin**); la entidad que solicita la necesidad (**entidad_que_solicita**); así como el trabajador responsable de la misma (**responsable_de_necesidad**). (Ver Fig.11)

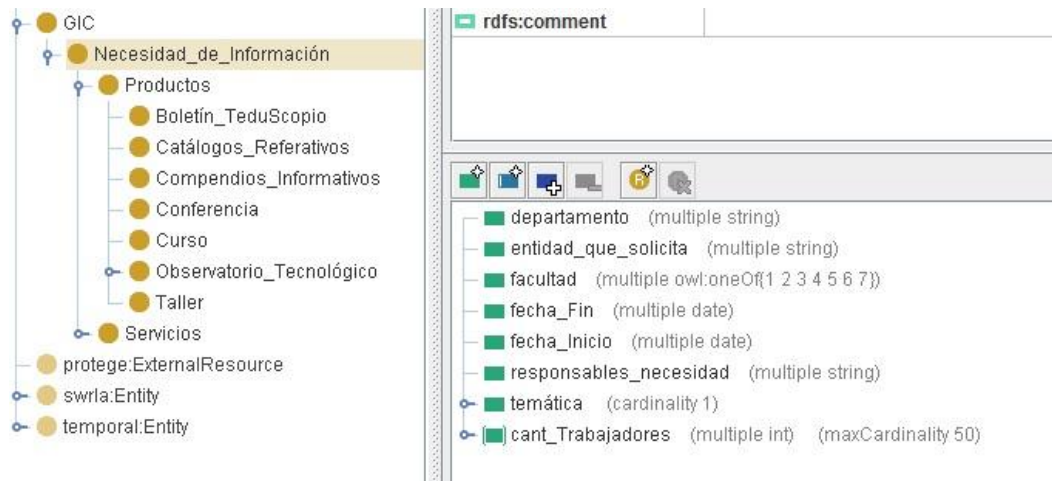


Fig.11 Propiedades de la clase **Necesidad_de_Información**

Toda clase tendrá como propiedades las heredadas de las clases padres. Ejemplo de ello son: **Producto** y **Servicio**, hijas de **Necesidad de Información**, por tanto adquieren todas sus propiedades, independientemente de las suyas propias. (Ver **Fig.12**)

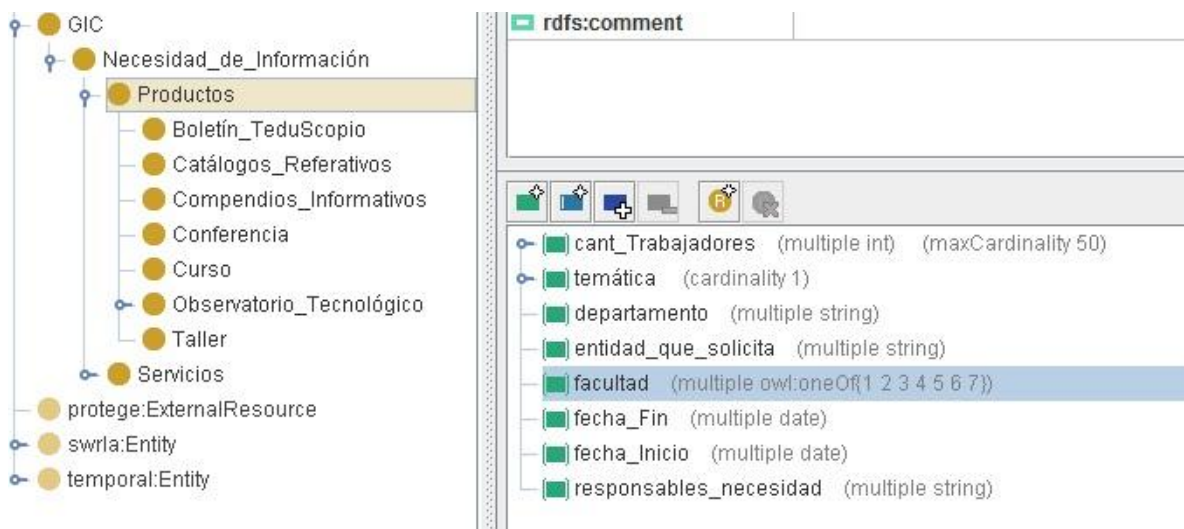


Fig.12 Propiedades de la clase **Productos**

6. Definir las facetas de las propiedades

En este paso se procede a definir el rango de cada una de las propiedades: string, int, float, date, datetime; para luego definir su cardinalidad y restricciones. Ejemplo de ello es en la clase **Productos**, la propiedad referida a la cantidad de trabajadores tiene valor restringido de 50 y el número de la facultad un rango de 1 hasta 7. (Ver **Fig.13**)

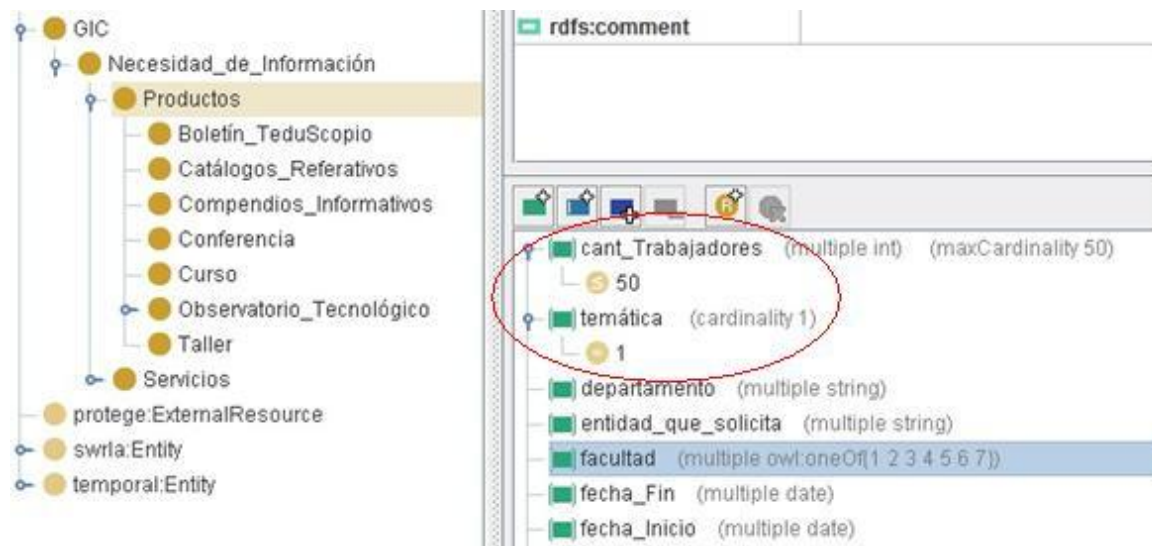


Fig.13 Propiedades de la clase **Productos**

7. Crear instancias

Este último paso consiste en crear instancias de las clases de la jerarquía propuesta. Las instancias son utilizadas para representar objetos determinados de un concepto (Ver **Fig.14**). La definición de una instancia individual de una clase requiere:

- Elegir una clase.
- Crear una instancia individual de la clase.
- Rellenar los valores de las propiedades.

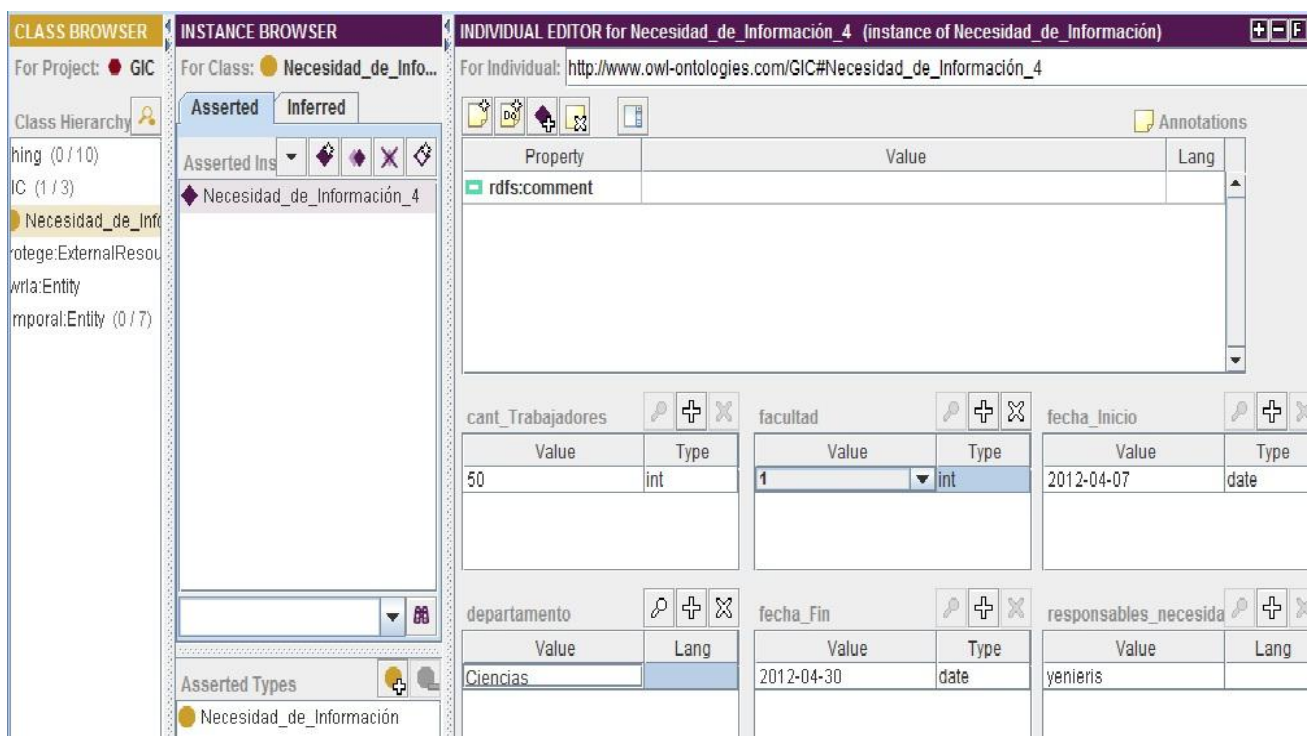


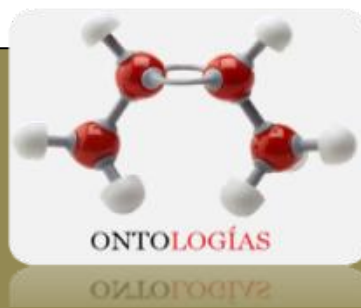
Fig.14. Instancia de la clase **Necesidad_de_Información**

Para la clase **Necesidad_de_Información** que se muestra en la figura anterior se crea una instancia denominada **Necesidad_de_Información_4**.

Conclusiones del capítulo

Durante el desarrollo de la propuesta de solución, se arribó a las siguientes conclusiones:

- El análisis realizado acerca de las diferentes tendencias tecnológicas actuales permitió que se escogieran las herramientas adecuadas para el modelado e implementación de la ontología.
- Se determinaron 31 conceptos clave a partir de una entrevista realizada a cada miembro del grupo GIC.
- Se diseñó una ontología que proporcionó la interrelación de la información gestionada por el grupo GIC permitiendo mejoras en el proceso de búsqueda de información.



Capítulo 3: Prueba

Introducción

La evaluación de una ontología resulta una actividad de suma importancia ya que proporciona un alto grado de confianza y seguridad del producto y de los resultados que se obtendrán luego que sea implantado. En la actualidad existen varios criterios para evaluar y analizar una ontología, ya sea con la finalidad de integrarla a un nuevo proyecto o de medir la calidad y funcionalidad de la misma. Este proceso es una tarea que se realiza a medida que se va realizando la ontología, para poder medir las potencialidades de ella e ir detectando posibles errores en cuanto a prestaciones o funcionalidades.

3.1: Evaluación de la ontología

Uno de los tipos fundamentales de evaluación de una ontología es por medio de razonadores. Los razonadores son herramientas que aprovechan el poder semántico de las ontologías, de forma que, aplicándose sobre éstas pueda extraer el conocimiento no expresado de forma explícita en su construcción. Estas herramientas son las que finalmente propician en gran parte el potencial de las ontologías como herramientas de gestión de la información. A la hora de definir una ontología se encuentran tres partes principales componentes en la misma:

- Clases e instancias, que son de forma básica los objetos o elementos que la componen.
- Propiedades, que representan las relaciones entre los anteriores.
- Reglas, que se emplean para modelar el conocimiento que no puede recogerse a través de los elementos anteriores.

De esta forma se podría decir que las ontologías recogen más conocimiento del que realmente se presenta en un principio y este puede extraerse, haciendo deducciones o respondiendo a consultas

lógicas, mediante el empleo de los razonadores. En la presente investigación se utilizó el razonador Pellet²⁶. Unas de sus funcionalidades principales son:

- **Comprobación de la consistencia:** Comprueba que no existen contradicciones en la ontología. La semántica de OWL define una especificación formal para la definición de la consistencia en una ontología mediante el uso Pellet.
- **Corrección de los conceptos:** Verifica si es posible que se definan instancias para una clase; si no es así, definir una instancia para una clase provocaría que toda la ontología fuera inconsistente.
- **Clasificación:** Observa la relación entre cada clase y crea la jerarquía de clases completa.
- **Realización:** Encuentra las clases más específicas a las que pertenece una instancia; en otras palabras, determina la clase a la que pertenece cada uno de los individuos.

Otro medio de evaluación para probar ontologías para llegar a determinar su reusabilidad, es mediante la métrica denominada Ontometric. Este método permite seleccionar entre un conjunto de ontología la que resulte más apropiada para lo que se necesite, además de la más idónea para ser utilizada en una aplicación, basándose en una medida cuantitativa. Agrupa 5 criterios fundamentales para la evaluación: contenido, lenguaje, metodología, herramientas, y costos. El presente trabajo busca evaluar una única ontología de un único dominio, por lo que este método no serviría para la evaluación.

Como otro paso, las ontologías se evalúan mediante el Esquema para evaluar ontologías únicas para un dominio de conocimiento. Este método fue producto de una investigación financiada por el Consejo de Desarrollo Científico y Humanístico de la Universidad Central de Venezuela. Los criterios a evaluar en cada fase del ciclo de vida del desarrollo de la ontología son: el uso correcto del lenguaje utilizado para la codificación, la exactitud de la estructura taxonómica, el significado de los términos y conceptos representados y la adecuación a los requerimientos especificados al inicio del desarrollo. (Casañas & otros, Esquema para evaluar ontologías únicas para un dominio de conocimiento, 2009) Mediante una investigación realizada por estos autores, los criterios más importantes que deben ser evaluados son (**Ver Fig.15**):

²⁶ editor de ontologías y frameworks de bases de conocimiento, gratuito open-source.

Criterios de evaluación de ontologías

Criterio Autor	Taxonomía	Lenguaje	Aplicación	Vocabulario	Arquitectura Requerimientos	Aceptación Social	Razonamiento Automático	Software
(1)	√	-	√	√	-	-	-	-
(2)	-	√	√	√	-	-	√	-
(3)	√	-	√	√	-	-	-	-
(4)	-	√	√	√	-	√	-	-
(5)	√	√	√	√	√	-	-	-
(6)	√	√	√	√	√	-	-	√

(1) Brewster y cols, 2004 (2) Obrst y cols, 2007 (3) Porzel y Malaka, 2004
 (4) Burton-Jones y cols, 2005 (5) Brank y cols, 2005 (6) Lozano-Tello, 2002 √ Considera - No considera

Fig. 15 Esquema de criterios de evaluación de ontologías

Fuente: (Casañas & otros, 2009)

Como es posible ver, los criterios que predominan son: taxonomía, lenguaje, el rendimiento de la aplicación y el vocabulario. Es por ello que los autores se basan en estos 4 criterios para desarrollar el método.

Sobre la base de este método de evaluación se trabajará en este trabajo. El esquema está compuesto por 4 fases y las actividades que lo constituyen; cada una de las fases en correspondencia con cada criterio de evaluación.

3.1.1 Fases:

1. Uso correcto del lenguaje

Se recomienda que el lenguaje seleccionado sea sólido y completo; y que la escritura esté libre de errores o defectos. El lenguaje escogido fue OWL permitiendo máxima expresividad sin perder completitud computacional. Para su evaluación se aplicarían métodos de razonamiento sobre la ontología. Se utilizó el razonador Pellet para la evaluación sintáctica de la ontología permitiendo el uso correcto del lenguaje.

Se concluye que no presenta incongruencias y cumple con todas las restricciones definidas. (**Ver Fig. 16**).

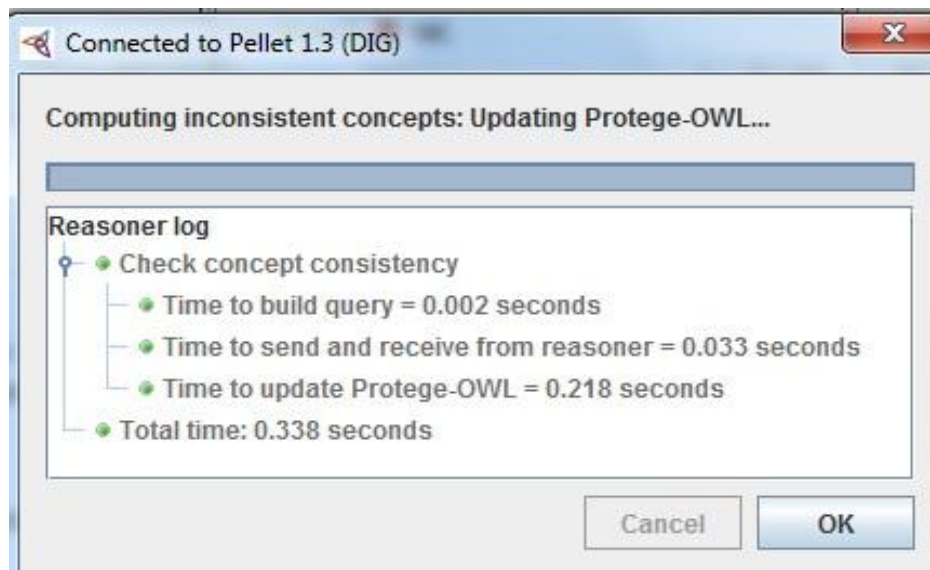


Fig.16 Chequeo de consistencia mediante el razonador Pellet

2. Exactitud de la estructura taxonómica

La evaluación taxonómica permite considerar el chequeo de inconsistencias, completitud y redundancia de los términos. Los errores más comunes son (Casañas & otros, 2009):

- ✓ La clasificación de conceptos como subclase de una clase a la que no pertenecen.
- ✓ Clases e instancias con diferentes nombres pero con definiciones similares.
- ✓ Omisión de conocimiento disjunto entre clases.
- ✓ Ausencia de conceptos.
- ✓ Clases con más de una relación de subclase.
- ✓ Clases definidas como generalización o especializaciones de sí misma.

Para corroborar todos estos errores se utilizó el razonador Pellet, mediante el cual se llevo a cabo el chequeo de redundancias en clases y relaciones. (**Ver Fig.17**)

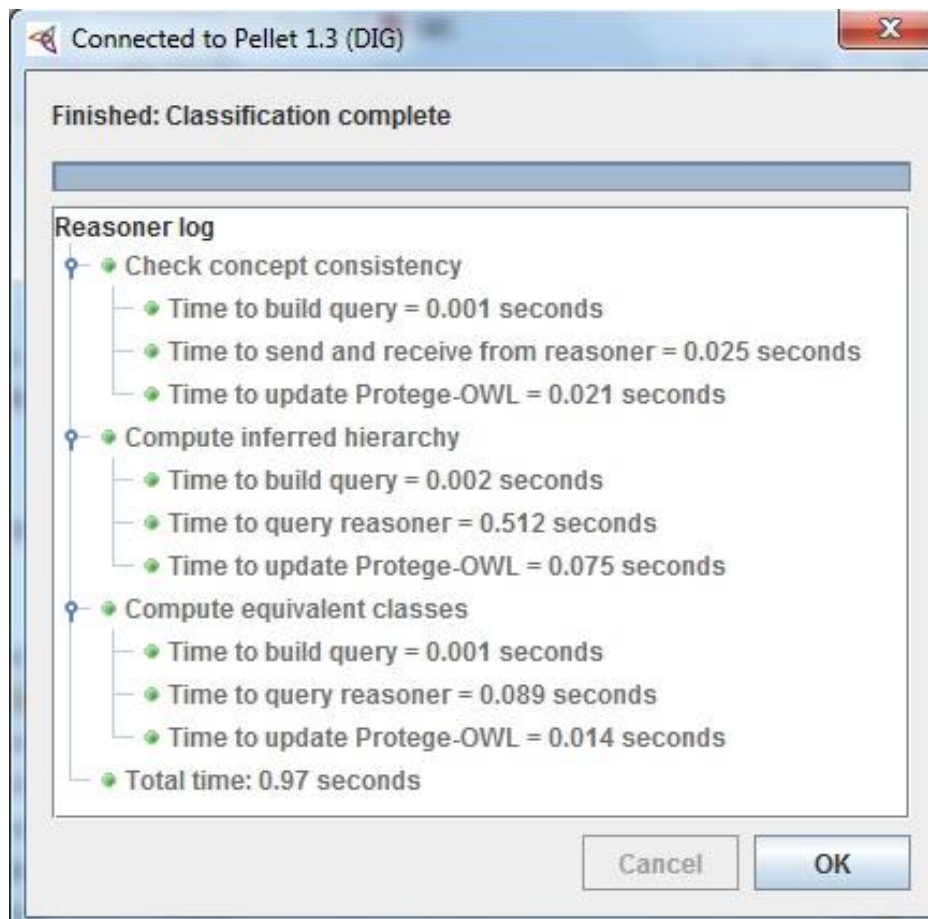


Fig.17 Chequeo de la taxonomía mediante el razonador Pellet

3. Validez del vocabulario

Evalúa el significado de los términos y conceptos a partir del conocimiento de expertos, recopilación de textos o cualquier fuente de conocimiento del dominio. Se verifica que cada término representado en la ontología exista y sean significativos en otras fuentes de conocimiento independientes. Por lo que se verificará el conocimiento almacenado en el corpus ²⁷ del dominio.

Las actividades en esta fase son:

- ✓ Analizar el corpus del dominio, donde se identifica y organiza en una tabla los términos significativos del dominio. (**Ver Tabla 4**)
- ✓ Evaluar el vocabulario mediante medidas de calidad como el recall (exhaustividad) y la precisión. Para ello se utilizará la tabla realizada en la actividad anterior.

Precisión: porcentaje de los términos de la ontología que aparecen el corpus con relación a la cantidad total de términos de la ontología mediante la siguiente expresión matemática:

Precisión = $CO-C / COnto$ donde:

CO-C = Cantidad de términos que se solapan entre la ontología y el corpus.

²⁷ conjunto más extenso y ordenado posible de datos o textos científicos, literarios, etc., que puedan servir de base para una investigación. <http://rae.es>

COnto = Cantidad total de términos de la ontología.

Recall: porcentaje de términos del corpus que aparecen en la ontología con relación al total de términos en el corpus, utilizando la siguiente expresión:

$$\text{Recall} = \text{CO-C} / \text{CCorp}$$

CCorp = Cantidad total de términos del corpus

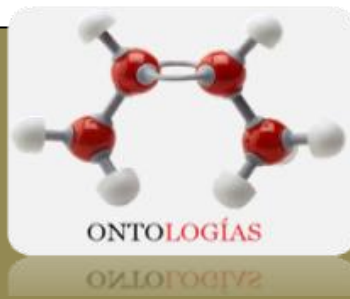
De acuerdo a los resultados se analiza que la precisión se indica un valor de 0,70 de los términos codificados en las ontologías que existen en el corpus; el resultado del Recall refiere que el 0,77 de los términos del corpus existen en la ontología. El valor del Recall se debe a que los términos del dominio son difíciles de identificar.

4. Adecuación a requerimientos

Se validan los requerimientos que la ontología debe de cumplir. La validación de los requerimientos se realiza durante todo el ciclo de desarrollo de la ontología, verificando que los objetivos planteados en el documento se cumplan. Con esta ontología se logró interrelacionar la información gestionada por el grupo GIC, y por consiguiente, mejoras en el proceso de búsqueda de la información.

Conclusiones del capítulo

La evaluación de la ontología de dominio para el grupo GIC fue posible gracias al esquema propuesto anteriormente; los resultados obtenidos fueron satisfactorios. La aplicación del esquema propuesto en cada fase del ciclo de desarrollo de las ontologías permitió identificar errores e inconsistencias sintácticas en el lenguaje OWL, redundancia y ambigüedad en las clases, así como omisiones en el vocabulario utilizado para representar el dominio de conocimiento.



Conclusiones Generales

- El estudio del estado del arte de la WS y las ontologías, así como el estudio de los principales elementos de las ontologías: metodologías, lenguajes y herramientas, contribuyó de manera importante en la creación de la solución de este trabajo de diploma.
- La utilización de la metodología seleccionada Noy y McGuinness facilitó el seguimiento y monitoreo del proceso de desarrollo que permitió mantener la calidad en la ontología.
- La herramienta Protegé mediante el lenguaje OWL, puestos en práctica permitió un desarrollo claro, fluido y definido de la propuesta de solución.
- La evaluación de las ontologías constituyó un requisito de obligatorio cumplimiento en el proceso de diseño de la solución propuesta, garantizando la integridad conceptual y la facilidad de operación sobre la ontología creada.
- Con el desarrollo de la solución se garantizará la interrelación de la información gestionada por el grupo, mejorando el proceso de búsqueda.

Recomendaciones

- Integrar la ontología al Observatorio Tecnológico.

- A, G.-P., M, F.-L., & M., C. (2004). *Ontological Engineering*. London. Inglaterra: Springer Verlag London.
- Abián, M. A. (2005). *El futuro de la Web.XML,RDF/RDFS, ontologías y la Web Semántica*.
- Agents, F. f. (2000). *FIPA Ontology Service Specification*.
- Agm. (2006). *mineriaweb*. Retrieved 2012, from *mineriaweb*: <http://cs.uns.edu.ar/~agm/mineriaweb/downloads/Outlines/clase15-outline.pdf>.
- Arano, S. (2005). *hipertext.net*. Retrieved 2012, from *hipertext.net*: <http://www.hipertext.net/web/pag260.htm>
- Báez, Y. R. (2011). *Estado de la producción científica en Web Semántica en el período 2005 – 2009 según la disponibilidad de Computers & Applied Sciences Complete*. La Habana: Universidad de las Ciencias Informáticas.
- BERNERS-LEE, T., HENDLER, J., & LASSILA, O. (2001). *Scientific American*. Retrieved 2012, from *Scientific American*: <http://www.sciam.com>
- Berners-Lee, T., Hendler, J., & ORA, L. (2001). The Semantic Web. *Scientific American* , 4.
- BRAVO, S., & M., R. (2009). *Módulo de Inferencia para un Repositorio Semántico*. Cuba: Universidad de las Ciencias Informáticas (UCI).
- Cabrera Pupo, K., & Martínez Gamboa, A. (2007). *Propuesta metodológica para la gestión del conocimiento basada en otología*. Universidad de las Ciencias Informáticas (UCI), Habana.Cuba.
- CALVO, R. A., GHIGLIONE, E., & ELLIS, R. (2006). *The OpenACS e-learning infrastructure*. Sydney,: The University of Sydney.
- Cantera Fonseca, J. M., Hierro Sureda, J. J., & Romo Zabala, P. Á. (2011). *Fundación Telefónica*. Retrieved 12 10, 2011, from *Fundación Telefónica*: http://sociedadinformacion.fundacion.telefonica.com/DYC/SHI/seccion=1188&idioma=es_ES&id=2009100116310013&activo=4.do?elem=4299
- Casañas, R., & otros, y. (2009). Esquema para evaluar ontologías únicas para un dominio de conocimiento. *Revista Venezolana de Información, Tecnología y Conocimiento* , 71.
- Casañas, R., & otros, y. (2009). Esquema para evaluar ontologías únicas para un dominio de conocimiento. *Revista Venezolana de Información, Tecnología y Conocimiento* , 71.
- Ceccaroni, L. (2001). *Onto WEDSS- An ontology-Based environmental decisión-support system for the management of wastewater treatment plants*.
- Centelles, M. (2005). *Hipertext.net*. Retrieved 02 09, 2009, from *Hipertext.net*: <http://www.hipertext.net/web/pag264.htm>
- Claudia Milena Rodríguez A*, W. C. (2009). Razonadores semánticos:un estado del arte. *INGENIUM: Revista de la Facultad de Ingeniería* .
- Contreras, J., & Martínez Comeche, J. A. (2002). *TUTORIAL ONTOLOGÍAS*. Madrid: ISOCO & Universidad Complutense de Madrid.

- Copyright Knowledge Master Corporation. (2007). *Knowledge Master*. Retrieved noviembre 2011, from Knowledge Master: <http://www.conceptmaps.it/KM-ConceptualKnowledgeBases-esp.htm>
- Devedzic, V. (2002). *Understanding ontological Engineering*. University of Belgrade. (Yugoslavia Communications of the ACM ed., Vol. Volume 45). Yugoslavia: School of business Administration Department of Information Systems.
- DIAZ, E. (2006). "Web Semántica para recuperar información". Universidad Carlos III, Madrid.
- Dürsteler, J. C. (2001, 01 22). *La revista digital de InfoVis.net*. Retrieved 10 2011, from La revista digital de InfoVis.net: <http://www.infovis.net/printMag.php?num=26&lang=1>
- Escrig, M. T., Pacheco, J., & Toledo, F. (2001, julio). *El Lenguaje de Programación de Prolog*.
- FIPA, F. f. (2000). *FIPA Ontology Service Specification* (Vol. XC00086C).
- Gómez Pérez, A. (1995). *Some ideas and examples to evaluate ontologies*. 1043-0989/95 IEEE.
- Gómezy otros, G.-P. A.-L. (2004). *Ontological Engineering*. Springer Verlag London.
- Graells, D. P. (1999). *LOS RIESGOS DE INTERNET*.
- Gruber, T. (1993). *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*. Knowledge Systems Laboratory, Stanford University.
- Guarino, N. (1998). *Formal Ontologies and Information Systems*.
- GUARINO, N. (1996). Understanding, Building, and Using Ontologies. In N. GUARINO, *Knowledge Acquisition Workshop*.
- Horridge, M. (2011, March 24). A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools Edition 1.3. *A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools Edition 1.3*. Manchester, Manchester, Inglaterra: Copyright The University Of Manchester.
- Instituto de Robótica. (2005). *Estado del Arte. Estudio comparativo de soluciones para el desarrollo de ontologías, y análisis de los diferentes lenguajes, herramientas y métodos que permiten representar una ontología*. Universidad de Valencia. D2.Versión 1.0.
- Interficto (equipo de link builders, S. y. (2010, 09 28). *articulo.org*. Retrieved 01 15, 2011, from *articulo.org*: http://www.articulo.org/articulo/25407/que_es_el_ciberespacio.html
- LABRADA, O. L. (2003). *Web Semántica y Ontologías, desafíos para el Profesional de la Información en el siglo XXI*. Cuba.
- Lapuente, M. J. (2011). *Hacia la Web Semántica*.
- Lapuente., M. J. (2010). *HIPERTEXTO: EL NUEVO CONCEPTO DE DOCUMENTO EN LA CULTURA DE LA IMAGEN*. 14: 07.
- Lee, B. (2006). Inventor de la Web y del HTML. (E. Anbarasan, Interviewer)
- Lee, T. B. (1999). *Weaving the Web*. Estados Unidos: Copyrighted Material.
- López, F. J. (2002). *Introducción a las Ontologías*. Escuela Universitaria Politécnica de Albacete.
- Lujani, F. (2010). *Maestros del Web*. Retrieved 2012, from Maestros del Web: <http://www.maestrosdelweb.com/editorial/tim-berners-web2/>

- marcialgalindoitt. (2010, mayo). *Administración del conocimiento*. Retrieved 2012, from Administración del conocimiento: http://marcialgalindoitt-admonconocimiento.blogspot.com/2010_05_01_archive.html
- Masana, S. (2002). *El ciberterrorismo: ¿una amenaza real para la paz mundial?*
- Noy, N. F., & McGuinness, D. L. (2005). *Desarrollo de Ontologías-101: Guía Para Crear Tu Primera Ontología*. (T. d. Antezanal, Ed.) Stanford: Universidad de Stanford, CA, 94305.
- Noy, N., & McGuinness, D. (2001). *Ontology development 101: A Guide to creating your first ontology* (Technical Report KSL-01-05 ed.). Stanford: Stanford knowledge Systems Laboratory. Stanford University.
- Ortiz, A. M. (2000). "SANTOVENI@ + ASISTENCIA BIBLIOGRAFICA" ,*Proyecto de referencia virtual y asistencia bibliográfica*. Retrieved from "SANTOVENI@ + ASISTENCIA BIBLIOGRAFICA" ,Proyecto de referencia virtual y asistencia bibliográfica.
- Paquette, G., & Magnan, F. (2008). *From a Conceptual Ontology to the TELOS Operational System*. Canada: ICE Research Chair, LICEF Research Center, Télé-université.
- Peñalvo, F. J. (2004). *Web Semántica y Ontologías*.
- Pérez, D. T. (2005, 12 14). *Puleva Salud*. Retrieved 01 15, 2011, from Puleva Salud: http://www.pulevasalud.com/ps/subcategoria.jsp?ID_CATEGORIA=103905&RUTA=1-3-73-103905
- PEROJO, K. R. (2006). *Web Semántica: un nuevo enfoque hacia la Organización de Información en los Sistemas de Gestión de Contenidos*. Cuba.
- Ramos, E., & Nuñez, H. (2007). *ONTOLOGÍAS: componentes, metodologías, lenguajes, herramientas y aplicaciones*. Universidad Central de Venezuela, Facultad de Ciencias, Caracas, Venezuela.
- Refsnes. (2012). *W3Schools*. Retrieved from W3Schools: www.W3Schools.com
- Reytor, M. A. (2009). *Módulo de Inferencia para un Repositorio Semántico*.
- Ros-Martín, M. (2004-2011). *El documentalista enredado*. Retrieved abril Miércoles, 5 , 2006 , from El documentalista enredado: <http://www.documentalistaenredado.ne>
- Soltero, A. P., Navarro Hernández, R. F., Sánchez Schmitz, G., & Barceló Valenzuela, M. (2007). *Avances hacia la Aplicación de las Tecnologías de la Web*. Revista Ingeniería Informática.
- Steve, G, Gangemi, A., & Pisanelli, D. (1998). Integrating Medical Terminologies with ONIONS. Retrieved from <http://saussure.irmkant.rm.cnr.it>: <http://saussure.irmkant.rm.cnr.it>
- SUÁREZ, Y., & PAZ, C. M. (2008.). *Mecanismos Semiautomáticos y Automáticos para la Generación de Ontologías de Dominio*. Cuba: Universidad de las Ciencias Informáticas (UCI).
- Swartout, W. (1999). *Ontologies*. *IEEE Intelligent Systems*.
- Tallarico, M. (2003). *Uso de ontologías en tareas de recuperación de información*. Tesina de Licenciatura.
- Tramullas, J. (1999). *Agentes y ontologías para el tratamiento de información: clasificación y recuperación en Internet* (p.247-252 ed.). España, España: IV Congreso ISKO España.

- Uschold, M., & Gruninger, M. (1996). *Ontologies: Principles, Methods and Applications* (Vol. Vol 11 Number 2). Knowledge Engineering Review.
- Vallez, M. (2009). *La web semántica y el procesamiento del lenguaje natural*.
- VAN HEIJST, G., SCHEREIBER, A., & Y WIELINGA, B. (1996). Using Explicit Ontologies in KBS Development. In G. VAN HEIJST, A. SCHEREIBER, & B. Y WIELINGA, *International Journal of Human and Computer Studies*.
- Velásquez Pérez, T., Puentes Velásquez, A. M., & Guzmán Luna, J. A. (2011). Ontologías: una técnica de representación de conocimiento. *Revista Avances en Sistemas e Informática* , Vol 8 (No.2), 216.
- W3C. (2010, 05 06). *Sitio Web de la Oficina Española del W3c*. Retrieved 10 06, 2011, from Sitio Web de la Oficina Española del W3c: <http://www.w3c.es/divulgacion/guiasbreves/websemantica>
- Zamora, P. L. (2003). *DEONTOLOGÍA Y AUTORREGULACIÓN EN LA*. Madrid: ISBN: 84-669-2869-3.