



Facultad 15

Universidad de las Ciencias Informáticas.

Título: Mecanismos de búsqueda, recuperación e inferencia en megaontologías. Aplicación en un Repositorio Semántico.

Trabajo de Diploma para optar por el título de Ingeniero en Ciencias Informáticas.

Autor(es): Yeniset Soto Isalgué

Yonosleydis Zamora Leyva

Tutor(es): MSc. David Leyva Leyva

Ing. Maikel Aparicio Reytor

Curso 2009-2010



Declaramos ser autores de la presente tesis y reconocemos a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

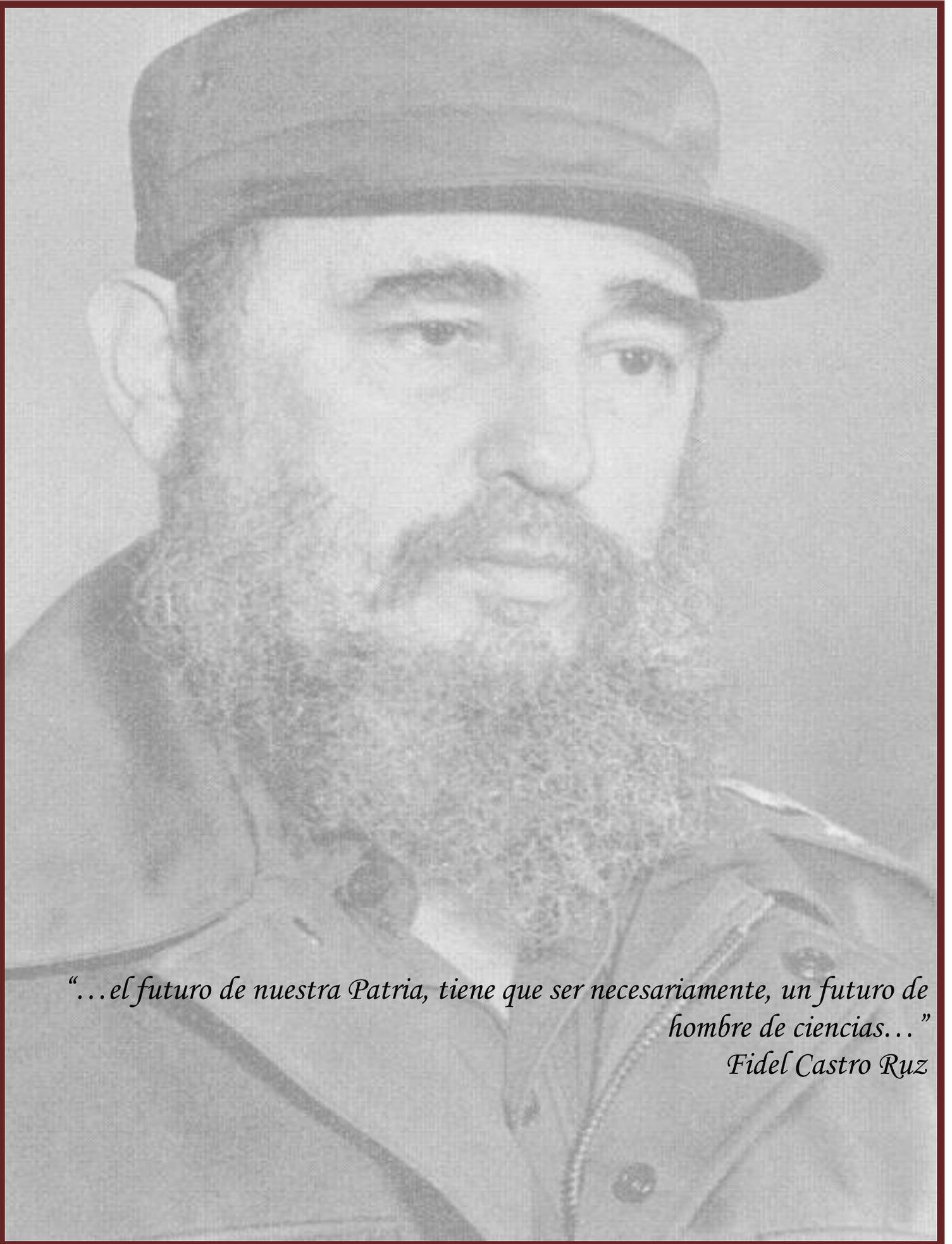
Para que así conste firmamos la presente a los ____ días del mes de _____ del año _____.

Yeniset Soto Isalgué

Yonosleydis Zamora Leyva

MSc.David Leyva Leyva

Ing.Maikel Aparicio Reytor



“...el futuro de nuestra Patria, tiene que ser necesariamente, un futuro de hombre de ciencias...”

Fidel Castro Ruz

Síntesis de los Tutores

David Leyva Leyva: Jefe de Departamento de Producción de Herramientas Educativas, perteneciente al Centro de Tecnologías para la Formación. Facultad 8, Universidad de Ciencias Informáticas (UCI). Teléfono: (53) (07) 837 2453. davidl@uci.cu Licenciado en Cibernética-Matemática, Universidad de Las Villas (UCLV), 1988. Máster en Computación Aplicada, Universidad de Las Villas (UCLV), 1995. Profesor de la Universidad de Holguín (UHo) desde 1990. Jefe de Departamento de Informática (2003-2005). Participó en un Proyecto de Educación a Distancia en la UHo del 2001 hasta el 2005, año en que comenzó a trabajar en la UCI. Ha participado en un gran número de eventos nacionales e internacionales. Le fueron concedidas sendas becas Intercampus en La Universidad de Castilla-La Mancha (1998) y La Universidad Autónoma de Madrid (2000). Trabajó como profesor invitado en la University of Belize (Belice, de 2001 a 2003) y de la Universidad Nacional de Ingeniería (Managua, Nicaragua, 2005). En la UCI ha sido Jefe de Departamento de la Especialidad, Vice decano de Producción e Investigación, Jefe de Departamento de Técnicas de Programación y Jefe del Polo Teleformación, en todos los casos en la Facultad 10.

Maikel Aparicio Reytor: Profesor de Departamento Implantación y Soporte Técnico, perteneciente al Centro de Tecnologías para la Formación. Facultad 8, Universidad de Ciencias Informáticas (UCI). Graduado como Ingeniero en Ciencias Informáticas en la Universidad de Ciencias Informáticas (UCI). Carretera a San Antonio de los Baños Km. 2 ½, Torrens, Boyeros, Ciudad de La Habana. Cuba. Teléfono: (53) (07) 835 8282. maparicio@uci.cu Categoría Docente: Instructor Recién Graduado en Adiestramiento. Su trabajo de diploma fue desarrollado en la línea investigativa del tema que se propone. Como estudiante trabajó desde segundo año como miembro del antiguo Polo Teleformación.

Agradecimientos

Compartidos

A la Revolución y principalmente a la UCI por brindarnos la posibilidad de realizar tan importante sueño.

A nuestros tutores por exigimos, apoyarnos y confiar en nosotras, gracias.

A todos los que de una manera u otra contribuyeron en nuestra formación profesional.

Trabajo de Diploma



Agradecimientos

Yeniset

A mi madre por todo el amor, dedicación y confianza brindada durante toda mi vida, eres única mamita, gracias.

A mi padre que a pesar de no estar ya mi lado me brindó mucho amor y grandes enseñanzas durante su vida.

A mi hermana por todo su amor y apoyo incondicional, te quiero mucho tata.

A mi familia por apoyarme y estar pendientes siempre de mí, a todos gracias.

A Yonita por su amistad y por haber compartido conmigo todo el desarrollo de este trabajo.

A mis grandes amigas Dayana y Daliana por ser tan buenas amigas, y haber permanecido siempre a mi lado en los buenos y malos momentos.

A mis amigos en general, a los encontrados en la universidad, a quienes recordaré siempre con mucho cariño por todos los momentos vividos en esta gran casa, a todos gracias por su amistad.



Agradecimientos

Yonosleydis

A mi madre por ser tan especial, brindarme su amor incondicional, su amistad y confianza siempre.

Te quiero mucho mimi, gracias

A mi padre que siempre me ha apoyado, gracias por brindarme tu amor y confianza.

Te quiero papá.

A mi familia por apoyarme, especialmente a mi hermano, mi abuela, mi tía Paula, mi prima Yanisel que me han acompañado en los momentos más difíciles de mi vida.

A Yeni por su amistad y por haber compartido conmigo todo el desarrollo de este trabajo.

A mi novio Carlos por soportar mi carácter y por estar siempre a mi lado, gracias por todo.

Te quiero così.

A mejor amigo que me acompañó durante esta etapa de mi vida, por su comprensión y su amistad, gracias

Te extraño mucho Jean Carlos.



Dedicatoria

Yeniset

*Al recuerdo de mi padre, que siempre ocupará un gran lugar en mi corazón.
A mi madre y a mi hermana por ser las personas más importantes de mi vida.*

Yonosleydis

A mis padres Emilia y Gerardo que son mi guía, por todo su esfuerzo y sacrificio y a mi hermanito Lisban.

Resumen

Con el creciente desarrollo de las tecnologías se ha incrementado el uso de Internet, el cual se ha convertido en un instrumento de uso masivo en la sociedad actual. La Web actual presenta varios problemas relacionados con la búsqueda y organización de la información. La Web Semántica es una posible solución a estos problemas ya que es una extensión de la actual que permite encontrar, compartir y combinar la información más fácilmente. Una ontología es un pilar fundamental en las tecnologías orientadas a la Web Semántica, estas definen los términos que se usan para describir y representar un cierto dominio.

En este trabajo se hace una propuesta para gestionar el conocimiento utilizando megaontologías. Inicialmente se analizan algunas de las metodologías existentes para la creación de ontologías y se propone una serie de pasos con el fin de crear una megaontología, también se proponen dos posibles soluciones, luego se explica la realización de los procesos de recuperación de Información mediante Fusión, Referencias a través de NameSpaces, Referencias directas por URL y por último se realizan las pruebas a los mecanismos propuestos.

Se logró probar una parte de la propuesta con el apoyo de las siguientes tecnologías: el lenguaje OWL-DL, Protégé como editor antológico, para gestionar el contenido de la Ontología se aprovecha Jena, el razonador Pellet y por último se utiliza el lenguaje de consulta SPARQL; obteniéndose resultados satisfactorios.

Palabras Claves: Web Semántica, Ontología, megaontología, Recuperación de información, Metodología, Referencias.



Tabla de contenido

INTRODUCCIÓN 14

CAPÍTULO 01. FUNDAMENTOS TEÓRICOS 20

 1.1 Introducción 20

 1.2 La Web actual 20

 1.3 La Web Semántica 21

 1.3.1 Tecnologías asociadas a la Web Semántica 22

 1.3.1.1 Lenguajes asociados a la Web Semántica 23

 Extensible Markup Language (XML) y XML Schema 23

 RDF y RDF-Schema 24

 1.3.1.2 Las Ontologías como Soporte de la Web Semántica 26

 1.3.1.2.1 Aplicaciones de las ontologías para la Web 28

 1.3.1.2.2 Tipos de Ontologías 30

 Alcance de su aplicabilidad 30

 Granularidad de la conceptualización 30

 Según el tipo de agente al que va destinada 31

 Según el grado de abstracción o nivel de razonamiento 31

 1.3.1.2.3 Componentes de una ontología (14) 31

 1.3.1.2.4 Consideraciones en el diseño de ontologías 31

 1.3.1.2.5 Lenguajes para escribir Ontologías 32

 SHOE (Simple HTML Ontology Extensions.) 32

 OIL 32



<i>DAML-OIL (DARPA's Agent Markup Language + Ontology Inference Layer)</i>	32
OWL (Ontology Web Language) (18)	33
KIF (Knowledge Interchange Format)	34
1.3.1.3 Lenguaje de reglas SWRL	35
1.3.1.4 Herramientas Ontológicas.	36
1.3.1.4.1 Editores ontológicos	36
Protégé	36
WebOnto	37
OilEd	37
Swoop	38
Jess (Java ExpertSystem Shell)	39
1.3.1.4.2 Razonadores Ontológicos.	40
Pellet	40
RACER	40
Cerebra	40
BOR	41
KAON 2	41
1.3.1.4.3 Sistemas de Almacenamiento	42
Sesame	42
Jena 2	42
1.3.2 Metodologías para el desarrollo de ontologías	42
Características generales	42
Metodologías más utilizadas	43



Methontology.....	43
Kactus.....	43
Sensus.....	43
OntoClean.....	43
CyC.....	44
On-to-Knowledge.....	44
1.3.3 Fusión de ontologías.....	45
1.3.3.1 Metodologías para la fusión de ontologías.....	46
1.3.3.2 Herramientas para el uso de estas metodologías.....	46
1.3.4 Recuperación de Información.....	47
1.3.4.1 Lenguajes de Recuperación de Información.....	48
1.4 E-Learning y Repositorios.....	50
1.4.1 Objetos de Aprendizaje (OA).....	52
1.4.2 Repositorios de OA.....	53
1.4.3 Repositorio Semántico.....	54
1.4.4 Repositorio Semántico de OA.....	54
1.4.5 Repositorio Semántico de OA en la UCI.....	55
1.5 Conclusiones.....	55
CAPÍTULO 2. PROPUESTA DE MECANISMOS MEGAONTOLÓGICOS.....	58
2.1 Introducción.....	58
2.2 Proceso de creación de la megaontología.....	58
2.2.1 Fusión de Ontologías.....	58
2.2.2 Referencias Ontológicas.....	59



2.3 Proceso de recuperación de la información	60
2.3.1 Proceso de recuperación de la información cuando se fusionan ontologías.....	60
2.3.2 Proceso de recuperación de la información a través de referencias por Namespaces.....	62
2.3.3 Proceso de recuperación de la información a través referencias directas por URL.....	62
Solución #1	62
Solución #2	63
CAPÍTULO 3 PRUEBAS	66
3.1 Introducción.....	66
3.2 Ontología base como apoyo al proceso de recuperación de la información.....	66
3.3 Pruebas a los procesos de creación de la megaontología.....	68
3.3.1 Fusión.....	68
3.3.2 Referencias por NameSpaces	69
3.3.3 Referencias directas por URL	71
3.4 Pruebas a los procesos de recuperación de la información	73
3.4.1 Fusión.....	73
3.5 Conclusiones.....	73
CONCLUSIONES GENERALES.....	74
RECOMENDACIONES	75
BIBLIOGRAFÍA	76
REFERENCIAS BIBLIOGRÁFICAS.....	78
ANEXOS.....	81
GLOSARIO DE TÉRMINOS	84



Introducción

Hoy día el desarrollo de las tecnologías se hace cada vez más creciente, provocando dependencias en la mayoría de las esferas sociales. Lograr el próspero avance de una empresa implica casi indispensablemente el uso de las tecnologías. Aparejado a esto viene la utilización de Internet como medio valioso de obtención de información.

Actualmente, en la sociedad, es cada vez más común el uso de Internet, creando en muchas personas el concepto de que desconocer la red es como estar excluido de la realidad actual. La Web es uno de los diversos servicios que ofrece Internet. Es un sistema que facilita el acceso a los diferentes recursos que Internet proporciona. Ha alcanzado gran éxito y popularidad al utilizar programas dotados de interfaces gráficas amigables que permiten a cualquier persona acceder a los documentos de la red.

El uso masivo de la web debido a su fácil acceso ha creado el mito de que algo solamente existe si nuestro buscador es capaz de retornarlo como resultado de una búsqueda. Uno de los principales problemas a la hora de realizar búsquedas en la web es que no siempre obtenemos la información específica que solicitamos. Esto es debido a la falta de organización que presenta la Web actual, provocando que se arrojen como resultados de las búsquedas informaciones innecesarias y carentes de significado para quien solicita el servicio, además del largo tiempo que emplean los usuarios recorriendo las páginas mostradas para localizar la información que realmente necesitan.

Según estudios realizados, se ha demostrado que una forma de solucionar estos problemas es la utilización de la Web Semántica, que *“trata de dotar de una estructura semántica a los contenidos significativos de la Web, creando un entorno en el que agentes software naveguen por las páginas realizando complejas tareas para los usuarios. La Web Semántica no es una nueva Web segregada de la actual. Es una extensión de la actual en la que la información se ofrece con un significado bien definido, permitiendo a ordenadores y personas trabajar de forma cooperativa”*. (1)

“La Web semántica mantiene los principios que han hecho de la web actual un éxito, como son los principios de descentralización, compartición, compatibilidad, máxima facilidad de acceso y contribución, o la apertura al crecimiento y uso no previstos de antemano. En este contexto, un problema clave es



alcanzar un entendimiento entre las partes que han de intervenir en la construcción y explotación de la Web: usuarios, desarrolladores y programas de muy diverso perfil. La web semántica rescata la noción de ontología del campo de la Inteligencia Artificial como vehículo para cumplir este objetivo.” (2)

Una ontología es un elemento fundamental en las tecnologías orientadas a la Web Semántica (WS). El término ontología en informática hace referencia a la formulación de un exhaustivo y riguroso esquema conceptual dentro de uno o varios dominios dados; con la finalidad de facilitar la comunicación y el intercambio de información entre diferentes sistemas y entidades.

La Web Semántica permite que la información que se encuentra en internet pueda interpretarse no solo por humanos sino también por los ordenadores. Entre las tecnologías básicas de la WS se encuentran *Extensible Markup Lenguaje (XML)* se utiliza para presentar, manipular y transmitir documentos y datos estructurados, *Resource Description Framework (RDF)* se utiliza para formalizar los meta-datos, *Ontology Web Lenguaje (OWL)*, considerado el lenguaje clave para el modelado de ontologías.

En la Universidad de las Ciencias Informáticas (UCI) desde hace varios años el antiguo Polo de Teleformación comenzó a desarrollar un Repositorio de Objetos de Aprendizaje (ROA), el cual gestiona paquetes bajo el estándar SCORM ¹ 1.2 y el esquema de metadatos LOM (Learning Object Metadata). Dicho repositorio ya se encuentra instalado para su explotación durante el presente curso.

Los Objetos de Aprendizaje (OA) se pueden definir como *“cualquier recurso con una intención formativa, compuesto de uno o varios elementos digitales, descrito con **metadatos**², que pueda ser utilizado y*

¹ SCORM (Shareable Content Object Reusable Model) es el estándar de e-learning más utilizado a nivel mundial. Está orientado además a que se utilicen repositorios de píldoras formativas, que son pequeñas unidades de información independientes. De esta manera, podría crearse un curso entero reuniendo un determinado número de estas unidades, que pueden formar parte a su vez de otros cursos distintos.

² **Metadatos:** son “datos sobre datos”. “descripciones estructuradas y opcionales que están disponibles de forma pública para ayudar a localizar objetos” o “datos estructurados y codificados que describen características de instancias conteniendo informaciones para ayudar a identificar, descubrir, valorar y administrar las instancias descritas”. Esta clase de definiciones hace mayor hincapié en los metadatos en relación con la recuperación de información.



*reutilizado dentro de un entorno e-Learning*³ (3). Los objetos de aprendizaje tienen como objetivo lograr la reutilización de los contenidos. Un elemento fundamental en este proceso son los repositorios de objetos de aprendizaje que posibilitan la gestión de los mismos.

Desde hace varios cursos en el propio polo se iniciaron los trabajos alrededor de la WS y la Gestión del Conocimiento basada en ontologías. En el curso recién concluido se implementó un Módulo de Inferencia para un Repositorio Semántico que constituye una extensión de ROA.

Como resultado de la investigación se crearon ontologías para modelar la organización de la meta-información de los Objetos de Aprendizaje y un módulo de inferencia para la gestión de información en las mismas, formando con esto la base para hacer una extensión semántica del ROA, que contará con otros módulos.(4)

El mencionado módulo ha sido probado con ontologías asociadas al e-Learning; sin embargo, cuando el repositorio esté en explotación a plena capacidad, éstas deben extenderse a otros dominios e interconectarse con las ontologías ya existentes, formando las llamadas megaontologías, es decir, la relación de múltiples ontologías en una misma búsqueda.

Una de las funcionalidades que le aporta mayor valor al repositorio son las búsquedas (simples o extendidas). La búsqueda simple se realiza sobre una serie de campos fijos preseleccionados y la extendida, el usuario selecciona los campos donde quiere buscar. Las búsquedas semánticas en el repositorio se basarían en ontologías construidas a partir de los metadatos de los paquetes SCORM.

Es deseable que la recuperación de información (en las ontologías construidas, o en el modelo ontológico) mantenga el mismo nivel de respuesta independientemente del crecimiento de las ontologías. Las búsquedas inferidas permiten agilizar este nivel de respuestas en el modelo ontológico. Del mismo modo, los mecanismos de inferencia en el Repositorio Semántico deben mantener también un aceptable nivel de respuesta cuando este maneje megaontologías.

³ El e-learning es un sistema de educación electrónico o a distancia en el que se integra el uso de las tecnologías de la información y otros elementos pedagógicos (didácticos) para la formación, capacitación y enseñanza de los usuarios o estudiantes en línea, es decir, se puede entender como una modalidad de aprendizaje dentro de la educación a distancia.



La inferencia es la forma en la que se obtienen conclusiones sobre la base de datos y declaraciones establecidas (5). Consecuentemente, podemos decir que la representación del conocimiento se logra con la definición de una taxonomía y sus reglas de inferencia relacionadas, estas le añaden una capacidad adicional a las taxonomías, ya que les permiten generar conocimiento. Las ontologías resultan muy útiles para facilitar el razonamiento automático, es decir, sin intervención humana. Partiendo de unas reglas de inferencia, un motor de razonamiento puede usar los datos de las ontologías para inferir conclusiones de ellos. (6)

En un trabajo de diploma anterior (4), se desarrolló un módulo de inferencia para un repositorio semántico. A medida que la información contenida en dicho repositorio comienza a crecer, aumenta de igual forma el número de ontologías. De ahí que en este trabajo se intente proponer mecanismos para realizar búsquedas semánticas en el Repositorio de Objetos de Aprendizaje (ROA), en la medida que se extienden las ontologías hasta conformar megaontologías, es decir, las búsquedas se realizan en los modelos ontológicos por separado, y es necesario hacerlas de una forma centralizada, esto sería posible mediante la unión de la información de los modelos ontológicos, dando lugar a las llamadas megaontologías, lo cual sería de extrema utilidad para que los usuarios tengan mayor facilidad para acceder a los servicios que brinda el repositorio.

Problema a resolver: ¿Cómo realizar la recuperación de información en el ROA cuando se manejen megaontologías?

Se define como **Objeto de estudio** la Ingeniería ontológica y como **campo de acción** la recuperación de información en un ROA semántico.

Objetivo general: Proponer mecanismos que garanticen la recuperación de información en el ROA en la medida que se obtienen megaontologías.

Objetivos específicos:



- Identificar procedimientos, tecnologías y/o metodologías para la creación de megaontologías así como la Gestión de Conocimiento (GC)⁴ en las mismas.
- Proponer mecanismos que garanticen la gestión e inferencia de conocimiento en megaontologías.
- Realizar pruebas separadas y a pequeña escala de los mecanismos propuestos.

Idea a defender:

La utilización de una estrategia adecuada para el acoplamiento de ontologías, y el uso de mecanismos para la recuperación de información en la megaontología resultante deben garantizar un nivel de respuesta similar al de las ontologías simples que la componen.

Tareas de investigación

- Estudiar los conceptos relativos a la WS y lenguajes utilizados para la implementación de la WS.
- Revisar Trabajos de Diploma previos sobre el tema de la WS desarrollados en la UCI.
- Estudiar las herramientas utilizadas para la construcción de ontologías, así como otras complementarias.
- Estudiar mecanismos utilizados para gestionar el conocimiento en las ontologías y megaontologías.
- Estudiar los tipos de reglas para procesos de inferencia en megaontologías.
- Plantear una estrategia para la creación de una megaontología.
- Crear una megaontología.
- Proponer y evaluar los mecanismos de recuperación de información e inferencia en megaontologías.
- Realizar pruebas a los mecanismos propuestos.

Durante el desarrollo de la investigación es necesaria la combinación de los métodos **teóricos y empíricos**. Entre los teóricos se utilizará el de **Análisis y Síntesis**, pues permite el estudio de conceptos

⁴ Gestión de Conocimiento (GC) “es el proceso sistemático de buscar, organizar, filtrar y presentar la información con el objetivo de mejorar la comprensión de las personas en un área específica de interés”. (Figuerola, y otros, 2006).



y teorías de la Ingeniería Ontológica, para sintetizar los elementos más importantes de dicho análisis y contribuir a la solución del problema.

Para la elaboración de diagramas relacionados con ontologías se utiliza el método de **Modelación** y el **Histórico-Lógico** para estudiar el transcurso de la evolución sobre el crecimiento de las ontologías, y de las tecnologías para la WS, además de estudiar los trabajos realizados anteriormente.

Se utiliza el método de la **Observación** el cual se encuentra dentro de los métodos Empíricos, pues permite conocer el funcionamiento de herramientas de software como editores ontológicos y razonadores, para hacer una valoración adecuada de ellas que posibilite el desarrollo de los mecanismos deseados.

El presente trabajo consta de una introducción, tres capítulos, conclusiones generales, referencias bibliográficas, anexos y glosario de términos, donde se encuentra la descripción de los términos no comunes.

Capítulo 1 Fundamentación teórica: Se abordan de forma general los principales conceptos presentes en la investigación, entre los que se encuentran la WS y sus tecnologías, los repositorios, los ROA y los Repositorios Semánticos (RS).

Capítulo 2 Propuesta de mecanismos megaontológicos: Se describe la propuesta de mecanismos para garantizar la recuperación de información en ROA en la medida que se obtienen megaontologías.

Capítulo 3 Prueba: En este capítulo se realizan las pruebas de los mecanismos propuestos, en el sistema.



Capítulo 1. Fundamentos Teóricos

1.1 Introducción.

Se abordan de forma general los principales conceptos presentes en la investigación, entre los que se encuentran la WS y sus tecnologías, los repositorios, los ROA y los RS.

1.2 La Web actual

En poco más de una década desde su aparición, la **World Wide Web(WWW)** se ha convertido en un instrumento de uso cotidiano en nuestra sociedad, comparable a otros medios tan importantes como la radio, la televisión o el teléfono, a los que aventaja en muchos aspectos. La Web es hoy un medio extraordinariamente flexible y económico para la comunicación, el comercio y los negocios, ocio y entretenimiento, acceso a información y servicios, difusión de cultura, etc. Paralelamente al crecimiento espectacular de la Web, las tecnologías que la hacen posible han experimentado una rápida evolución. La generación dinámica de páginas, el acoplamiento con bases de datos, la mayor interactividad con el usuario, la concepción de la Web como plataforma universal para el despliegue de aplicaciones, la adaptación al usuario, son algunas de las tendencias evolutivas más marcadas de los últimos años.

Hoy casi todo está representado de una u otra forma en la Web, y con la ayuda de un buen buscador, podemos encontrar información sobre casi cualquier cosa que necesitemos. La Web está cerca de convertirse en una enciclopedia universal del conocimiento humano. Por otra parte, la Web nos permite realizar diferentes actividades de nuestra vida diaria con una comodidad, economía y eficiencia sin precedentes. A pesar de todas estas posibilidades y facilidades que nos ofrece, presenta también sus limitaciones, pues la información contenida en ella carece de organización. Muchas veces cuando realizamos búsquedas en la Web, no obtenemos los resultados esperados, y empleamos demasiado tiempo en la búsqueda de la información deseada. Esto se debe a la falta de capacidad de las representaciones en que se basa la Web actual para expresar significados. Los contenidos y servicios en la Web se presentan en formatos (HTML) e interfaces (formularios) comprensibles por personas, pero no por máquinas.



Para darle solución a estos problemas que presenta la Web actual se han realizado estudios acerca de la WS o Web del futuro llamada así por Tim Berners-Lee en 1999, quien fue el creador de la Web actual y pensador de esta Web futura.

1.3 La Web Semántica

La Web Semántica es 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. Al dotar a la Web de más significado y, por lo tanto, de más semántica⁵, se pueden obtener soluciones a problemas habituales en la búsqueda de información gracias a la utilización de una infraestructura común, mediante la cual, es posible compartir, procesar y transferir información de forma sencilla. Esta Web extendida y basada en el significado, se apoya en lenguajes universales que resuelven los problemas ocasionados por una Web carente de semántica en la que, en ocasiones, el acceso a la información se convierte en una tarea difícil y frustrante.

El creador de la Web actual Tim Berners-Lee define la WS como: *“El primer paso es colocar los datos en la Web de un modo en que las máquinas puedan entenderlos naturalmente convertirlos a esa forma. Esto crea lo que yo llamo una Web Semántica: una red de datos que pueden ser procesados directa o indirectamente por máquinas.”* (4)

“La Web semántica es una extensión de la actual Web en la cual la información se da mediante un significado bien definido, lo que facilita que los ordenadores y la gente trabajen en cooperación.” (4)

La Web semántica propone superar las limitaciones de la Web actual mediante la introducción de descripciones explícitas del significado, la estructura interna y la estructura global de los contenidos y servicios disponibles en la WWW. Frente a la semántica implícita, el crecimiento caótico de recursos, y la ausencia de una organización clara de la Web actual, la Web semántica aboga por clasificar, dotar de estructura y anotar los recursos con semántica explícita procesable por máquinas.

⁵ El término **semántica** se refiere a los aspectos del significado, sentido o interpretación del significado de un determinado elemento, símbolo, palabra, expresión o representación formal.



Esta Web mantiene los principios que han hecho un éxito de la Web actual, como son los principios de descentralización, compartición, compatibilidad, máxima facilidad de acceso y contribución, o la apertura al crecimiento y uso no previstos de antemano.

La WS se basa en la idea de añadir metadatos semánticos y ontológicos a la WWW. Esas informaciones adicionales que describen el contenido, el significado y la relación de los datos se deben proporcionar de manera formal, para que así sea posible evaluarlas automáticamente por máquinas de procesamiento. El objetivo es mejorar Internet ampliando la interoperabilidad entre los sistemas informáticos usando "agentes inteligentes". Agentes inteligentes son programas en las computadoras que buscan información sin operadores humanos.

Dicha Web proporcionará un salto cualitativo sobre el potencial de la Web. Las principales ventajas de esta nueva revolución en Internet serán el desarrollo de aplicaciones con esquemas de datos comunes, fomento de las transacciones entre empresas por comercio electrónico y búsqueda de información con inferencias. Para poder lograr estos objetivos se necesita unificar los contenidos semánticos por medio de ontologías (de las que abordará más adelante) que formalicen este conocimiento de forma consensuada y reutilizable. Estas constituyen una de las tecnologías que sustentan la WS.

1.3.1 Tecnologías asociadas a la Web Semántica

Los principales componentes de la Web Semántica son los metalenguajes y estándares de representación, por ejemplo los ya famosos: XML, XML Schema, RDF, RDF Schema y OWL. La OWL Ontology Web Language describe la función y relación de cada uno de estos componentes de la Web semántica. Más tarde daremos una detallada descripción de estos lenguajes y tecnologías que facilitan la implementación de la WS, que como hemos dicho anteriormente dota de semántica a la Web ya existente, además propone nuevas técnicas para la representación de la información y el conocimiento que faciliten la localización, compartición, integración y recuperación de recursos a través de la Web.



1.3.1.1 Lenguajes asociados a la Web Semántica

Extensible Markup Language (XML) y XML Schema.

XML, es un metalenguaje que define la sintaxis utilizada para definir otros lenguajes de etiquetas estructurados desarrollado por el World Wide Web Consortium (W3C), no es más que un conjunto de reglas para definir etiquetas semánticas que nos organizan un documento en diferentes partes. (7)

XML no ha nacido sólo para su aplicación en Internet, sino que se propone como un estándar para el intercambio de información estructurada entre diferentes plataformas. Se puede usar en bases de datos, editores de texto, hojas de cálculo y casi cualquier cosa imaginable. Es una tecnología sencilla que tiene a su alrededor otras que la complementan y la hacen mucho más grande y con unas posibilidades mucho mayores. Algunos de estos lenguajes que usan XML para su definición son XHTML, MathML etc. Tiene un papel muy importante en la actualidad ya que permite la compatibilidad entre sistemas para compartir la información de una manera segura, fiable y fácil.

Ventajas de XML

- Es extensible, lo que quiere decir que una vez diseñado un lenguaje y puesto en producción, igual es posible extenderlo con la adición de nuevas etiquetas de manera que los antiguos consumidores de la vieja versión todavía puedan entender el nuevo formato.
- El analizador es un componente estándar, no es necesario crear un analizador específico para cada lenguaje. Esto posibilita el empleo de uno de los tantos disponibles. De esta manera, se evitan bugs y se acelera el desarrollo de la aplicación.
- Si un tercero decide usar un documento creado en XML, es sencillo entender su estructura y procesarlo. Mejora la compatibilidad entre aplicaciones.

Un "Schema XML" es algo similar a un DTD (Document Type Definition) es decir, define qué elementos puede contener un documento XML cómo están organizados y qué atributos y de qué tipo pueden tener sus elementos. (8)



Entre los beneficios que ofrece XML Schema respecto a las DTDs se encuentran los siguientes:

- Permite crear tipos definidos por el usuario
- Permite que el texto que aparece en los elementos esté restringido a tipos específicos tales como tipos numéricos en formatos específicos o incluso tipos más complicados como listas o uniones.
- Permite restringir los tipos para crear tipos especializados, por ejemplo especificando valores mínimo y máximo.
- Permite la extensión de tipos complejos mediante el uso de una forma de herencia.
- Es un superconjunto de DTDs.
- Permite restricciones de unicidad y de clave externa.
- Está integrado con espacios de nombres para permitir a diferentes partes de un documento adaptarse a un esquema diferente.
- Él mismo se especifica mediante sintaxis XML

Básicamente el lenguaje XML Schema nos permite definir elementos globales y locales. XML Schema tiene un sistema de tipos de datos que permiten a un elemento en concreto, especificar si deberá contener un entero o si podría contener alguna cadena. En otras palabras, XML Schema nos va a permitir tener más control sobre los contenidos de un documento XML, mientras que con los DTD no se puede. (9)

Aunque XML se incluye aquí como un lenguaje de especificación de ontologías, en realidad no lo es, no está orientado a cumplir ese objetivo sino el intercambio de datos. Este lenguaje permite únicamente extraer datos, pero no su contenido semántico. Desde este punto de vista, XML presenta una serie de problemas por basar su entendimiento en la estructura de los documentos, ya que cuando un esquema XML cambia, por ejemplo, introduce nuevos elementos intermedios que pueden invalidar alguna consulta anterior que haya sido basada en la estructura. (4)

RDF y RDF-Schema.

Según el W3C, El Resource Description Framework (RDF) es un fundamento para el procesamiento de metadatos; proporciona interoperabilidad entre aplicaciones que intercambian información legible por



máquina en la Web, utiliza XML para intercambiar descripciones de recursos Web, puede utilizarse en diferentes áreas de aplicación, por ejemplo: en la recuperación de recursos para proporcionar mejores capacidades a los motores de búsqueda etc. (10)

Si bien suele decirse que RDF es un lenguaje, lo cierto es que resulta más exacto describirlo como un modelo de datos para las instancias de metadatos o, por abreviar, como un modelo de metadatos. En RDF, la construcción básica es la tripleta (sujeto, propiedad, objeto). Toda tripleta RDF corresponde a una sentencia RDF: la parte que identifica a qué se refiere la sentencia es el sujeto; la parte que identifica la característica del sujeto a la que se refiere la sentencia es la propiedad o el predicado; la parte que identifica el valor de la propiedad es el objeto.

RDF no se asocia a ningún dominio en particular: se puede emplear en cualquier campo. Cada persona u organización puede definir su propia terminología o vocabulario mediante lo que se conoce como RDFS (RDF Schema, esquema RDF), que se define en función de sentencias RDF. Además, RDFS permite especificar las entidades a las que pueden aplicarse los atributos del vocabulario.

¿Cómo está diseñado RDF? RDF está cuidadosamente diseñado para tener las siguientes características:

Independencia - Dado que una propiedad es un recurso, toda organización independiente o incluso cada persona puede inventarlas.

Intercambio - Ya que las sentencias RDF se escriben en XML pueden ser fácilmente usadas para intercambiar información.

Escalabilidad - Las sentencias RDF son simples, registros con tres campos (Recurso, propiedad, valor) por lo que son fáciles de manejar y de usar para buscar objetos incluso en volúmenes realmente grandes. La Web ya es lo suficientemente grande y continúa creciendo. Es probable que tengamos en algún momento miles de millones de RDFs flotando a nuestro alrededor algún día. Por eso la escalabilidad es importante.

Las propiedades son recursos - Las propiedades pueden tener sus propias propiedades y pueden ser encontradas y manipuladas como cualquier otro recurso. Esto es importante porque tendremos muchísimos recursos que manejar. Demasiados como para buscarlos uno por uno.



Los valores pueden ser recursos - Por ejemplo, la mayoría de las páginas web podrían tener una propiedad llamada "home" que apunte al home del sitio. Por lo tanto, los valores de sus propiedades que podrían incluir el título y autor de la página también tienen que incluir recursos.

Las sentencias pueden ser recursos - Las sentencias también tienen propiedades. Dado que no hay un estándar para todos los recursos posibles y dado que la Web es demasiado grande como para que cada uno provea el suyo tendremos que realizar búsquedas basadas en los metadatos de otras personas.

RDF provee una forma para expresar enunciados simples acerca de los recursos usando propiedades y valores. Sin embargo, se necesitaba indicar algunas veces que lo que estaban describiendo eran tipos o clases específicas de recursos. RDF por sí sólo no proporciona tal vocabulario, por lo que las clases y propiedades se describen en RDF Vocabulary (también conocido como RDF Schema). RDF Schema no proporciona un vocabulario sobre aplicaciones orientadas a clases, sino que provee de mecanismos para especificar que tales clases y propiedades son parte de un vocabulario y de cómo se espera su relación. También permite definir a los recursos como instancias de una o más clases. Además, permite que las clases puedan ser organizadas en forma jerárquica. Por lo tanto, RDF Schema extiende a RDF para incluir un amplio vocabulario con un significado adicional.

¿Resulta verdaderamente necesario definir un esquema para RDF? Sí: un esquema RDF permite comprobar si un conjunto de tripletas (sujeto, propiedad, objeto) RDF (un conjunto de metadatos, en definitiva) resulta válido o no para ese esquema, RDFS permite controlar la validez de los valores y restringir las entidades a las cuales pueden aplicarse ciertas propiedades. (4)

A partir de lo anteriormente analizado podemos percibir que XML proporcionaría la interoperabilidad sintáctica y RDF/RDF-S, la interoperabilidad semántica; es decir estos metalenguajes poseen semántica pero no la suficiente como para dotar de conocimiento a la Web, esto se lograría con el uso de otros lenguajes escritos a partir de ellos, como los lenguajes que se utilizan para escribir ontologías.

1.3.1.2 Las Ontologías como Soporte de la Web Semántica

Las ontologías actualmente son materia de investigación, desarrollo, y aplicación en temas relacionados con la computación, ingeniería del conocimiento, procesamiento de lenguaje natural y gestión del



conocimiento Empiezan a ser utilizadas a finales de los años ochenta en el ámbito de la Inteligencia Artificial como una forma para compartir y reutilizar el conocimiento existente, y es a mediados de los noventa donde se empieza a darle aplicabilidad en la Web. Hoy en día ha sufrido un impulso debido al desarrollo de la Web Semántica, donde la idea principal es que la información almacenada en la Web sea interpretada no solo por humanos sino también por un computador. La literatura de inteligencia artificial contiene varias definiciones de ontología muchas de ellas contradicen otras, dentro de los cuales podemos citar los siguientes:

Una de las definiciones más extendidas es la dada por Tom Gruber (*Gruber, 1993*):

“Una ontología es una especificación explícita de una conceptualización. El término proviene de la filosofía, donde una ontología es un recuento sistemático de la existencia. En sistemas de Inteligencia Artificial, lo que existe es lo que puede ser representado. Cuando el conocimiento de un dominio se representa mediante un formalismo declarativo, el conjunto de objetos que puede ser representado se llama universo del discurso. Esos conjuntos de objetos, y las relaciones que se establecen entre ellos, son reflejados en un vocabulario con el cual representamos el conocimiento en un sistema basado en conocimiento. Así, en el contexto de IA, podemos describir la ontología de un programa como un conjunto de términos. En tal ontología, las definiciones asocian nombres de entidades del universo del discurso con textos comprensibles por los humanos que describen el significado de los nombres, y axiomas formales que limitan la interpretación y buen uso de dichos términos. Formalmente, una ontología es una teoría lógica”.

La siguiente definición (33) es la que más se ajusta al tema:

“Estructura jerárquica que define formalmente las relaciones semánticas de un conjunto de conceptos. Se usa para crear vocabularios estructurados para la recuperación o el intercambio de información”

El consorcio W3C⁶ define el concepto de ontología de la siguiente manera:

⁶W3C : (‘World Wide Web Consortium’) disponible en : <http://www.w3.org/>



“Una ontología define los términos utilizados para describir y representar un área de conocimiento. Las ontologías son utilizadas por personas, por bases de datos y por aplicaciones que necesitan compartir la información de un dominio (un dominio simplemente es un área temático específico o de conocimiento, como la Medicina, etcétera). Las ontologías incluyen definiciones de conceptos básicos del dominio, utilizables por ordenadores y las relaciones entre los mismos (debe mencionarse que el término no es utilizado en el sentido de la lógica). Las ontologías codifican el conocimiento en un dominio y también el conocimiento que lo extiende. De esta manera, permiten la reutilización de ese conocimiento.”

Como podemos apreciar el concepto de ontología es una pieza fundamental para soportar la representación del conocimiento que necesita la Web semántica. Para que esto pueda llevarse a cabo, se necesita que el conocimiento de la Web esté representado de forma que sea comprendido por los ordenadores y sea reutilizable. Las ontologías proporcionan la vía para representar este conocimiento.

Existen muchas ventajas en la utilización de ontologías, entre ellas: (12)

- Constituyen una herramienta para la adquisición de conocimiento.
- Permite compartir conocimiento
- Crear una red de relaciones entre conceptos.
- La reutilización de conocimiento.

1.3.1.2.1 Aplicaciones de las ontologías para la Web

- Servir de herramienta para la adquisición de información.
- Servir de herramientas de referencia en la construcción de sistemas de bases de conocimiento que aporten consistencia, fiabilidad y falta de ambigüedad a la hora de recuperar información.
- Normalizar los atributos de los metadatos aplicables a los documentos.
- Crear una red de relaciones que aporte especificación y fiabilidad.
- Permitir compartir conocimiento.



- Posibilitar el trabajo cooperativo al funcionar como soporte común de conocimiento entre organizaciones, comunidades científicas, etc.
- Permitir la integración de diferentes perspectivas de usuarios.
- Permitir el tratamiento ponderado del conocimiento para recuperar información de forma automatizada.
- Permitir la construcción automatizada de mapas conceptuales y mapas temáticos.
- Permitir la reutilización del conocimiento existente en nuevos sistemas.
- Permitir la interoperabilidad entre sistemas distintos.
- Establecer modelos normativos que permitan la creación de la semántica de un sistema y un modelo para poder extenderlo y transformarlo entre diferentes contextos.
- Servir de base para la construcción de lenguajes de representación del conocimiento.

En muchas ocasiones el término “ontología” es reemplazado por el término “taxonomía”⁷. En forma general no está muy clara la diferencia que existe entre ellos, debido a que algunos expertos concuerdan en que una ontología no es más que una taxonomía, ya que describe una representación del conocimiento de manera poco profunda. Pero si analizamos la definición de Neches, una ontología no es solamente la definición de términos básicos y relaciones sino también las reglas que permitirán la inferencia, por ello Tallarico M. (2008), enfatiza que una ontología difiere de una taxonomía, ya que una ontología tiene una estructura interna mucho más explícita que una taxonomía, además de que refleja cierto consenso entre las clases que la conforman.

Las ontologías típicas que se pueden encontrar y que son utilizadas en la Web, están conformadas por taxonomías conjuntamente con reglas de inferencia⁸.

⁷ **Taxonomía.-** permite la organización y representación de clases, subclases, relaciones, permitiendo agregar propiedades a las clases, así como su respectiva herencia entre clase y subclase. Éstas son una poderosa forma de representación del conocimiento.

⁸ **Reglas de Inferencia.-** estas le añaden una capacidad adicional a las taxonomías, ya que les permiten generar conocimiento. Por lo general el formato de una regla es ‘si’, “entonces”. Ej.

{Si estudia psicología entonces recomendar artículos de humanidades}



Consecuentemente, podemos decir que la representación del conocimiento se logra con la definición de una taxonomía y sus reglas de inferencia relacionadas.

1.3.1.2.2 Tipos de Ontologías

Las ontologías se pueden clasificar teniendo en cuenta diferentes criterios:(13)

Alcance de su aplicabilidad

- Ontologías de dominio: Proporcionan un vocabulario necesario para describir un dominio dado. Incluyen términos relacionados con los objetos del dominio y sus componentes.
- Ontologías de tarea: Proporcionan un vocabulario para describir términos involucrados en los procesos de resolución de problemas los cuales pueden estar relacionados con tareas similares en el mismo dominio o en dominios distintos. Incluyen nombres, verbos, frases y adjetivos relacionados con la tarea.
- Ontologías generales: Representan los datos generales que no son específicos de un dominio. Son las ontologías de nivel más alto ya que describen conceptos generales (espacio, tiempo, materia, objeto, etc.)
- Ontologías específicas: Son ontologías especializadas que describen los conceptos para un campo limitado del conocimiento o una aplicación concreta.

Granularidad de la conceptualización

- Ontologías terminológicas: Especifican términos que son usados para representar conocimiento en el universo de discurso. Suelen ser usadas para unificar vocabulario en un dominio determinado.
 - Ontologías de información: Ofrecen un marco para el almacenamiento estandarizado de información.
 - Ontologías de modelado del conocimiento: Especifican conceptualizaciones del conocimiento. Poseen una rica estructura interna y suelen estar ajustadas al uso particular del conocimiento que describen (términos y semántica).
-



Según el tipo de agente al que va destinada

- Lingüísticas, vinculados a aspectos gramáticos y semánticos.
- No lingüísticas, a diferencia de las lingüísticas que están destinadas para el entendimiento del ser humano, esta clase se orienta hacia los robots y agentes inteligentes
- Mixtas, que combinan características de las dos anteriores.

Según el grado de abstracción o nivel de razonamiento

- Descriptivas, llamadas así debido a que permiten relaciones entre conceptos.
- Lógicas, a diferencia de las descriptivas permiten realizar inferencias lógicas, mediante el uso de reglas o axiomas.

1.3.1.2.3 Componentes de una ontología (14)

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. Suelen formar la taxonomía del dominio. Por ejemplo: subclase-de, parte-de, parte-exhaustiva-de, conectado-a, etc.

Funciones: son un tipo concreto de relación donde se identifica un elemento mediante el cálculo de una función que considera varios elementos de la Ontología. Por ejemplo, pueden aparecer funciones como categorizar-clase, asignar- fecha, etc.

Instancias: se utilizan para representar objetos determinados de un concepto.

Axiomas: son teoremas que se declaran sobre relaciones que deben cumplir los elementos de la Ontología. Por ejemplo: "Si A y B son de la clase C, entonces A no es subclase de B", "Para todo A que cumpla la condición C1, A es B", etc.

1.3.1.2.4 Consideraciones en el diseño de ontologías

Cuando se desea representar un dominio mediante una ontología, se deben tomar diversas decisiones de diseño. Entre las más importantes, según (15), están:



- Claridad: una ontología debe comunicar de manera eficaz el significado de los términos. Además, las definiciones deben ser objetivas y estar expresadas en lenguaje lo más natural posible.
- Coherencia: una ontología es incoherente si una frase que se puede deducir a partir de axiomas se contradice con una definición o ejemplo dado de manera informal.
- Extensibilidad: debería ser posible definir nuevas expresiones a partir del vocabulario existente en las ontologías, de manera que no se requiera la descripción de nuevos términos.

1.3.1.2.5 Lenguajes para escribir Ontologías.

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 Web Semántica. Las Ontologías y las etiquetas se incrustaban en archivos HTML. Este lenguaje permite definir clases y reglas de inferencia, pero no negaciones o disyunciones. A su albur se desarrollaron muchos editores, buscadores, *APIs* (Application Programming Interface), etc. El proyecto fue abandonado a medida que se desarrollaron Ontology Inference Layer (OIL) y DAML; aunque también existe una socialización de este lenguaje en XML.

OIL

Este lenguaje, derivado en parte de SHOE, fue impulsado también por el proyecto de la Unión Europea On-To-Knowledge. Utiliza la sintaxis del lenguaje XML y está definido como una extensión de RDFS. Se basa tanto en la lógica descriptiva (declaración de axiomas) como en los sistemas basados en frames (taxonomías de clases y atributos). OIL posee varias capas de sub-lenguajes, entre ellas destaca la capa base que es RDFS, a la que cada una de las subsiguientes añade alguna funcionalidad y mayor complejidad. La principal carencia de este lenguaje es la falta de expresividad para declarar axiomas.

DAML-OIL (DARPA's Agent Markup Language + Ontology Inference Layer)

Nació fruto de la cooperación entre OIL y DARPA y unifica los lenguajes DAML y OIL. Se basa ya en estándares del W3C. El lenguaje DAML se desarrolló como una extensión del lenguaje XML y de RDF y para extender el nivel de expresividad de RDFS DAML- OIL 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.

Por lo tanto *DAML+OIL* es un lenguaje que nos ofrece más expresiones sofisticadas para las descripciones de clasificaciones y propiedades de los recursos que las que ofrecía *RDF* y *RDFS*. *DAML+OIL* es la evolución de *RDFS*, en el que se redefinen muchas de sus descripciones y se añaden muchas otras para mejorar el lenguaje y aportar propiedades y mecanismos para que el lenguaje defina ontologías que después pueden ser empleadas por sistemas de razonamiento para poder inferir sobre la información.(17)

OWL (Ontology Web Language) (18)

OWL es uno de los lenguajes recomendados por la organización W3C que incorpora lecciones aprendidas en el diseño y aplicación del lenguaje DAML+OIL. Al igual que los anteriores lenguajes, es usado cuando la información contenida en los documentos necesita ser procesada por las aplicaciones. Este lenguaje provee un mayor conjunto de primitivas para representar el significado de los elementos y sus relaciones con otros elementos en una ontología, permite definir las mismas. OWL se puede formular en RDF, por lo que se suele considerar una extensión de éste. Además, incluye toda la capacidad expresiva de RDF y la extiende con la posibilidad de utilizar expresiones lógicas. Permite la creación de clases y permite atribuir ciertas propiedades a las relaciones, como cardinalidad, simetría, transitividad o relaciones inversas

OWL se divide en tres sublenguajes **OWL-Lite**, **OWL-DL** y **OWL-Full**, cada uno de los cuales queda definido a continuación:

- **OWL Lite**: Tiene restricciones simples (cardinalidad sólo 0 o 1), facilita compatibilidad con otros modelos/paradigmas, facilita desarrollo de herramientas de autor. Tiene un razonamiento eficiente, facilita desarrollo de herramientas, compatibilidad con otros modelos. Es el menos expresivo, se usa en situaciones donde sólo existe una simple jerarquía de clases y donde se necesitan pocas restricciones.
- **OWL DL (Description Logics)**: Es de término medio en cuanto a la expresividad; está basado en lógica descriptiva, la cual es una lógica basada en primer orden; es favorable para el razonamiento automático, para la clasificación de jerarquías y para detectar las inconsistencias en las ontologías.



Limita la expresividad intentando conseguir decidibilidad y máxima expresividad manteniendo completitud. Usa restricciones para asegurar la existencia de un procedimiento de razonamiento decidible. No puede haber restricciones de cardinalidad (locales ni globales) en propiedades transitivas, ni sus inversas, ni sus superpropiedades. No pueden faltar definiciones. Los axiomas de desigualdad sólo se pueden aplicar a instancias con nombre.

- OWL Full: Es el más expresivo, se usa en situaciones donde es más importante una extrema expresividad que la capacidad de razonamiento. Realiza unión de sintaxis OWL y RDF (sin restricciones).

OWL Full se considera la más completa de todas y se supone una extensión de DL que a su vez es una extensión de Lite, por lo que toda ontología correcta en OWL Lite es una ontología correcta en OWL DL, y toda conclusión correcta en OWL Lite es una conclusión correcta en OWL DL (pero no a la inversa). De la misma manera esto también ocurre con OWL DL y OWL Full respectivamente. (17)

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, metaconocimientos, conjuntos, razonamientos no monotónicos, etc.; y pretende ser capaz de representar la mayoría de los conceptos y distinciones actuales de los lenguajes más recientes de representación del conocimiento. Está diseñado para intercambiar conocimiento entre distintos sistemas de computación. (19)

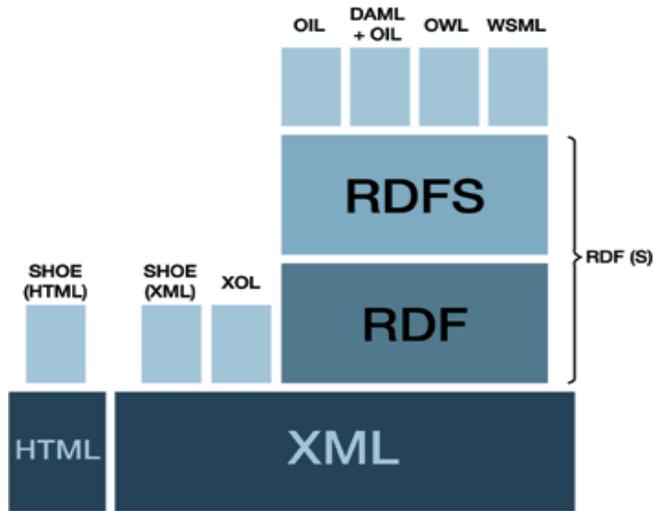


Fig. 1. Relaciones entre los lenguajes de ontologías de la Web Semántica.

1.3.1.3 Lenguaje de reglas SWRL

El significado en inglés para SWRL es “Semantic Web Rules Language”, lo que traducido significa “Un lenguaje de Reglas para la Web Semántica”. Se trata de un nuevo lenguaje basado en una combinación de los sublenguajes OWL DL y OWL Lite con los sublenguajes RuleML⁹ de DataLog¹⁰. El propósito de este nuevo lenguaje es extender los axiomas OWL para incluir reglas. Por tanto, permite la combinación de dichas reglas con una base de conocimiento.

Las reglas se representan en SWRL de forma parecida a una lógica de primer orden; es decir, un antecedente, implicación y un consecuente. En lenguaje natural, dichas reglas se podrían leer como: “Cuando se den las condiciones del antecedente, deben cumplirse también las reglas en el consecuente”.

Una regla tiene la siguiente forma:

antecedente => consecuente

Reglas

⁹ <http://ruleml.org/>

¹⁰ El core de RuleML es Datalog, que constituye un sublenguaje de la lógica de Horn



- Una ontología en sintaxis abstracta de OWL contiene una secuencia de **axiomas** y **hechos**.
- Las **reglas** se forman con un **antecedente** y un **consecuente**.
- Cada **consecuente** puede estar formado por un conjunto de átomos (incluso el conjunto vacío).
- Se pueden utilizar **referencias URI**, que sirven para identificar una regla.

SWRL se ha convertido en el lenguaje de reglas recomendado por la comunidad de WS. Para conocer más acerca de la sintaxis concreta de SWRL, se puede consultar en su Web oficial de la w3c: <http://www.w3.org/Submission/SWRL>

1.3.1.4 Herramientas Ontológicas.

Dentro las herramientas más utilizadas para la construcción de ontologías se encuentran los siguientes:

1.3.1.4.1 Editores ontológicos

Protégé

Es un editor de Ontologías escrito en Java, gratuito y de código abierto. Tras él existe una gran comunidad de desarrolladores y de usuarios universitarios, empresariales y gubernamentales. Actualmente permite trabajar con RDFS y dispone de una extensión para OWL. Su sencillez y su buena documentación lo hacen ideales para los principiantes en Ontologías. (21)

Permite (20)

- Edición de Ontologías
- Creación de herramientas de adquisición de conocimiento mediante formularios relacionados con las ontologías descritas.
- Creación de bases de conocimiento mediante la entrada de instancias particulares de los datos de la ontología.
- Ejecución de aplicaciones que operen sobre la base de conocimiento.



El modelo de Protégé está basado en un simple metamodelo comparable con sistemas orientados a objetos y sistemas basados en marcos. Éstos básicamente pueden representar ontologías que consisten en clases, propiedades, características de propiedades e instancias. La ventaja de Protégé es que su metamodelo es una ontología Protégé. La estructura del metamodelo de Protégé habilita fácilmente la extensión y adaptación a otras representaciones. Usando las vistas de la interface de usuario de Protégé, los diseñadores de ontologías básicamente crean clases y les asignan propiedades. Usando la función *resulting ontologies*, Protégé es capaz de generar automáticamente interfaces de usuario que soporten la creación de instancias. En cada clase de la ontología el sistema crea un formulario con componentes de edición para cada propiedad de la clase. (18)

WebOnto

WebOnto fue diseñado para facilitar la navegación, creación y edición cooperativa de ontologías sin sufrir problemas de interfaz. En particular se diseñó para proporcionar una interfaz de manipulación directa que muestra expresiones ontológicas. Es fácil de usar y tiene facilidades de escalabilidad. Soporta la navegación, creación y edición colaborativa de ontologías representadas en OCML¹¹(Operational Conceptual Modelling Language). El enfoque cooperativo usado en WebOnto es similar al usado en Ontolingua pues los usuarios modifican la misma ontología. Cuando un usuario está modificando una ontología, los otros usuarios sólo pueden explorarla. Las características principales de WebOnto son:

- Gestión de ontologías usando un interfaz gráfico.
- Generación automática de instancias a partir de definiciones de clases, soporte al modelado de PSMs¹² (Platform-specific models).
- Inspección de elementos, teniendo en cuenta la herencia de propiedades y el chequeo de consistencia.

OiEd

¹¹ <http://kmi.open.ac.uk/projects/ocml/>

¹² PSMs (Modelo Específico de la Plataforma) es uno de los modelos utilizados en la Arquitectura Basada en Modelos (MDA), este especifica el sistema en más detalle en términos de la tecnología elegida para la implementación.



Es la herramienta bautizada como el “notepad” de los editores de ontologías. Basado inicialmente para el desarrollo de ontologías *OIL* y *DAML + OIL* se han ido realizando numerosas actualizaciones para que acepte la mayoría de los lenguajes de especificación actuales. Es un editor bastante utilizado por los investigadores porque aporta la posibilidad de interactuar con un razonador como *FACT* o *RACER* que permiten comprobar la consistencia de una ontología. Una de las desventajas que presenta este editor es la carencia de recursos para soportar ontologías grandes, migración e integración de otras ontologías y diferenciación de versiones.

Permite construir el conocimiento y luego ser exportado tanto a RDFS como DAML + OIL u OWL. Permite establecer jerarquías, chequear las consistencias de las ontologías y añadir relaciones de manera implícita. Además, actualmente mediante OilEd es posible acceder a ontologías OWL almacenadas en repositorios Sesame, lo que permitirá migrar sin la necesidad de abordar nuevas herramientas. (17)

Swoop¹³ sirve para crear, editar y depurar ontologías OWL. Fue desarrollado por los laboratorios MIND en la universidad de Maryland, College Park. Es gratuito y su código es abierto. Su desarrollo ha pasado recientemente a manos de Google Code.

Características de Swoop (22)

- El aspecto de Swoop es similar a los navegadores Web; tiene su barra de direcciones para cargar ontologías desde la Web (o desde la maquina local), hipervínculos en los elementos, como describimos anteriormente, y botones (Atrás, Adelante) para una navegación transversal. Además, tiene el menú Bookmarks donde almacenar ontologías favoritas para próximas llamadas.
- Modo de edición amigable, con distintos colores y fuentes (por ejemplo, tipos y formatos de letra para añadir o eliminar), que enfatizan los cambios y acciones, y que hacen más cómodo el trabajo con las ontologías.
- Utiliza la Manchester OWL API.
- Distintos modos de vista.
- Soporta múltiples ontologías simultáneas.

¹³ <http://code.google.com/p/swoop>



- Ejecuta consultas a la caja A, escritas en RDSL sobre Pellet
- Particiona automáticamente la ontología transformándola al modo E-connections1.
- Una ontología puede referirse a otra sin tener que importar todo su contenido.
- Modo de depuración amigable con el usuario, explicando las inconsistencias encontradas de la ontología de un modo claro, y diciendo sus causas.
- Control de versiones.
- Soporte para anotar la ontología con la colaboración de los usuarios.

Jess (Java ExpertSystem Shell)¹⁴

Jess es un lenguaje de programación basado en CLIPS. Fue creado por Ernest Friedman-Hill de Sandia National Laboratories en Livermore, CA. Su funcionamiento se basa en reglas y usa el algoritmo Rete (Algoritmo de Redundancia Temporal). Se puede usar de dos modos:

- Por un lado se puede usar como un **Sistema Experto**¹⁵ (Sistema Basado en Conocimiento). Un Sistema Experto ejecuta ciertas reglas (definidas por el programador) sobre el conocimiento que se tiene.
- Por otro lado, JESS se ha escrito en JAVA y se puede integrar perfectamente en él accediendo directamente a sus clases y librerías para conseguir un desarrollo más rápido de las aplicaciones.

En JESS como en cualquier otro lenguaje podemos encontrar elementos básicos como palabras reservadas, variables locales o globales, funciones y estructuras como if-then, while, foreach, for, try-catch, Y cuyo funcionamiento es bastante similar a Java.

JESS trabaja en base a tres puntos clave:

- **Hechos:** son la forma que utiliza JESS para representar el conocimiento.
- **Reglas:** nos indican la acción que se debe llevar a cabo cuando surge una determinada condición.

¹⁴ Disponible en <http://herzberg.ca.sandia.gov/jess/>

¹⁵ Un **Sistema Experto** es un Sistema Basado en Conocimiento con información precisa de un dominio y con capacidad de aprendizaje sobre él y que llega a trabajar de forma similar a una persona especializada en ese dominio.



- **Preguntas:** se utilizan para obtener respuestas.

1.3.1.4.2 Razonadores Ontológicos.

Pellet: es un razonador de OWL-DL basado en Java. Puede ser utilizado conjuntamente con bibliotecas del API de Jena o del OWL. Mediante su uso es posible validar, comprobar la consistencia de Ontologías, clasificar la taxonomía y contestar a un subconjunto de consultas RDQL (conocido como consultas a ABox en terminología del DL). Se trata de un razonador DL basado en los algoritmos tableaux desarrollados para DL expresiva. Soporta todas las construcciones del OWL DL incluyendo las relacionadas con los nominales, es decir, owl: oneOf y owl: hasValue. (17)

Unas de las funcionalidades principales de este razonador son las de:

- **Comprobación de la consistencia:** Comprobando que no existen contradicciones en la ontología. La semántica de OWL define una especificación formal para la definición de la consistencia en una ontología empleando 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.

RACER (Renamed ABox and Concept Expression Reasoner): desarrollado por Ralf Möller y Volker Haarslev en 1999, pero que ha sido renovado periódicamente hasta la fecha. Es un razonador diseñado para la Web Semántica. Permite la inferencia tanto en conceptos como en instancias, soporta Ontologías escritas en RDF/RDFS/Daml/OWL (apenas tiene restricciones con estos lenguajes) y posee un lenguaje de consulta sencillo para la inferencia de instancias. Puede ser utilizado por OilEd y Protégé para comprobar la consistencia de la Ontología y para hacer consultas sobre el conocimiento. (17)

Cerebra: es un motor de inferencia desarrollado por Network Inference. Parecido al razonador de lógica descriptiva FaCT, sólo que en este caso sí aporta soporte para la inferencia sobre instancias y tipos de



datos. Soporta Ontologías RDF, DAML+OIL y OWL, y como lenguaje de consulta para razonar emplea XQuery [XQuery04]. Tiene versiones para trabajar conjuntamente con los editores OilEd y Protégé. (17)

BOR: es un razonador desarrollado por el Laboratorio Sirma del proyecto On-To-Knowledge. Tiene soporte para ambos tipos de inferencia, tanto sobre instancias como sobre conceptos: chequeo de la consistencia y del modelo de la Ontología, construcción de la jerarquía de conceptos, clasificación de conceptos definidos. El razonador puede ser usado con Ontologías escritas en DAML + OIL, con algunas restricciones, y con Ontologías escritas en la especificación OWL Lite. Además, se puede incorporar a la aplicación Sesame, para dar soporte a Ontologías DAML+OIL en este tipo de repositorios y poder inferir conocimiento o simplemente recuperarlo. (17)

KAON 2 es un razonador semántico desarrollado por la Universidad de Manchester y la Universidad de Karlsruhe, que permite la manipulación y ejecución de inferencias con ontologías representadas en OWL DL, SWRL etc, y que ha sido utilizado en proyectos europeos como DIP¹⁶, SEKT¹⁷ y OntoGov¹⁸. A diferencia de lo que sucede con otros razonadores basados en lógicas descriptivas, tales como RACER o Pellet, KAON2 no implementa cálculo de tableaux como base de su proceso de razonamiento, sino que utiliza una serie de algoritmos que reducen una base de conocimiento a un programa datalog disjunto lo que aumenta de forma considerable el rendimiento del razonador a la hora de ejecutar los procesos de inferencia.

Entre las principales características destacamos que:

- Ofrece una librería de programación para la manipulación ontologías en OWL DL
- Es capaz de resolver consultas conjuntivas que están expresadas en el lenguaje SPARQL
- Un interfaz DIG¹⁹ que permite el acceso a la funcionalidad de KAON2 de otras herramientas, tales como Protégé.

¹⁶ <http://dip.semanticweb.org>

¹⁷ <http://www.sekt-project.org>

¹⁸ <http://www.ontogov.com>

¹⁹ <http://dig.sourceforge.net/>



- Un módulo para extraer datos de bases de datos relacionales.

1.3.1.4.3 Sistemas de Almacenamiento

Sesame es un repositorio para RDF-Schema desarrollado por AIdministrador Nederland bv. Tiene funciones para añadir y eliminar información escrita en RDF en los repositorios, para ser almacenada en cualquier tipo de base de datos (MySQL, Oracle etc.). Soporta los lenguajes de consulta RQL, RDQL y SeRQL, para acceder al conocimiento.

Jena 2 colección de herramientas desarrollado por Hewlett-Packard para la Web Semántica. En esta colección, hay un parser para RDF, un API, el lenguaje de consulta RDQL, soporte para ontologías RDFS, DAML+OIL y OWL y un sistema de almacenamiento basado en bases de datos BerkeleyDB.

1.3.2 Metodologías para el desarrollo de ontologías

Características generales

Para diseñar una ontología es necesario seguir una metodología que defina los pasos precisos en la representación de la información. Cada metodología tiene sus propias características, sin embargo, la gran mayoría de ellas comparten una serie de propiedades:

- Ciclo de vida: es el tiempo transcurrido durante la concepción, elaboración y mantenimiento de la ontología.
- Herramientas de apoyo: las metodologías deben ser apoyadas por herramientas ontológicas que faciliten la creación de la ontología.
- Nivel de modelado: da una idea sobre cómo se realiza la extracción de los datos.
- Nivel de abstracción: da una idea del nivel de relación existente entre la información concreta y la forma en que se extrae dicha información. Cuanto mayor sea el nivel de abstracción, más fácil será para las aplicaciones extraer los datos.



- Uso de ontologías base: normalmente las ontologías reutilizan términos o características de otras ya creadas. (23)

Metodologías más utilizadas

Entre las metodologías más utilizadas podemos destacar las siguientes:

Methontology: Se trata de un modelo a seguir para especificar ontologías tanto a nivel de conocimiento como a nivel de conceptualización. Este método cubre todo el ciclo de vida de la construcción de la ontología, el cual está basado en prototipos evolutivos porque esto permite agregar, cambiar y renovar términos de una nueva versión (prototipo). Los principales pasos de este método son:(17)

- Especificación: de la meta, el alcance y la granularidad de la ontología.
- Conceptualización: en este paso se realiza el modelo conceptual de la ontología. Estructura el conocimiento adquirido usando lenguajes de representación independientes de los lenguajes de implementación (tablas, UML, jerarquías).
- Implementación: Formalizar e implementar el modelo conceptual con lenguajes formales como RDF, OWL, etc.
- Evaluación: Comprobar el funcionamiento de la ontología.

Kactus: 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. (17)

Sensus: Se trata de una metodología que combina recursos para crear nuevas ontologías. (17)

OntoClean: Metodología para validar la adecuación ontológica de las relaciones taxonómicas. Desarrollado por N. Guarino y C. Welty. Analiza el modelo conceptual en los sistemas de información, a partir de las nociones para el análisis ontológico utilizadas en la filosofía. (17)



CyC: Esta metodología fue el resultado de un proyecto de MCC (Microelectronics and Computer technology). Tiene mecanismo de inferencia y control. Se caracteriza por organizar las ontologías en conjuntos modulares llamados microteorías. En cada microteoría está almacenado el conocimiento y razonamiento para un determinado dominio específico como tiempo, espacio, casualidad o agentes. Pueden existir múltiples microteorías para un solo dominio reflejando diferentes perspectivas. (17)

On-to-Knowledge: Esta metodología incluye la identificación de los objetivos alcanzados por las herramientas de gestión del conocimiento y está basado en el análisis de escenarios de uso. Los pasos propuestos por la metodología son: (17)

- Punto de partida: donde los requerimientos de la ontología son especificados y capturados, preguntas de habilidades son identificadas, ontologías potencialmente reusables son estudiadas y se construye una primera versión borrador de la ontología.
- Refinamiento: donde se produce una ontología madura y orientada a aplicación.
- Evaluación: donde los requisitos y las preguntas de habilidades son chequeadas y la ontología es testeada en el entorno de la aplicación.
- Mantenimiento de la ontología.

Las ontologías constituyen la herramienta fundamental para el desarrollo de la WS. Tienen características que permiten la reutilización del conocimiento, haciendo que el trabajo en la creación de ontologías tome un mayor auge.

Se han establecido algunas características propias de las ontologías que las diferencian de otros sistemas:

- Pueden existir ontologías múltiples: El propósito de una ontología es hacer explícito algún punto de vista, por lo que, a veces, será preciso combinar dos o más ontologías. Cada ontología va a introducir conceptualizaciones específicas.
- Podemos identificar niveles de abstracción de las ontologías: Estos niveles de generalización o abstracción nos dan una topología de ontologías. La idea es caracterizar una red de ontologías usando multiplicidad y abstracción, y puesto que no podemos aspirar a tener una descripción



completa del mundo, podemos pensar en una estrategia de construcción gradual de abajo hacia arriba.

- Multiplicidad de la representación: Un concepto puede ser representado de muchas formas, por lo que pueden coexistir múltiples representaciones de un mismo concepto.
- Mapeo de ontologías: Establecer relaciones entre los elementos de una o más ontologías, para establecer conexiones, especializaciones, generalizaciones, etc.

Estudios de diferentes autores coinciden en que los principales pasos para la creación de una ontología se pudieran resumir en los siguientes: (24)

- 1.- Determinar el dominio y alcance de la ontología, además del propósito u objetivo de la misma.
- 2.- Considerar la reutilización de ontologías existentes
- 3.- Enumerar términos importantes de la ontología
- 4.- Definir las clases y la jerarquía de clases
- 5.- Definir las propiedades/slots de las clases
- 6.- Definir las facetas o restricciones de las propiedades o spots
- 7.- Crear instancias

1.3.3 Fusión de ontologías.

Las ontologías permiten la recuperación y acceso a la información al igual que aportan conocimiento sobre un dominio concreto de forma genérica y formal a través de metadatos con la finalidad de ser reutilizado y compartido por diferentes grupos de personas, mediante aplicaciones específicas para ello. (34)

Es habitual encontrar diferentes ontologías que modelan el mismo conocimiento, pero que difieren en el criterio seguido para clasificar los términos generales de la taxonomía. La fusión de ontologías no es más que la elaboración de una ontología a partir de otras existentes sobre un mismo dominio.

Noy y Musen (1999) definen la **fusión de ontologías** y su alineamiento de la siguiente manera:

- **Alineamiento:** consiste en establecer diferentes tipos enlaces entre dos ontologías, manteniendo las originales
- **Fusión de ontologías :** generar una única ontología a partir de las originales



Ventajas de la fusión de ontologías

- Reutilización de **ontologías** existentes.
- Semántica más completa de los términos de la ontología resultante a través de los metadatos.
- Aumento del número de relaciones semánticas por concepto.
- La **recuperación y acceso a la información** en los motores de búsqueda sería más efectiva ya que se obtendrían mejores resultados relacionados con la consulta.

1.3.3.1 Metodologías para la fusión de ontologías.

Estas ventajas que aporta la fusión de ontologías para la recuperación y acceso a la información, ha conllevado a la realización de investigaciones acerca de este tema, debido a la importancia de su uso. De estas investigaciones han surgido metodologías como:

- **FCA - Merge**: La fusión se realiza identificando los documentos que contienen información que se identifica como instancias que pertenecen a conceptos de ambas ontologías.
- **Onions**: con este método se puede crear una biblioteca de ontologías provenientes de diferentes fuentes.
- **Prompt**: se basa en la fusión y alineamiento de ontologías semiautomáticas siguiendo el formato OKBC²⁰.

Esta metodología (Prompt) tiene desarrollado un plug-in para trabajar con Protégé, el cual se utiliza para fusionar ontologías, además esta metodología luego de realizar la fusión verifica la consistencia de la ontología mediante el algoritmo de Tableaux²¹.

1.3.3.2 Herramientas para el uso de estas metodologías

- **Onto-merge**²²

²⁰ Open Knowledge Base Connectivity (OKBC), es un protocolo y una API para acceder al conocimiento en los sistemas de representación del conocimiento, como repositorios de ontologías o base de datos relacionales.

²¹ Es un conjunto preescrito de instrucciones o reglas bien definidas, ordenadas y finitas que permite realizar una actividad mediante pasos sucesivos que no generen dudas a quien lo ejecute. Dado un estado inicial y una entrada, siguiendo los pasos sucesivos se llega a un estado final y se obtiene una solución.

²² <http://cs-www.cs.yale.edu/homes/dvm/daml/ontology-translation.html>



- Chimaera²³
- Semantic Matching²⁴
- Prompt²⁵

1.3.4 Recuperación de Información

La recuperación de información es el conjunto de tareas mediante las cuales el usuario localiza y accede a los recursos de información que son pertinentes para la resolución del problema planteado. En estas tareas desempeñan un papel fundamental los lenguajes documentales, las técnicas de resumen, la descripción del objeto documental, etc. (38)

Concepto de sistema de recuperación de información

Proceso donde se accede a una información previamente almacenada, mediante herramientas informáticas que permiten establecer ecuaciones de búsqueda específicas. Dicha información ha debido de ser estructurada previamente a su almacenamiento. (39)

La recuperación de la información en la WS

La WS para recuperar información pretende la creación de una forma de intercambio de datos que se fundamente en representaciones de las informaciones contenidas en los recursos de la Web, de manera que puedan ser entendidos por las máquinas.

Se busca la mejora en la comunicación entre las máquinas y los humanos, reduciendo la acción de estos segundos sobre los primeros considerablemente. La forma de conseguirlo es dando a la Web un mayor significado, con lo cual se favorece la compartición, procesamiento y transferencia de información de forma sencilla. (40)

Ventajas (40)

²³ <http://www-ksl.stanford.edu/software/chimaera/>

²⁴ <http://www.semanticmatching.com/>

²⁵ <http://protege.stanford.edu/plugins/prompt/prompt.html>



- La WS para recuperar información permite al usuario dejar en manos del software tareas de procesamiento o razonamiento de un determinado contenido.
- Por otra parte, el introducir significado a una Web mejora la eficacia de los buscadores, ya que facilita su tarea de estudio del documento.
- De esta manera, se puede afirmar que la WS facilita la localización de recursos, la comunicación entre programas y personas y el procesamiento y manipulación de datos e información automáticamente en lugar de manualmente, como hasta ahora.

Cómo plantear una estrategia de búsqueda (39)

No existen pautas que nos indiquen cómo hacer con exactitud todas las búsquedas debido a que cada consulta es distinta. Por eso es conveniente definir un procedimiento básico de trabajo:

- Planteando el tema desde distintos puntos de vista
- Determinando qué se sabe del tema
- Formulando nuestra búsqueda mediante:
 - La selección de palabras clave que representen lo que busco (utilizar diccionarios, sinónimos, tesauros, ontologías, etc.)
 - La traducción de las palabras importantes a otros idiomas (inglés)
- Seleccionando las herramientas de búsqueda (índices, motores, metabuscadores). Se recomienda usar distintas herramientas a la vez.
- Aplicando las palabras clave en las herramientas de búsqueda seleccionadas.

1.3.4.1 Lenguajes de Recuperación de Información

Los lenguajes de recuperación (también query languages) son lenguajes informáticos utilizados para recuperar información de sus almacenes. Puede definirse un lenguaje de recuperación como un conjunto de órdenes, operadores y estructuras que, con una organización según unas normas lógicas, permiten la consulta de fuentes y recursos de información electrónica.

Los lenguajes de recuperación pueden dividirse en dos grandes grupos, distinguiendo si su uso se enfoca a las bases de datos relacionales o a la recuperación de información. Los ejemplos son muy variados, incluyendo a **SeRQL** y **SPARQL**. (35)



- SQL, lenguaje de recuperación reconocido como estándar y ampliamente utilizado para bases de datos relacionales.
- XQuery, lenguaje de recuperación para XML.
- OQL (Object Query Language) lenguaje de consulta estándar para bases de datos orientadas a objetos.

SparQL es uno de los lenguajes de consulta específicos para ontologías. Es un lenguaje estándar para recuperación de información. La importancia de SparQL radica en gran parte en su capacidad de expresar sus consultas en un lenguaje de alto nivel sin importar el formato en que estos datos han sido almacenados.

Tim Berners Lee expresó:

“Pretender usar la Web Semántica sin SparQL es como pretender usar una base de datos relacional sin SQL”, “SparQL hace posible consultar información desde bases de datos y otros orígenes de datos en sus estados primitivos a través de la Web”.

SPARQL es un lenguaje de recuperación basado en RDF; su nombre es un acrónimo recursivo del inglés SPARQL Protocol And RDF Query Language. Se trata de una recomendación para crear un lenguaje de consulta dentro de la Web semántica que está ya implementada en muchos lenguajes y bases de datos.

(36)

El **lenguaje de recuperación SPARQL** ha sido diseñado para un uso a escala de la Web, así permite hacer consultas sobre orígenes de datos distribuidos, independientemente del formato. A la hora de recuperar información la creación de una sola consulta a través de diferentes almacenes es mejor que múltiples consultas, además de tener un coste menor y de ofrecer unos resultados mejores. (37)

Componentes

El **lenguaje SPARQL** posee tres componentes importantes: URIs, literales y variables procedentes del lenguaje RDF.

- URIs sirven para especificar las URLs su especificación está en la RFC 3987.
- Literales se describen como una cadena de caracteres encerradas entre " ".



- Variables estas variables son globales, además debe de ser prefijadas por "?" o "\$" no formando parte del nombre de la variable.

Sintaxis

La sintaxis de **SPARQL** es similar a la de RQL, añadiendo algunas modificaciones para facilitar el análisis sintáctico (parsing) del lenguaje. Como RQL, SeRQL se basa en una interpretación formal del grafo de RDF, diferenciándose de RQL en que se basa directamente en RDF Model Theory.

El lenguaje **SPARQL** tiene cuatro tipos de consultas, estas con sus respectivas descripciones:

- **SELECT**: Devuelve todo, o un conjunto de, las variables que coinciden con el patrón de búsqueda.
- **CONSTRUCT**: Devuelve un grafo RDF construido por la sustitución de variables en un conjunto de tres plantillas.
- **ASK**: Devuelve un boolean indicando si los patrones de la consulta coinciden o no.
- **DESCRIBE**: Devuelve un grafo RDF que describe los recursos encontrados.

1.4 E-Learning y Repositorios

Por definición, el e-Learning es el suministro de programas educacionales y sistemas de aprendizaje a través de medios electrónicos. El e-Learning se basa en el uso de una computadora u otro dispositivo electrónico (por ejemplo, un teléfono móvil) para proveer a las personas de material educativo. La educación a distancia creó las bases para el desarrollo del e-Learning, el cual viene a resolver algunas dificultades en cuanto a tiempos, sincronización de agendas, asistencia y viajes, problemas típicos de la educación tradicional.

Así mismo, el e-Learning puede involucrar una mayor variedad de equipo que la educación en línea. El término de e-Learning o educación electrónica abarca un amplio paquete de aplicaciones y procesos, como el aprendizaje basado en Web, capacitación basada en computadoras, salones de clases virtuales y colaboración digital. (25)

En un artículo escrito por el Ing. Jorge A. Mendoza, publicado el 10 de junio del 2003, hace alusión a diversas definiciones más comunes de e-learning, entre ellas: (25)



Técnicamente, el e-Learning es la entrega de material educativo vía cualquier medio electrónico, incluyendo el Internet, Intranets, Extranets, audio, vídeo, red satelital, televisión interactiva, CD y DVD, entre otros medios.

Para los educadores, e-Learning es el uso de tecnologías de redes y comunicaciones para diseñar, seleccionar, administrar, entregar y extender la educación.

Siendo breves y prácticos, el e-Learning es el aprendizaje basado en tecnología.

Para los elocuentes, el e-Learning es el empleo del poder de la red mundial para proporcionar educación, en cualquier momento, en cualquier lugar.

Siendo descriptivos, la educación electrónica es la capacitación y adiestramiento de estudiantes y empleados usando materiales disponibles para Web a través de Internet, llegando a ofrecer sofisticadas facilidades como flujo de audio y vídeo, presentaciones en PowerPoint, vínculos a información relativa al tema publicada en la Web, animación, libros electrónicos y aplicaciones para la generación y edición de imágenes.

Para los epigrafistas, el e-Learning representa la convergencia del aprendizaje y el Internet.

Y finalmente, para los visionarios y futuristas, el e-Learning es a la educación convencional lo que el e-Business a los negocios ordinarios.

Entre las principales ventajas que ofrece la educación virtual podemos encontrar (26):

- La reducción de costos para dar cursos a más número de participantes que lo tradicional en un aula de clases.
- Ahorro en seminarios y capacitación de empresas muy descentralizadas como los bancos y la flexibilidad de horarios, factor de suma importancia pues permite al estudiante calendarizar el curso de la mejor forma posible.



- La interacción que generan los cursos, despertando el interés del estudiante y ayudando a aquellos tímidos a ser de los más activos en clases por medio de foros de discusión y otros medios de participación.

Los repositorios o repositorios digitales constituyen una herramienta de apoyo para el E-Learning.

Un repositorio es un lugar donde se almacena información, la cual puede o no estar estructurada. Cuando un usuario busca información en un repositorio, describe de manera breve el recurso que quiere obtener como respuesta. Los buscadores son la herramienta primaria para realizar búsquedas en repositorios. (27)

Un Repositorio Digital es un sistema en red formado por hardware, software, datos y procedimientos que contiene objetos digitales, contiene metadatos, asegura la identificación persistente del objeto mediante un identificador único persistente. Ofrece funciones de gestión, archivo y preservación de los objetos, proporciona un acceso fácil, controlado y estandarizado a los objetos. Ofrece los sistemas adecuados de seguridad para los objetos y los metadatos, sostenible en el tiempo. (28)

A continuación se abordaran los temas acerca de los Objetos de Aprendizaje (OA) y los ROA, pues el E-Learning utiliza estos recursos digitales con fines educativos, los cuales se almacenan en los mencionados repositorios.

1.4.1 Objetos de Aprendizaje (OA)

El Learning Technology Standards Committee (LTSC)²⁶ estableció en el momento de su creación, en el año 1996, que objeto de aprendizaje es “cualquier entidad, digital o no digital, que pueda ser utilizada, reutilizada o referenciada durante un proceso de aprendizaje mediado por la tecnología”.

David Wiley, editor de un libro muy conocido sobre los objetos de aprendizaje (2001), propone una definición más cerrada: “Cualquier recurso digital que pueda ser reutilizado para favorecer el aprendizaje”.

²⁶ <http://ieeeltsc.org/>



Un objeto de aprendizaje es un conjunto de recursos digitales, autocontenible y reutilizable, con un propósito educativo y constituido por al menos tres componentes internos: Contenidos, actividades de aprendizaje y elementos de contextualización. El objeto de aprendizaje debe tener una estructura de información externa (metadatos) que facilite su almacenamiento, identificación y recuperación. El presente trabajo se basa en esta definición. (29)

1.4.2 Repositorios de OA

La gestión de contenidos en los entornos **e-learning** está generando cambios en diversos sentidos. Entre estos cambios, los contenidos educativos se están construyendo como objetos de aprendizaje y se están recopilando en contenedores, conocidos como Repositorios de Objetos de Aprendizaje, que los organizan y los hacen accesibles para diferentes aplicaciones y perfiles de usuarios.

En un artículo escrito por Clara López Guzmán llamado Los repositorios de objetos de aprendizaje como soporte para los entornos e-learning, hace referencia a algunas definiciones sobre los ROA tales como: (30)

“Sobre OA se ha insistido en que la reutilización debe ser una de sus principales bondades, es entonces casi intrínseco que se tenga un lugar destinado para su almacenamiento y clasificación para facilitar posteriormente su mantenimiento, localización y, posiblemente, también compartir ese OA con otros sistemas para aplicaciones diversas. Este almacén es lo que se conoce como ROA y que ya ha sido mencionado en capítulos anteriores. Los sistemas de repositorios son la infraestructura clave para el desarrollo, almacenamiento, administración, localización y recuperación de todo tipo de contenido digital”.

“Un repositorio es un concepto tan amplio que va desde sencillos sistemas de almacenamiento hasta complejos entornos que incorporan, además de los sistemas de almacenamiento, conjuntos de herramientas que ayudan al proceso de reutilización”.

“Los repositorios de objetos de aprendizaje son bases de datos con búsquedas que alojan recursos digitales y/o metadatos que pueden ser utilizados para el aprendizaje mediado”.



“Un ROA es una colección de OA que tienen información (metadatos) detallada que es accesible vía Internet. Además de alojar los OA los ROA pueden almacenar las ubicaciones de aquellos objetos almacenados en otros sitios, tanto en línea como en ubicaciones locales”.

Un componente importante para la reutilización de un objeto es que esté adecuadamente descrito, a través de sus correspondientes metadatos. Si los metadatos utilizados se apegan a los estándares propuestos por los grupos importantes y se basan en tecnologías abiertas también la reutilización de estos toma un papel de gran valor para compartir información entre repositorios. La consistencia semántica en la descripción de los recursos y la normalización de los datos que cada elemento contenga, también son factores clave para la posible reutilización a través de sistemas.

1.4.3 Repositorio Semántico

Los RS son aquellos que utilizan tecnologías de la WS para optimizar varios procesos, que se realizan de forma ineficiente en los repositorios tradicionales, tales como la recuperación/búsqueda, la organización y el almacenamiento de la información contenida en los mismos, y para esto se utilizan esencialmente las ontologías, pues se pueden crear varias ontologías de diversos dominios para la organización de la información por áreas, y los razonadores, para mejorar los métodos de recuperación de la información mediante las inferencias que estos permiten realizar. (4)

1.4.4 Repositorio Semántico de OA.

Al añadirle significado y estructura al contenido de los ROA, se logra mejorar los procesos que éstos llevan a cabo, o sea, al añadirle semántica a los ROA mediante las tecnologías de la WS pasarían a ser ROA Semántico. Las ontologías permiten darle este significado a los ROA, de forma tal que se puede diseñar una ontología donde cualquier concepto enlazado a la representación de una actividad de aprendizaje es considerado un objeto de aprendizaje. Este nuevo modelo debe proporcionar una serie de funcionalidades adaptadas a cada conceptualización particular de “objeto de aprendizaje” y no necesariamente restringidas a una única definición. La flexibilidad de este nuevo esquema, nos permitirá almacenar en un repositorio cualquier tipo de información sobre un objeto de aprendizaje, normalizada o sin normalizar, y si se desea (depende de la concepción particular del creador) con la especificación del propósito educativo del mismo.



1.4.5 Repositorio Semántico de OA en la UCI.

El primer intento realizado en Cuba para construir un Repositorio de Objetos de Aprendizaje fue en la Universidad de Holguín, en donde se “propuso desarrollar una plataforma Web dinámica para el e-Learning que posibilitara la configuración y administración de la misma, además de la creación, actualización y oferta de cursos de una forma flexible y agradable, sustentada sobre un sistema código abierto”. (31)

En la Universidad de las Ciencias Informáticas se implementó un Repositorio de Objetos de Aprendizaje (ROA). “Este sistema permite disponer de un lugar común donde profesores y estudiantes almacenen los recursos empleados en el proceso de enseñanza aprendizaje” (8). Actualmente ROA cuenta con las cuatro funcionalidades básicas de cualquier repositorio: subir, actualizar, explorar paquetes y buscar recursos. (32)

Con el desarrollo de esta herramienta y la puesta en marcha en la Universidad, se espera que su principal problema ocurra al crecer el número de OA, lo que traerá consigo que las búsquedas se harán ineficientes, lo que traerá consigo buscar un modo de mejorar el sistema de búsquedas y a la vez el acceso a la información del mismo, y conjuntamente organizar los OA del repositorio teniendo en cuenta las descripciones de sus datos. (4)

1.5 Conclusiones.

En este capítulo se expusieron los principales conceptos relacionados con esta investigación, así como los aspectos relacionados con las Ontologías, Web Semántica, Repositorios de Objetos de aprendizaje y metodologías de las ontologías. Se mostraron las principales herramientas y lenguajes utilizados para construir y representar Ontologías, de los cuales para la elaboración del trabajo se ha decidido utilizar las siguientes tecnologías:

OWL debido a qué está diseñado para usarse cuando la información contenida en los documentos necesita ser procesada por programas o aplicaciones, permite atribuir ciertas propiedades a las



relaciones, como cardinalidad, simetría, transitividad, o relaciones inversas. Además, incluye toda la capacidad expresiva de RDF y la extiende con la posibilidad de utilizar expresiones lógicas, es decir, posee mayor poder expresivo que éste. Es un lenguaje que surge como una revisión al lenguaje DAML-OIL, resultando ser mucho más potente.

OWL-DL: Se utilizará este lenguaje ya que acepta la mayor cantidad de reglas semánticas, es estricto, permite que los procesos de inferencia sean óptimos; es más exacto y ofrece mayor cantidad de contenido.

Prompt: Esta metodología tiene una herramienta con su mismo nombre, la cual tiene desarrollado un plug-in para trabajar con Protégé, se basa en el lenguaje OWL, permite manejar múltiples ontologías, tiene en cuenta diferentes características, identifica conflictos como restricciones que violen la herencia.

SPARQL como lenguaje para realizar consultas a los datos explícitamente representados en OWL, es un lenguaje estándar para la recuperación de información, basado en RDF. Permite expresar sus consultas en un lenguaje de alto nivel, sin importar el formato en que estos datos han sido almacenados. Además este lenguaje está incluido entre las interfaces que posee Protégé (editor ontológico que se utilizará).

Pellet como razonador para validar sintácticamente la Ontología, está dentro de la línea de Java y es capaz de procesar ontologías escritas en OWL auxiliándose en la API de desarrollo Jena. Este se utilizará en la realización de las pruebas a la solución propuesta.

Jena API de desarrollo para visualizar el contenido de las ontologías, permite trabajar con OWL, está dentro de la línea Java, posee una interfaz que permite integrar cualquier motor de inferencia, para realizar los procesos de inferencia.

Protégé para construir las Ontologías debido a que es uno de los editores ontológicos más utilizado actualmente, cuenta con un entorno abierto y fácil de entender, las contribuciones en forma de plug-ins están haciendo de esta herramienta un entorno sumamente potente. Contiene además una serie de



interfaces, que facilitan el trabajo del usuario, como SPARQL un “lenguaje de consulta”, y una interfaz para conectarse a diferentes razonadores como Pellet, que es el que se utilizara para realizar las pruebas.

Trabajo de Diploma



Capítulo 2. Propuesta de mecanismos megaontológicos.

2.1 Introducción.

En este capítulo se muestra una propuesta para el trabajo con megaontologías en la cual se han aprovechado las facilidades que brindan las herramientas descritas en el capítulo anterior. Como parte de la propuesta se tratarán temas relacionados con la creación, fusión y recuperación de información en las megaontologías en ROA.

2.2 Proceso de creación de la megaontología.

Para la creación de la megaontología, se analizaron diferentes metodologías para la fusión de ontologías, y quedo definido en el capítulo anterior que se utilizará la metodología Prompt, pues tiene desarrollado un plug-in en el editor ontológico que se usará (Protégé). Se empleará otra forma para relacionar las ontologías, o sea, para conformar la llamada megaontología, esto será a través de referencias de una ontología base a una auxiliar.

Para este proceso se proponen los siguientes pasos, con los cuales no se pretende crear un patrón a seguir, si no una guía de apoyo para trabajar con megaontologías, como parte de la propuesta inicial.

1. Para la creación de la megaontología es necesario contar con **dos ontologías** como mínimo, estas se van a identificar como ontología base y ontología auxiliar, la elección de la ontología base es la que esté más completa en número de conceptos y atributos.
2. Se debe verificar **el dominio** de las ontologías ya que con esto se restringe el área a modelar.

2.2.1 Fusión de Ontologías

Realizar la fusión de las ontologías con la ayuda de la metodología Prompt la cual propone las siguientes actividades:

- 1) **Crear una lista inicial de operaciones sugeridas:** El conjunto de operaciones de fusión sugeridas incluyen la de la fusión y alineamiento :
 - a. Crear una nueva clase en la nueva ontología fusionando las clases comunes de las ontologías originales.



- b. Crear un nuevo atributo en la nueva ontología fusionando los atributos de las ontologías originales.
 - c. Crear una nueva agrupación con las clases comunes y los atributos que estas contienen, mediante la fusión de las agrupaciones de las ontologías originales.
 - d. Crear una nueva clase en la ontología "deep-copy" que consiste en copiar todos los padres de una clase hasta la raíz de la jerarquía y todas aquellas relaciones (atributos) que estén contenidas en dichas clases.
 - e. Crear una nueva clase en la nueva ontología "shallow-copy" que consiste en copiar únicamente la clase y no sus padres o las clases y atributos a los que se refiere.
- 2) **Seleccionar la siguiente operación:** El experto elige una operación de la lista, esta decisión depende de la semántica que desee reflejar.
 - 3) **Ejecutar la operación seleccionada:** Consiste en ejecutar la operación seleccionada por el experto en la actividad 2. Después de este punto, el experto debe indicar los cambios adicionales asociados al propósito de la operación.
 - 4) **Encontrar conflictos:** Esta operación se lleva a cabo analizando los resultados de la ontología. Los siguientes conflictos pueden aparecer en la ontología resultante: Conflictos de nombres: más de una clase puede tener el mismo nombre Referencias pendientes: una clase puede estar relacionada con otra clase que no exista Redundancia en la jerarquía de clases: dos relaciones diferentes referencian a un mismo padre. Valor de los atributos: el valor de los atributos resultantes puede que no represente la semántica que se quiere reflejar con la fusión.
 - 5) **Actualizar la lista de operaciones:** Para actualizar la lista de operaciones, el algoritmo considera similitud semántica.

2.2.2 Referencias Ontológicas



Los autores definieron referenciar ontologías a través de NameSpaces²⁷ (NS) y referencia directa por las URL de las ontologías.

Cuando se hace la referencia a través de NS se le asocia a la URL de una ontología un prefijo para acceder a todos los conceptos y propiedades de la misma.

La referencia directa por URL consiste en definir en la ontología base un concepto que permita acceder directamente al fichero que contiene la ontología.

2.3 Proceso de recuperación de la información

Recuperación de la información: Es la forma de localizar y acceder a la información en cualquier base de conocimiento haciendo uso de varias tecnologías. Incluye además procesos de inferencia.

Para el desarrollo del proceso de recuperación de la información en megaontologías, se tuvieron en cuenta diferentes aspectos que arrojaron posibles vías de solución. Estas posibles soluciones se aplicarán en dependencia de la eficiencia y facilidad de uso que proporcionen, las mismas serán explicadas a continuación.

2.3.1 Proceso de recuperación de la información cuando se fusionan ontologías.

Cuando se fusionen ontologías mediante la metodología Prompt, ya explicada anteriormente, la recuperación de la información se realizará mediante búsquedas clásicas auxiliadas por lenguajes de consultas como SPARQL. Pues hace posible realizar las consultas haciendo uso de los NS, que se encuentran en la ontología resultado de la fusión, a través de los cuales se puede acceder a todos los componentes de la ontología.

²⁷ namespace: un conjunto de nombres en el cual todos los nombres son únicos, es un contexto en el que un grupo de uno o más identificadores pueden existir.



Para el desarrollo de este proceso se plantearán una serie de pasos a seguir después de fusionar las ontologías, ellos son:

- 1- Recuperar Información: Para recuperar la información en la ontología fusionada se realizaran varios subprocesos que se auxilian en lenguajes y herramientas necesarias para este proceso.
 - 1.1- Utilizar una API de desarrollo (Jena), para almacenar, extraer y recuperar la información.
 - 1.1.1- Extraer el Modelo Ontológico.
 - 1.1.2- Incorporar o extraer información del Modelo Ontológico, propiciando el proceso de recuperación de la información.
 - 1.2- Una vez extraído el Modelo Ontológico, también se puede recuperar información sobre el mismo, utilizando lenguajes tales como: SPARQL para realizar consultas a los datos explícitamente representados en OWL.

El mencionado proceso de recuperación de la información puede mejorarse realizando inferencias sobre la ontología para hacer más accesible la información de la ontología, así como obtener nuevos conocimientos. Esto se logra de la siguiente manera:

- 1- Realizar inferencias: las reglas de inferencia permiten generar conocimiento, a través de éstas un motor de razonamiento puede usar los datos de las ontologías para inferir conclusiones de ellos.
 - 1.1- Cargar la ontología fusionada haciendo uso de la API de desarrollo para extraer el modelo ontológico.
 - 1.2- Utilizar un motor de inferencia (Pellet), para procesar ontologías escritas en OWL, el cual se conecta a la API de desarrollo Jena, pues esta posee una interfaz que permite integrar cualquier motor de inferencia, para realizar los procesos de inferencia.
 - 1.3- Aplicar procesos de razonamiento para verificar la taxonomía, es decir, verificar la consistencia del modelo ontológico.
 - 1.4- Ejecutar las reglas en el Modelo Ontológico, mediante SWRL, haciendo uso del motor de inferencia. La ontología base debe contener las reglas que permitan relacionarla con la auxiliar.



1.5- Se obtiene como resultado un modelo inferido, sobre el cual se aplican consultas con el SPARQL, o mediante Jena, se extrae o incorpora información de este Modelo Ontológico, con lo cual se lograría una mejor calidad y velocidad de los procesos de recuperación de la información.

2.3.2 Proceso de recuperación de la información a través de referencias por Namespaces.

Para el desarrollo de este proceso se realizarán los mismos pasos descritos en el proceso anterior. La diferencia entre ambos radica en que a través de referencias por NS la ontología base sólo puede acceder a la información de la auxiliar mediante los NS que se importan de la auxiliar a la base, la base no contiene el conocimiento de la auxiliar, es decir, sólo puede consultar dicho conocimiento, estas referencias unen las ontologías en una misma estructura pero no las fusiona. Mientras que la fusión pasa la estructura de la ontología auxiliar a la base, formando una megaontología con el conocimiento implícito de ambas.

Los NS están compuestos por un prefijo que hace alusión a la URL de la ontología. Tanto en la fusión como en la referencia están estructurados de la siguiente forma:

- En la ontología base NS= URL de la ontología base# recurso al que quiere acceder de la base.
- En la ontología auxiliar NS= URL de la ontología auxiliar # recurso al que quiere acceder de la auxiliar.

Ej:<http://www.semanticweb.org/ontologies/2010/2/Ontology1268059072311.owl#Buscar>.

2.3.3 Proceso de recuperación de la información a través referencias directas por URL.

Este proceso consiste en definir en la ontología base un concepto que permita acceder directamente al fichero que contiene la ontología. Se presentarán dos soluciones, las cuales pueden ser utilizadas en el ROA.

Solución #1

- 1- Recuperar información: la recuperación de información en esta solución propone, crear un único Modelo Ontológico partiendo de dos ontologías independientes (base y auxiliar), sobre el cual se aplicarán los procesos para recuperar su información.



- 1.1- Cargar la ontología base, la cual cuenta con una propiedad url para acceder directamente al fichero que contiene la ontología auxiliar, haciendo uso de la API de desarrollo Jena para extraer los respectivos modelos ontológicos y crear un nuevo modelo ontológico.
- 1.2- Utilizar un motor de inferencia (Pellet), para procesar ontologías escritas en OWL, el cual se conecta a la API de desarrollo Jena.
- 1.3- Aplicar procesos de razonamiento a cada modelo para verificar la taxonomía, y la consistencia de los modelos ontológicos.
- 1.4- Ejecutar las reglas escritas en las ontologías mediante SWRL, haciendo uso del motor de inferencia. La ontología base debe contener las reglas que permitan relacionarla con la auxiliar.
- 1.5- Se obtiene como resultado un nuevo modelo inferido, sobre el cual se aplican consultas con el SPARQL, o mediante Jena, se extrae o incorpora información de este Modelo Ontológico, con lo cual se lograría una mejor calidad y velocidad de los procesos de recuperación de la información.

Solución #2

- 1- Recuperar información: la recuperación de información en esta solución propone, realizar referencias desde el modelo de la ontología base al modelo de la auxiliar, a través de un concepto o propiedad que contiene directamente la URL de la ontología.
 - 1.1- Cargar la ontología base, la cual cuenta con una propiedad url, que permite acceder directamente al fichero que contiene la ontología auxiliar, haciendo uso de la API de desarrollo Jena para extraer los respectivos modelos ontológicos.
 - 1.2- Utilizar un motor de inferencia (Pellet), para procesar ontologías escritas en OWL, el cual se conecta a la API de desarrollo Jena.
 - 1.3- Aplicar procesos de razonamiento a cada modelo para verificar la taxonomía, y la consistencia de los modelos ontológicos.
 - 1.4- Verificar las relaciones entre los modelos ontológicos, a través de la referencia, esta toma como vínculo entre ambos modelos el concepto o propiedad que contiene la URL de la ontología auxiliar, y las reglas que se definieron en la ontología base.
 - 1.5- Ejecutar las reglas escritas en las ontologías mediante SWRL, haciendo uso del motor de inferencia. La ontología base debe contener las reglas que permitan relacionarla con la auxiliar.



1.6- Inferir los modelos ontológicos de las ontologías por separado (base y auxiliar).

1.7- Aplicar el proceso de recuperación de la información de estos modelos ontológicos, auxiliándose en la API de desarrollo Jena, para extraer o incorporar información de los mencionados modelos ontológicos.

Trabajo de Diploma



Trabajo de Diploma



Capítulo 3 Pruebas

3.1 Introducción.

En este capítulo se realizarán las pruebas a los procesos descritos en el capítulo 2, para verificar la validez de dichos procesos y garantizar la eficiencia de su aplicación.

3.2 Ontología base como apoyo al proceso de recuperación de la información.

Para apoyar el proceso de recuperación de la información se encuentra en fase de construcción una ontología base llamada Historial, la cual permite el control estadístico de los recursos del ROA, dicha ontología se utilizará para referenciar a la Ontolom, ontología auxiliar que contiene toda la información almacenada en el repositorio.

Historial, modela tres conceptos con sus propiedades que permitirán gestionar la información en el repositorio. En la siguiente imagen se muestran los conceptos y propiedades de esta ontología base.





Fig.2 Ontología Base

Buscar: Simula una tabla Hash²⁸, para facilitar las búsquedas sin modificar la consistencia del dominio.

Propiedades:

- propiedad: nombre de los conceptos del estándar LOM.
- valor: valor asociado a la propiedad.

Historia: Contiene la información cuantitativa de la ontología, o sea actualiza la cantidad de OA que son añadidos y eliminados de la misma.

Para esto utiliza la propiedad:

- cant OA: posee la cantidad de OA agregados y la cantidad de OA eliminados.

Está compuesta además por una propiedad llamada Ontolom_Anterior, la cual contiene la URL de la ontología anterior a la que está en uso.

OntoLom_caracteristicas: Contiene la cantidad real de OA de la ontología que esté utilizando el ROA, en dependencia de las actualizaciones que se le hagan al repositorio.

Propiedades:

- cant OA: cantidad real de OA.
- fecha: es un identificador (id) para diferenciar una ontología de otra, según las actualizaciones que se realicen en el repositorio.
- observaciones: cualquier observación que se desee agregar.

²⁸ Una tabla **hash** o **mapa hash** es una estructura de datos que asocia *llaves* o *claves* con *valores*. La operación principal que soporta de manera eficiente es la búsqueda: permite el acceso a los elementos (teléfono y dirección, por ejemplo) almacenados a partir de una clave generada (usando el nombre o número de cuenta, por ejemplo). Funciona transformando la clave con una **función hash** en un hash, un número que la tabla hash utiliza para localizar el valor deseado.

- url: contendrá la dirección url, que permitirá acceder directamente al fichero de la ontología auxiliar, o sea esta propiedad permite realizar el proceso de recuperación de la información a través de referencias directas por url.

3.3 Pruebas a los procesos de creación de la megaontología.

3.3.1 Fusión

En la realización de este proceso se aplicó el procedimiento explicado en el capítulo 2 para la fusión de ontologías. Se fusionaron las ontologías Historial y OntoLom, base y auxiliar respectivamente, dando como resultado una megaontología con el conocimiento íntegro de ambas, es decir, con todos los conceptos y propiedades contenidos en ellas. Se obtuvieron los resultados esperados, pues el proceso se desarrolló satisfactoriamente.

La siguiente imagen representa la estructura de conceptos de la megaontología obtenida en el Protégé.

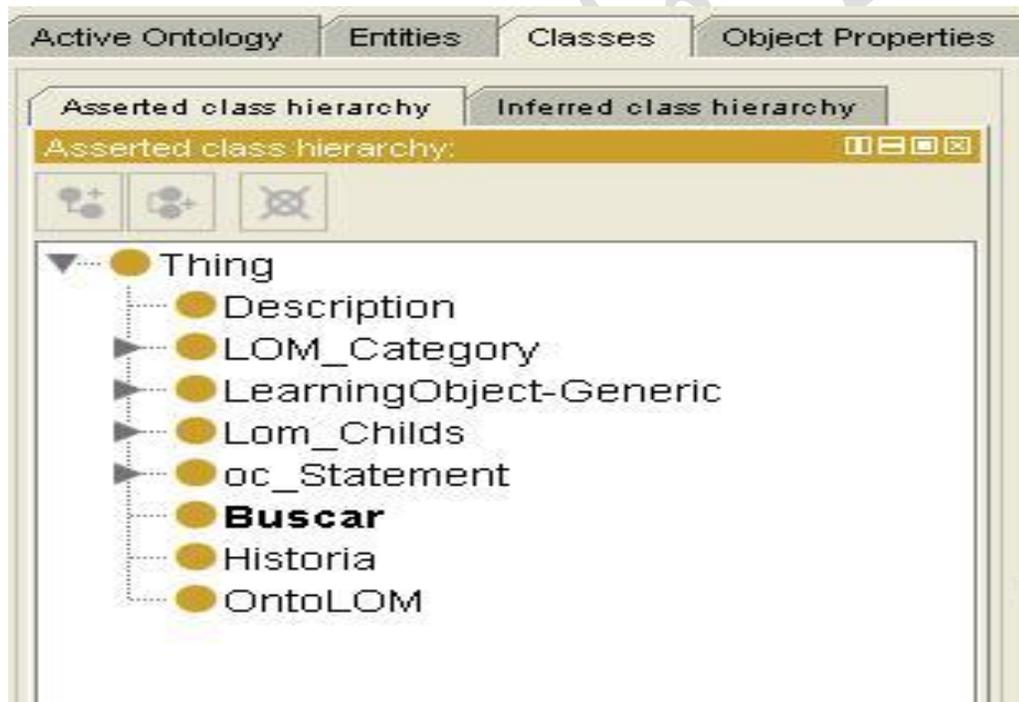


Fig.3 Estructura de la megaontología



Para validar la megaontología se utilizó un servicio brindado por el sitio oficial MINDSWAP (<http://www.mindswap.org/2003/pellet/demo.shtml>), que es la empresa que se dedica al desarrollo del razonador Pellet, dicho sitio permite verificar la consistencia de las ontologías, es decir, si son válidas o no. Se arrojó como resultado la consistencia de la megaontología, demostrando su validez. En los procesos de Referencias por NS y Referencias por URL, se empleó el mismo sitio para validar las megaontologías. La siguiente imagen muestra los resultados obtenidos.

Results

Input file: Text area

OWL Species: Lite

DL Expressivity: ALCHF(D)

Consistent: Yes

Time: 1702 ms (Loading: 1599 Species Validation: 85 Consistency: 18)

Estos resultados demuestran que la megaontología es consistente, inicialmente estaba escrita en el lenguaje OWL-DL, perdió expresividad semántica quedando en el sublenguaje Lite. Contiene una expresividad de la lógica descriptiva ALCHF (D), demorando en cargar el modelo ontológico 1599 ms, en validar las especies 85 ms, y en verificar la consistencia 18 ms, con un tiempo total de 1702 ms.

3.3.2 Referencias por NameSpaces

Como se explicó en el capítulo 2, las referencias a través de NS, consisten en asociarle a la URL de una ontología un prefijo para acceder a todos los conceptos y propiedades de la misma. En este caso desde la ontología base Historial se puede acceder a la auxiliar OntoLom mediante el NS asociado a esta. En este proceso ambas ontologías se unen en una misma estructura, pero a diferencia de la fusión, el conocimiento de OntoLom no pasa íntegro a Historial, esta última solo puede acceder a dicho conocimiento a través de su NS. La siguiente figura muestra una imagen del NS de OntoLom importado desde Historial en el Protégé.

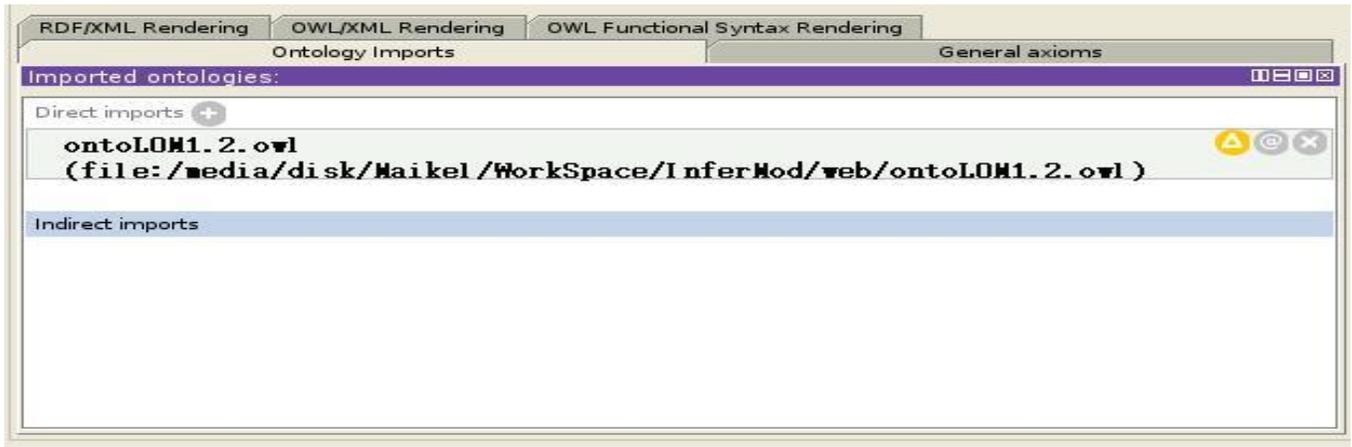


Fig.4

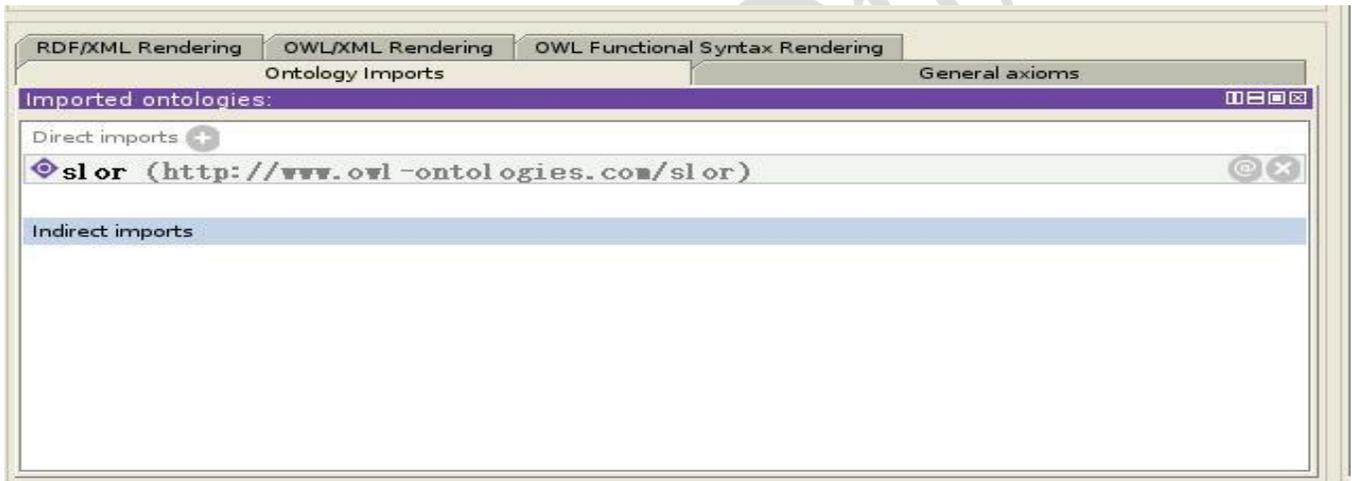


Fig.5

Los resultados obtenidos en la validación fueron los siguientes.

Results

Input file: Text area

OWL Species: Lite

DL Expressivity: ALCHF(D)

Consistent: Yes

Time: 1695 ms (Loading: 1597 Species Validation: 81 Consistency: 17)



Estos resultados demuestran que la megaontología es consistente, inicialmente estaba escrita en el lenguaje OWL-DL, perdió expresividad semántica quedando en el sublenguaje Lite. Contiene una expresividad de la lógica descriptiva ALCHF (D), demorando en cargar el modelo ontológico 1597 ms, en validar las especies 81 ms, y en verificar la consistencia 17 ms, con un tiempo total de 1695 ms.

3.3.3 Referencias directas por URL

Para la realización de este proceso, se creó en Historial una propiedad llamada url que contiene la dirección URL de OntoLom, que permite acceder directamente al fichero donde se encuentra esta ontología auxiliar. La siguiente figura muestra la imagen de esta propiedad en el Protégé.

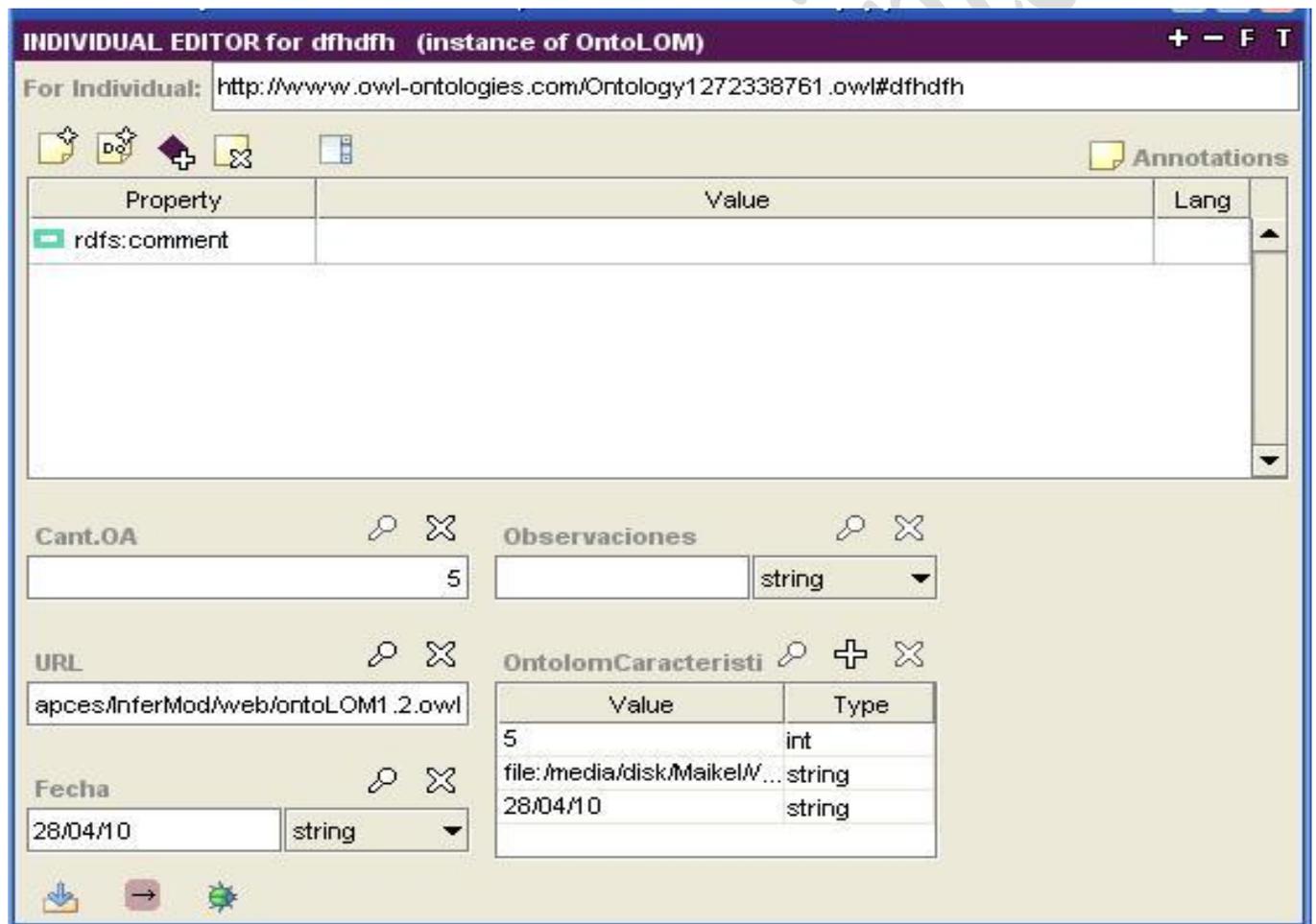


Fig.6



Los resultados obtenidos en la validación fueron los siguientes.

Results

Input file: Text area

OWL Species: Lite

DL Expressivity: ALCHF(D)

Consistent: Yes

Time: 1725 ms (Loading: 1607 Species Validation: 100 Consistency: 17)

Estos resultados demuestran que la megaontología es consistente, inicialmente estaba escrita en el lenguaje OWL-DL, perdió expresividad semántica quedando en el sublenguaje Lite. Contiene una expresividad de la lógica descriptiva ALCHF (D), demorando en cargar el modelo ontológico 1607 ms, en validar las especies 100 ms, y en verificar la consistencia 17 ms, con un tiempo total de 1725 ms.

En la tabla (ver anexo 1) se muestra que entre los procesos de creación de la megaontología, el de referencias a través de NameSpaces es el más rápido, incluso comparado con una ontología simple, es el que mantiene un mejor nivel de respuesta en la recuperación de la información, ya que esta es proporcional al proceso de inferencia y al tiempo total de carga de la megaontología. No siendo así en el proceso de referencias a través de URL, este es mucho más lento que los demás procesos haciendo que la recuperación de la información sea de la misma forma.

Además de utilizar el sitio oficial MINDSWAP para hacer las pruebas, también se realizaron pruebas haciendo uso de una interfaz para inferencia que posee el módulo de inferencia INFERMOD realizado en la tesis anterior (2008-2009), mediante la cual se puede verificar la consistencia de las megaontologías así como la cantidad de elementos que posee, es decir, la cantidad de propiedades y conceptos de las megaontologías. (Ver anexo 2, 3 ,4). Los resultados de las pruebas fueron los siguientes: en el caso de la fusión y la referencia a través de NameSpaces ambas resultaron consistentes y con un total de 15 elementos entre propiedades y conceptos, la referencia por URL también fue consistente y 16 elementos en total.



3.4 Pruebas a los procesos de recuperación de la información

3.4.1 Fusión

El proceso de recuperación de la información cuando se fusionen las ontologías es similar al realizado sobre una sola ontología, la diferencia es que en la megaontología aumenta el número de conceptos y recursos, lo que provoca que el proceso de inferencia sea un poco más lento. Cuando se extrae el modelo ontológico, al mismo se le aplican consultas con el SPARQL o se recupera información de él a través de Jena, manteniéndose el mismo nivel de respuesta.

El desarrollo de la investigación va más avanzado que la realización de las pruebas a los procesos propuestos en la misma, por lo que aún no se han probado los procesos de recuperación de información a través de referencias por NS y referencias directas por URL. Se seguirá trabajando para lograr la realización de dichas pruebas satisfactoriamente.

3.5 Conclusiones.

En este capítulo se probó una parte de los mecanismos propuestos, aunque estos no se llevaron a cabo completamente, sirvieron para verificar una buena porción de ellos y a pesar de que funcionan se debería concluir con estas pruebas siguiendo los pasos propuestos. Finalmente, se determinó que aunque estos mecanismos no se aplicaron en su totalidad, sí se pudo comprobar que en la fusión se mantiene el nivel de recuperación de información.



Conclusiones Generales.

Con el cumplimiento de los objetivos trazados en el presente trabajo se ha arribado a las siguientes conclusiones:

- ❖ Como resultado de esta investigación surgió una propuesta para el trabajo con megaontologías, que permitió realizar los procesos de creación, fusión y recuperación de información en las mismas, aprovechando para ello las facilidades que brindan las Ontologías
- ❖ Se estudiaron las diferentes herramientas que sirven para construir Ontologías.
- ❖ Se creó un Procedimiento para la creación de megaontologías, el cual se puso en práctica en la creación de las megaontologías que se utilizaron para el desarrollo de las pruebas.
- ❖ Se propusieron procesos para la recuperación de información sobre las megaontologías creadas.
- ❖ Esta investigación sirve como guía o punto de partida para el desarrollo de trabajos posteriores relacionados con las megaontologías y puesta en marcha de los procesos representados en la misma.



Recomendaciones

A partir de los resultados que proporciona este trabajo de diploma y para maximizar las utilidades que puede generar:

- ❖ Que sea consultada la bibliografía de la investigación, por parte de todas aquellas personas que se apoyen en este trabajo, para que amplíen así la información obtenida en el mismo.
- ❖ Se recomienda automatizar los procesos propuestos en la investigación.
- ❖ Enriquecer la investigación con la búsqueda de nuevas herramientas para el trabajo posterior con megaontologías.
- ❖ Realizar una investigación sobre los procesos de inferencia en megaontologías, incluyendo las extensiones fuzzy.

Bibliografía

1. **Antonio Martín, Sonsoles Celestino, Adela Valdenebro, Julia Mensaque.** Asociación Andaluza de Bibliotecarios. [Online] [Cited: 01 5, 2010.] http://www.aab.es/pdfs/jba15/10-Ontologia_e_inteligencia_artificial.pdf.
2. **Casafont, Merce Lorente.** Hipertext.net. [Online] [Cited: 01 21, 2010.] <http://www.hipertext.net/web/pag259.htm>.
3. **Cedeño, Alberto Barrón.** [Online] [Cited: 01 28, 2010.] <http://www.matem.unam.mx/~rajsbaum/cursos/web/ontologias1.pdf>.
4. **Rafael Pedraza Jiménez, Lluís Codina, Cristòfol Rovira.** lluiscodina.com. [Online] [Cited: 02 4, 2010.] <http://www.lluiscodina.com/webSemanticaOntologias2007.pdf>.
5. **Andrea Delgado, Ignacio García Rodríguez de Guzmán.** Sociedad de Ingeniería del Software y Tecnologías de Desarrollo de Software . [Online] [Cited: 02 4, 2010.] <http://www.sistedes.es/TJISBD/Vol-3/No-3/articulos/pnis-09-delgado-ontologia.pdf>.
6. **Francisco Arias, Julián Moreno, Demetrio Ovalle.** Universidad Nacional de Colombia. [Online] [Cited: 02 5, 2010.] <http://pisis.unalmed.edu.co/avances/archivos/ediciones/Edicion%20Avances%202007%202/02.pdf>.
7. **Lea en Binario.** [Online] <http://leaenbinario.blogspot.com/2009/05/herramientas-para-trabajar-con.html>.
8. **Herramientas Gestion del Conocimiento - Herramientas Ontologias - Herramientas Tesauros.** [Online] [Cited: 02 15, 2010.] <http://leaenbinario.blogspot.com/2009/05/herramientas-para-trabajar-con.html>.
9. **Modelado de Colaboración uso de ontologías .** [Online] [Cited: 03 2, 2010.] <http://www.computers.com/es/consejos-509155.htm>.
10. **Tramullas, Dr. Jesús.** tramullas.co. [Online] [Cited: 03 4, 2010.] <http://tramullas.com/papers/isko99.pdf>.
11. **STI Innsbruck.** [Online] [Cited: 03 10, 2010.] <http://www.sti-innsbruck.at/fileadmin/documents/papers/47.pdf>.
12. **Díaz, Alicia.** Universidad Nacional de La Plata. [Online] [Cited: 03 10, 2010.] http://www.unlp.edu.ar/uploads/docs/tecnologias_para_la_web_semanticas___informatica.pdf.



13. **Graciela Barchini, Margarita Álvarez, Susana Herrera y Melina Trejo.** Departamento de Ingeniería Informática y Ciencias de la Computación. [Online] [Cited: 03 11, 2010.] <http://www.inf.udec.cl/~revista/ediciones/edicion14/barchini.pdf>.
14. **Aparicio Reytor, Maikel y Bravo Zalazar, Susana .** *Módulo de Inferencia para un Repositorio Semántico.* 2009.
15. **Samper Zapater, José Javier.** *ONTOLOGÍAS PARA SERVICIOS WEB SEMÁNTICOS DE INFORMACIÓN DE TRÁFICO: DESCRIPCIÓN Y HERRAMIENTAS DE EXPLOTACIÓN.* 2005.
16. **Pablo Velázquez Muñoz.** Unos métodos en el desarrollo de ontologías. 2010 titan.tel.uva.es/wikis/.../TrabajoIndividualFinal_PabloVelazquezMuñoz.pdf
17. Estudio Comparativo de Lenguajes para la Búsqueda y Recuperación de Información Semántica. [Online] 02 23, 2010. <http://www.sociedadelainformacion.com/15/lenguajes.pdf>.



Referencias Bibliográficas

1. **T.Berners-Lee, J.Hendler , O.Lassila.** The Semantic Web. Scientific American. [En línea] 2001.
2. **Castells.** La Web Semántica . [En línea] 2003. www.ii.uam.es/~castells/publications/castells-uclm03.pdf.
3. **Lopez, C.G.** Los Repositorios de Objetos de Aprendizaje como soporte a un entorno e-Learning, Tesina doctoral, Universidad de Salamanca. [En línea] 2005.
4. **Aparicio Reytor, Maikel y Bravo Zalazar, Susana .** *Módulo de Inferencia para un Repositorio Semántico.* 2009.
5. <http://www.mitecnologico.com/Main/ReglasDeInferencia>. . [En línea] [Citado el: 12 de 01 de 2010.]
6. [En línea] [Citado el: 12 de 01 de 2010.]
<http://www.utpl.edu.ec/semanticlab/http://www.utpl.edu.ec/semanticlab/mvagila/ontologias-un-elemento-clave-de-la-web-semantica/14>.
7. geneura.ugr.es. [En línea] [Citado el: 01 de 02 de 2010.]
<http://geneura.ugr.es/~maribel/xml/introduccion/index.shtml>.
8. manuales.dgsca.unam.mx . [En línea] 2010.
<http://manuales.dgsca.unam.mx/xml/Qu%E9%20es%20DTD.htm>.
9. www.mitecnologico.com. [En línea] [Citado el: 01 de 02 de 2010.]
<http://www.google.com/cu/search?hl=es&client=firefox-a&rls=org.mozilla%3Aes-ES%3Aofficial&q=xml+y+xml+schema+%2B+caracteristicas+&btnG=Buscar&meta=&aq=o&oq=>.
10. www.abcdatos.com. [En línea] 06 de 08 de 2009.
<http://www.abcdatos.com/webmasters/tutorial/v505.html>.
11. www.sidar.org. [En línea] 07 de 10 de 2007. <http://www.sidar.org/recur/desdi/traduc/es/rdf/rdfsched.htm>.
12. www.tucamon.es. [En línea] 10 de 11 de 2008. [Citado el: 28 de 01 de 2010.]
<http://www.tucamon.es/contenido/creacion-de-ontologias-con-protege>.
13. **Thomas, Gruber.** *A translation approach to portable ontology specifications.* Knowledge Acquisition. 1993.
14. **Carrera.C, S.** Adquisición semi-automática del conocimiento: una arquitectura preliminar. [En línea] 2007. <http://www.grupocole.org/cole/library/ps/Car2007a.pdf>.



15. **Gruber, Thomas.** *Towards Principles for the Design of Ontologies used for Knowledge Sharing.* *International Journal of Human-Computer Studies.* 1995.
16. **Vazquez, Cristian.** www.dcc.uchile.cl. [En línea] 12 de 2002.
<http://www.dcc.uchile.cl/~cvasquez/meta/lenguajes.html>.
17. **Samper Zapater, José Javier.** *ONTOLOGÍAS PARA SERVICIOS WEB SEMÁNTICOS DE INFORMACIÓN DE TRÁFICO: DESCRIPCIÓN Y HERRAMIENTAS DE EXPLOTACIÓN.* 2005.
18. **Flórez Fernández, Héctor Arturo.** *Construcción de Ontologías OWL.* Bogotá : s.n., 2008.
19. Barrón Cedeño, Alberto. *Ontologías en acción Protégé OWL.* (20 de septiembre del 2005).
<http://theory.lcs.mit.edu/~rajsbaum/cursos/web/ontologías2.pdf>.
20. www.scribd.com. [En línea] 2008. [Citado el: 5 de 02 de 2010.]
21. Abián, Miguel Ángel. *El Futuro de la Web XML, RDF/RDFS, Ontologías y la web Semántica.* Disponible en: <http://www.javahispano.org/licencias/>.
22. José A. Alonso. *Introducción a la programación funcional* creado 29/11/2007
http://www.cs.us.es/~jalonso/trabajos_dirigidos/2008-tesis-AP.pdf?searchterm=ontolog%C3%ADa
23. Pablo Velázquez Muños. *Unos métodos en el desarrollo de ontologías.* 2010
titan.tel.uva.es/wikis/.../TrabajoIndividualFinal_PabloVelazquezMuñoz.pdf
24. Anisleiby Fernández Hernández. *Ontologías nuevos retos.* Marzo 2009.
http://dialnet.unirioja.es/servlet/fichero_articulo?codigo=2924584&orden=0
25. (Por: Ing. Jorge A. Mendoza, Publicado: Junio 10 del 2003, e-Learning, el futuro de la educación a distancia, consultado: 10/2/10, <http://www.informaticamilenium.com.mx/paginas/mn/articulo78.htm>).
26. **Henst, Christian Van Der.** *E-learning. maestros de la web.* [En línea] 16 de 08 de 2002. [Citado el: 10 de 02 de 2010.]
27. **(Cabrera Pupo, Katuska y Martínez Gamboa, Aylén. 2007.** *Propuesta metodológica para la gestión del conocimiento basada en ontología.* Habana: UCI, 2007.)
28. [En línea] [Citado el: 11 de 02 de 2010.] <http://www.slideshare.net/lhuaroto/repositorio-digital>.
29. [En línea] [Citado el: 11 de 02 de 2010.]
<http://www.colombiaaprende.edu.co/html/directivos/1598/article-99543.html>.



30. [En línea] [Citado el: 11 de 02 de 2010.]
http://www.biblioweb.dgsca.unam.mx/libros/repositorios/bibliotecas_digitales.htm#421r
31. **Tamayo, Daymy y Querejeta, Andrey. 2005.** Plataforma para el desarrollo de e-learning en la Universidad de Holguín Oscar Lucero Moya. . *Plataforma para el desarrollo de e-learning en la Universidad de Holguín Oscar Lucero Moya.* . Holguín: Universidad de Holguín Oscar Lucero Moya, 2005.
32. **Leyva, David, y otros. 2006.** HERRAMIENTAS PARA LA CREACIÓN Y GESTIÓN DE OBJETOS DE APRENDIZAJE REUTILIZABLES. *Informática 2009.* [En línea] 06 de 2006. [Citado el: 12 de 02 de 2010.]
http://www.informaticahabana.com/evento_virtual/files/SWL08.pdf.
33. **Woodley, Mary S.** www.sedic.es. [Online] 09 15, 2003. http://www.sedic.es/glosario_DCMI.pdf.
34. Fusion de ontologías de metadatos, FAC, Onions, Prompt. [Online] [Cited: 02 18, 2010.]
telefonica.net/web2/fusionontologias/docs/fusionOntologias.doc.
35. <http://serqlsparql.50webs.com/>. [Online] 04 09, 2007. [Cited: 02 18, 2010.]
36. Estudio Comparativo de Lenguajes para la Búsqueda y Recuperación de Información Semántica. [Online] 02 23, 2010. <http://www.sociedadelainformacion.com/15/lenguajes.pdf>.
37. [Online] 2008. [Cited: 02 24, 2010.]
<http://www.balearsinnovacio.com/blog/wpcontent/uploads/2008/06/sparql.pdf>.
38. [Online] 7,03,2008. [Cited: 4,03, 2010.] <http://tramullas.com/documatica/3-1.html>
39. [Online] 13,02,2009. [Cited: 4,03, 2010.] http://www.mariapinto.es/e-coms/recu_infor.htm
40. [Cited: 24,03, 2010.] <http://websemantica.iespana.es/>



Anexos

Anexo 1.

Proceso de creación de la Megaontología	Lenguaje	Consistencia	Tiempo de carga	validar las especies	Verificar consistencia	Tiempo Total
Megaontologías						
Fusión	Lite	Si	1599 ms	85 ms	18 ms	1702 ms
Referencias por NameSpaces	Lite	Si	1597 ms	81 ms	17 ms	1695 ms
Referencias directas por URL	Lite	Si	1607 ms	100 ms	17 ms	1725 ms
Ontología Simple						
Historial	Lite	Si	1587 ms	82 ms	17 ms	1697 ms

Fig.1 tabla comparativa entre megaontología y ontología.

**Anexo 2.**

```
Classifying finished in 00:00:00
Classifying 16 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
Classifying 16 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
Classifying 16 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
Classifying 145 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
Classifying 15 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
Classifying 15 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
```

Consola
maikel : mc - KDE Terminal Emulator

Fig. 2. Pruebas interfaz INFERMOD Fusión.**Anexo 3.**

```
Classifying finished in 00:00:00
Classifying 16 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
Classifying 16 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
Classifying 16 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
Classifying 145 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
Classifying 15 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
Classifying 15 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
```

Consola
maikel : mc - KDE Terminal Emulator

Fig. 3. Pruebas interfaz INFERMOD Referencias a través de NameSpaces.



Anexo 4.

```
modularity
query
realize
trans-tree
unsat

Pellet is a tool for managing and reasoning with OWL ontologies.
For more information, see http://clarkparsia.com/pellet
root@lekiam:/media/disk/Maikel/Herramientas/pellet-2.0.0-rc5/lib# java -jar pellet-cli.jar dig
INFO [main] (HttpServer.java:729) - Version Jetty/5.1.5rc1
INFO [main] (Container.java:74) - Started HttpContext[/,/]
INFO [main] (SocketListener.java:204) - Started SocketListener on 0.0.0.0:8081
INFO [main] (Container.java:74) - Started org.mortbay.http.HttpServer@a61164

PelletDIGServer Version 2.0.0-rc5 (March 3 2009)
Port: 8081
Classifying 16 elements
Classifying: 100% complete in 00:00:00
Classifying finished in 00:00:00
█

+ Consola rgada: .....
maikel : mc - KDE Terminal Emulator
```

Fig. 4. Pruebas interfaz INFERMOD Referencias directas por URL.



Glosario de Términos

Web Semántica: La Web Semántica es una Web extendida y dotada de mayor significado, apoyada en lenguajes universales, que van a permitir que los usuarios puedan encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida.

Semántica: La Semántica se refiere a todo aquello que está vinculado o pertenece a la significación de las palabras. La misma está asociada al significado, interpretación y sentido de las palabras, de los símbolos y expresiones. Por ello, es que Semántica se llama también a la parte de la Lingüística que se ocupa justamente de estudiar el significado de los signos lingüísticos y sus combinaciones.

Metalinguaje: En lógica y filosofía del lenguaje, un metalinguaje es un lenguaje que se usa para hablar acerca de otro lenguaje. Al lenguaje acerca del cual se está hablando se lo llama el lenguaje objeto. El metalinguaje puede ser idéntico al lenguaje objeto, por ejemplo cuando se habla acerca del español usando el español mismo. Un metalinguaje a la vez puede ser el lenguaje objeto de otro metalinguaje de orden superior, y así sucesivamente. Distintos metalinguajes pueden hablar acerca de diferentes aspectos de un mismo lenguaje objeto.

Estándares de representación: XML, XML Schema, RDF, RDF Schema y OWL.

Ontología: Estructura jerárquica que define formalmente las relaciones semánticas de un conjunto de conceptos. Se usa para crear vocabularios controlados/estructurados para la recuperación o el intercambio de información.

Objetos de aprendizaje: materiales digitales creados como pequeñas piezas de contenido o de información, con la finalidad de maximizar el número de situaciones educativas en que las que el recurso pueda ser utilizado.



Repositorio: Lugar físico donde se almacenan los documentos en forma digital.

WWW: abreviatura de World Wide Web, es básicamente un medio de comunicación de texto, gráficos y otros objetos multimedia a través de Internet, es decir, la web es un sistema de hipertexto que utiliza Internet como su mecanismo de transporte o desde otro punto de vista, una forma gráfica de explorar Internet.

W3C: son las siglas de World Wide Web Consortium, un consorcio fundado en 1994 para dirigir a la Web hacia su pleno potencial mediante el desarrollo de protocolos comunes que promuevan su evolución y aseguren su interoperabilidad. El consorcio está compuesto por un grupo de programadores, desarrolladores web, ejecutivos de la industria y usuarios que ayudan a definir las especificaciones para el desarrollo de la tecnología web.

API (Application Program Interface): es un conjunto de convenciones internacionales que definen cómo debe invocarse una determinada función de un programa desde una aplicación. Cuando se intenta estandarizar una plataforma, se estipulan unas APIs comunes a los que deben ajustarse todos los desarrolladores de aplicaciones.