

Universidad de las Ciencias Informáticas
Facultad 10



**PROPUESTA METODOLÓGICA PARA LA GESTIÓN
DE CONOCIMIENTO BASADA EN ONTOLOGÍAS.**

**TRABAJO DE DIPLOMA PARA OPTAR POR EL TÍTULO DE
INGENIERO EN CIENCIAS INFORMÁTICAS**

Autores

Aylen Martínez Gamboa.
Katuska Cabrera Pupo.

Tutores

Ing. Daymy Tamayo Ávila
MSc. David Leyva Leyva

Ciudad de La Habana
Julio 2007

Haz primero lo que es necesario, luego lo que es posible y de pronto te encontrarás haciendo lo imposible.

San Francisco de Asís.

Declaración de Autoría.

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 firmo la presente a los ____ días del mes de _____ del año _____.

Aylen Martínez Gamboa.
Firma de la Autora.

Katuska Cabrera Pupo.
Firma de la Autora.

Ing. Daymy Tamayo Ávila.
Firma de la Tutora

MSc. David Leyva Leyva.
Firma del Tutor

Datos de Contacto.

MSc. David Leyva Leyva: Vicedecano de Producción e Investigación, Facultad 10, Universidad de Ciencias Informáticas. Teléfono: (53) (07) 837 2503. davidl@uci.cu
Licenciado en Cibernética-Matemática, Universidad de Las Villas (UCLV), 1988. Master en Computación Aplicada, Universidad de Las Villas (UCLV), 1995. Profesor de la Universidad de Holguín desde 1990. Jefe de Departamento de Informática (2003-2005). Participó en un Proyecto de Educación a Distancia en la Universidad del 2001 hasta el 2005, año en el que comenzó a trabajar en la Universidad de Ciencias Informáticas. 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). Actualmente es miembro de un grupo de investigación y desarrollo orientado a la extensión de la plataforma Moodle y el desarrollo de otras herramientas para la teleformación.

Ing. Daymy Tamayo Ávila: Especialista de la Dirección de Teleformación, Profesora de Ingeniería de Software Facultad X, Universidad de Ciencias Informáticas. Carretera a San Antonio de los Baños Km. 2 ½, Torrens, Boyeros, Ciudad de La Habana. Cuba. Teléfono: (53) (07) 837 2453. daymy@uci.cu Ingeniera Informática, Universidad de Holguín, 2005. Culminó sus estudios con índice académico de 4.94 puntos, obteniendo Título de Oro. Fue Alumna de Alto Aprovechamiento Académico y Alumna Ayudante en la Disciplina de Técnicas de Programación de Computadoras desde el segundo año de su carrera. Participó en un Proyecto de Educación a Distancia en la Universidad del 2001 hasta el 2005, año en el que comenzó a trabajar en la Universidad de Ciencias Informáticas. Ha participado en un gran número de eventos nacionales e internacionales. Ha sido tutora de varios Trabajos de Curso y Diploma, y de Alumnos Ayudantes. Actualmente es líder de un grupo de investigación y desarrollo orientado al desarrollo y soporte de herramientas para la teleformación.

Agradecimientos

Para el creador de este proyecto nuestro Comandante en Jefe por su eterno ejemplo Muchas Gracias. A mis padres quienes con su ejemplo me enseñaron ha ser quien soy, sin ellos este sueño no sería una realidad. A mí hermano por ser mi eterno confidente y mi amigo en todo momento. A mí cuñada por enseñarme el valor que tienen los sueños. A mis sobrinos, quienes con su ternura e inocencia iluminaron todas estas noches de ausencia. A mi novio por estar a mi lado cuando parecía que el mundo se derrumbaba sobre mí. A toda mi familia por estar siempre a mi lado y hacerme sentir especial. A las hermanas maravillosas que el preuniversitario me regaló y de las que nunca me podré separar. A todo el equipo de desarrollo EROS. A Lachi, Marta, Nilda y familia por su amistad y acogerme como una hija más. A toda la familia de Aylen y Dianly, a los profesores, los cómplices del aula, laboratorios, los compañeros de trabajo de mis padres, mis vecinos y amigos. Muchas gracias queridos amigos.

Katiuska Cabrera Pupo.

Ante todo a mi familia, en especial a mis padres por creer en mí, por apoyarme en todo momento, por brindarme todo el amor del mundo.

A la revolución, a Fidel. A nuestros tutores David Leyva Leyva y Daymy Tamayo Ávila por tener tanta paciencia y dedicación en nuestro trabajo. A todas las personas que de una forma u otra apoyaron nuestra idea, a todos los profesores, a todos los amigos. Gracias por hacer de este sueño una realidad.

Aylen Martínez Gamboa.

Queremos agradecer especialmente a Maikel Aparicio y David Escalona por todo el esfuerzo y dedicación con el que contribuyeron a la realización de este trabajo. A nuestros tutores por la confianza depositada, su constancia y por adentrarnos en este tema tan maravilloso.

Resumen.

En la actualidad la tecnología avanza vertiginosamente y dentro de este desarrollo Internet ha sido un resultado muy importante debido a las ventajas informativas que trae consigo. Aprovechando este progreso y las necesidades que tienen las empresas surgen sistemas para la Gestión del Conocimiento. Los métodos de búsqueda utilizados en estos sistemas no son eficientes debido a los problemas que presenta la Web Actual, los cuales son causados principalmente por la existencia de mucha información y la falta de organización de la misma. Una posible solución a dichos problemas es la Web Semántica. El componente fundamental que utiliza esta web es la Ontología, esta le aporta la semántica a su contenido, factor que favorece a la Gestión del Conocimiento.

En este trabajo se hace una propuesta para gestionar el conocimiento utilizando Ontologías. Primeramente se analizan algunas de las metodologías existentes para construir Ontologías y se propone una serie de pasos con este fin, luego se especifican los elementos necesarios para la fase de inferencia y por último se plantea cómo lograr que interactúe el usuario con el sistema. Esta propuesta se pone en práctica con una versión previa de la implementación de un prototipo de repositorio de objetos de aprendizaje semántico, en donde se define como dominio de la Ontología los estándares educativos y se emplean como herramientas: el lenguaje OWL, Protégé como editor, para visualizar el contenido de la Ontología se aprovecha Jena, el razonador pellet y por último para hacer consultas se utiliza SPARQL.

Palabras Claves: Gestión del Conocimiento, Web Semántica, Ontologías, Estándares Educativos, Metodología, Razonadores y Editores de Ontologías

Índice

Agradecimientos	I
Resumen.....	II
Introducción	1
Capítulo 1. Fundamentos Teóricos	7
1.1. Introducción al Capítulo	7
1.2. Gestión del Conocimiento.....	7
1.3. La Web Actual.....	10
1.3.1. Surgimiento de la Web Semántica.....	14
1.3.2. Componentes de la Web Semántica.....	16
1.4. Ontologías.....	17
1.4.1. Tipos de Ontologías.....	18
1.4.2. Componentes de una Ontología.....	19
1.4.3. Aplicaciones y beneficios.....	20
1.4.4. Lenguajes para representar Ontologías.....	22
1.4.5. Herramientas para Construir Ontologías.....	25
1.4.6. Sistemas de Almacenamiento.....	26
1.4.7. Razonadores de Ontologías.....	27
1.5. e-Learning.....	29
1.5.1. Objetos de Aprendizaje.....	30
1.5.2. Los metadatos.....	31
1.6. Estándares Educativos.....	33
1.6.1 Estándares de e-Learning.....	35
1.6.2. Ventajas del uso de estándares.....	38

1.7. Los Repositorios.	39
1.7.1. Repositorios de Objetos de Aprendizaje.....	39
1.7.2. Repositorios en la Web.	40
1.7.3. Caso Cuba.	42
1.7.4. Repositorio Semántico.	44
1.8. Conclusiones del Capítulo.	45
Capítulo 2. Propuesta Metodológica.....	47
2.1. Introducción al Capítulo	47
2.2. Primera Fase: Proceso de Construcción de la Ontología.	47
2.3. Segunda Fase: Inferencia.....	52
2.4. Tercera Fase: Interacción entre el Usuario y el Sistema.	55
2.5. Conclusiones del Capítulo.	62
Capítulo 3. Caso Práctico.	63
3.1. Introducción al Capítulo.	63
3.2. Primera Fase: Proceso de Construcción de la Ontología.....	63
3.3. Segunda Fase: Inferencia.....	70
3.4. Tercera Fase: Interacción entre el Usuario y el Sistema.	77
3.5. Conclusiones del Capítulo.	78
Conclusiones.	79
Recomendaciones.	81
Referencias Bibliográficas.	82
Bibliografía.	87
Anexos.	92
Glosario de Términos.....	99

Índice de Figuras.

Figura 1: Ejemplo de una propiedad utilizando Protégé.	50
Figure 2 Ejemplo de una instancia.....	51
Figura 3: Jerarquía de clases sobre estándares educativos desarrollada en Protégé.	68
Figura 4: Ejemplo de Propiedades de una clase.	69
Figure 5: Diferencia entre la Web Actual y la Web Semántica.	92
Figura 6: Estructura de la Web Semántica.	93
Figura 7: Usando Ontologías como metadatos.....	93
Figura 8: Estructura del metadato de LOM.....	94
Figura 9: Comparación entre metodologías existentes para construir Ontologías.	95
Figura 10: Ejemplo de Tesauro.....	96
Figura 11: Representación de la web actual usando Ontologías.....	97
Figura12. Persona registrada en varias comunidades con una misma cuenta de acceso.	97
Figura 13: Funcionamiento del sistema.	98
Figura 14: Estadísticas sobre el crecimiento de internet.	98

Introducción

En estos últimos años la *tecnología* ha avanzado vertiginosamente y con ello la forma de manejar la *información*, llegando a ser casi imposible que una empresa no haga uso de estos avances para la realización de sus actividades cotidianas.

En todo este progreso tecnológico ha sido *Internet*^{*} un resultado importante debido a las ventajas informativas que trae consigo, ya que se logra contar con más información, la cual se encuentra disponible las 24 horas del día sin tener que trasladarnos a lugares distantes para encontrarla [CICEI 1995]. Internet viene siendo una gran biblioteca, con millones de documentos esperando ser consultados a través de la Web [Merlo 1998], la cual permite realizar diferentes actividades de una manera cómoda, eficiente y económica.

El surgimiento de Internet ha hecho posible que exista mucha información acerca de distintos temas y que la misma sea accesible desde cualquier parte del mundo, provocando que las personas acudan a este medio para obtener y compartir recursos.

Al existir pérdidas de información en las empresas y aprovechando este avance tecnológico, surgen diferentes sistemas para la gestión de conocimiento, la cual consiste en asegurarse de que la información que posee cada uno de los empleados de las empresas sea codificada y almacenada en algún soporte electrónico. De esta manera los datos se encuentran accesibles dentro de una red local a la cual tienen acceso todos los trabajadores de la misma, evitándose la posibilidad de perder cualquier información valiosa de la empresa.

El crecimiento de la red de redes como también se le conoce, se hace cada día mayor (Ver Anexo 10). Y aunque es sumamente difícil medir el tamaño de la Web, ya en el 2001

***Internet** es una red mundial de computadoras formada por la interconexión de otras redes menores. [Wikipedia, 2007].

un estudio [Bergman 2001] señalaba que “unos 10^9 usuarios utilizan la web, y que ésta contiene del orden de $4 \cdot 10^9$ documentos, un volumen de información equivalente a entre 14 y 28 millones de libros.

Como dato comparativo, la asociación American Research Libraries, que agrupa unas 100 bibliotecas en EE.UU., tiene catalogados unos 3.7 millones de libros. La biblioteca de la Universidad de Harvard, la mayor de EE.UU., contiene en torno a 15 millones de libros. Estas cifras incluyen sólo lo que se ha dado en denominar la web superficial, formada por los documentos estáticos accesibles en la web.” [Castells]

La Web actual a pesar de haber alcanzado un gran éxito presenta algunas limitaciones, las cuales si se pudiesen mejorar permitirían un mayor grado de aceptación por parte de los usuarios. La existencia de mucha información en la web y la falta de organización de la misma dificulta el trabajo a la hora de realizar alguna búsqueda sobre un tema determinado, pues en ocasiones se busca algo sobre un tema y se obtiene como resultado documentos cuyo contenido no guarda relación alguna con lo que se quiere encontrar o se encuentra lo que se buscaba sólo después de visitar varias páginas.

Muchas veces ocurre que el término utilizado para realizar nuestra búsqueda posee disímiles significados, arrojando resultados ajenos al deseado, pues se obtienen páginas Web relacionadas con sus otras acepciones.

La Universidad de las Ciencias Informáticas (UCI), no está ajena a este problema, debido a la gran cantidad de información y recursos tecnológicos con los que se trabaja a diario. Los estudiantes de esta institución en su quehacer cotidiano acuden a Internet para realizar búsquedas que complementen la bibliografía relacionada con los temas tratados en clases y para su superación personal. Además algunos proyectos productivos cuentan con un *servidor* local en donde se almacena información referente a dicho proyecto, la cual es consultada por los trabajadores del mismo cuando la necesiten. La universidad cuenta con una *intranet* que permite acceder a Internet y a la biblioteca digital del centro. Una posible solución a estos problemas que presenta la web actual es la *Web Semántica* (WS), o Web del futuro, la cual no es más que “Una visión de una futura Web en la cual

la información que en la versión actual de la Web es comprensible solamente por los seres humanos también esté disponible de una manera formal para sistemas inteligentes". [Sánchez 2005]

Esta Web vendría siendo la solución más factible a los problemas que presenta la Web actual debido a que posibilita que los investigadores de cualquier disciplina puedan compartir sus conocimientos, permite el procesamiento automático de la información, se mejora el comercio electrónico y la búsqueda de información de manera eficaz y precisa, factores que contribuirían a un mejor funcionamiento de los sistemas para la gestión del conocimiento [Abián 2005], tema que será abordado más adelante.

Pero el uso de esta nueva Web no resolvería todo el problema pues para ello es necesario que la información que se guarda en los *repositorios*^{*} se encuentre estructurada de forma tal que facilite el razonamiento automatizado de la misma. Una forma de estructurar la información es mediante *Ontologías*. Esta palabra viene del griego *ontos* "el ser" y *logos* "estudio de". [Diccionario Océano]

Según el diccionario de la Real Academia de la Lengua Española (DRAE) el término Ontología es la "Parte de la metafísica que trata del ser en general y de sus propiedades transcendentales".

La Ontología tiene varias acepciones según la ciencia en la que se emplea, en filosofía, la ontología "es el estudio de lo que es en tanto que es y existe. Por ello es llamada la teoría del ser, es decir, el estudio de todo lo que es: qué es, cómo es y cómo es posible. La Ontología se ocupa de la definición del ser y de establecer las categorías fundamentales o modos generales de ser de las cosas a partir del estudio de sus propiedades". [Ontología 2007]

"Las Ontologías son para la Inteligencia Artificial recursos construidos que permiten

^{*}Un **repositorio** es un sitio centralizado donde se almacena y mantiene información, habitualmente bases de datos o archivos informáticos" [Wikipedia 2007].

representar el conocimiento compartido y común sobre algo". [Arano 2005]. En la rama de la Informática: *"Una Ontología define los términos a utilizar para describir y representar un área de conocimiento. Las Ontologías son utilizadas por las personas, las bases de datos, y las aplicaciones que necesitan compartir un dominio de información (un dominio es simplemente un área de temática específica o un área de conocimiento, tales como medicina, fabricación de herramientas [...], etc.). Las Ontologías incluyen definiciones de conceptos básicos del dominio, y las relaciones entre ellos, que son útiles para los ordenadores*". [Hendler 2004]

Las Ontologías permiten la representación del conocimiento que necesita la WS, definiendo formalmente los conceptos de los diferentes dominios y sus relaciones, con capacidad para realizar deducciones con este conocimiento. [Lozano]

Debido a las deficiencias que presenta la Web Actual y al surgimiento de la WS como solución a los problemas mencionados anteriormente se hace necesario el estudio y profundización de este tema en aras de contribuir a su futura realización. La presencia de una Ontología con un dominio estructurado y relacionado con la información de acuerdo a su significado permite construir un *buscador*^{*} donde los usuarios especifican sus criterios de búsquedas en función de los conceptos y atributos.

A diferencia de buscadores tradicionales los buscadores semánticos[†] devuelven solamente la información solicitada [Castells 2006]. Se hace necesario que los sistemas informáticos vayan más allá de una simple lectura de las letras que componen las palabras y que intenten en cierto modo descifrarlas factor que puede ser aprovechado para la gestión del conocimiento. [Réjean 1999]

* Un **buscador** es un programa, ubicado en un sitio de Internet, que recibe un pedido de búsqueda, lo compara con las entradas de su base de datos y devuelve el resultado. [Clarín]

† Un **buscador semántico** entiende el significado de la consulta y al buscar la información, esta se encuentra organizada por metadatos, lo que permite devolver resultados más eficientes.

De lo analizado anteriormente llegamos a plantearnos como **problema científico**:
¿Cómo lograr una eficiente gestión del conocimiento en contextos con información estructurada?

Definimos como **objeto de estudio** la Gestión del Conocimiento en la Web Semántica y enmarcando como **campo de acción** el *proceso* de búsqueda *semántica* en dominios específicos.

Es por ello que el **objetivo general** de este trabajo, es hacer una propuesta metodológica para gestionar el conocimiento basada en Ontologías y su utilización en diferentes contextos.

Preguntas Científicas:

- I- ¿Cuáles conceptos, tecnologías y herramientas permiten gestionar conocimiento en ambientes Web?
- II- ¿Cómo recuperar conocimiento y hacer inferencias en ambientes Web semánticamente estructurados?
- III- ¿Podrá ser elaborado y aplicado un procedimiento basado en Ontologías para la gestión de conocimiento en contextos con información estructurada?

Para responder a estas preguntas se proponen las siguientes tareas:

1. Estudiar cómo representar conocimiento con potencialidad de Web Semántica y cómo recuperarlo en ambientes de información estructurada.
2. Analizar detalladamente los componentes de la Web Semántica y el uso de lenguajes de especificación de Ontologías.
3. Estudiar el proceso de implementación de Ontologías para dominios específicos.
4. Examinar las diferentes herramientas que sirven para construir Ontologías, interactuar y realizar inferencias sobre ellas.
5. Investigar qué mecanismos son empleados para la recuperación de la información en la Web Semántica.
6. Elaborar un procedimiento genérico para la gestión del conocimiento basado en Ontologías y que además haga uso de las herramientas adecuadas.

Para la realización de este proyecto investigativo es necesario la combinación de los métodos teóricos y empíricos. Dentro de los teóricos, el Análisis y Síntesis pues se analizan los conceptos asociados a la Web Semántica y su comprensión permite hacer la integración de elementos diversos en una propuesta metodológica para la gestión del conocimiento. De los empíricos utilizamos la Observación sistemática para conocer cómo funcionan los sistemas en uso que implementan ideas de la Web Semántica y cuál es su tendencia de desarrollo. La revisión de documentos se utiliza como una técnica subordinada al método de análisis y síntesis para comprender los conceptos fundamentales y el funcionamiento de las aplicaciones documentadas en la bibliografía consultada.

El presente trabajo consta de introducción, tres capítulos, conclusiones, recomendaciones, bibliografía, referencias bibliográficas, glosario de términos y anexos.

En el Capítulo 1, Fundamentos Teóricos, se abordan ampliamente algunos aspectos necesarios para nuestra investigación, se explica qué es la Gestión de Conocimiento, se estudian las herramientas que se emplean para construir la Web Semántica, así como los aspectos relacionados con el dominio en que se piensa aplicar.

En el Capítulo 2, Propuesta Metodológica, se hace referencia a las etapas que componen esta propuesta. La primera es el proceso de construcción de la Ontología, para ello inicialmente se realiza una observación sobre algunas metodologías existentes para construir Ontologías, la segunda, es la etapa de inferencia y en la última se propone la interacción entre el usuario y el sistema.

En el Capítulo 3, Caso Práctico, se pone en práctica la propuesta, con una versión previa de la implementación de un prototipo de repositorio de *objetos de aprendizaje* semántico.

Capítulo 1. Fundamentos Teóricos

1.1. Introducción al Capítulo

En este capítulo se explican los elementos en que se basará la propuesta metodológica, se empezará por conceptualizar la Gestión del Conocimiento (GC) y por definir lo que se denomina Repositorio de Objetos de Aprendizaje. Además se verá cómo está compuesta la Web Semántica, los problemas que se resolverían con su uso. Una breve referencia a la Web 3.0, así como las Ontologías, clasificaciones y herramientas para su construcción.

1.2. Gestión del Conocimiento.

El conocimiento está en las mentes de las personas, por tanto, no siempre está disponible cuando es necesario para la organización. Para tratar este problema surge la GC cuya principal misión es crear un ambiente en el que el conocimiento y la información disponible en una organización sean accesibles a todos, permitiendo ser usados para estimular la innovación y mejorar las decisiones [Simón 2006].

La 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 2006]

Cuando se hace referencia a la GC no se está hablando de gestión de la información (GI). *“El objetivo básico de esta última es organizar y usar los recursos de información de la organización para permitirle operar, aprender y adaptarse a los cambios del ambiente. La GI constituye el fundamento de la fase inicial de la gestión del conocimiento”*. [Rojas 2006]

Resulta importante aclarar que existen diferencias entre conocimiento, datos e información. Los datos son una representación simbólica* que por sí solo no tiene un

* Cuando se menciona una **representación simbólica** se refiere que pueden ser números, letras, etc.

valor semántico y constituyen la base de la pirámide del conocimiento. Al conjunto de datos organizados y analizados en un contexto determinado se le denomina información, que no es lo mismo que conocimiento. Recopilar y organizar datos, es algo que puede hacer un software informático. El conocimiento es un paso adelante, es identificar, estructurar y sobre todo utilizar la información para obtener un resultado. Requiere aplicar la intuición y la sabiduría, propios de la persona a la información. La capacidad de interpretar esos datos es lo que provoca que la información se convierta en conocimiento. [Figuroa 2006]

Las personas depositan, representan y graban conocimiento en objetos que emplean diariamente. Por lo tanto, el conocimiento es lo que se tiene en la mente más el de los objetos que se manipulan, que pueden ser físicos o no físicos como son las bases de datos. [Canals 2003]

En una organización, aparte de las personas y de los objetos que se utilizan, también existen otras cosas que contienen conocimiento, como por ejemplo los procesos. Un *proceso* es lo que permite tener establecidas ciertas maneras de hacer las cosas, estrategias que a lo largo del tiempo se ha comprobado que funcionan y que ahorran el trabajo de tener que repensarlas cada vez que se hace lo mismo. [Canals 2003]

En la GC hay dos procesos fundamentales, la creación y la transmisión de conocimiento. La transmisión se puede ver cuando se intenta poner de forma explícita el conocimiento en una base de datos, lo que se hace es ponerlo allí para que al cabo del tiempo alguien pueda recogerlo, de esta forma se está transmitiendo en el tiempo. Cuando se utilizan herramientas de comunicación se intenta transmitir el conocimiento en el espacio. Algunos de estos instrumentos son las intranets, los portales, las bases de datos relacionales y las documentales. [Canals 2003]

Estos dos procesos que pueden pensarse por separado también están totalmente interrelacionados, pues la creación de conocimiento no es algo que se hace partiendo de la nada, sino que para crear conocimiento se utiliza conocimiento que proviene de otras

personas y de otros lugares, por lo tanto, ha habido un proceso de transmisión previo. Son procesos que están muy interrelacionados y que juntos hacen que el conocimiento dentro de las organizaciones mejore y se utilice. Otro elemento importante que ayuda a hacer que el conocimiento funcione es el contexto, este va a permitir interpretar el conocimiento y transmitirlo. [Canals 2003]

Existen dos tipos de conocimiento: explícito y tácito, el primero se ha definido como la noción objetiva y racional que puede ser expresada con palabras, números, fórmulas, etc. El segundo es aquel que una persona, comunidad, organización o país, tiene incorporado o almacenado en su mente, en su cultura y es difícil de explicar. Es necesario explicar que este conocimiento puede estar compuesto por: ideas, experiencias, destrezas, habilidades, costumbres, valores, historias y creencias. [Figuroa 2006]

Cuando estos conocimientos permiten actuar se llaman competencias o conocimiento en acción. El problema que presenta este tipo de conocimiento es su dificultad a la hora de transmitirlo, por ello es necesario gestionarlo creando códigos que faciliten su transmisión. [Figuroa 2006]

Los mapas conceptuales (MC) facilitan la organización y representación del conocimiento de forma gráfica, formando una red de conceptos, en la que los nodos representan los conceptos y los enlaces las relaciones entre ellos.

Las siguientes herramientas son ejemplos de *plataformas* para la GC que emplean MC:

- ❖ GECOSOFT, incluye mecanismos para importar y exportar conocimiento hacia Ontologías, las que pueden ser usadas por otros sistemas.
- ❖ SERVIMAP, hace función de servidor. La información se almacena de forma clasificada.
- ❖ SMART IDEA, es utilizado para la construcción de MC [Simón 2006] más rápido, constante y eficaz [Gracia 1999].

- ❖ open-EIS: Sistema multiplataforma que posibilita la construcción de una infraestructura óptima para la administración de informaciones no estructuradas. La información es recogida una sola vez y puede ser reutilizada en los diferentes módulos como *e-Learning*, gestión de contenidos o gestión documental [Compañía Community4you].
- ❖ CMAP TOOLS, permite generar MC en diferentes formatos, fácil de utilizar, presenta una ayuda bastante ilustrada y sencilla, llena de ejemplos. Es hasta el momento el software más completo dentro del área de los MC.
- ❖ MOCOSOFT, fue desarrollada en la CUJAE, Cuba. Aunque se queda por debajo de CMAP TOOLS brinda varias funcionalidades, permite gráficamente dibujar mapas, seleccionar nodos y a partir de ellos la relación con otros nodos. Se pueden definir distintos tipos de nodos, crear un MC a partir de uno que ya esté hecho y permite guardarlos en diversos formatos.

Existen herramientas, plataformas y servidores para potenciar una adecuada GC, sistemas a los cuales se les perfecciona el diseño, los gráficos, se le adicionan funcionalidades pero los métodos de búsqueda en los repositorios siguen siendo los mismos y tienden a colapsar [Lucas 2006]. Por otro lado, la incapacidad de las máquinas para comprender el lenguaje natural las hace inflexibles.

1.3. La Web Actual.

Mundialmente, miles de personas y empresas acuden diariamente a la Web con el fin de encontrar o publicar información, comercializar con diferentes personas o para realizar tareas que forman parte de sus actividades cotidianas. Esto ha sido producto a la gran expansión que ha tenido la World Wide Web (WWW), de la cual se puede decir que surge en 1989 cuando Tim Berners-Lee su creador, presenta este proyecto en el Instituto Europeo de Investigación de la Física de Partículas (CERN) en Ginebra, Suiza, aunque no fue hasta 1990 que recibe su nombre definitivo.

Su objetivo inicialmente era compartir los trabajos que realizaban los físicos de este instituto a través de un entorno de hipertexto (HTML) y basado en arquitectura cliente-servidor (HTTP). [WWW 1996]. En 1993 con el lanzamiento de Mosaic, el primer navegador de dominio público, compatible con Unix, Windows, y Macintosh, por el National Center for Supercomputing Applications (NCSA), marca el momento en que la WWW se da a conocer al mundo, desplegándose primero en universidades y laboratorios, y en cuestión de meses al público en general, iniciando así lo que sería su vertiginoso crecimiento. Los usuarios acogieron con entusiasmo la facilidad con que se podían integrar texto y gráficos y saltar de un punto a otro del mundo en una misma interfaz y la extrema sencillez para aportar contenidos a una web mundial.

En 1994 algunos de los miembros que desarrollaron Mosaic desarrollan Netscape, otro navegador con algunas mejoras que impulsan la programación en la web, en ese mismo año se constituye el consorcio W3C el cual es presidido por Tim Berners-Lee. Con la aparición de lenguajes como JavaScript y otros, las páginas web contienen programas enteros, permitiendo una mayor eficiencia, capacidad dinámica y capacidad de interacción entre las páginas. [Castells]

La web actual ha alcanzado un avance tecnológico significativo, el número de usuarios se incrementa diariamente en grandes cantidades resultando difícil conocer su tamaño exacto. La extensión que ha alcanzado la web ha sido sin lugar a dudas su clave para alcanzar el éxito pero a la vez ha venido siendo un problema pues la información se encuentra desorganizada, los contenidos de las páginas web se encuentran en formatos e interfaces comprensibles para las personas pero no para las máquinas [Castells], a este problema se le conoce como carencia de *interoperabilidad** completa de los sistemas de información (SI). *“La web no facilita la creación de una comprensión común y compartida de un dominio de forma que esta pueda ser usada por organismos,*

* **Interoperabilidad** *“es la capacidad de un sistema o un producto para trabajar con otros sistemas o productos sin especial esfuerzo por parte del cliente”* [Abián 2005].

personas y máquinas”[Abián 2005]. Según la IEEE* se establecen 3 tipos de interoperabilidad.

“Interoperabilidad Técnica: Se refiere a la capacidad de los SI para intercambiar señales. Exige de una conexión física (cable, fibra óptica, ondas electromagnéticas) entre los sistemas y un conjunto de protocolos de comunicaciones (como la pila TCP/IP). Si dos ordenadores no pueden intercambiar señales, no puede existir interoperabilidad alguna... Para intercambiar datos se precisa que los SI usen el mismo formato para el intercambio de datos.... ”

“Interoperabilidad sintáctica: Se refiere a la capacidad de los SI para leer datos procedentes de otros SI y obtener una representación que puedan usar. Significa que en todos los SI involucrados, tanto la codificación de los datos como los protocolos de acceso son compatibles.”

“Interoperabilidad Semántica: Es la capacidad de los SI para intercambiar información basándose en un común significado de los términos y expresiones que usan [...]. La interoperabilidad semántica no debe confundirse con la sintáctica: esta se refiere a la dificultad de procesar automáticamente los datos, no a su contenido. Aquella se refiere a la dificultad de relacionar los términos desconocidos que aparecen en los datos con los ya conocidos o definidos. Esta interoperabilidad no puede existir antes de la interoperabilidad técnica y la sintáctica“. [Abián 2005]

El principal problema que presenta la web actual y del cual son víctimas la mayoría de los usuarios es la dificultad para encontrar información, actualmente el método más utilizado para esto es el uso de motores de búsqueda como Google o Yahoo. En los buscadores el usuario introduce en un cuadro de texto algunas palabras sobre lo que quiere encontrar y luego obtiene como resultado páginas web o documentos que

* IEEE “Institute of Electrical and Electronics Engineers, una asociación técnico-profesional mundial dedicada a la estandarización”. [Wikipedia 2007]

guardan relación con los términos introducidos en el buscador [Sánchez 2005]. Para mejorar estos resultados se utilizan técnicas de recuperación de la Información.

“La recuperación de la información (RI) consiste en buscar documentos que exhiban un mayor parecido a la pregunta formulada. Se puede definir el objetivo de la recuperación como la identificación de una o más referencias de páginas web que resulten relevantes para satisfacer una necesidad de información”. [Delgado 1998]

Los buscadores actuales son relativamente buenos aunque presentan determinados problemas a la hora de mostrar sus resultados debido a que se buscan las palabras claves en los documentos existentes en la web sin atender a la sinonimia o la polisemia, esto representa una molestia para el cliente que en muchas ocasiones tendrá que consultar decenas de páginas para encontrar la información deseada y en algunos casos puede que no encuentre nada de lo que buscaba [Abián 2005]. La polisemia se puede apreciar cuando *“realizamos una consulta en el buscador introduciendo un término y encontramos páginas web que contienen tal término pero con un significado distinto de aquello que nos interesa”*, la sinonimia cuando *“realizamos una consulta en el buscador introduciendo un término y no encontramos páginas web relacionadas con aquello que nos interesa porque tales páginas web contienen un sinónimo del término (pero no el propio término)”* y la última barrera de la Web es el idioma. A ésta se le conoce como Multilingüismo y se puede apreciar siempre que *“realizamos una consulta en el buscador introduciendo términos en inglés y nos encontramos páginas Web relacionadas con lo que nos interesa pero escritas en español (o cualquier otro lenguaje)”*. [Sánchez 2005]

Si la web tuviese asociadas anotaciones formales que identificasen inequívocamente los conceptos y entidades principales contenidos en las páginas, un buscador podría evitar cometer los errores descritos anteriormente y localizar sólo los documentos relacionados exactamente con lo que se quería encontrar [Sánchez 2005] , además sería bueno si se pudiesen automatizar algunas tareas que en la actualidad resultan engorrosas tales como la organización, actualización y la inserción de nuevos documentos en la web.

Para alcanzar esos logros se están haciendo varios estudios, hay varias personas esforzándose para lograr crear la llamada Web Semántica o Web del Futuro como la llamara en 1999 Tim Berners-Lee Lee, el creador de la web actual y el pensador de esta futura web.

1.3.1. Surgimiento de la Web Semántica.

Según el Diccionario de la Real Academia Española (DRAE), la palabra semántica es el “*Estudio del significado de los signos lingüísticos y de sus combinaciones*” [Diccionario 2001]. Después de analizar este vocablo se puede entender lo que se quiere decir cuando se habla de Web Semántica. En términos del propio Tim Berners-Lee:

“Tengo un sueño en dos partes para la Red. Primero veo que se convierte en un medio muy poderoso de comunicación entre los hombres. Luego [...] las computadoras cooperan. Las máquinas pasan a ser capaces de analizar todos los datos que circulan en la red: contenidos, enlaces y transacciones entre personas y computadoras. [Anbarasan 2000]

Esta futura web no suplantaría la actual sino que vendría siendo una extensión de esta, mejoraría todos los problemas que presenta la web mediante la introducción de descripciones explícitas del significado. Actualmente la web se asemeja a un grafo formado por nodos del mismo tipo y arcos (hiperenlaces) igualmente indiferenciados, por el contrario en la web semántica cada nodo (recurso) tiene un tipo (profesor, tienda, pintor, libro), y los arcos representan relaciones explícitamente diferenciadas (pintor – obra, profesor – departamento, libro – editorial). (Ver Anexo 1)

El éxito de la Web Semántica está en que las máquinas puedan comprender el significado de la información que se encuentra disponible en ella, para los humanos comprender el significado de una palabra no requiere de ningún esfuerzo, pero ¿Cómo lograr que las máquinas piensen?, las máquinas no pueden pensar, lo que se puede

lograr es que estas infieran o deduzcan conclusiones mediante algoritmos lógicos y matemáticos. [Abián 2005]

La Web Semántica hace uso de Ontologías comunes para que los participantes de la misma, contribuyendo o consumiendo recursos, puedan trabajar de forma autónoma con la garantía de que las piezas encajen [Castells]. Para proporcionar una semántica entendible de la información se emplean *metadatos* (Ver Anexo 3).

A pesar de que existan herramientas y se tengan claramente definidos los componentes necesarios para hacer realidad la Web Semántica, con los resultados obtenidos hasta el momento no se puede afirmar si esta web será un éxito o un fracaso, pero sí se puede asegurar que para lograr hacerla realidad hace falta mucho estudio y más tiempo para poder eliminar algunas de las limitaciones a las que se enfrenta. Esta futura web tiene que resolver los problemas existentes en la actual y además los que podrían aparecer debido a su surgimiento.

Uno de los conflictos evidentes de la web actual y que debe ser tenido en cuenta es el poder proporcionar mecanismos para acceder a la información en diferentes idiomas, permitiendo la creación y el acceso al contenido independientemente del idioma original de los proveedores de contenido o los usuarios. Se hace necesario realizar un esfuerzo significativo para organizar el contenido de la Web Semántica, almacenarlo y proporcionar los mecanismos para encontrarlo. Todas esas tareas deben ser realizadas y coordinadas de forma escalable, ya que las soluciones deben estar preparadas para el enorme crecimiento de la Web Semántica. Por otra parte sobre esta web existen disímiles criterios, hay quienes la ven muy buena y esperan que muy pronto se pueda utilizar, hay otros que no la creen necesaria, otros que plantean que será un profundo fracaso y que todo el esfuerzo que se está haciendo es inútil. [Abián 2005]

El futuro de la tecnología de la web se basará en la capacidad de la red para servir no sólo a las personas sino también a la información de agentes al servicio de los humanos. Aunque ya hay quienes hablan de la web 3.0 se debe destacar que esta es sólo una

mejora de la 2.0, es decir que el salto más significativo será lograr la Web Semántica. Por otra parte la 3.0 exige de mucha tecnología y recursos los cuales no son accesibles para todos los países principalmente los del tercer mundo.

1.3.2. Componentes de la Web Semántica.

Aunque esta web no se ha logrado hacer realidad, existen varias personas trabajando en aras de lograr hacerla realidad. Para su construcción se hace necesario el uso de todas las combinaciones de varios de los siguientes componentes: (Ver Anexo 2) [W3C 2006], [Castells], [Woodley 2003] y [Abián 2005].

XML: (eXtensible Markup Language). Es un lenguaje con una importante función en el proceso de intercambio, estructuración y envío de datos en la Web. Describe los datos de tal manera que es posible estructurarlos utilizando para ello etiquetas, como lo hace HTML, pero que no están predefinidas, delimitando de esta manera los datos y favoreciendo la interoperabilidad de los mismos.

XML SCHEMA: Es un lenguaje cuyo objetivo principal es definir la estructura en bloques de un documento XML pero de una forma mucho más precisa. El propósito de un esquema es definir y describir una clase de documentos XML usando estas construcciones para restringir y documentar el significado, uso y relaciones de las partes constituidas: tipo de datos, elementos y su contenido, atributos y sus valores, entidades y su contenido, y anotaciones. Los esquemas documentan su propio significado, uso y función.

RDF: Lenguaje para la definición de Ontologías y metadatos en la web. El elemento de construcción básica en RDF es el "triple" o sentencia, que consiste en dos nodos (sujeto y objeto) unidos por un arco (predicado), donde los nodos representan recursos, y los arcos propiedades.

RDF Schema (RDFS): Se pueden definir jerarquías de clases de recursos, especificando las propiedades y relaciones que se admiten entre ellas.

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

Como último componente se debe agregar los metadatos en los cuales se hace referencia a algunos datos dentro de las páginas web, así como información sobre cómo hacer deducciones con ellos, es decir, axiomas que podrán aplicarse en los diferentes dominios que trate el conocimiento almacenado.

Otros elementos que deben estar presentes en la Web Semántica son:

Lógica: El razonamiento lógico permite determinar si los datos son correctos e inferir conclusiones a partir de ellos.

Prueba: Las pruebas explican y verifican los pasos de los razonamientos lógicos.

Confianza: Técnicas que aseguran la identidad y fiabilidad de los datos y servicios.

1.4. Ontologías.

Las Ontologías no son explícitas, en el sentido de que no se detallan en un documento ni se organizan de forma jerárquica o matemática. Su uso es inherente a las personas, todos conocen que una familia se compone de varios miembros, que un hijo no puede tener más de un padre y una madre biológicos, que los padres tienen o han tenido padres. Estas relaciones no necesitan ser explicadas, pues forman parte del conocimiento. Sin embargo, cuando se trata de términos poco comunes o cuando se quiere que estos términos sean procesados por máquinas, se precisa explicitar las Ontologías; esto es, desarrollarlas en un documento o darles una forma que sea entendible para los ordenadores.

En la construcción de la Web Semántica las Ontologías se convertirán en una pieza clave, ya que permiten adicionarle semántica al contenido de la web. Se debe realizar un

gran esfuerzo en su creación, en proporcionar infraestructuras adecuadas para el desarrollo, gestión y mapeo de las mismas y en el apropiado control de la evolución.

Este término se empieza a utilizar a finales de los años 80 en el campo de la inteligencia artificial donde se emplean como vocabularios comunes para las personas y aplicaciones que trabajan en un dominio. Luego fue extendiéndose llegando a introducirse dentro del campo de la Informática en la segunda mitad de los años 90, donde se aplica dentro de la web para la inclusión de descripciones semánticas explícitas de recursos (contenidos y servicios).

Toda Ontología modela mediante conceptos y relaciones lo que se conoce sobre un dominio ó un área de conocimiento [Abián 2005]]. Estas permiten trabajar con conceptos en lugar de palabras claves. Desde el punto de vista de las fuentes de información describen el contenido de los repositorios de datos independientemente de la representación sintáctica de los mismos, posibilitando su integración semántica.

Con la aparición de las Ontologías en la web los usuarios pueden organizar la información de forma tal que los sistemas de software interpretan el significado de la misma. Las aplicaciones extraen automáticamente datos de las páginas web, los procesan y sacan conclusiones de ellos, pueden tomar decisiones y negociar con otros agentes o personas. Esto se debe a que las Ontologías catalogan y definen los tipos de cosas que existen en un cierto dominio al igual que sus relaciones, sus propiedades y sus componentes. *"Por ejemplo, una Ontología del mundo empresarial usará conceptos como Venta, Compra, Transferencia, Pago, etc.; y relaciones como "Una Transferencia corresponde a una Venta o a una Compra", "Un Pago corresponde a una o varias Transferencias", etc."* [Castells 2003]. Las Ontologías se representan mediante clases, propiedades y atributos de las clases, relaciones entre clases y restricciones sobre los atributos y propiedades de las clases.

1.4.1. Tipos de Ontologías.

Las Ontologías se pueden clasificar de acuerdo a la cantidad y tipo de estructura de la

conceptualización en: [Castells 2003]

Ontologías terminológicas: Especifican los términos que son usados para representar conocimiento en el universo del discurso. Suelen ser usadas para unificar vocabulario en un campo determinado.

Ontologías de información: Especifican la estructura de almacenamiento de bases de datos. Ofrecen un marco para almacenamiento estandarizado de información.

Las Ontologías también son clasificadas según su dependencia y relación con una tarea específica desde un punto de vista. La clasificación es la siguiente:

Ontologías de Alto nivel o Genéricas: Describen conceptos más generales.

Ontologías de Dominio: Describen un vocabulario relacionado con un dominio genérico.

Ontologías de Tareas o Técnicas básicas: Describen una tarea, actividad o artefacto, por ejemplo componentes, procesos o funciones.

Ontologías de Aplicación: Describen conceptos que dependen tanto de un dominio específico como de una tarea específica, y generalmente son una especialización de ambas.

1.4.2. Componentes de una Ontología.

Las Ontologías tienen los siguientes componentes que servirán para representar el conocimiento de algún dominio [Gruber 1993]:

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. [Lozano]

1.4.3. Aplicaciones y beneficios.

Las Ontologías se usan para favorecer la comunicación entre personas, organizaciones y aplicaciones, lograr la interoperabilidad entre sistemas informáticos, razonar automáticamente y para la ingeniería de software.

En la ciencia y en la tecnología los problemas derivados de la falta de comprensión común entre personas, organizaciones y aplicaciones adquieren una gran importancia. Las Ontologías proporcionan una comprensión común de un determinado dominio y eliminan estas confusiones de conceptos y términos.

La primera aproximación de una Ontología es el tesoro. Un tesoro es "un tipo de lenguaje documental que representa la estructuración conceptual de un determinado campo del conocimiento (Ver Anexo 6). El tesoro proporciona una organización semántica principalmente a través de la explicitación de las relaciones establecidas entre dichos conceptos y, eventualmente, a través de un significado restringido de los términos que los representan". [Arano 2005]

Se trata de una estructura en la que se guardan términos que se ven relacionados por su cercanía semántica. Son varios los autores que han profundizado en las diferencias existentes entre las Ontologías y los tesoros. Algunos de ellos entienden que las

primeras son superiores a los segundos por varias razones: en primer lugar, presentan un nivel más alto de concepción y de descripción del vocabulario. Además, las Ontologías se caracterizan por un desarrollo semántico más profundo para las relaciones del tipo clase/subclase y para las relaciones cruzadas, lo que supone la ampliación de éstas y un mayor cuidado en su descripción, por supuesto explícita. En tercer lugar, destacan el uso de la lógica de descripción empleada en la descripción de situaciones. Y por último, hacen hincapié en la reusabilidad de las Ontologías y en la posibilidad del trabajo en sistemas heterogéneos, al describir formalmente objetos en el mundo, sus propiedades, y las relaciones entre estos Objetos. [García 2004]

Las Ontologías serán imprescindibles en el desarrollo de la Web Semántica y en los futuros sistemas de gestión empresarial porque permitirán que las aplicaciones estén de acuerdo en los términos que usan cuando se comunican. Mediante ellas, será mucho más fácil recuperar información relacionada temáticamente, aun cuando no existan enlaces directos entre las páginas web.

Por ejemplo, una Ontología puede usarse para especificar que las termitas son un tipo de isóptero. De este modo, un buscador que use esa Ontología mostrará páginas web sobre termitas cuando un usuario busque información sobre los isópteros.

Las Ontologías también permiten que los sistemas interoperen. Dos sistemas son interoperables si pueden trabajar conjuntamente de una forma automática, sin esfuerzo por parte del usuario.

Las Ontologías resultan muy útiles para facilitar el razonamiento automático, es decir, sin intervención humana, usando para ello reglas de inferencia. Un motor de razonamiento puede usar los datos de las Ontologías para inferir conclusiones de ellos. Una de las aplicaciones más importantes del razonamiento automático es la validación de datos.

También se usa para establecer relaciones entre Ontologías y así poder descubrir relaciones ocultas o inesperadas entre los datos e integrar esquemas de bases de datos.

Cuando se trabaja con bases de datos federadas* (vistas unificadas de bases de datos independientes) esta integración de esquemas resulta imprescindible pues estas bases de datos aparentan ser una sola, pero se componen de información extraída de bases de datos independientes, que pueden estar alejadas miles de kilómetros.

En la ingeniería del software, las Ontologías ayudan a la especificación de los sistemas de software. Como la falta de un entendimiento común conduce a dificultades en identificar los requisitos y especificaciones del sistema que se busca desarrollar, las Ontologías facilitan el acuerdo entre desarrolladores y usuarios.

Los beneficios de utilizar Ontologías se pueden resumir de la siguiente forma:

- ❖ Proporcionan una forma de representar y compartir el conocimiento utilizando un vocabulario común.
- ❖ Permiten usar un formato de intercambio de conocimiento.
- ❖ Proporcionan un protocolo específico de comunicación.
- ❖ Permiten una reutilización del conocimiento. [Wikipedia 2007]

1.4.4. Lenguajes para representar Ontologías.

Las Ontologías para ser expresadas requieren de un lenguaje lógico y formal. En la inteligencia artificial se han desarrollado numerosos lenguajes para este fin, algunos basados en la lógica de predicados, como KIF y Cycl que ofrecen poderosas primitivas de modelado, y otros basados en frames (taxonomías de clases y atributos), que tienen un mayor poder expresivo, pero menor poder de inferencia; e incluso existen lenguajes orientados al razonamiento como Description Logic y Classic. Todos estos lenguajes han

* Una **BD federada** “aparenta ser una BD normal y corriente, pero no tiene existencia física: es una vista lógica. Las bases de datos federadas son muy importantes en la Web, pues dan una vista común de los datos procedentes de fuentes muy distintas (agencias de noticias, portales, foros, periódicos y revistas electrónicas, etc.)”. [Wikipedia 2007]

servido para desarrollar otros lenguajes aplicables a la Web. En un lenguaje de Ontologías se pretenderá un alto grado de expresividad y uso.

Dentro de los principales lenguajes de Ontologías se destacan los siguientes:

- ❖ **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*, etc; el proyecto fue abandonado a medida que se desarrollaron OIL y DAM; aunque también existe una socialización de este lenguaje en XML.

- ❖ **OIL**: Ontology Inference Layer. Este lenguaje, derivado en parte de SHOE, fue impulsado también por el proyecto de la Unión Europea On-To-Knowledge. Utiliza 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.

- ❖ **OWL**: Web Ontology Language o Lenguaje de Ontologías para la Web es un lenguaje de etiquetado semántico para publicar y compartir Ontologías en la Web. Se trata de una recomendación del W3C, y puede usarse para representar Ontologías de forma explícita, es decir, permite definir el significado de términos en vocabularios y las relaciones entre aquellos términos (Ontologías). En realidad, OWL es una extensión del lenguaje RDF y emplea las tripletas de RDF, aunque es un lenguaje con más poder expresivo que éste. Se trata de un lenguaje diseñado para usarse cuando la información contenida en los documentos necesita ser procesada por programas o aplicaciones, en oposición a situaciones

donde el contenido solamente necesita ser presentado a los seres humanos. Al igual que OIL, OWL se estructura en capas que difieren en la complejidad y puede ser adaptado a las necesidades de cada usuario, al nivel de expresividad que se precise y a los distintos tipos de aplicaciones existentes (motores de búsqueda, agentes, etc.). Existen 3 sublenguajes de OWL, los cuales van creciendo respecto al nivel de expresión:

- ❖ OWL Lite: Útil para la creación de jerarquías y restricciones simples, sólo permite valores de cardinalidad 0 y 1, pierde en expresividad. [Hernández]
- ❖ OWL DL: (Description Logic), es el lenguaje más sencillo e indicado para los usuarios que requieren el máximo de expresividad y decibilidad (todos los cálculos acaban en un tiempo finito). Una clase puede ser a la vez subclase de muchas clases, no puede ser una instancia de otra clase.
- ❖ OWL Full: Máximo nivel de expresión y la libertad sintáctica de RDF. Permite expresiones de segundo orden, pero sin decibilidad. Por ejemplo, una clase puede ser tratada simultáneamente como una colección de individuos y como un individuo por sí mismo.[Barrón 2005] Es el más completo por lo que se necesita mucho poder computacional para poder hacer inferencias es por eso que se dice que no tiene garantía computacional.
- ❖ **KIF**: Knowledge Interchange Format es un lenguaje para representar Ontologías basadas en la lógica de primer orden. KIF está basado en la lógica de predicados con extensiones para definir términos, metaconocimiento, conjuntos, razonamientos no monotónicos, etc.; y pretende ser 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. [Barrón 2005]

Las Ontologías son vocabularios comunes que, junto con otras tecnologías que proveen de herramientas y lenguajes para generar marcado y procesamiento semántico, harán posible la Web Semántica. Es necesario que los documentos generados para la Web posean una semántica formalizada en Ontologías con el fin de que este conocimiento sea intercambiado por los agentes de software. [Wikilearning 2006]

1.4.5. Herramientas para Construir Ontologías.

Existen herramientas para representar Ontologías, es decir, el conocimiento en entorno de páginas web con lenguajes de marcado propios. La mayoría permiten describir el contenido de los documentos en forma de metadatos, estos son soportados sobre una Ontología representada en RDF Schema (RDFS) o basados en grafos conceptuales.

Dentro de las herramientas más utilizadas por la comunidad y usualmente recomendadas para la construcción de Ontologías se destaca Protégé.

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. [Abián 2005]

Este editor incluye además SPARQL, ofrece a los desarrolladores y usuarios finales un camino para presentar y utilizar los resultados de búsquedas a través de una gran variedad de datos, los cuales pueden ser personales, redes sociales y metadatos sobre recursos digitales como música e imágenes, siendo útil para recuperar y organizar la información. SPARQL también proporciona un camino de integración sobre recursos diferentes. Permite:

- ❖ Extraer información en diversas formas, incluyendo URIs.
- ❖ Extraer subgrafos RDF.

- ❖ Construir nuevos grafos RDF basados en la información de los grafos consultados.

OiEd: 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. [Samper 2005]

WebODE: es una herramienta que basa la construcción de Ontologías en el método Methontology, del cual se hablará más adelante, permite exportar el conocimiento a diferentes lenguajes de especificación (RDFs, OWL, OIL, DAML + OIL). [Samper 2005]

1.4.6. Sistemas de Almacenamiento.

Los sistemas de almacenamiento posibilitan mantener las Ontologías en bases de datos e ir añadiendo nueva información, y con la ayuda de razonadores probar la consistencia de la Ontología. Aunque existen varios sistemas entre ellos el más utilizado es Jena.

JENA: es un framework desarrollado por los laboratorios de HP. Su fin es manipular metadatos desde aplicaciones escritas en Java. En su versión 2, JENA incluye un analizador sintáctico de RDF, una API para RDF, una API de Ontologías que permite trabajar con OWL y RDFS, un subsistema de razonamiento y un sistema de persistencia. Permite trabajar con el lenguaje de consultas de RDF (RDQL). Haciendo uso de Java proporciona clases para representar *modelos*, recursos, propiedades, etc.

Sesame: es un repositorio para RDF-Schema que permite añadir y eliminar información, 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. [Samper 2005]

KAON Tool: implementa una interfaz independiente del sistema en el que se almacenarán las Ontologías, ya sean cualquier base de datos o un fichero texto. Implementa un API para leer las descripciones de los recursos, emplea RQL para realizar consultas y soporta tanto Ontologías DAML + OIL como RDF. [Samper 2005]

1.4.7. Razonadores de Ontologías.

Una de las herramientas más utilizadas para trabajar con las Ontologías son los razonadores, que sirven para realizar inferencia, a través de los conceptos y en algunos casos las instancias, obteniendo nuevo conocimiento. Generalmente difieren en el lenguaje formal en el que se especifica el conocimiento, así como los lenguajes de consulta que puedan utilizar. En la actualidad, son varios los razonadores o sistemas deductivos basados en lógica descriptiva que permiten el razonamiento y la inferencia en las Ontologías. Los principales son:

- ❖ **FaCT** (Fast Classification of Terminologies): desarrollado por Ian Horrocks, que puede ser usado para chequear el grado de satisfacción de las descripciones. Permite reglas transitivas, inversas, restricciones cualificadas, jerarquías etc. Es lo suficientemente expresivo para soportar y razonar sobre cualquier base de conocimiento. Escrito en Common Lisp, fácilmente ejecutable por cualquier programa Lisp de forma local, tiene escrita una versión servidor FaCT para ser usada vía interfaz CORBA sobre cualquier sistema con acceso a la red. Actualmente es el razonador empleado por defecto en el editor de Ontologías OilEd para clasificar los conceptos en una jerarquía según las descripciones que tengan. [Samper 2005]

- ❖ **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. [Samper 2005]

- ❖ **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. [Samper 2005]

- ❖ **DamJessKB:** es el razonador para DAML + OIL descrito por National Science Foundation (NSF), Knowledge and Distributed Intelligence in the Information Age (KDI) Initiative. Implementado usando Jess (Java Expert System Shell), basa su mecanismo en parsear la información expresada en DAML + OIL en reglas para el sistema experto como es Jess. Tiene funcionalidades para soportar la clasificación de clases definidas a partir de las reglas lógicas como las disyunciones, restricciones de valores, etc. También tiene soporte para los tipos de datos de XML Schema y para las propiedades transitivas. Actualmente se está migrando a OWL.

- ❖ **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. [Samper 2005]

- ❖ **JTP** (Java Theorem Prover): es una arquitectura para razonar sobre conocimiento descrito en DAML + OIL. Escrito en Java, soporta el modelo de axiomas de DAML + OIL, realiza una precomputación cuando se carga la Ontología, y ha añadido un clasificador para realizar la jerarquía de conceptos de la Ontología. Por el momento, es el único razonador que permite el lenguaje de consulta DQL específico para DAML + OIL. [Samper 2005]

- ❖ **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. [Samper 2005]

Los razonadores basados en lógica descriptiva deben poseer la suficiente expresividad para poder permitir que el conjunto de constructores que forman parte de los lenguajes de Ontologías sean soportados, y permitir inferencias, tanto de clasificación como de chequeo de instancias.

1.5. e-Learning.

Existen diferentes conceptos de e-Learning pues depende del uso que de él hacen las personas. Según Manuel Castells, sociólogo y profesor universitario español, catedrático de Sociología y de Urbanismo en la Universidad de California [Wikipedia 2007]. "Técnicamente, el e-learning es la entrega de material educativo vía cualquier medio electrónico, incluyendo el Internet, Intranets, Extranet, audio, vídeo, red satelital,

televisión interactiva, CD y DVD, entre otros medios". [Castells 1996]

Para los educadores "es el uso de tecnologías de redes y comunicaciones para diseñar, seleccionar, administrar, entregar y extender la *educación*". [Cánepa 2005]

La educación a distancia *"Tuvo su origen en Alemania, Inglaterra y Norteamérica. Nació con el propósito de ayudar en la instrucción de personas que por circunstancias diversas les era muy difícil asistir a la enseñanza presencial. La educación por correspondencia, reemplaza la lección oral del profesor por medio de cursos impresos que se convierten en objetivos del aprendizaje que han de cumplirse con la ayuda del profesor"*. [Cánepa 2005]

Luego con el surgimiento de Internet se abre una nueva puerta para esta modalidad, la expansión, la posibilidad de que los cursos lleguen rápidamente a sus casas, de control al generarse reportes sobre el progreso de los alumnos, pudiendo establecer métodos de evaluación efectivos, la facilidad de actualizarse y que las mismas lleguen inmediatamente a los alumnos son ventajas que presenta la educación a distancia asistida por ordenadores que traducida al inglés es e-Learning.

1.5.1. Objetos de Aprendizaje.

Dentro de la enseñanza centrada en el aprendizaje del alumno en los entornos de *formación* on-line (e-Learning), el concepto de Objeto de Aprendizaje (OA) juega un papel importante en la construcción y distribución personalizada de contenidos, así como la reutilización de los mismos en nuevos contextos. [Del Moral 2005]

Los Objetos de Aprendizaje (OA) o como se puede encontrar en diferentes bibliografías Learning Object (LO) o Reusable Learning Object (RLO), están basados en programación orientada a objetos y son piezas individuales autocontenidas y reutilizables de contenido que tienen propósitos educacionales en ambientes basados en web. *"Constituidas por paquetes de información multiformato y con carácter interactivo, dotado*

de las siguientes características: Orientado a presentar información para lograr un único objetivo educativo a través de micro-unidades didácticas que contemplen: contenidos, recursos, actividades y evaluación; Extrapolable a otros contextos por su potencial reusable; Relevante como experiencia de aprendizaje significativo que sirve de anclaje para adquirir conocimientos posteriores; Compatible técnicamente para ser visualizado independiente del formato y dispositivo; Identificable a través de metadatos; Adaptable a las situaciones y necesidades específicas de los estudiantes; Durable frente a los cambios tecnológicos sin necesidad de rediseño o cambio de código importante.” [Del Moral 2005]

Un objeto de aprendizaje se compone de dos partes fundamentales: una caracterización que define al componente como recurso educativo y el contenido que puede ser cualquier tipo de archivo, texto, imagen, video, etc. [Arias 2007]. El contenido de los OA puede organizarse de diferentes formas para construir módulos, lecciones, tópicos, etc. Para lograr una mayor utilidad de los OA se requiere que estos objetos sean compatibles con diversos ambientes y sistemas de administración de aprendizajes, fáciles de migrar de una plataforma a otra, de forma tal que sea fácil de localizar, acceder, archivar y re-utilizar el contenido. La satisfacción de estos requisitos dará una vida útil más larga a nuestros materiales didácticos electrónicos y su valor será mayor. De allí la importancia del establecimiento de estándares para el diseño y descripción de los Objetos de aprendizaje. Los potenciales beneficios de reutilización, interoperatividad, durabilidad y accesibilidad sólo pueden ser alcanzados con un amplio consenso en torno a estándares apropiados.

Para facilitar la utilización de los OA en otros sistemas debe existir un esquema de metadatos que contenga información específica sobre su contenido.

1.5.2. Los metadatos.

Este término existe desde antes de la aparición de Internet pero no fue hasta hace unos años que se hizo popular [López 2006]. Desde el punto de vista social: *“El concepto de metadatos es análogo al uso de índices para localizar objetos en vez de datos. Por*

ejemplo, en una biblioteca se usan fichas que especifican autores, títulos, casas editoriales y lugares para buscar libros. Así, los metadatos ayudan a ubicar datos.” [Wikipedia 2007]

El término metadatos *viene del griego meta, «después de» y latín datum, «lo que se da», «dato», literalmente «sobre datos», son datos que describen otros datos. En general, un grupo de metadatos se refiere a un grupo de datos, llamado recurso* [Wikipedia 2007]. Un registro de metadatos consiste en un registro de elementos y atributos necesarios para describir un recurso. [López 2006]

“HTML permite definir metadatos para una página web a través de su etiqueta <meta>. Esos metadatos (author, keywords...) caracterizan la página y describen su contenido. Los metadatos, utilizados tradicionalmente en el entorno bibliotecario, están resultando de gran utilidad en la Web, tanto en Sistemas de Recuperación de Información como en Sistemas de Navegación”. [Montero 2003]

En el campo de la informática *“los metadatos son datos altamente estructurados que describen información, el contenido, la calidad, la condición y otras características de los datos. Es información sobre información o datos sobre los datos”. [Metadatos 2007]*

Los metadatos se pueden clasificar según los siguientes criterios:

“Contenido. Subdividir metadatos por su contenido es lo más común. Se puede separar los metadatos que describen el recurso mismo de los que describen el contenido del recurso. Es posible subdividir estos dos grupos más veces, por ejemplo para separar los metadatos que describen el sentido del contenido de los que describen la estructura del contenido o los que describen el recurso mismo de los que describen el ciclo vital del recurso.

Variabilidad. Según la variabilidad se puede distinguir metadatos mutables e inmutables. Los inmutables no cambian, no importa qué parte del recurso se vea, por ejemplo el nombre de un fichero. Los mutables difieren de parte a parte, por ejemplo el contenido de

un vídeo.

Función. Los datos pueden ser parte de una de las tres capas de funciones: subsimbólicos, simbólicos o lógicos. Los datos subsimbólicos no contienen información sobre su significado. Los simbólicos describen datos subsimbólicos, es decir añaden sentido. Los datos lógicos describen cómo los datos simbólicos pueden ser usados para deducir conclusiones lógicas, es decir añaden comprensión" [Wikipedia 2007].

Como se vio anteriormente los metadatos son empleados junto a los OA para garantizar su reutilización pues posibilitan la comunicación declarando cómo están relacionados los datos y convirtiéndolos de un formato a otro, posibilidad que es empleada en la representación del conocimiento y en las consultas a buscadores donde se obtienen resultados más precisos.

Los metadatos aplicados a objetos de aprendizaje permiten:

- ❖ Crear descripciones bien estructuradas de recursos de aprendizaje, que faciliten el descubrimiento, localización, evaluación y adquisición de recursos de aprendizaje por parte de estudiantes, profesores o procesos de software automáticos.
- ❖ Compartir descripciones de recursos entre sistemas de búsqueda de recursos, que conlleven una reducción en los costes de los servicios proveedores basados en descripciones de recursos de alta calidad.
- ❖ Adaptación de las descripciones de recursos a las necesidades concretas de una comunidad, que puede incluir la elección de vocabularios controlados y adecuados para la clasificación, reducción o ampliación del número de elementos de descripción.
[Rodríguez]

1.6. Estándares Educativos.

Un estándar se define como un conjunto de reglas o normas que especifican cómo debe realizarse un servicio específico, producirse un determinado producto o efectuarse un proceso dado de modo que se garantiza una cierta calidad y compatibilidad con otros productos o servicios. Estos estándares son generados por organizaciones internacionales, ya sean públicas o privadas e incluso gubernamentales.

Para las plataformas e-Learning se hace necesario el uso de estándares bien definidos en aras de asegurar, en primer lugar, la eficiencia de los contenidos desarrollados *online*, además de facilitar su gestión, en segundo lugar incrementar la cantidad de los contenidos así como permitir personalizarlos y reutilizarlos, como tercer paso es fundamental asegurar la compatibilidad entre diferentes plataformas y por último la más importante de las funcionalidades de estas plataformas es el permitir realizar seguimientos a los alumnos en los cursos. [Navarro 2004]

Es necesario aclarar la diferencia entre un estándar y una especificación dado a que en este contexto en ocasiones se suele llamar estándares a algunas especificaciones que funcionan como estándares de facto^{*}. [Burgos 2006]

Un estándar es algo normativo, que tiende a evolucionar de manera muy lenta, manifestándose como conclusión sobre una teoría y es completo, cubriendo cada área de la teoría y debe estar respaldada por una autoridad respetada mundialmente en el campo en el que trabaja, como ISO (International Standards Organization), BSI (British Standards Institute), CEN (Centre European de Normalization) o IEEE (Institute of Electrical and Electronics Engineers) mientras que una especificación no tiene características normativas, tendiendo a ser experimentos apoyados por la industria y que evolucionan de manera muy rápida. [Hernández]

* Un **estándar de facto** “es aquel patrón que se caracteriza por no haber sido consensuada ni legitimada por un organismo de estandarización al efecto, se trata de una norma generalmente aceptada y ampliamente utilizada por iniciativa propia de un gran número de interesados”. [Wikipedia 2007].

1.6.1 Estándares de e-Learning.

Existen varias instituciones reconocidas internacionalmente encaminadas a la propuesta de estándares para los distintos componentes que forman la arquitectura de las plataformas de formación. Se destacan: AICC (Aviation Industry CBT Committee), IEEE (Institute of Electrical and Electronics Engineers), IMS (Instruction Management Systems) y ADL (Advanced Distributed Learning).

IEEE.

Es un cuerpo multinacional que desarrolla estándares internacionales para sistemas eléctricos, electrónicos, computacionales y comunicacionales. Está organizado en distintos comités que se juntan y analizan las distintas tecnologías, entregando como resultado una especificación o recomendación en forma de estándar.

AICC.

Es un grupo internacional de profesionales del entrenamiento y capacitación basada en tecnología. Fue creado cuando no existía Internet, ni la web, sólo se hablaba del CBT (Computer-Based Training) o Capacitación Basada en Computadoras. Actualmente ha extendido su ámbito de actuación y trabaja conjuntamente con IEEE, IMS y ADL generando directrices y recomendaciones en diversas áreas desde el hardware hasta la interoperabilidad. [Burgos 2006]

IMS.

Es un consorcio que se enfoca completamente en desarrollar especificaciones en formato XML como el lenguaje seleccionado para realizar los intercambios de información entre aplicaciones en Internet. [Burgos 2006]

Esta especificación cubre un amplio rango de características que se persiguen hacer interoperables entre plataformas, que van desde los metadatos, la interoperabilidad de intercambiar el diseño instruccional entre plataformas, hasta la creación de cursos online.

ADL.

Nace con el objetivo de investigar y desarrollar especificaciones que fomenten la adopción y el avance del e-Learning.

La principal misión de ADL es buscar mecanismos que aseguren la educación y los materiales de capacitación de alta calidad que puedan ajustarse a las necesidades de cada institución. Como resultado de estas investigaciones y recomendaciones ADL se está transformando en un gran impulsor para convertir las especificaciones en estándares. La publicación que más ha repercutido de ADL es el Modelo de Referencia de Objetos de Contenido Compartibles, o Shareable Content Object Reference Model, conocido ya en el ambiente como SCORM. [Burgos 2006]

LOM.

El estándar IEEE LTSC Learning Object Metadata especifica la sintaxis y la semántica para objetos de aprendizaje. Este estándar se centra en proporcionar el conjunto mínimo de atributos necesarios para que estos objetos de aprendizaje sean gestionados, localizados y evaluados de manera apropiada. LOM se integra con otros estándares ya definidos como SCORM y ADL que complementen la definición y utilización precisas de cada objeto de aprendizaje. Sus objetivos abarcan también la compartición e intercambio de objetos de aprendizaje entre profesores y pedagogos utilizando cualquier tecnología que desarrolle sistemas de aprendizaje. *“La actual versión de LOM se compone de 9 categorías y 47 elementos”.* (Ver Anexo 4) [Navarro 2004]

Categorías:

General. *“Recoge algunos de los principales elementos de identificación del documento descrito: código de identificación, título del recurso, idioma, breve descripción de su contenido, palabras claves, cobertura temporal o geográfica, estructura e información sobre su granularidad o nivel de agregación”.*

Ciclo de vida. *“Recoge información relativa a la autoría, fecha de creación, versión y estado del recurso descrito”.*

Metadatos. *“Proporciona información sobre el esquema de metadatos empleado en la descripción del recurso, fecha, nombre del creador e idioma del registro”.*

Técnica. *“Recoge información relativa al formato, tamaño, URI, duración y requisitos técnicos para la utilización del recurso”.*

Uso educativo. *“Describe el uso educativo del recurso, tipo de recurso de que se trata, tipo y nivel del usuario al que se dirige, contexto de utilización, tipo y nivel de interactividad que presenta, densidad semántica, dificultad, idioma y descripción de su uso”.*

Derechos. *“Recoge aspectos relativos a las restricciones de uso asociadas al recurso: coste, protección de los derechos de autor y otras restricciones de uso”.*

Relación. *“Proporciona información sobre las relaciones, en caso de que las haya, establecidas entre el recurso descrito y otros recursos”.*

Anotación. *“Recoge los comentarios del catalogador sobre el uso pedagógico del recurso”.*

Clasificación. *“Descripción del contenido del recurso a partir de uno o varios sistemas de clasificación, vocabularios y palabras claves”.* [Marzal 2006]

En este estándar cada elemento es opcional. La lista de elementos requeridos difiere dependiendo del componente al cual describa. [Marzal 2006]

SCORM.

La especificación SCORM logra combinar de excelente forma los elementos de IEEE, AICC e IMS en un único documento consolidado de fácil implementación. SCORM se diseñó para facilitar la portabilidad de cursos e información de una plataforma a otra y para permitir que los contenidos sean reutilizados en varios cursos, empaquetándolos en

objetos modulares. El objetivo de SCORM es conseguir contenidos reutilizables, interoperables, duraderos y accesibles, independientemente del sistema de gestión y entrega de contenidos utilizado. [Burgos 2006]

QTI.

IMS Question & Test Interoperability Specification (QTI) es una especificación para conseguir una manera estándar de compartir evaluación y datos. Es una especificación XML interoperable para describir preguntas y tests. Con el fin de permitir a los usuarios importar y exportar preguntas, tests y resultados entre diferentes aplicaciones, IMS QTI tiene que ser claro y conciso de esta manera se evitan ambigüedades. Esta especificación fue construida de forma que es posible soportar preguntas y tests tanto simples como complejos [Burgos 2006]. Así se dispone de almacenes de preguntas y bases de datos con los resultados obtenidos por los alumnos a los que cualquier sistema de enseñanza electrónica puede acceder.

“Los autores de las evaluaciones pueden crear sus propias preguntas o incluir preguntas diseñadas por otros usuarios de IMS QTI, haciendo más fácil crear repositorios de preguntas reutilizables en diferentes sistemas”. [Burgos 2006]

QTI permite el desarrollo de varios tipos de preguntas como: respuesta múltiple, verdadera o falsa, rellenar los espacios en blanco, seleccionar texto, desplazar objetos, ordenar, relacionar ítems, conectar puntos. [Burgos 2006]

1.6.2. Ventajas del uso de estándares.

El uso de los estándares favorece el crecimiento, la expansión y la generalización de cualquier tecnología. Para el caso específico del e-Learning aún están en proceso de adopción y según se vayan introduciendo, se obtienen beneficios en contenidos, infraestructura y funcionalidad en cuanto a la interoperabilidad entre las distintas plataformas de formación y entre otras cosas su reutilización y manejabilidad.

La aplicación de estándares no es obligatoria, no se trata de una serie de leyes de estricto cumplimiento; estos son documentos que definen características de productos, servicios o procesos de acuerdo a criterios técnicos, de ahí que su utilización sea voluntaria. El empleo de estándares para diseñar materiales de e-Learning conlleva a una serie de ventajas, entre las cuales cabría resaltar:

- ❖ Incrementar la calidad y cantidad de los contenidos.
- ❖ Posibilitar el intercambio de cursos y contenidos.
- ❖ Personalizar contenidos.
- ❖ Asegurar la compatibilidad.
- ❖ Permitir la búsqueda de contenidos.
- ❖ Fomentar la profesionalización en la creación de contenidos.
- ❖ Garantizar la viabilidad de la inversión en e-Learning.
- ❖ Aumentar la oferta de cursos.
- ❖ Los productos no quedan obsoletos.

1.7. Los Repositorios.

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

Los repositorios según su contenido se pueden clasificar en: Repositorios de Objetos de Aprendizaje, de documentación, de bases de datos, actualizaciones de sistemas operativos como Ubuntu, Debian, etc.

1.7.1. Repositorios de Objetos de Aprendizaje.

Los Repositorios de Objetos de Aprendizaje surgen de la necesidad de compartir y guardar organizadamente recursos educativos provenientes de disímiles fuentes para luego ser utilizados.

“Muchos autores plantean que la idea de Repositorio es intrínseca a los Objetos de Aprendizaje. No es posible pensar en objetos de aprendizaje si no se los concibe albergados en repositorios. Como objetos aislados no tienen ninguna relevancia ni significado real. Una manera de comprender los repositorios, es imaginar una combinación entre una biblioteca digital y un buscador como Yahoo o Google, pero mucho más sofisticado que ambos. Por una parte, los objetos son de naturaleza diversa (al contrario de una biblioteca) y por otra parte, los criterios de búsqueda deben considerar bastante más que títulos, autores o palabras claves. El tipo de componentes albergados en un repositorio, que deben tener sus propias identidades y ser por lo tanto localizables, son tan variados como gráficos, imágenes, textos, videos, documentos e integración de ellos como capítulos de un curso o hasta cursos completos. Un aspecto muy importante de los repositorios es que no necesariamente albergan físicamente los objetos que contienen; les basta con apuntar a ellos”. [Varas]

Estos repositorios agrupan y almacenan los OA de dos formas: la primera, los recursos educativos y sus correspondientes metadatos dentro de un mismo sistema e incluso dentro de un mismo servidor, y la segunda, sólo se encuentran los metadatos, en este caso el repositorio contiene sólo los descriptores y se accede al objeto a través de una referencia a su ubicación física que puede ser en otro sistema o repositorio de objetos. En ambos casos el contenido puede ser accedido por otros sistemas o usuarios mediante interfaces de búsquedas. [López 2006]

1.7.2. Repositorios en la Web.

La iniciativa de Objetos de Aprendizaje a nivel mundial permite que cualquier docente o estudiante pueda acceder al material publicado en los diferentes bancos de Objetos de Aprendizaje (LOR - Learning Object Repository).

Una de las formas de hacer posible el acceso a recursos de aprendizaje consiste en contar con repositorios que almacenen objetos de aprendizaje y sus respectivos

metadatos. Algunos proyectos sobre repositorios de objetos de aprendizaje son: Merlot^{*} y Careo[†] los cuales conducen a algunos de los bancos más representativos y con mayor cantidad de objetos almacenados para el apoyo de las labores de enseñanza y aprendizaje.

Merlot (Multimedia Educational Resource for Learning and Online Teaching)

Repositorio comunitario, de recursos educativos para utilizarse en la educación en línea. Es un recurso abierto de libre acceso diseñado primordialmente para estudiantes y docentes de educación superior. Colecciona enlaces a materiales de aprendizaje en línea con anotaciones, revisiones y tareas.

Características:

- ❖ Número de objetos de aprendizaje: más de 10.000.
- ❖ Repositorio libre y gratuito diseñado principalmente para estudiantes de educación superior o universitaria.
- ❖ Contenidos relacionados con negocios, arte, humanidades, ciencia y tecnología, ciencias sociales.
- ❖ Se pueden realizar búsquedas utilizando algunos (pocos) metadatos basados en el estándar IEEE LOM.
- ❖ El contenido físico generalmente es un enlace al contenido alojado en la Web del autor.
- ❖ No ofrece la posibilidad de descargar objetos de aprendizaje en formato empaquetado estándar.

Careo (Campus Alberta Repository of Educational Objects)

Su principal objetivo es la creación de una colección de material de aprendizaje

^{*} Disponible en: <http://www.careo.org>.

[†] Disponible en: <http://www.merlot.org>.

multidisciplinario basado en Web, que pueda ser fácilmente recuperado.

Características:

- ❖ A diferencia de MERLOT, los metadatos de cada objeto de aprendizaje están mucho más detallados.
- ❖ Muchos de los objetos de aprendizaje se pueden descargar en forma de paquete comprimido. Pero el paquete no es estándar, porque no contiene los metadatos.
- ❖ Metadatos utilizados: IEEE LOM, CanCore (una adaptación canadiense de LOM)
- ❖ Interoperabilidad con otros repositorios: especificación IMS DRI.

Los repositorios actuales de objetos de aprendizaje, tales como MERLOT o CAREO, por lo general describen los distintos recursos didácticos existentes en la Web, estos almacenan registros de metadatos asociados a los objetos descritos y de este modo se garantiza una búsqueda mucho más estructurada del conocimiento existente.

No obstante, la búsqueda no es la única ventaja de la existencia de repositorios, permiten revisiones cooperativas de los objetos de aprendizaje, de modo que la calidad del contenido es revisada y cuestionada por los distintos participantes que acceden al repositorio a través de Internet. [Soto 2006]

1.7.3. Caso Cuba.

La Universidad Virtual de Salud (UVS) constituye uno de los proyectos de la red cubana Infomed e integra herramientas e-Learning estandarizadas (Scout Portal Toolkit) con otras de fuentes abiertas para el trabajo colaborativo, la sindicación y la gestión docente descentralizada. Esta herramienta *“permite optimizar la manera de gestionar los catálogos de fuentes de información en ambiente digital de los que dispone una organización”*. [Presno]

La incorporación de esta herramienta a la UVS de Cuba proporciona las siguientes ventajas:

- ❖ Posibilita salvar las búsquedas efectuadas por cada usuario en un tópico determinado, así como el establecimiento de servicios de retroalimentación y diseminación selectiva de temas de interés personal vía correo electrónico y RSS (Really Simple Syndication) lo que hace posible establecer un ranking de las últimas actualizaciones en la página principal.
- ❖ Permite entrar a diversos fórum para discusiones de tópicos, comentarios, recomendaciones, servicio de Diseminación Selectiva, entre otras opciones.
- ❖ Facilita la Clasificación de los recursos y la posibilidad a los usuarios de conocer cantidad de registros según temáticas. [Presno]

“El protocolo RSS para referenciar y syndicar respectivamente los recursos educativos en la UVS utiliza el estándar de IEEE-LOM. Mediante este protocolo los estudiantes reciben información y acceden a los objetos de aprendizaje de forma personalizada, ellos mismos definen la dinámica de su aprendizaje. A la vez pueden acceder y compartir los nuevos recursos desde su propia página Web o escritorio del ordenador con total inmediatez”. [Jardines 2006]

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”.* [Tamayo 2005]

Con la puesta en práctica de este sistema se logró una extensa matrícula, constituida fundamentalmente por profesionales vinculados a centros laborales y estudiantes de la universidad, quienes finalizaron sus cursos exitosamente. *“Este repositorio facilita la creación de los cursos y la reutilización de materiales educativos, aumentando el*

valor de estos al ser utilizados en distintos contextos, disminuyendo el costo de elaboración de los cursos". [Tamayo 2005]

Basado en la experiencia de la Universidad de Holguín, en la Universidad de las Ciencias Informáticas se trabaja en la construcción de su propio 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"* [Leyva 2006]. Actualmente ROA cuenta con 4 funcionalidades básicas subir, actualizar, explorar paquetes y buscar recursos, [Leyva 2006] se debe señalar que se está trabajando en función de perfeccionar este sistema, el cual se piensa poner en práctica para toda la universidad en septiembre del curso 2007-2008.

1.7.4. Repositorio Semántico.

Existe una nueva generación de repositorios flexibles en la que todas las conceptualizaciones del término objeto de aprendizaje tienen lugar: el repositorio semántico.

Los repositorios son semánticos cuando incluyen la propiedad de analizar la semántica de su contenido, es por eso que la web 2.0 se llamaría web semántica.

Repositorio de Objetos de Aprendizaje Semántico (SLOR).

El Repositorio de Objetos de Aprendizaje Semántico (SLOR) proporciona 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 esquema permite almacenar en un repositorio cualquier tipo de información sobre un objeto de aprendizaje.

Este repositorio tiene como dominio de una Ontología los estándares LOM y SCORM. En ella cualquier concepto enlazado con la representación de una actividad de aprendizaje es considerado un OA.

Los clientes del repositorio pueden entre otras funcionalidades, añadir, recuperar, modificar y buscar objetos de aprendizaje sin importar la conceptualización utilizada por el creador. Por ejemplo, un cliente software utilizando un modelo basado en LOM, podría recuperar objetos de tales repositorios, incluso aquellos objetos que estén albergados en otro tipo de sistemas con otras especificaciones de ejecución (como por ejemplo sistemas de administración de contenidos de aprendizaje (LCMS) basados en SCORM).

El e-Learning se basa en estándares educativos para permitir la interoperabilidad y reutilización de sus contenidos. Es por ello que todo el contenido que se encuentra en un repositorio de objetos de aprendizaje debe cumplir con estos estándares, los que hacen posible la comunicación entre las diferentes plataformas.

En el Capítulo 3 con la implementación de una versión previa, se pone en práctica un prototipo de repositorio semántico de objetos de aprendizaje, en donde se define como dominio de la Ontología los estándares educativos.

1.8. Conclusiones del Capítulo.

En este capítulo se expusieron los principales conceptos relacionados con esta investigación, así como las posibles soluciones a los problemas planteados en la situación problemática. Fueron estudiados los aspectos relacionados con las Ontologías, Web Semántica y la gestión del conocimiento llegando a la conclusión de que no existen plataformas para la GC que se basen propiamente en Ontologías. Se mostraron las principales herramientas y lenguajes utilizados para construir y representar Ontologías.

Para la realización de este trabajo se ha decidido utilizar las siguientes tecnologías:

OWL debido a que está diseñado para usarse cuando la información contenida en los documentos necesita ser procesada por programas o aplicaciones, permite definir clases mediante condiciones sobre sus miembros, mediante combinación booleana de clases o por enumeración de las instancias que pertenecen a la clase. Además OWL permite atribuir ciertas propiedades a las relaciones, como cardinalidad, simetría, transitividad, o relaciones inversas. [Abián 2005]

Protégé para construir las Ontologías debido a la transparencia de su trabajo con OWL a pesar de tener su propio lenguaje interno para definir Ontologías, cuenta con un entorno abierto y fácil de entender, que ha generado en torno suyo toda una comunidad que contribuye activamente a ampliar el entorno con todo tipo de contribuciones en forma de plug-ins y está haciendo de esta herramienta un entorno sumamente potente. [Castells]

SPARQL como lenguaje para realizar consultas a los datos explícitamente representados en OWL.

Jena para visualizar el contenido de las Ontologías, permite trabajar con OWL, está dentro de la línea de Java, es gratuito y de código abierto.

Pellet como razonador para validar sintácticamente la Ontología, está dentro de la línea de Java y se puede trabajar utilizando OWL y Jena.

Capítulo 2. Propuesta Metodológica.

2.1. Introducción al Capítulo

En este capítulo se muestra una propuesta de desarrollo para la cual se han aprovechado las facilidades que brindan las herramientas seleccionadas en el capítulo anterior.

Como parte de esta metodología se han identificado tres fases que son fundamentales para la obtención de la misma. Cada una de estas etapas contiene un conjunto de pasos que modelan el orden necesario para llevarla a cabo. La primera etapa corresponde al proceso de construcción de la Ontología, la segunda, consiste en la etapa de inferencia, como última fase se propone la interacción entre el usuario y el sistema.

2.2. Primera Fase: Proceso de Construcción de la Ontología.

El desarrollo de las Ontologías es un proceso evolutivo e iterativo a través del cual se van refinando y completando sus detalles. Es necesario recordar que ellas modelan la realidad del mundo y todos los conceptos relacionados deben estar claramente reflejados en su representación. Para su elaboración no existen estándares por los cuales se pueda regir. La mejor solución es aquella que cubra todas las necesidades de la aplicación teniendo en cuenta la definición más exacta del dominio en cuestión.

Como se menciona anteriormente la primera fase de la propuesta la constituye la construcción de la Ontología. Existen varias metodologías que son aplicables en dependencia del objetivo de la aplicación o proceso en el que se utilice, entre ellas se destaca: *“Methontology, que lleva a cabo su cometido mediante las tareas de especificación, adquisición de conocimiento, conceptualización, integración, implementación, evaluación y documentación de las Ontologías”*.

Otra metodología orientada a proceso es *“On-To-Knowledge (OTK) la cual incluye una identificación de metas, objetivos, que son logrados con herramientas de soporte para el*

manejo de conocimiento. Las distintas fases que promueven son: estudio de factibilidad, fase inicial, refinamiento, inferencia, evaluación y aplicación-evolución”.

“Cyc KB: consiste en la codificación de información extraída a mano y la adición posterior de más información usando herramientas de soporte y formalización.”

“Uschold y King: se basa en identificar el propósito para posteriormente construir, evaluar y documentar las Ontologías.”

“Grüninger y Fox: basada en la identificación de escenarios y la formulación de preguntas de competencia (Competency Questions), extracción de conceptos y relaciones relevantes y la formalización en Lógica de Primer Orden” [Samper 2005].

Varios autores han comparado estas metodologías atendiendo a diferentes criterios (Ver anexo 7), lo cierto es que no se puede decir qué tan buena es una y qué tan mala es otra, pues cada una puede ser perfectamente usada según el proceso para el que fueron concebidas.

Después de ver las metodologías anteriores, se proponen los siguientes pasos como parte de la primera etapa de la propuesta, la cual no pretende crear un patrón sino una guía que sirva de apoyo para trabajar con las mismas.

Nº 1.

Para la construcción de Ontologías primeramente se debe **determinar el dominio**. Una correcta determinación del dominio de aplicación de un problema ayuda a conocer sus fronteras y mejora la comunicación entre todas las partes involucradas. [The Reuse Company]

Las respuestas a las siguientes preguntas pueden ser de gran ayuda a la hora de definir su dominio y alcance:

1. ¿Cuál es el dominio que cubrirá la Ontología y para qué se usará?

2. ¿Quién la usará y mantendrá?
3. ¿A qué tipo de preguntas la información contenida en las Ontologías dará respuestas?

Las preguntas que responden al alcance de la Ontología se les conocen como preguntas de competencia. La base de conocimiento debe ser capaz de responder a ellas, además, serán utilizadas para futuras pruebas de control de calidad.

Nº 2.

Después de seleccionado el dominio es recomendable **buscar si existen otras Ontologías** que guarden relación con la que se quiere construir. Hay muchas Ontologías documentadas en Internet* par diversos dominios.

Nº 3.

Resulta importante definir una lista en la que estén incluidos los **términos más formulados por el usuario** y que estén relacionados con el dominio. La lista debe estar escrita en un lenguaje que sea entendible por los clientes, conteniendo para cada término diversos sinónimos.

Nº 4.

La información presente en las Ontologías se **estructura mediante relaciones entre clases y subclases**. A través de esta jerarquía se describen los conceptos y la interacción entre ellos. Este proceso se puede realizar de tres formas, la primera es definiendo desde los conceptos más generales hasta los más específicos dentro del dominio (top-down). La segunda consiste en el proceso inverso, o sea, ir desde lo más específico hasta lo más general (bottom-up) y la tercera vía es la combinación de los dos procesos anteriores, primero se detallan los conceptos más relevantes y luego se generalizan y especifican apropiadamente.

* SWoogle es un sistema para la indización y recuperación de documentos escritos en RDF, OWL y DAML. Disponible en: <http://swoogle.umbc.edu>.

Nº 5.

En estos momentos se tienen reflejadas en el modelo todas las clases y sus relaciones, ahora se deben **definir las propiedades o SLOTS** de las mismas. Los SLOTS son los que describen las cualidades internas de los conceptos. Representan las propiedades distintivas de los objetos.

Por ejemplo una Ontología que tenga definido como dominio los animales tendrá como propiedades la edad, el área en donde habitan, tipo de comida de la que se alimentan, con qué otros animales se relacionan, número de patas, etc.

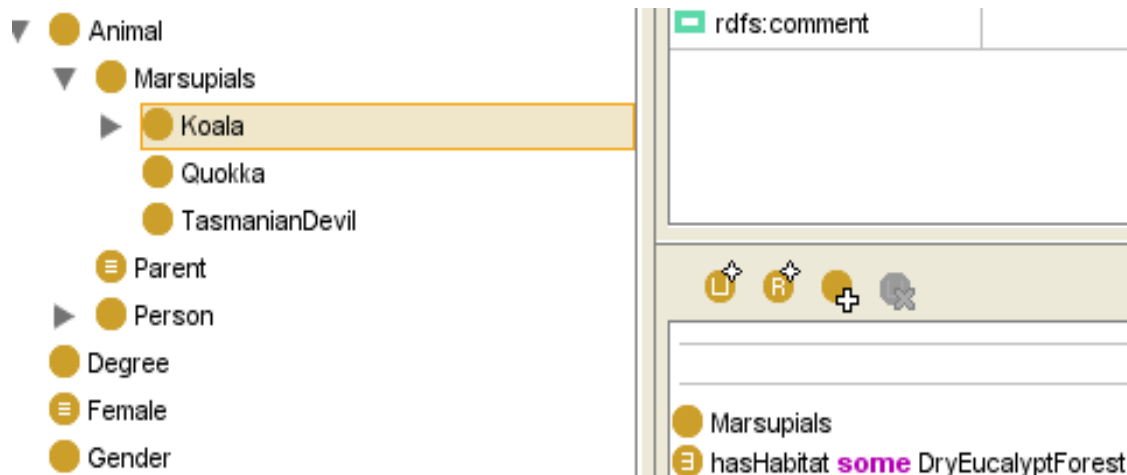


Figura 1: Ejemplo de una propiedad utilizando Protégé.

En la figura se muestra cómo se puede definir que los koalas habitan en bosques donde haya eucalipto mediante la propiedad hasHabitat (tiene hábitat).

Después de definidas las propiedades y la **cardinalidad** entre estas y las clases, se procede a **crear instancias**, las cuales definen los valores de sus propiedades. Incluir ejemplo de código OWL.



Figure 2 Ejemplo de una instancia.

La figura anterior muestra cómo se puede declarar que el género puede ser femenino (female) o masculino (male). Mediante la declaración de instancias de la clase género (gender).

Nº 6.

Para saber si la Ontología está correcta se debe hacer una **evaluación del modelo conceptual**, analizando detenidamente la taxonomía de la ontología, comprobando si mediante ella se logra representar el mundo real, exactamente igual a como lo ve el usuario final, verificar que no existan contradicciones ni redundancias con la semántica en alguna parte del modelo.

No existe una herramienta que permita evaluar este modelo, pero sí se puede examinar periódicamente la ontología por un experto o mediante la interacción de la ontología con el sistema, en ambos casos se deben plasmar en un documento todas las observaciones encontradas.

Nº 7.

Después de realizados los pasos anteriores se obtiene como resultado una ontología bien estructurada que cuenta con conceptos y relaciones entre ellos, esto permite que se puedan hacer inferencias y dar respuestas con mayor objetividad a las interrogantes formuladas por los usuarios o sistemas.

La principal característica de la web semántica es la capacidad de inferir información a través de conocimientos básicos. Mediante la **confección de axiomas y reglas de inferencia** se hace posible el razonamiento de esta futura web, es importante aclarar que el uso de estas no es necesario para el caso en el que sólo se necesite consultar los datos, a esto se le conoce como ontología de concepto.

Nº 8.

Para **relacionar la Ontología ya terminada con bases de datos**, se puede utilizar el Lenguaje R2O (relational two Ontology), que propone el enlace entre los datos almacenados en la base de datos y las instancias correspondientes a la Ontología.

Paralelamente a los pasos anteriores se debe **ir documentando** el por qué de las decisiones tomadas, la evaluación realizada y el conocimiento adicional para usarla. Además las Ontologías deben ser indexadas y colocadas con las existentes, posibilitando su futura reutilización.

Antes de terminar con esta fase se recomienda tener un especial cuidado con los nombres de las clases, instancias, propiedades, evitar las faltas de ortografía y establecer una guía sobre el uso de mayúsculas y minúsculas. Muchas veces no se le presta atención a estos aspectos que embellecen o empobrecen el trabajo, es válido recordar que la ontología puede ser usada por varios desarrolladores, a los que esto les puede causar confusión, desagrado y hasta repudio.

2.3. Segunda Fase: Inferencia.

En este período se van a vincular la mayoría de las herramientas seleccionadas en el capítulo anterior, de ahí que sea esta la fase de mayor complejidad. En ella se pueden detectar errores de la etapa anterior, en caso de detectarlos se tendría que retroceder para corregirlos. Sólo cuando se llegue al final de esta fase, se puede decir que ha concluido la fase anterior.

Esta etapa comienza cuando mediante Jena se crea un modelo para cargar la Ontología y para ser utilizado posteriormente por el razonador. Paralelamente al proceso de construcción del modelo, se indica qué razonador es el que se va a utilizar. Jena toma del modelo los axiomas, las propiedades, el dominio de las mismas y las direcciones físicas de las clases de la Ontología en forma de URL. Por otra parte, Pellet extrae del modelo los Abox, Tbox y una colección de subgrafos más específicos que modelan las

instancias, propiedades, clases, axiomas e incluso la unión de algunos subgrafos. Con esta información el razonador inferirá nuevas reglas para leer las clases y subclases de la Ontología, las que junto a los Abox y Tbox formarán parte de la Base de Conocimiento.

Un Motor de Razonamiento en base a Lógica Descriptiva, asocia dos mecanismos internos en su entendimiento del conocimiento: Abox y TBox. El primero contiene aserciones sobre objetos, específico para las instancias del dominio de discurso. El segundo contiene el conocimiento normativo en forma de una terminología y es construido a través de declaraciones que describen propiedades generales de conceptos.

Ejemplo [Samper 2005]:

Tbox: Hombre: Humano \wedge Masculino; Padre_Feliz: Hombre \wedge tener_Hijos_Hembras

Abox: Juan: Padre_Feliz; (Juan, Mari): Tener_Hijo_Hembra

Las bases de conocimiento (KB) son la evolución lógica de los sistemas de bases de datos tradicionales (BD), en un intento de plasmar no ya cantidades inmensas de datos, sino elementos de conocimiento, normalmente en forma de hechos y reglas, así como la manera en que éste ha de ser utilizado. [SAMPER 2005] Las KB proveen un mayor soporte para la inferencia, encontrando sobre el modelo respuestas que no han sido dadas explícitamente e involucrando menor número de datos pero con una mayor complejidad.

Cada uno de estas tecnologías tiene su forma para acceder a la Ontología:

Cuando Jena carga el modelo puede acceder a las clases, subclases, instancias y propiedades mediante direcciones URL que apuntan a la localización de estos elementos dentro de la Ontología.

Pellet por su parte accede a la Ontología a través de grafos e intersecciones de estos que están almacenados dentro de la base de conocimiento e indican cómo acceder a la Ontología.

En ocasiones Jena y Pellet se comportan como si fueran un sólo componente, pues el razonador toma las clases que tiene Jena, las especializa y las utiliza como si fueran suyas para acceder a la Ontología.

SPARQL primeramente necesita que le declaren los espacios de nombre o namespace con los cuales va a trabajar, estos son especificados en forma de URIs como se puede ver a continuación:

RDF: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
XML Schema: <http://www.w3.org/2001/XMLSchema#>
RDFS: <http://www.w3.org/2000/01/rdf-schema#>
OWL: <http://www.w3.org/2002/07/owl#>

Este lenguaje para acceder a la Ontología utiliza URIs, a las que se le especifica el lugar donde se efectuará la búsqueda^{*}.

El siguiente ejemplo muestra cómo se hace referencia a una Ontología llamada Periodic Table:

<http://www.daml.org/2003/01/periodictable/PeriodicTable#>

Luego se formula la consulta de la siguiente forma: sujeto-predicado-objeto o sea ?element table:name ?name.

```
PREFIX table: <http://www.daml.org/2003/01/periodictable/PeriodicTable#>
SELECT ?name
FROM <http://www.daml.org/2003/01/periodictable/PeriodicTable.owl>
```

^{*}Guía rápida sobre SPARQL. Disponible en: <http://www.dajobe.org/2005/04-sparql/SPARQLreference-1.8.pdf>

WHERE { ?element table:name ?name. }

Esta consulta devuelve el nombre de todos los elementos de la tabla periódica.

En este momento ya se pueden realizar consultas a la Ontología utilizando conjuntamente Jena con Pellet quienes a través de SPARQL acceden a la Ontología para efectuar las consultas o pueden hacerlo directamente. De esta manera se verifica si la Ontología responde a las preguntas para las cuales fue creada y por ende su consistencia, dándose por concluida la primera y segunda fase, sólo nos resta incluirle al sistema una interfaz donde el usuario pueda formular correctamente su pregunta y obtener resultados.

2.4. Tercera Fase: Interacción entre el Usuario y el Sistema.

Para realizar búsquedas e interactuar con la Ontología se debe contar con una interfaz similar a un buscador para realizar las consultas. Para hacer una búsqueda se introducen las palabras clave relacionadas con lo que se desea buscar. Estos términos pueden estar entre comillas, separados por espacios, guión bajo o medio*, operadores lógicos, operadores de existencia y de aproximación.

La expresión formulada por el usuario debe ser previamente analizada por el sistema antes de efectuar la búsqueda. Cuando además de vocablos se utilizan operadores para formar las expresiones de una consulta, se tiene mayor riqueza en las búsquedas. Algunos buscadores incluyen asistentes para realizar consultas más avanzadas mediante la elaboración de expresiones que permiten una mejor definición del espacio de búsqueda. Por ejemplo, si se quiere encontrar una página que hable de frutas, que no trate de manzanas, pero sí de uvas y peras, que el precio sea menor que 300€, se formularía de la siguiente manera: "Fruta" –manzana +uva +pera precio<300€ y se obtendría el resultado esperado.

* Algunas personas al guión bajo y medio le denominan underscore y score.

La delimitación de fechas pudiera ser otra forma de hacer búsquedas avanzadas, en la que se restringe el dominio de la búsqueda. Si el usuario quisiera ver qué documentos fueron publicados del 1^o al 9 de mayo se puede establecer un rango de fechas o se puede buscar la última, las dos o las tres semanas antes, el último mes o año, etc.

Los usuarios que interactúen con este tipo de sistemas podrán realizar búsquedas de diversas formas:

- ❖ Dominios: En caso de utilizar más de un dominio, se puede especificar en cuál de ellos se desea buscar. Si se quiere utilizar más de uno se especificarán los dominios y el sistema buscará según los indicados por el usuario. En caso de que sólo se cuente con un solo dominio no será necesario especificarlo.
- ❖ Nombre: Si se conoce cómo se llama el recurso que se desea encontrar será suficiente con poner el nombre del mismo.
- ❖ Fecha: Ya descrito.
- ❖ Expresiones usando Operadores: Debido a la gran utilidad que tienen las expresiones que emplean operadores a la hora de buscar, resulta imprescindible su uso en este tipo de aplicación.

La información almacenada en la base de datos tiene que estar correctamente estructurada. Cada recurso tiene anexo un metadato en donde se describen las propiedades del mismo, esto permite que se puedan realizar las siguientes búsquedas:

- ❖ Categorías: Para el caso en que la información esté ordenada de esta forma, el recurso será guardado y consultado por esa categoría cada vez que el usuario lo especifique.
- ❖ Metadatos: Se puede buscar por cada una de las etiquetas de los metadatos.

Después de formulada la consulta y seleccionadas las palabras fundamentales para la búsqueda, estas serán enviadas mediante Jena al razonador Pellet, quien ya tiene definidas las reglas necesarias para acceder a las clases y subclases de la Ontología a través de SPARQL. La Ontología las tomará, si es necesario inferirá nuevos datos, los que examinará en los recursos relacionados que se encuentran en la base de datos, utilizando para ello R2O (Ver Anexo 9). Para devolver los resultados se hará el proceso inverso.

Para el caso que se esté buscando documentos, una vez obtenidas las respuestas relacionadas con los criterios introducidos por el usuario se procede a organizarlos. Para ello se debe implementar algún algoritmo de ordenamiento, definido por el desarrollador del sistema. Una posible forma de hacerlo sería consultando un archivo que contenga el número de veces que se encuentra el término en un documento, el mismo tendría la estructura Término-Documento-Frecuencia y serviría para saber el grado de relevancia que tiene el documento. Esta frecuencia definirá el orden en que serán mostrados los resultados, contendrá el contenido, una pequeña descripción del mismo, la fecha en que fue publicado, el formato y el tamaño en kb*.

Los resultados de la búsqueda pueden ser mostrados en forma de párrafo, tabla, lista, la que puede estar ordenada alfabéticamente, por el grado de relevancia, por fecha de publicación o por alguna etiqueta de los metadatos asociados al recurso. Si es una imagen puede estar de un lado el dibujo y del otro los datos, todas estas formas pueden ser seleccionadas por el usuario o por el administrador del sistema.

La búsqueda semántica se diferencia de la tradicional en que la primera no trata de encontrar un literal entre toda la información del sistema, sino que ofrece al usuario la posibilidad de explicar qué significa ese literal. Por ejemplo, si el usuario está interesado en localizar ferias de calzado en Alicante, con el buscador semántico seleccionará el contenido Feria, se le mostrarán los campos (atributos) de ese concepto, y rellenará los

*Un **kilobyte** "es una unidad de medida común para la capacidad de memoria de las computadoras. Es equivalente a 1024 bytes. Se abrevia como KB, K, Kbyte." [Wikipedia 2007].

campos pertinentes, a diferencia del sistema tradicional en el que introduciría simplemente el literal "ferias calzado Alicante", ofrecería información variada sobre feria, calzado y Alicante, los resultados mostrados no tienen nada que ver con lo que realmente el usuario necesita. Para encontrar lo que realmente desea tendrá que realizar una búsqueda manual entre las opciones que aparecen, ocasionándole dificultad y pérdida de tiempo.

Para el caso en que la búsqueda semántica no sea tan precisa se puede recurrir a la búsqueda por palabras clave, estas son las que se hacen hoy en la web. Básicamente existen dos formas de realizar búsquedas: por temas, en la que se organiza el contenido en áreas generales y sub-temas, se puede buscar en ellos ingresando en cada categoría, por ejemplo, para buscar información sobre el actor Antonio Banderas, primero habrá que seleccionar la sección Entretenimiento; dentro de ésta, Cine y después Actores. Por palabras clave, en esta se le solicita al visitante uno o más términos para utilizar como criterio de la búsqueda, luego retornan las páginas halladas en cuyos textos se incluyen estas palabras, por ejemplo, textos que contengan la palabra "actor".

Entre los buscadores más conocidos por palabras clave se destacan Google, Altavista y Lycos y por temas Yahoo, Olé e Infospace. Cada uno de ellos incluye su propio mecanismo de RI el que modifican, adaptan y aplican a sus sistemas.

AltaVista

Ordena los documentos según una puntuación que obedece a los siguientes criterios: Las palabras o frases requeridas se encuentran en las primeras palabras del documento (por ejemplo el título). Cuanto más cerca se encuentren entre sí las palabras o frases requeridas. Cuantas más veces aparezcan las palabras o frases solicitadas. En la página de resultados de búsquedas simples no aparece la puntuación, mientras que en las avanzadas sí.

Altavista busca documentos que contengan una o varias de las palabras clave, ordenados de forma que los documentos que contienen mayor número de palabras clave son presentados antes.

Altavista considera las frases como palabras separadas por espacios. Para que considere una frase completa, debe encerrarse entre comillas, o bien sustituir los espacios entre las palabras por ";".

Cuando las palabras clave se introducen en minúsculas, Altavista buscará esas palabras tanto en minúsculas como en mayúsculas; mientras que si se introduce alguna letra en mayúscula, buscará la coincidencia exacta.

Se puede forzar con palabras requeridas o prohibidas, poniendo delante de la palabra un + o un - respectivamente. Asimismo se puede utilizar el * como comodín, por ejemplo para buscar palabras derivadas de una misma raíz.

Yahoo jerarquiza de acuerdo a lo siguiente:

Número de Palabras Claves Encontradas: A mayor número de palabras claves encontradas, mayor la calificación que obtendrá el documento.

Peso de Acuerdo a la Sección en el Documento: Los documentos que tengan las palabras claves en el título tienen mayor calificación que los que se encuentran en el cuerpo o en su URL.

Generalización de Categorías: Las palabras que se encuentren en categorías superiores de acuerdo a la jerarquización tipo árbol de Yahoo tendrán mejor calificación que las encontradas en categorías más restringidas [Aste]

Algunos de los mecanismos de RI utilizados en la WWW son:

Los **multibuscadores** o **metabuscadores**, que proporcionan la posibilidad de buscar en un número determinado de motores de búsqueda de forma simultánea. *"Existen multibuscadores que presentan los resultados de forma concatenada, es decir para cada motor interrogado se presenta una lista de los resultados obtenidos; y otros que permiten obtener los resultados de forma integrada, eliminando los duplicados e indicando para cada resultado qué buscador o buscadores lo han proporcionado. En el primer caso el servicio es más rápido ya que a medida que cada uno de los buscadores va devolviendo sus resultados éstos son presentados inmediatamente al usuario; en el segundo caso, hay que esperar que todos los buscadores hayan devuelto sus resultados y que el multibuscador confeccione su propia lista, generalmente ordenada por algún criterio de relevancia".* [Delgado 1998]

Otra forma de RI son los sistemas **agentes**, estos son programas que buscan información a intervalos definidos sin requerir la presencia del usuario. Los agentes pueden realizar funciones de búsqueda, discriminación y selección. Así se puede distinguir los siguientes tipos de agentes: Agentes vigilantes, Agentes aprendices, Agentes compradores, Agentes de recuperación de información y Agentes ayudantes. [Delgado 1998]

Los **canales** son agentes de búsqueda que no están implementados en el cliente sino en el servidor. El servidor publica su información de tal manera que los agentes sean capaces de encontrarla y reconozcan su valor para sus usuarios, los que pueden personalizar lo que se les entrega y elegir cómo y cuándo quieren ver esa información, disponiendo además de capacidad de almacenamiento. Un usuario puede configurar un perfil personal que puede usar para asegurarse de que sólo se le entregará la información que le es de interés. Se le pueden proporcionar enlaces para obtener más información sobre un tema. También puede especificar cuan frecuentemente desea las actualizaciones. Además puede obtener una variedad de objetos: aplicaciones Java, objetos multimedia, texto, etc.

Para lograr la interoperabilidad entre sistemas se debe tener presente la compatibilidad entre los metadatos, de esta forma será posible organizar, localizar, recuperar, procesar, evaluar, intercambiar y reutilizar su contenido. Además el lenguaje de etiquetado semántico OWL va a permitir la comunicación entre las aplicaciones.

Una vez analizados los aspectos necesarios para establecer una relación adecuada entre el usuario y el sistema damos por terminada esta fase y con ello esta metodología. Se debe destacar que la existencia de un *portal* apoyado por un navegador semántico aportaría una mejor GC.

Ejemplo de navegadores semánticos son:

Tabulator, trabaja como un navegador dentro de otro navegador, se quiere que con el paso del tiempo la evolución lo convierta en un software independiente. Es un programa de código abierto, basado en Ajax, funciona dentro de Firefox aunque es necesario solventar un pequeño problema de seguridad y el widget de Opera. La idea es navegar por recursos de forma escalable, lo que supone que no necesita cargar en memoria toda la información contenida dentro del fichero que está visualizando (generalmente en RDF, pero no de forma excluyente). La información se va mostrando conforme el usuario la va necesitando. [Senso 2007]

Piggy Bank, se trata de una extensión escrita en Java para el navegador Firefox que hace posible extraer determinados elementos clave de una página web y almacenarlos en RDF. Dependiendo de la información que encontremos en una página web, Piggy actuará de dos maneras diferentes, si el sitio tiene un fichero RDF o meta tags del HTML, el programa capturará esa información y la integrará en un repositorio, a modo de base de datos local, organizada en función de la estructura descrita. Si por el contrario, el sitio no dispone de información de este tipo, el software invocará a un scraper para que extraiga esta información y la estructure. [Senso 2007]

Los scrapers son programas capaces de trabajar con cualquier texto para procesarlo y estructurarlo. De hecho, son muy empleados por los buscadores de Internet como anexo al trabajo realizado por sus arañas. [Senso 2007]

2.5. Conclusiones del Capítulo.

En este capítulo se presentó una propuesta metodológica para la GC mediante el empleo de Ontologías y se determinaron las fases que se deben seguir para llevar a cabo un correcto cumplimiento de la misma.

Se presentaron los principales aspectos para el desarrollo de esta metodología, la cual se piensa que pueda ser modificada en un futuro según se vaya profundizando más en el tema, en aras de lograr su utilización en diferentes sectores. Se ha investigado acerca de las metodologías existentes para la creación de Ontologías y algunos mecanismos de recuperación de la información utilizados en la web actual.

Para plantear cómo se realizan las búsquedas se ha indagado sobre los navegadores actuales y algunos intentos de programas para transformarlos a semánticos. Se abordaron los elementos fundamentales en el proceso de inferencia así como la vinculación de las herramientas empleadas en esta etapa.

Capítulo 3. Caso Práctico.

3.1. Introducción al Capítulo.

En el presente capítulo se pondrá en práctica la propuesta planteada en el capítulo 2. Se ejemplificará la propuesta metodológica a través de una Ontología para un dominio específico, en este caso, estándares educativos. Se abordarán algunas de las características principales relacionadas con el contenido del dominio.

3.2. Primera Fase: Proceso de Construcción de la Ontología.

Siguiendo los pasos propuestos en este trabajo primeramente se selecciona el dominio de la Ontología, como está enmarcada en contextos pedagógicos se decidió definir como dominio los estándares educativos. Para medir el alcance serán respondidas las siguientes preguntas:

1. ¿Quién usará y mantendrá la Ontología?

Esta Ontología formará parte de un repositorio de objetos de aprendizaje que será utilizada por los usuarios del mismo así como por los interesados en la educación a distancia, su mantenimiento será llevado a cabo por estudiantes de la propia universidad.

2. ¿Para qué se usará la Ontología?

La Ontología posibilitará inferir nuevos conocimientos a partir de la información que se encuentra en el repositorio. Esto permitirá que exista un mayor grado de exactitud a las preguntas formuladas por el usuario y una mayor satisfacción de este.

3. ¿A qué tipo de preguntas la información contenida en las Ontologías dará respuestas?

Los usuarios obtendrán respuestas a preguntas como:

¿Qué son los estándares educativos?

¿Para que se usan?

¿Con que otros estándares se relacionan?

Resulta necesario establecer un vocabulario que sea entendible por el usuario como planteamos en el paso 3. Existen términos que en el caso nuestro no deben faltar:

Vocabulario

Aprendizaje	Educación
Empaquetamiento	XML
<i>Entidad digital</i>	Sintaxis
Entidad no digital	Semántica
Entrenamiento	Repositorio
Especificaciones	Plataforma
Estándares	Ontología
e-Learning	Online
Formación	Objetos de aprendizaje
Información	Metadatos
Interoperabilidad	Estándares de facto

Como se plantea en el paso 4 a continuación se debe determinar la jerarquía de las clases pero para ello es importante tener presente la información sobre cada uno de los siguientes estándares: IEEE, ADL, AICC, IMS y SCORM.

IEEE.

Dentro de la IEEE se encuentra la LTSC (Learning Technology Standards Committee). Es el Comité de Estándares para Tecnología del Aprendizaje, encargado de desarrollar estándares técnicos, recomendaciones y guías para la tecnología educativa. Su principal aporte es LOM y su base principal es el estándar IMS

LOM.

El objetivo de LOM es la creación de descripciones estructuradas de recursos educativos. Su modelo de datos especifica qué aspectos de un objeto de aprendizaje deberían ser descritos y qué vocabularios se pueden utilizar en dicha descripción. Esto

se hace en un fichero XML, el cual contiene como etiquetas las 9 categorías y más de 70 elementos de LOM. LOM tiene una versión 1.0.

IMS.

Por su parte, el estándar IMS utiliza de LOM las siguientes categorías:

General.

Ciclo de Vida.

Meta información.

Técnica.

Educativa.

Clasificación.

“La información que se debe dar a conocer en la categoría general es el identificador del OA que está dado por la entrada del catálogo, el idioma y la estructura del OA. Las características del ciclo de vida es necesario ingresar el estado del OA, ya sea, borrador, final, en revisión o no disponible, para la meta información solamente se requiere el identificador de de meta información que viene dado por la entrada del catálogo, en cuanto a las características técnicas únicamente se requiere el formato del OA, para el aspecto educativo es necesario contar con el tipo de interactividad de que se trata y el usuario principal al que va destinado el OA y finalmente un identificador de clasificación para el O”. [Priego]

Además, IMS contiene las siguientes especificaciones [Baltasar]:

- ❖ IMS LIP (Learner Information Packaging): Esta especificación define estructuras XML para el intercambio de información de los alumnos entre sistemas de gestión de aprendizaje, sistemas de recursos humanos, sistemas de gestión del conocimiento, y cualquier otro utilizado en el proceso de aprendizaje. LIP tiene dos versiones: 1.0 y 1.0.1.

- ❖ IMS QTI (Question and Test Interoperability): El IMS QTI propone una estructura de datos XML para codificar preguntas y test online. El objetivo de esta especificación es permitir el intercambio de estos tests y datos de evaluación entre distintas plataformas. Versiones: 1.2, 1.2.1, 2.0 y 2.1.
- ❖ IMS LD (Learnig Design): Se ocupa de describir y codificar el diseño pedagógico, es decir las metodologías educativas implícitas en un proceso de enseñanza de forma que sean procesables por un LMS. Versiones: 1 y 1.1.0.
- ❖ IMS CP (Content Packaging): El objetivo de esta especificación es permitir la distribución de contenidos reutilizables e intercambiables, es decir, describe el modo en el que se debe empaquetar el contenido educativo para que pueda ser procesado por otro sistema LMS diferente. Un documento fundamental es el manifiesto, dicho documento es un fichero XML en el que se describe la estructura de los contenidos incluidos en el paquete y el contenido que pueden ser cursos individuales, conjuntos de cursos, o cualquier tipo de recurso necesario en el proceso educativo. Versiones: 1.1.2, 1.1.3 y 1.1.4.
- ❖ IMS SS (Simple Sequencing): Se ocupa de la definición de los mecanismos que permitan la secuenciación de los recursos educativos dentro de un sistema e-Learning. El objetivo es poder definir, por ejemplo, el orden en el que se presentan los objetos de aprendizaje o las reglas para seleccionar un objeto de aprendizaje entre varios posibles en función del comportamiento o de las respuestas del alumno. Versión 1.0.

ADL.

ADL (Advanced Distributed Learning): La misión de ADL es la de proveer acceso de la más alta calidad en educación y entrenamiento, en cualquier lugar y momento, para lo cual crean el modelo SCORM. Es un perfil de aplicación ya que combina muchas especificaciones (IMS, AICC, IEEE) y las particulariza para un caso concreto.

SCORM.

Este modelo es un conjunto de estándares y especificaciones para compartir, reutilizar, importar y exportar OA. SCORM recomienda el uso de los vocabularios definidos por LOM e indica que para una mayor capacidad de búsqueda y reutilización, algunos metadatos deben ser obligatorios. [Rebollo 2004]

Para el estándar SCORM

, las categorías requeridas son las siguientes:

General.

Ciclo de vida.

Meta información.

Técnica.

Derechos.

“Para la categoría general es requerido conocer el identificador que esta dado por un catálogo de OA y una entrada de catalogo, por ejemplo, la entrada para el catálogo ISBN sería 84-481-3339-0, también es requerido el titulo, descripción y palabras clave del OA. Para el ciclo de vida es necesario saber la versión y el estado del OA. En la categoría de meta información se requiere saber que esquema de meta información se ha usado, es decir, bajo que criterio se utilizan los metadatos. Para la categoría técnica se debe de proporcionar información que describa el formato del OA así como su ubicación. La categoría de derechos requiere que se conozca si el OA esta sujeto a costo y copyright” [Priego]. Versiones: 1.0, 1.1, 1.2 y 1.3.

AICC.

El AICC (Aviation Industry CBT Committee) es el primer organismo con normas para el intercambio de cursos. Su principal aportación es la propuesta de CMI (Computer Managed Instruction), que proporciona un conjunto de normas para el intercambio de contenidos entre plataformas de formación y para la gestión y el seguimiento de los resultados de aprendizaje. [Rebollo 2004]

Después de analizar estos estándares, las relaciones entre ellos y teniendo en cuenta que en e-Learning los estándares son estándares de facto y a su vez estos son especificaciones se obtiene como resultado el siguiente modelo:

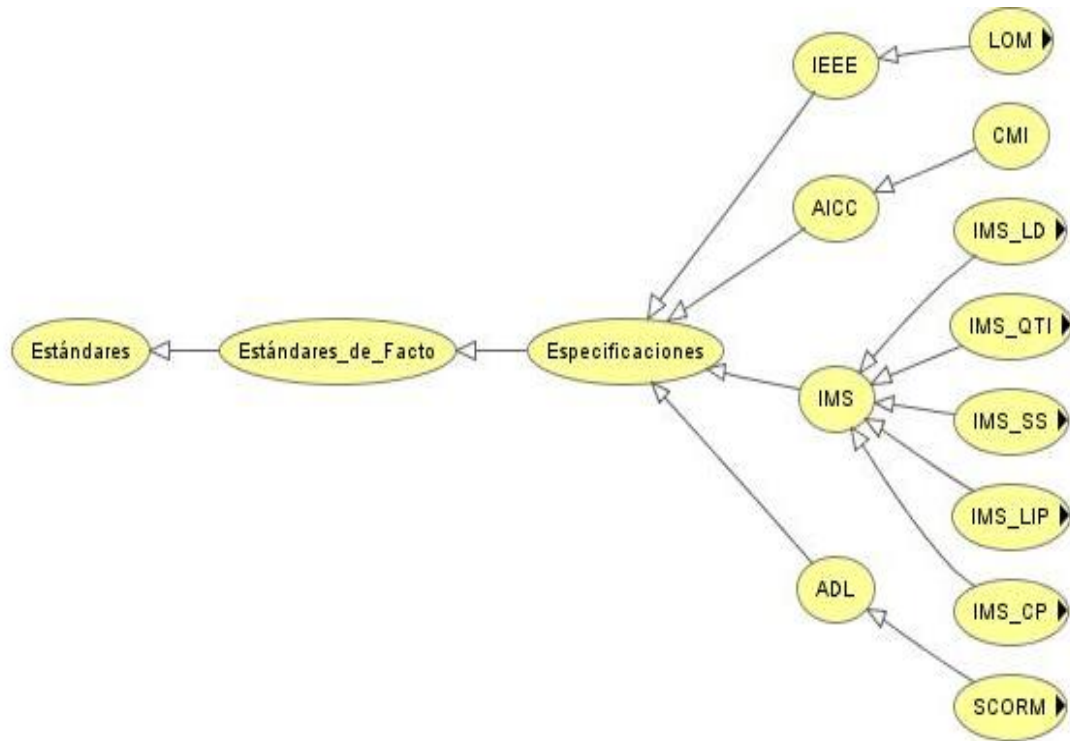


Figura 3: Jerarquía de clases sobre estándares educativos desarrollada en Protégé.

En estos momentos se cuenta con la jerarquía de clases, ahora se definen las propiedades para establecer las relaciones entre ellas como se plantea en el paso número 5.

- IMS
- ⊗ Contiene **only** Paquete
- = Contiene **exactly** 1
- = Tiene_Nombre **exactly** 1
- ⇒ Tiene_Sinonimo **has** "IMS Content Packaging Specification"
- ⇒ Tiene_Sinonimo **has** "Instructional Management Systems CP "
- ⇒ Tiene_Sinonimo **has** "Instructional Management Systems Content Packaging Specification"
- ⇒ Tiene_Sinonimo **has** " Content Packaging Specification"
- ⇒ Tiene_Sinonimo **has** "IMS Content Packaging "
- ⇒ Tiene_Sinonimo **has** "Instructional Management Systems Content Packaging "
- ⇒ Tiene_Sinonimo **has** " CP"

Figura 4: Ejemplo de Propiedades de una clase.

En la figura anterior se muestra las propiedades de la clase IMS CP, así como la cardinalidad. Esta clase solamente puede contener un paquete. Seguidamente se le adicionan valores a las instancias, concluyendo así con este paso.

Para evaluar si lo que se tiene hecho hasta el momento está correcto, se compara si la forma en que están estructuradas las clases y las propiedades permite responder las preguntas necesarias para el usuario final, como bien se plantea en el paso 6.

Preguntas:

¿Qué son los estándares de facto?

¿Qué es una especificación?

Dado un estándar ¿Con qué otro se relaciona?

¿Qué es IMS o SCORM o LOM?

¿Qué versiones existen de cada uno de los estándares?

Cada vez que se crea una nueva propiedad debe comentarse el por qué de la misma y cómo es que funciona, de esta forma si alguien hiciera uso de esta Ontología podría entender mejor cómo es que funciona. En el paso número 8 se plantea la necesidad de crear axiomas y reglas para hacer inferencias en la Ontología. En este caso es

necesario hacerlas si se conectase con la base de datos de ROA (revisar paso número 9), pero esta es sólo una prueba para verificar la metodología y las respuestas a las preguntas que se formularán pueden ser respondidas desde la propia Ontología.

3.3. Segunda Fase: Inferencia.

El objetivo de esta etapa es vincular la Ontología con la aplicación mediante el uso de las herramientas antes mencionadas. El siguiente código fue desarrollado por un grupo de estudiantes, quienes siguiendo los pasos de esta fase lograron acceder y consultar la Ontología.

Primeramente es necesario especificar la ubicación de la Ontología, en el caso que nos ocupa, esta se encuentra dentro de un servidor Apache. La manera en que va a ser referenciada es la siguiente:

```
private String url = "http://localhost/Ont/Ayuda.rdf-xml.owl"
```

Después de localizarla se puede cargar la Ontología leyendo el modelo desde Jena:

```
public void leerOntologia(String url){  
    this.modelo.read(this.getUrl());  
}
```

Una vez que fue leído se carga de la siguiente forma:

```
public void cargarOntologia(){  
    modelo = ModelFactory.createOntologyModel();  
    try {  
        manager = OWLManager.createOWLOntologyManager();  
    } catch (OWLException e) {  
        // TODO Auto-generated catch block  
        e.printStackTrace();  
    }  
}
```

Para cargar el razonador que vamos a utilizar (Pellet) se procede del siguiente modo:

```
public void cargarRazonador(){
    com.hp.hpl.jena.reasoner.Reasoner reasoner =
PelletReasonerFactory.theInstance().create();
    razonador = reasoner;
}
```

El siguiente fragmento permite validar sintácticamente la Ontología en forma de reporte:

```
public void reportesValidacion(){
    ValidityReport report = modelo.validate();
    printIterador( report.getReports(), "Validation Results" );
}
```

Si nuestra Ontología está correcta sintácticamente, una vez cargada la Ontología con Jena y Pellet se crea automáticamente la base de conocimiento correspondiente al modelo con el que se está trabajando. En estos momentos sólo se necesita definir las búsquedas que se quieren realizar.

Para ello se debe llamar a la clase Consultas, la cual contiene los métodos necesarios para la realización de este proceso.

```
public class Consultas {

    private String consulta;
    private String nameSpaces;
    private Query query1;
    private ResultSet resultados;
    private OntModel modelo;
    private leerModelo mod;
```



```
private OntClass claseUso;
private ArrayList<OntClass> listaClases;

/**
 * Aquí se crea el constructor en donde se cargan los namespaces que utiliza la Ontología
 */

public Consultas(OntModel temp){
    modelo = temp;
    mod = new leerModelo();
    listaClases = new ArrayList<OntClass>();
    nameSpaces = "PREFIX url:<" + mod.getUrl() + ">\r\n"+
        "PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>\r\n"+
        "PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>\r\n"+
        "PREFIX owl:<http://www.w3.org/2002/07/owl#>\r\n"+
        "PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>\r\n"+
        "PREFIX xsd:<http://w.w.w3.org/2001/XMLSchema#>\r\n";
}

/**
 * En el siguiente método se crea la consulta que se quiere realizar.
 */

public void crearConsulta2(String consult){
    this.query1 = QueryFactory.create(this.getNameSpaces() + consult);
}

/**
 * Una vez enviada la consulta se guardan los resultados en la variable resultados.
 */

public void enviarConsulta(OntModel model, leerModelo mod1){
    mod1.leerOntologia1();
    QueryExecution qexec =
```

```

        QueryExecutionFactory.create(this.query1,model);
        this.resultados = qexec.execSelect();
        ResultSetFormatter.out( System.out, this.resultados, this.query1 );
        System.out.println();
    }

    public void lanzarConsulta(String consulta,OntModel model, leerModelo mod1){
        this.crearConsulta2(consulta);
        this.enviarConsulta(model,mod1);
    }
/**
 * Para retornar todas las propiedades de un modelo será necesario hacer lo siguiente:
 */
    public void retornaPropiedades(){
        for (Iterator it = modelo.listOntProperties();it.hasNext();){
            OntProperty idPropiedad = (OntProperty) it.next();
            System.out.println(idPropiedad.getLocalName());
        }
    }

/**
 * Si solamente se quiere retornar todas las propiedades de una clase seleccionada y los
 * comentarios de sus propiedades se hace lo siguiente:
 */
    public ArrayList retornaPropiedadesxClass(OntClass clas){
        ArrayList<OntProperty> lisPropiedades = new ArrayList<OntProperty>();
        for (Iterator it = clas.listDeclaredProperties();it.hasNext();){
            OntProperty idPropiedad = (OntProperty) it.next();
            lisPropiedades.add(idPropiedad/* .getLocalName()* /);
            /*System.out.println("-----
            "+idPropiedad.getLocalName() +" :")

```

```

        );
        System.out.print("*****" +
            idPropiedad.getComment(""));
        System.out.println();*/
    }
    return lisPropiedades;
}

/**
 * Para retornar las clases y las propiedades de cada una
 */
public void retornaClasesAndPropiedades(){
    for (Iterator i = modelo.listClasses(); i.hasNext(); ) {
        OntClass c = (OntClass) i.next();
        if(c.getLocalName() != null){
            System.out.println( "Clase " + c.getLocalName());
            retornaPropiedadesxClass(c);
        }
        else continue;
    }
}

/**
 * El siguiente método retorna dada una propiedad el comentario de la misma
 */
public String retornaCommentDePropiedades(OntProperty prop){
    return prop.getComment("");
}

/**
 * Para retornar el dominio de la propiedad pasada como parámetro

```

```
*/
    public void retornarDominiosDePropiedades(OntProperty propiedad){
        System.out.println(propiedad.getDomain());
    }

/**
 * Retorna comentario de la clase que se le pasa como parámetro.
 */
    public String retornarComentarioDeClase(OntClass clase){
        return clase.getComment("");
    }

/**
 *Para listar los nombres de las primeras 82 clases
 */
    public ArrayList<OntClass> listarClases(){
        listaClases.clear();
        for (Iterator i = modelo.listClasses(); i.hasNext() && listaClases.size() <=82;)
        {
            OntClass c = (OntClass) i.next();
            if(c.getLocalName() != null)
                listaClases.add(c);
        }
        return listaClases;
    }

/**
 * Si dada una clase se quiere retornar las instancias de la misma se procedería de la
 * siguiente manera.
 */
```

```
public OntClass retornarClaseXNombre(String nombreClase){
    listarClases();
    OntClass p = null;
    for(int o = 0; o < listaClases.size(); o++){
if(nombreClase.equals(listaClases.get(o).getLocalName().toString())) {
        p = listaClases.get(o);
        break;
    }//while
    }return p;
}
```

/**

* Para buscar las subclases de una clase.

**/

```
public Iterator BuscarsubClases(String nombreClase){
    boolean bandera = false; //no se ha encontrado
    String nombre = null;
    OntClass c = null;

    for (Iterator i = modelo.listClasses(); i.hasNext();) {
        c = (OntClass) i.next();
        nombre = c.getLocalName();
        if(nombre != null)
            if(nombre.equals(nombreClase)){
                claseUso = c;
                bandera = true;
                break;
            }
        else continue;
    }
```

```
    }  
    return c.listSubClasses();  
}
```

Con la unión de todos estos métodos se puede consultar la Ontología desde una aplicación. Concluyendo de esta manera la segunda fase. Mediante la formulación de preguntas se pudo comprobar la consistencia de la Ontología y la metodología propuesta.

3.4. Tercera Fase: Interacción entre el Usuario y el Sistema.

Esta fase corresponde a la interacción entre el usuario y el sistema. Para lo que se debe analizar primeramente cómo es que ocurre este fenómeno actualmente en ROA. Las consultas se hacen a partir del nombre, las categorías y la etiqueta palabra clave, correspondiente al metadato del paquete SCORM que sube el usuario al repositorio. La nueva versión ROAS incluye las actuales consultas y añade la búsqueda por cada una de las etiquetas de los metadatos, dominios, fecha y formulación de expresiones utilizando operadores.

Cuando el usuario formula la consulta teniendo en cuenta los criterios de búsqueda antes mencionados, el sistema analiza cada uno de los términos de la expresión por separado y los compara con los datos que están en una tabla Hash que contiene los sinónimos de las clases y subclases de la Ontología. De esta forma se evita que se realice una consulta innecesaria a la Ontología para el caso en que la consulta esté fuera del dominio de la misma.

Para que un sistema de GC trabaje de forma eficaz debe ser local, personalizado, natural e interactivo. Debe tener un portal que responda a las necesidades e intereses del usuario, desde el acceso a la información específica justo a tiempo, hasta la participación en comunidades de práctica y la resolución de problemas complejos. Para el caso que nos ocupa, se cuenta con una plataforma de teleformación, un repositorio de objetos de aprendizaje y una herramienta de autor, las que se encuentran aisladas, sería mucho

más cómodo y útil si se contara con un portal que uniera estas comunidades, donde cada persona acceda a estas con una misma cuenta de acceso (Ver Anexo 8). Los usuarios de este portal seguirían siendo los estudiantes y profesores y su objetivo principal mejorar los resultados educativos.

3.5. Conclusiones del Capítulo.

En este capítulo se probó una parte de la metodología en un dominio específico, aunque esta no se llevó a cabo completamente sirvió para verificar una buena porción de ella y a pesar de que funciona se debería concluir con esta prueba siguiendo los pasos propuestos.

Finalmente se determinó que aunque la metodología no se aplicó en su totalidad, principalmente con el paso tres, puede ser aplicada en cualquier contexto.

Conclusiones.

Con la realización de esta investigación se ha arribado a la era de la gestión del conocimiento, espacio en donde la información recobra un nuevo significado, aproximándose cada vez más al pensamiento humano. La metodología propuesta está encaminada a gestionar el conocimiento en entornos donde la información esté estructurada aprovechando para ello las facilidades que brindan las Ontologías. En el presente trabajo se llegó a las siguientes conclusiones:

- ❖ Como resultado de esta investigación surgió una propuesta metodológica basada propiamente en Ontologías, la cual permite crear y modificar la información depositada en la base de conocimiento.
- ❖ Para la realización de este trabajo fue necesario investigar los elementos relacionados con la Web Semántica.
- ❖ Se estudiaron las diferentes herramientas que sirven para construir Ontologías.
- ❖ Fueron definidas diferentes vías para realizar búsquedas semánticas.
- ❖ Se analizaron los elementos fundamentales que componen la Web Semántica, centrándose esencialmente en la Ontología.
- ❖ Se logró representar el conocimiento utilizando mecanismos de Web Semántica.
- ❖ Se implementó una Ontología en un dominio específico.
- ❖ En la Universidad de las Ciencias Informáticas no existe una comunidad o grupo de desarrollo que se dedique al tema de la Web Semántica.
- ❖ La Universidad hasta este momento no contaba con un equipo de investigación

que se dedicara al estudio de la Gestión del Conocimiento utilizando Ontologías.

- ❖ Si se quiere lograr una mejor gestión del conocimiento se debe destacar la necesidad de crear un portal en el que se vinculen diferentes comunidades.
- ❖ Ninguno de los portales con que cuenta la Universidad tiene una estructura Semántica ni la información se encuentra organizada con metadatos.

Recomendaciones.

Esta propuesta está en sus primeros pasos, todavía falta desarrollar habilidades, conocimientos y prácticas en aras de lograr su puesta en práctica y perfeccionamiento para lo cual resulta necesario tener en cuenta las siguientes recomendaciones:

- ❖ Pensar en la idea de modificar la intranet de la Universidad para que sea semántica. De esta forma se mejoraría la eficiencia y la satisfacción de los usuarios.
- ❖ Introducir la Web Semántica dentro de los temas que pueden ser desarrollados en un futuro en la Universidad.
- ❖ Trabajar en la implementación de un Repositorio de Objetos de Aprendizaje Semántico.
- ❖ Profundizar e implementar cómo se conectan las bases de datos relacionales con las Ontologías mediante R2O.
- ❖ Utilizar diagramas UML para ilustrar las clases, subclases y relaciones entre ellas que puedan existir en la Ontología.

Referencias Bibliográficas.

1. [Abián 2005] Abián, Miguel Ángel. Ontologías: qué son y para qué sirven. (9 diciembre 2005). Disponible en: <http://www.wshoy.sidar.org/index.php?2005/12/09/30-ontologias-que-son-y-para-que-sirven>.
2. [Anbarasan 2000] Anbarasan, Ethirajan. Tim Berners-Lee, el señor de la red. (Septiembre 2000). Disponible en: http://www.unesco.org/courier/2000_09/sp/dires.htm#top.
3. [Arano 2005] Arano, Silvia. Los tesauros y las Ontologías en la Biblioteconomía y la Documentación. (2005). Disponible en: <http://www.hipertext.net/web/pag260.htm> núm. 3.
4. [Burgos 2006] Burgos Solans, Daniel. Tesis Doctoral. Estudio de la estructura y del comportamiento de las comunidades virtuales de aprendizaje no formal sobre estandarización del e-learning. (20 de abril del 2006). Disponible en: http://dspace.ou.nl/bitstream/1820/626/2/BURGOSDaniel_Resumen_Castellano_15Mayo06.pdf.
5. [Castells] Castells, Pablo. La web semántica .Disponible en: <http://www.ii.uam.es/~castells/publications/castells-uclm03.pdf>.
6. [Castells 2003] Castells, Pablo; Saíz, Francisco. Ontologías para la web semántica. (16 de octubre de 2003). Disponible en: <http://64.233.167.104/search?q=cache:qbTTDJdIHOMJ:www.ii.uam.es/~castells/docencia/semanticweb/apuntes/2-ontologias.pdf+tipos+de+Ontologias&hl=es&ct=clnk&cd=10&gl=cu>.
7. [Castells 1996] Castells, Manuel. La era de la información. Vol. 1. (1996). La sociedad en red. Alianza, editorial.
8. [Cánepa 2005] Cánepa, Carlos. Marco Conceptual para la Construcción de un Modelo de E-Learning. (2005) Disponible en: http://sisbib.unmsm.edu.pe/BibVirtualData/publicaciones/risi/N2_2005/a10.pdf.

9. [Diccionario 2001] Diccionario de la lengua española en su vigésima segunda edición. (2001). Disponible en: <http://buscon.rae.es/drael/>.
10. [Diccionario Océano] Diccionario Ilustrado Océano de la Lengua Española. Edición del Milenio. Disponible en: <http://www.oceano.com>. (Consultada el 4 de abril del 2007).
11. [Delgado 1998] Delgado Domínguez, Adelaida M. Mecanismos de Recuperación de Información en la WWW. (19 de junio de 1998). Disponible en: <http://servidorti.uib.es/adelaida/tice/modul6/memfin.pdf>.
12. [Del Moral 2005] del Moral, M^a Esther; Cernea, Doina Ana. Diseñando Objetos de Aprendizaje como facilitadores de la construcción del conocimiento. (19 de noviembre del 2005). Disponible en: <http://www.uoc.edu/symposia/spdece05/ppt/ID16.ppt>.
13. [Figuroa 2006] Figuroa, Liliana; Palavecino, Rosa. Aproximación a la diferencia entre Gestión de la Información y la Gestión del Conocimiento. (2006). Disponible en: <http://www.cibersociedad.net/congres2006/gts/comunicacio.php?id=618&llengua=es>.
14. [FCTIC 2006] Fundación CTIC, Departamento de I+D+i. Buscador semántico para un Boletín Oficial (26 de julio de 2006). Disponible en: <http://www.di.uniovi.es/~labra/cursos/ver06/pres/bopa-curso.pdf>.
15. [Hendler 2004] Hendler, Jim. Preguntas frecuentes sobre el Lenguaje de Ontologías Web (OWL) del W3C. (10 de febrero de 2004). Disponible en: <http://www.w3c.es/Traducciones/es/SW/2005/owlfaq>.
16. [Jardines 2006] Jardines Méndez, José. Educación en red: mucho más que educación a distancia. Experiencia de las universidades médicas cubanas (20 de febrero de 2006). Disponible en: http://bvs.sld.cu/revistas/ems/vol20_2_06/ems07206.htm (Consultado el 22 de Junio de 2007).

17. Leyva 2006] Leyva Leyva, David; Tamayo Avila, Daymy. HERRAMIENTAS PARA LA CREACIÓN Y GESTIÓN DE OBJETOS DE APRENDIZAJE REUTILIZABLES. (Junio de 2006)
18. [Lozano] Lozano Tello, Adolfo. Ontologías en la Web Semántica. Disponible en: <http://www.informandote.com/jornadasIngWEB/articulos/jiw02.pdf> (Consultado el 3 de febrero del 2007).
19. [Metadatos 2007] ¿Qué son los Metadatos? Disponible en: <http://antares.inegi.gob.mx/metadatos/metadat1.htm> (Consultada el 8 de febrero del 2007).
20. [Marzal 2006] Marzal García-Quismondo, Miguel Ánge; Calzada-Prado, Javier; Cuevas Cerveró, Aurora. Desarrollo de un Esquema de Metadatos para la Descripción de Recursos Educativos: El Perfil de Aplicación MIMETA. Disponible en: Repositorio del proyecto Desarrollo y soporte de herramientas para teleformación <\\10.7.1.15\documents\Manuales\LOM\DESARROLLO DE UN ESQUEMA DE METADATOS PARA LA DESCRIPCIÓN DE RECURSOS EDUCATIVOS EL PERFIL DE APLICACIÓN MIMETA.htm> (29, 4, OCTUBRE-DICIEMBRE, 551-571, 2006)
21. [Montero 2003] Montero, Yusef Hassan; Martín Rodríguez, Óscar; Martín Fernández, Francisco J. Clasificaciones Facetadas y Metadatos (I): Conceptos Básicos. (28 de Febrero de 2003). Disponible en: http://www.nosolousabilidad.com/articulos/clas_facetadas1.htm
22. [Navarro 2004] Navarro Buendía, Marcos. Proyecto fin de carrera. Evaluación de Plataformas de E-Learning de Licencia Pública. (Febrero 2004). Disponible en: <http://www.uv.es/ticape/docs/nabuen/PFC.pdf>.
23. [Ontología 2007] Ontología. (modificada por última vez el 27 enero 2007). Disponible en: <http://es.wikipedia.org/wiki/Ontolog%C3%ADa.htm>
24. [Priego] Priego, Leobardo; López, Fabiola. Metodología para el uso de estándares

- internacionales en la creación de objetos de aprendizaje. Disponible en: Repositorio del [proyecto \\10.7.1.15\documents\Manuales\LOM\Metodología.pdf](#) (Consultada en 16 de abril del 2007).
25. [Presno] Presno Quesada, Ileana; Rodríguez Perojo, Keilyn. Greenstone y Scout: dos soluciones para la gestión de contenidos implementadas en la Biblioteca Virtual en Salud de Cuba. Disponible en: <http://www.congresoinfo.cu/UserFiles/File/Info/Info2006/Ponencias/116.pdf> (Consultada el 20 de Junio de 2007).
26. [Rojas 2006] Rojas Mesa, Yuniét. De la gestión de información a la gestión del conocimiento. (16 de febrero del 2006). Disponible en: http://bvs.sld.cu/revistas/aci/vol14_1_06/aci02106.htm.
27. [Samper 2005] Samper Zapater, José Javier. Ontologías para Servicios Web Semánticos de Información de Tráfico: Descripción y Herramientas de Explotación. (5 de Julio de 2005). Disponible en: http://www.tesisexarxa.net/TESIS_UV/AVAILABLE/TDX-0628106-085805//SAMPER.pdf.
28. [Sánchez 2005] Sánchez Fernández, Luis y Fernández García, Norberto. La Web Semántica: fundamentos y breve "estado del arte". (Revista Novática Nº 178, noviembre-diciembre de 2005). Disponible en: <http://www.ati.es/novatica>.
29. [Tamayo 2005] Tamayo Avila, Daymy; Querejeta Rodríguez, Andrey. Plataforma para el desarrollo de e-learning en la Universidad de Holguín Oscar Lucero Moya. Trabajo de Diploma presentado en opción al título de Ingeniero Informático, Holguín, junio 2005.
30. [Vieyra 2000] Elizalde Vieyra, Guadalupe. Repositorios de información (Introducción). (23 de mayo del 2000). Disponible en: <http://www.fismat.umich.mx/~elizalde/tesis/node7.html>.
31. [Varas] Leonor Varas, María. Repositorios De Objetos De Aprendizaje. Disponible en: http://www.alejandria.cl/recursos/documentos/documento_varas.doc.

32. [W3C 2005] ¿Qué es el Consorcio World Wide Web (W3C)?. (10 de junio del 2005). Disponible en: <http://www.w3c.es/Consortio/>.
33. [W3C 2006] El W3C de la A a la Z. (6 de marzo del 2006). Disponible en: <http://www.w3c.es/divulgacion/a-z/#o>
34. [Wikipedia 2007] Tomado de la Wikipedia. (noviembre-junio del 2007). Disponible en: <http://es.wikipedia.org>.

Bibliografía.

1. [Abián 2005] Abián, Miguel Ángel. El Futuro de la Web XML, RDF/RDFS, Ontologías y la web Semántica. (15 de septiembre de 2005). Disponible en: <http://www.javahispano.org/licencias/>.
2. [AGTI 2007] Adhocnet Glosario de términos sobre Internet. (2007). Disponible en: <http://www.adhocnet.com/varis/glosario.htm>.
3. [Arias 2007] Dr. Sánchez Arias, Víctor Germán. Diseño de un patrimonio de recursos educativos basado en una red de acervos abiertos y distribuidos de objetos de aprendizaje. Disponible en: http://www.lania.mx/~victor/articulos/art_toa_victor_1.doc. (Consultada el 4 de febrero del 2007).
4. [Aste] Aste, Margarita. La Búsqueda en Internet. Disponible en: <http://www.quipus.com.mx/r14inter.htm> (Consultad el 12 de mayo del 2007).
5. [Baltasar] Fernández Manjón, Baltasar. Especificaciones y estándares en e-learning. Disponible en: http://reddigital.cnice.mec.es/6/Articulos/articulo_capitulo.php?articulo=2&capitulo=4 (Consultada el 14 de abril del 2007).
6. [Barrón 2005] Barrón Cedeño, Alberto. Ontologías en acción Protégé OWL. (20 de septiembre del 2005). Disponible en: <http://theory.lcs.mit.edu/~rajsbaum/cursos/web/ontologias2.pdf>.
7. [Bergman 2001] M. K. Bergman. The Deep Web: Surfacing Hidden Value. The Journal of Electronic Publishing. (Agosto del 2001). Disponible en <http://www.press.umich.edu/jep/07-01/bergman.html>.
8. [Canals 2003] Canals, Agustí. La gestión del conocimiento. (julio del 2003). Disponible en: <http://www.uoc.edu/dt/20251/>.

9. [Castells 2006] P. Castells, E. Pulido, C. Carranza, M. Rico, F. Perdrix, E. Piqué, J. Cal, R. Benjamins, J. Contreras, J. Lorés, T. Granollers. Neptuno: tecnologías de la web semántica para una hemeroteca digital. Disponible en <http://nets.ii.uam.es/neptuno/publications/neptuno-interaccion04.pdf> (consultada el 4 de noviembre del 2006).
10. [CICEI 1995] CICEI (Centro Informático y de Comunicaciones del Edificio de Ingenierías), Universidad de Las Palmas de Gran Canaria. Título: Ventajas que Proporciona Internet. (1995). Disponible en <http://www.ulpgc.es/otros/tutoriales/internet/tutor/ventajas.html>.
11. [Clarín] Clarín_com. Diccionario de términos informáticos. Disponible en: <http://www.clarin.com/suplementos/informatica/htm/glosario.htm#basededatos>. (Consultado el 10 de mayo 2007)
12. [Compañía Community4you] Compañía Community4you GmbH Asesoramiento en gestión de conocimientos y eLearning. open-EIS – soluciones prácticas para la gestión del conocimiento. Disponible en: [http://www.open-eis.com/gestion del conocimiento eLearning/oe 0 solucion.c4u infosystem46 899 detailtext,false_reiter_c4u_infosystem46_897,1.htm](http://www.open-eis.com/gestion%20del%20conocimiento%20eLearning/oe_0_solucion.c4u_infosystem46_899_detailtext,false_reiter_c4u_infosystem46_897,1.htm) (Consultada el 1 de febrero del 2007).
13. [DPCBD 2006] Diez pasos para conquistar la Biblioteca Digital. (2006).Disponible en: <http://biblioteca.unisabana.edu.co/>.
14. [Estadísticas 2007] Internet World Stats. (10 de marzo del 2007). Disponible en: <http://www.internetworldstats.com/stats.htm>.
15. [García 2004] García Jiménez, Antonio. Instrumentos de Representación del Conocimiento: TESAUROS Versus Ontologías. (2004). Disponible en: <http://www.um.es/fccd/anales/ad07/ad0706.pdf>.
16. [Glosario] Glosario de Términos de Education Training & Consulting http://www.etc.org.mx/esp/sec_6/glosario.htm (Consultado el 12 de mayo de 2007).

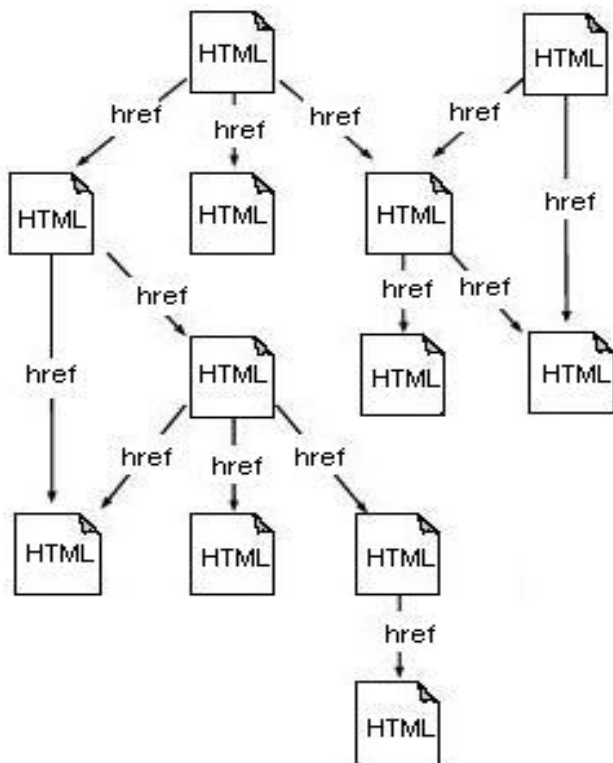
17. [Glosario 2004] Glosario. (7 de Abril del 2004). Disponible en: <http://www.w3.org/2005/03/DOM3Core-es/glosario.html>
18. [Gruber 1993] Gruber T. Toward Principles for the Design of Ontologies Used for Knowledge Sharing Technical Report KSL- 93-04, Knowledge Systems Laboratory, Stanford University, CA, 1993.
19. [GTEI 2006] Glosario de términos específicos de Internet_ Internet como recurso educativo. (19 de febrero de 2006). Disponible en: <http://internetrecursoeducativo.blogia.com/2006/021901-glosario-de-terminos-especificos-de-internet..php>.
20. [GTI 2007] Glosario de Términos Internet. Disponible en: http://www.graphicsperu.com/glosario_internet.htm (Consultado el 8 de mayo de 2007).
21. [GTM 2000] Glosario de términos de Marketing. (17 de octubre del 2000). Disponible en: <http://www.muñerdesempresa.com/marketing/marketing001001.shtml>.
22. [GTRGC 2007] Glosario de Términos Relacionados con la Gestión del Conocimiento. Disponible en: <http://www.innovapyme.com/conocimiento/diccionario.asp?inicio=0> (Consultada el 10 de mayo de 2007).
23. [Hernández] Hernández, Eduardo. Estándares y Especificaciones de E-learning: Ordenando el Desorden. Disponible en: <http://www.uv.es/ticape/docs/eduardo.pdf>. (Consultada el 12 de febrero del 2007).
24. [James 2005] Hendler, James, Knowledge is Power: the view from the Semantic Web, Robert Englemore Memorial Lecture delivered at AAAI (American Association for Artificial Intelligence). (Julio 2005). Disponible en: <http://www.cs.umd.edu/~hendler/presentations/englemore.pdf>.

25. [Kaplan 2005] Kaplan-Leiserson Eva. Glossary. <http://www.learningcircuits.org/glossary.html> (4/04/05)
26. [Lucas 2006] Carlos Lucas, Juan. Empresas 2.0 y gestión del conocimiento. (6 de mayo del 2006). Disponible en: <http://juancarloslucas.com.ar/?p=118>
27. [López 2006] López Guzmán, Clara; García Peñalvo, Francisco; Pernías Peco, Pedro. Desarrollo de repositorios de objetos de aprendizaje a través de la reutilización de los metadatos de una colección digital: de Dublin Core a IMS. Disponible en: <http://www.um.es/ead/red/M2/lopez27.pdf>. (Consultada en 4 de noviembre del 2006)..
28. [Merlo 1998] Merlo Vega, José Antonio. Suministro de documentos en Internet. (29 diciembre 1998). Disponible en <http://exlibris.usal.es/merlo/escritos/redcsumi.htm>.
29. [Piñero 2005] Piñero, Aldo. ¿Qué es la gestión del conocimiento? (2005). Disponible en: <http://www.losrecursoshumanos.com/gestion-del-conocimiento.htm>.
30. [Ramírez 2000] Ramírez Mejía, Fabián Iliusha. La Gestión del conocimiento en las Organizaciones de Sistemas. (2000). Disponible en: <http://www.tuobra.unam.mx/publicadas/010723131219-LA.html>.
31. [Rebollo 2004] Rebollo Pedruelo, Miguel. El estándar SCORM para EaD. (Diciembre 2004). Disponible en: <http://mrebollo.webs.upv.es/pubs/tesina.pdf>
32. [Réjean 1999] Gracia, Pagola y Réjean, Roy. La gestión del conocimiento y de la información textual en soporte electrónico. (Julio 1999). Disponible en: http://www.elprofesionaldelainformacion.com/contenidos/1999/julio/la_gestion_del_conocimiento_y_de_la_informacion_textual_en_soporte_electronico.html.
33. [Rodríguez] Rodríguez Echeverría, Roberto; Pablos Lamas, Felipe; Carretero Lavado, Moisés. Soporte de Metadatos en Moodle. Disponible en:

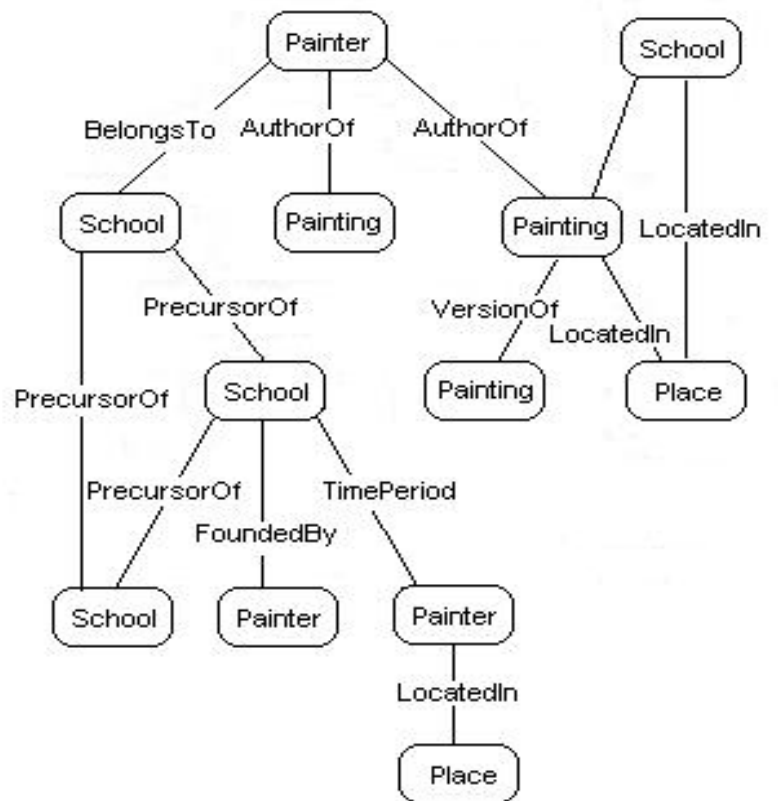
- http://campusvirtual.unex.es/cala/epistemowikia/index.php?title=Soporte_de_metadatos_en_Moodle (Consultada el 11 de noviembre del 2006).
34. [Senso 2007] Senso, José A. Navegadores semánticos o semantizar el navegador. (10 de Abril 2007). Disponible en: <http://www.thinkepi.net/repositorio/navegadores-semanticos-o-semantizar-el-navegador/>.
35. [Simón 2006] Simón, Alfredo; Estrada, Vivian; Rosete, Alejandro; Lara, Vladimir. GECOSOFT: UN Entorno Colaborativo para la Gestión del Conocimiento con Mapas Conceptuales. (2006). Disponible en: <http://cmc.ihmc.us/cmc2006Papers/cmc2006-p156.pdf>.
36. [The Reuse Company] The Reuse Company. Presente y Futuro de la Reutilización de Software. Disponible en: <http://www.webtaller.com/maletin/articulos/presente-futuro-reutilizacion-software.php>. (Consultada el 12 de abril del 2007).
37. [WWW 1996] Departamento de física aplicada I. (E.U.I.T.I. e I.T.T.). SURGIMIENTO WORLD WIDE WEB. (20 de mayo de 1996). Disponible en: <http://www.vc.ehu.es/wuagacaj/manual/web/origen.html>.
38. [Woodley 2003] Woodley, Mary S. Glosario DCMI. (15 de septiembre de 2003). Disponible en: <http://es.dublincore.org/documents/usageguide/glossary.shtml>.

Anexos.

Anexo 1.



a) Web actual



b) Web semántica

Figure 5: Diferencia entre la Web Actual y la Web Semántica.

Anexo 2.

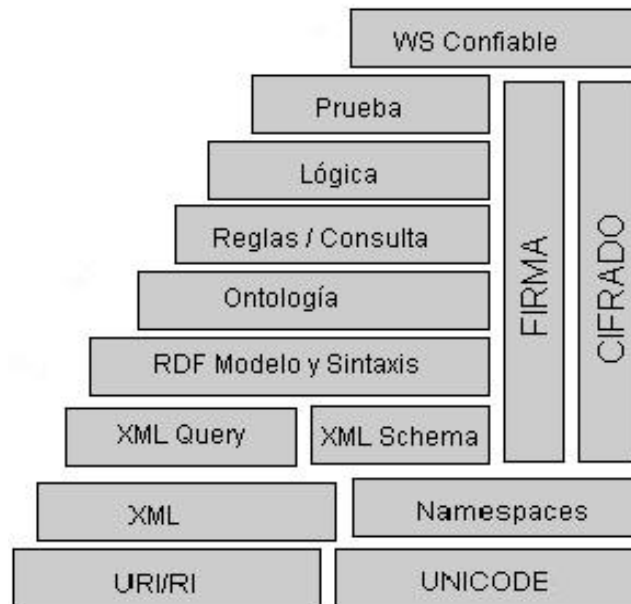


Figura 6: Estructura de la Web Semántica.

Anexo 3.

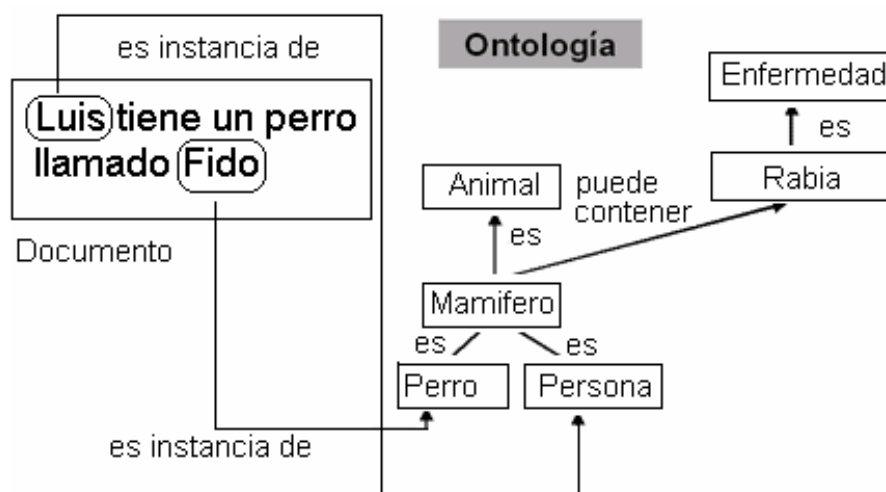


Figura 7: Usando Ontologías como metadatos.

Anexo 4.

categorias	Elementos	Valor
1. General	1.1 Identificador	
	1.2 Título	
	1.3 Entrada de catálogo	
	1.3.1 Catálogo	
	1.3.2 Entrada	
	1.4 Lengua	
	1.5 Descripción	
	1.6 Descriptor	
	1.7 Cobertura	
2. Ciclo de vida	1.8 Estructura	
	1.9 Nivel de agregación	
	2.1 Versión	
	2.2 Estatus	
	2.3 Otros colaboradores	
	2.3.1 Función	
	2.3.2 Entidad	
	2.3.3 Fecha	
	3. Meta- metainformación	3.1 Identificador
3.2 Entrada de catálogo		
3.2.1 Catálogo		
3.2.2 Entrada		
3.3 Otros colaboradores		
3.3.1 Función		
4. Técnica	3.3.2 Entidad	
	3.3.3 Fecha	
	3.4 Esquema de metadatos	
	3.5 Lengua	
	4.1 Formato	
	4.2 Tamaño	
	4.3 Ubicación	
	4.4 Requisitos	
	4.4.1 Tipo	
4.4.2 Nombre		
5. <i>Uso educativo</i>	4.4.3 Versión mínima	
	4.4.4 Versión máxima	
	4.5 Comentarios sobre la instalación	
	4.6 Otros requisitos para plataformas	
	4.7 Duración	
	5.1 Tipo de <i>interactividad</i>	activa, expositiva
	5.2 Tipo de recurso de aprendizaje	ejercicio, simulación, cuestionario, etc.
	5.3 Nivel de interactividad	0 - 4 (desde muy baja hasta muy alta)
	5.4 Densidad <i>semántica</i>	0 - 4
5.5 <i>Usuario principal</i>	profesor, autor, alumno, director	
5.6 <i>Contexto</i> [Nivel educativo]	Educación Primaria,... Formación Profesional	
6. Derechos	5.7 Edad	
	5.8 Dificultad	0 - 4
	5.9 <i>Tiempo previsto</i> de aprendizaje	
	5.10 Descripción	
	5.11 Lengua	
	6.1 Coste	
	6.2 Copyright y otras restricciones	
	6.3 Descripción	
	7. Relación [con otros recursos]	7.1 Tipo [naturaleza de la relación con el recurso principal]
7.2 Recurso [recurso principal al que se refiere esta relación]		
7.2.1 Identificador		
8. Observaciones	7.2.2 Descripción	
	8.1 Persona	
	8.2 Fecha	
9. Clasificación	8.3 Descripción	
	9.1 Finalidad	
	9.2 Nivel taxón (taxonómico)	
	9.2.1 Fuente	
	9.2.2 Taxón	
	9.2.2.1 Id	
	9.2.2.2 Entrada	
	9.3 Descripción	
	9.4 Descriptor	

Figura 8: Estructura del metadato de LOM.

Anexo 5.

<i>Característica</i>	CYC	Ushold & King	Grüniger & Fox	KACTUS	METHONTOLOGY	SENSUS	OTK
Ciclo de vida propuesto	Prototipos de desarrollo	No propuesta	Prototipos de desarrollo o incremental?	Prototipos de desarrollo	Prototipos de desarrollo	No propuesto	Incremental y ciclica con Prototipos de desarrollo
Estrategia con respecto a la aplicación	Independiente de aplicación	Independiente de aplicación	Semi-dependiente de aplicación	Dependiente de aplicación	Independiente de aplicación	Semi-dependiente de aplicación	Dependiente de aplicación
Estrategia para identificar conceptos	No especificada	Middle-out	Middle-out	Top-down	Middle-out	No especificada	Top-down, bottom-up, middle-out depende de la aplicación
Uso de una ontología base	Si	No	No	No	No	Si	Depende de los recursos disponibles para el Proyecto
Herramientas que dan soporte	Cyc	Ninguna específica	Ninguna específica	Ninguna específica	ODE WebODE OntoEdit Protégé2000	Ninguna específica (Usualmente Ontosaurus)	OntoEdit con sus plugins

Figura 9: Comparación entre metodologías existentes para construir Ontologías. [Samper 2005]

Anexo 6.

Término:

Tipo de búsqueda: Exacta

- Registros que contienen todos los términos y sus derivados
- Registros que contienen alguno de los términos o sus derivados

Resultado de la Búsqueda.

TESAURO

Término	Ingeniería
Término amplio:	Ciencias y técnicas
Término(s) específico(s):	Ingenierías técnicas

Figura 10: Ejemplo de Tesauro. [Barrón 2005]

Anexo 7.

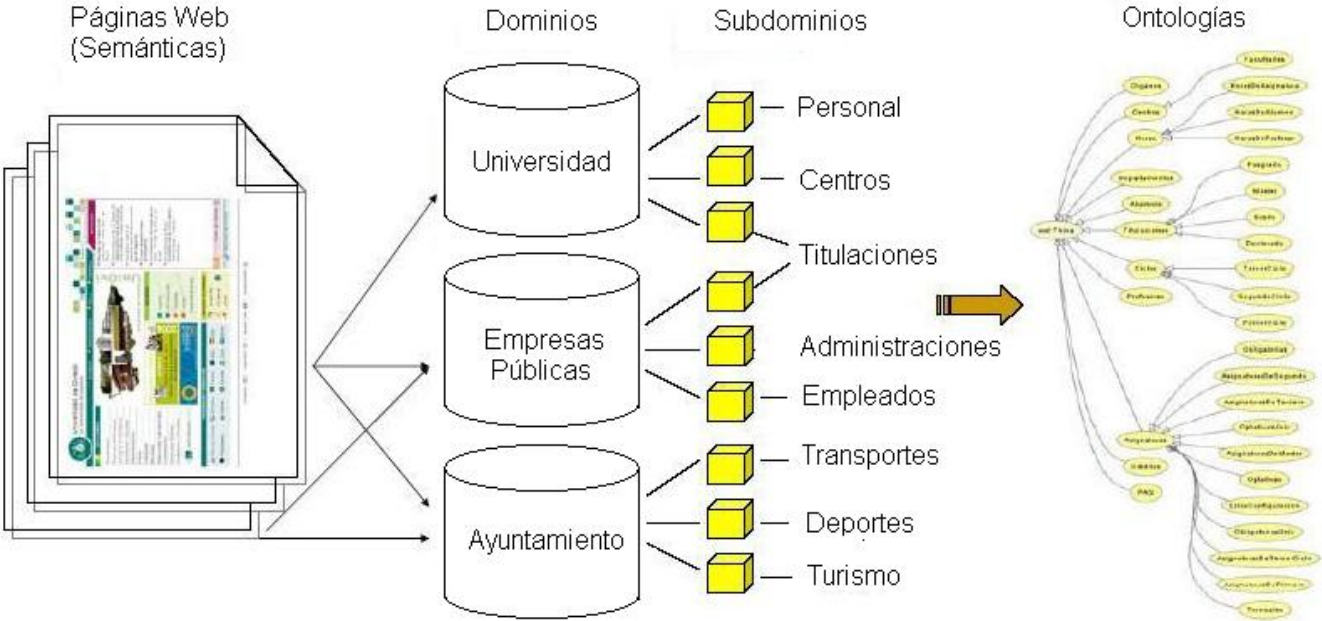


Figura 11: Representación de la web actual usando Ontologías.

Anexo 8.

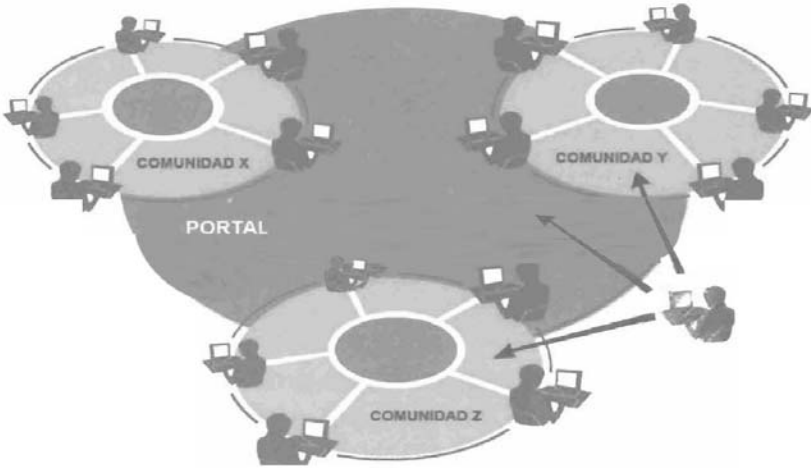


Figura12. Persona registrada en varias comunidades con una misma cuenta de acceso.

Anexo 9.

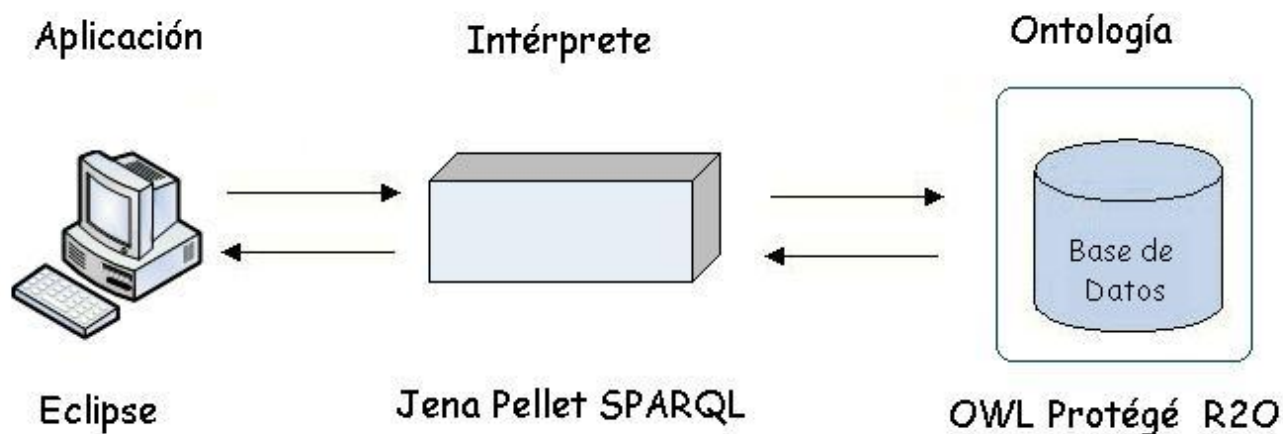


Figura 13: Funcionamiento del sistema.

Anexo 10.

WORLD INTERNET USAGE AND POPULATION STATISTICS						
World Regions	Population (2007 Est.)	Population % of World	Internet Usage Latest Data	% Population (Penetration)	Usage % of World	Usage Growth 2000-2007
Africa	933,448,292	14.2 %	33,334,800	3.6 %	3.0 %	638.4 %
Asia	3,712,527,624	56.5 %	398,709,065	10.7 %	35.8 %	248.8 %
Europe	809,624,686	12.3 %	314,792,225	38.9 %	28.3 %	199.5 %
Middle East	193,452,727	2.9 %	19,424,700	10.0 %	1.7 %	491.4 %
North America	334,538,018	5.1 %	233,188,086	69.7 %	20.9 %	115.7 %
Latin America/Caribbean	556,606,627	8.5 %	96,386,009	17.3 %	8.7 %	433.4 %
Oceania / Australia	34,468,443	0.5 %	18,439,541	53.5 %	1.7 %	142.0 %
WORD TOTAL	6,574,666,417	100.0 %	1,114,274,426	16.9 %	100.0 %	208.7 %

Figura 14: Estadísticas sobre el crecimiento de internet. [Estadísticas 2007]

Glosario de Términos.

1. **Accesibilidad:** Se refiere a la facilidad de acceso a las TIC y a la información que ofrece Internet, sin limitación alguna, para personas que poseen algún tipo de discapacidad. [GTEI 2006]
2. **API:** Es una Interfaz de Programación de Aplicaciones (Application Programming Interface), un conjunto de funciones o métodos usados para acceder a ciertas funcionalidades. [Glosario 2004]
3. **Aplicación:** Es el programa que el usuario activa para trabajar en el ordenador. Existen muchos programas de ordenador que pueden clasificarse como aplicación. Generalmente se les conoce como Software. [Kaplan 2005]
4. **Aprendizaje:** Es el proceso de adquirir conocimientos, habilidades, actitudes o valores, a través del estudio, la experiencia o la enseñanza. [Wikipedia 2007]
5. **Base de datos:** Conjunto de datos organizados de modo tal que resulte fácil acceder a ellos, gestionarlos y actualizarlos. [Clarín]
6. **Buscador, motor de búsqueda:** Es un programa, ubicado en un sitio de Internet, que recibe un pedido de búsqueda, lo compara con las entradas de su base de datos y devuelve el resultado. [Clarín]
7. **Educación:** Proceso bidireccional mediante el cual se transmiten conocimientos, valores, costumbres y formas de actuar. [Wikipedia 2007]
8. **Entidad digital:** Es un documento de cualquier tipo, ya sea imagen, texto, video, audio etc, que ha sido digitalizado, codificado e integrado con metadatos para su uso y recuperación posterior. [DPCBD, 2006]
9. **e-Learning:** Es aquella modalidad de formación a distancia no presencial o semipresencial que utiliza una metodología específica basada en las nuevas tecnologías de la información y la comunicación. [GTRGC, 2007]
10. **Formación:** Acción de suministrar a una persona, o a varias, la información y entrenamiento precisos para que conozca o aprenda a realizar y a desempeñar un nuevo papel, ejerciendo actividades y funciones nuevas. [GTRGC, 2007]
11. **Información:** Conjunto de datos interrelacionados, debidamente estructurados y asociados a una experiencia. [GTRGC, 2007]

12. Internet: Red mundial de ordenadores interconectados basada en el protocolo TCP/IP. Ofrece distintos servicios: e-mail, foros, FTP, chat, compartir ficheros, videoconferencia, etc. [GTEI, 2006]
13. Intranet: Red privada, local o no, que utiliza los mismos protocolos de Internet. [GTEI, 2006]
14. Interoperabilidad: La posibilidad de que distintos tipos de ordenadores, redes, sistemas operativos, y aplicaciones trabajen juntos de forma eficaz, sin comunicación previa, de tal forma que puedan intercambiar información de manera útil y con sentido. [Woodley, 2003]
15. Metabuscaor: Es una herramienta de búsqueda de información en la Web, que busca la información que se solicita en varios buscadores convencionales al mismo tiempo, lo cual puede ahorrarnos mucho tiempo a la hora de realizar una búsqueda, pues evita el trabajo de visitar y utilizar varios buscadores a la vez. [DPCBD, 2006]
16. Metadatos: Información estructurada que describe, explica, localiza, o maneja información especializada acerca de un recurso de información que tenga algún sentido para las máquinas o los humanos. [DPCBD, 2006]
17. Método: Es una operación o función que está asociada a un objeto y que tiene permiso para manipular los datos del objeto. [Glosario 2004]
18. Modelo: Es la representación real de los datos obtenidos a partir de la información disponible. Ejemplos son el modelo de estructura y el modelo de estilo de que representan la estructura analítica y la información de estilo asociada a un documento. El modelo podría ser un árbol, o un grafo orientado, o cualquier otra. [Glosario 2004]
19. 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. [DPCBD, 2006]
20. Online: En línea, conectado. Estado en que se encuentra una computadora cuando se conecta directamente con la red a través de un dispositivo, por ejemplo, un módem. También implica todas las actividades que realiza una persona o empresa en la red. [GTRGC, 2007]
21. 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. [Woodley, 2003]
22. Paquete (packet): La unidad de datos que se envía a través de una red. [AGTI, 2007]
23. Plataforma: Término general que normalmente designa una arquitectura de hardware, aunque a veces también se utiliza para sistemas operativos o para ambas cosas. [GTRGC, 2007]

24. Página web: Documento de hipertexto al que se accede desde Internet y que contiene textos, imágenes, sonidos, animaciones y enlaces o vínculos que permiten acceder a otras páginas. [GTEI, 2006]
25. Portal: Página utilizada para comenzar una sesión de Internet. Los portales se caracterizan por incluir información útil tal como noticias, el clima, servicio de correo electrónico y en general cualquier información relevante para el usuario. [GTI, 2007]
26. Proceso: Es un conjunto de actividades o eventos que se realizan o suceden con un determinado fin. [Wikipedia 2007]
27. Repositorio: Lugar físico donde se almacenan los documentos en forma digital. [DPCBD, 2006]
28. Semántica: Estudio del significado de los signos lingüísticos y de sus combinaciones desde el punto de vista sincrónico en la teoría lingüística generativa, componente de la gramática que interpreta la significación de los enunciados generados por la sintaxis y el léxico. [Diccionario Océano]
29. Servidor: Es un ordenador que trata las peticiones de datos, el correo electrónico, la transferencia de ficheros, y otros servicios de red realizados por otros ordenadores (clientes). [AGTI, 2007]
30. Schema (esquema): Define un conjunto de obligaciones estructurales y de valores aplicables a documentos XML. Los esquemas pueden ser expresados en lenguajes de esquema, como XML. [Glosario 2004]
31. Sintaxis: Reglas que gobiernan la estructura de un lenguaje.
32. Tecnologías de Información. Se refiere a todas las herramientas (software y hardware) disponibles para el procesamiento de datos, desde su captura, hasta la generación de información para la toma de decisiones, siendo este el objetivo final del proceso. [Glosario]
33. Web Semántica: Es una extensión de la Web actual que cuenta con información debidamente estructurada, lo que proporciona un significado bien definido. Mejora la forma en la que las máquinas y las personas trabajan en cooperación. [Samper 2005]
34. XML: El lenguaje extensible de marcas es un metalenguaje que nació como forma restringida del SGML (Standard Generalized Markup Language). [GTRGC, 2007]