

UNIVERSIDAD CENTROCCIDENTAL
"LISANDRO ALVARADO"

**VOCABULARIO BASADO EN XML Y ONTOLOGÍA PARA LA
COORDINACIÓN DE POSTGRADO DEL DECANATO DE CIENCIAS Y
TECNOLOGÍA DE LA UNIVERSIDAD CENTROCCIDENTAL LISANDRO
ALVARADO.**

CLORELBYS COROMOTO ESCOBAR QUERALES

Barquisimeto, 2010

UNIVERSIDAD CENTROCCIDENTAL “LISANDRO ALVARADO”
DECANATO DE CIENCIAS Y TECNOLOGIA
POSTGRADO EN CIENCIAS DE LA COMPUTACIÓN

**VOCABULARIO BASADO EN XML Y ONTOLOGÍA PARA LA
COORDINACIÓN DE POSTGRADO DEL DECANATO DE CIENCIAS Y
TECNOLOGÍA DE LA UNIVERSIDAD CENTROCCIDENTAL LISANDRO
ALVARADO.**

Trabajo presentado para optar al grado de
Magíster Scientiarum

Por: CLORELBYS COROMOTO ESCOBAR QUERALES

Barquisimeto, 2010

**VOCABULARIO BASADO EN XML Y ONTOLOGÍA PARA LA
COORDINACIÓN DE POSTGRADO DEL DECANATO DE CIENCIAS Y
TECNOLOGÍA DE LA UNIVERSIDAD CENTROCCIDENTAL LISANDRO
ALVARADO.**

Por: CLORELBYS COROMOTO ESCOBAR QUERALES

Trabajo de Grado Aprobado

<hr/> (Jurado 1)		<hr/> (Jurado 2)
<hr/> (Jurado 3)		

INDICE GENERAL

	PAG.
DEDICATORIA	vi
AGRADECIMIENTO	vii
INDICE DE TABLAS	viii
INDICE DE FIGURAS	ix
RESUMEN	x
INTRODUCCION	1
CAPITULO I	3
EL PROBLEMA	3
Planteamiento del Problema	3
Objetivo General	10
Objetivos Específicos	10
Justificación e Importancia	11
Alcances y Limitaciones	13
CAPITULO II	15
MARCO TEORICO	15
Antecedentes de la Investigación	15
Proyecto Galen	15
FrameWork Adoptando Reglas y Ontologías en los Servicios de Búsqueda Web	17
Proyecto SAW	21
Vocabulario en XML para la Gestión Catastral de Inmuebles Urbanos	23
Ontologías para Servicios Web Semántica de Información de Trafico	24
Ontología del Instituto Nacional de Cáncer	26
Bases Teóricas	29
Web Semantica	29
Funciones de la Web Semantica	31
RDF	32
OWL	34
Ontologias	36
Tipos de Ontologias	38
Aplicación de la Ontologia	39
XML	43
Vocabularios en XML	45
MethoOntology	48
Herramientas Protégé	52
Lenguaje de Programación JAVA	54
JENA Semantic Web FrameWork	55
Pellet OWL Razonador para JAVA	55
NetBeans	55
Bases Legales	56
CAPITULO III	58
MARCO METODOLÓGICO	58

Tipo de Investigación.....	58
Naturaleza de la Investigación.....	59
Diseño de la Investigación	59
Fases del Estudio.....	60
Fase I Diagnóstico.....	61
Universo y Muestra	61
Técnicas y Recolección de Datos.....	61
Fase II: Factibilidad.....	63
Factibilidad Técnica	63
Factibilidad Económica	64
Factibilidad Operativa	65
Fase III: Diseño de la Propuesta.....	66
Fase IV Fase Metodológica	66
CAPITULO IV.....	71
PROPUESTA DEL ESTUDIO.....	71
Introducción	71
Aplicación de la Metodología MethoOntology.....	73
Especificación	73
Dominio	74
Modelo Conceptual.....	75
Tarea 1 Glosario de Términos.....	76
Tarea 2 Taxonomía de Conceptos	87
Tarea 3 Diagrama de Relaciones Binarias Ad-Hoc.....	94
Tarea 4 Diccionario de Conceptos.....	97
Tarea 5 Detalle de Relaciones	100
Tarea 6 Instancias	103
Tarea 7 Atributos de Clase.....	103
Tarea 8 Constantes.....	105
Desarrollo Propuesta.....	110
Ontología Test.....	114
Arquitectura	116
CAPITULO V.....	118
CONCLUSIONES Y RECOMENDACIONES	118
Conclusiones.....	118
Recomendaciones.....	120
REFERENCIAS BIBLIOGRAFICAS	121
ANEXOS	125
ANEXO 1	126
Curriculum Vitae.....	130

DEDICATORIA

A Dios, por haberme dado la vida y darme la constancia necesaria para culminar esta investigación.

A mis padres, por ser pilares fundamentales en mi formación.

A Gustavo, por su amor, cariño y apoyo, gracias por estar allí.

A mi Abuela, por sus bendiciones.

A mi hermana, por su apoyo y cariño.

A mis tías, primos, padrinos, por sus consejos, y bendiciones.

A mis amigas María, Sara y Yoselin, por permitir que la amistad permanezca en el tiempo.

AGRADECIMIENTO

Al Ing. Edgar González, por darme las directrices necesarias a lo largo de la ejecución de la investigación.

A todos los compañeros de estudio y demás personas, que contribuyeron al bien del trabajo.

A la Universidad Centroccidental Lisandro Alvarado, por permitirme ser parte de esta casa de estudios.

INDICE DE TABLAS

	PAG.
Tabla 1 Glosario de Términos	76
Tabla 2. Diccionario de Conceptos	97
Tabla 3. Detalle de Relaciones.....	100
Tabla 4. Atributos de Clase	103
Tabla 5. Constantes	105

INDICE DE FIGURAS		Página
Figura 1. Framework de Búsqueda Web		19
Figura 2. Query de procesamiento soportado por el Modelo de Ontología		20
Figura 3. Modelo de Ontología		21
Figura 4. Componentes de Saw y su Interacción		22
Figura 5. Jerarquía de Clases para el Proyecto de NCI		27
Figura 6. Propiedades de la Clase Torax de NCI.....		28
Figura 7. Capas de la Web Semántica		31
Figura 8. Ejemplo de Ontología		38
Figura 9. Tipos de Ontologías según el nivel de dependencia.....		44
Figura 10. Arquitectura en la Creación de una ontología.....		49
Figura 11. Tareas Conceptualización de la Ontología		51
Figura 12. Ciclo de Vida de una Ontología.....		52
Figura 13. Bienvenido a Protégé.....		67
Figura 14. Creación de un nuevo Proyecto.....		67
Figura 15. URI en Protégé.....		68
Figura 16. Tipo de Lenguaje en Protégé		68
Figura 17. Vista en Protégé		69
Figura 18. Pantalla Principal en Protégé		70
Figura 19. Vocabulario Basado en Ontología con la Herramienta Protégé		87
Figura 20. Taxonomía Nro. 1.....		89
Figura 21. Taxonomía Nro. 2.....		90
Figura 22. Taxonomía Nro. 3.....		90
Figura 23. Taxonomía Nro. 4.....		91
Figura 24. Taxonomía Nro. 5.....		91
Figura 25. Taxonomía Nro. 6.....		91
Figura 26. Taxonomía Nro. 7.....		92
Figura 27. Taxonomía Nro. 8.....		93
Figura 28. Diagrama de Relación Binaria Ad-Hoc.....		94
Figura 29. Diagrama de relación Binaria Nro 1.....		95
Figura 30. Diagrama de relación Binaria Nro 2.....		96
Figura 31. Diagrama de relación Binaria Nro 3.....		96
Figura 32. Diagrama de Relación en Protégé		102
Figura 33. Diagrama de las Relaciones entre los Conceptos de la Ontología		111
Figura 34. Ejemplo de un archivo RDF.....		111
Figura 35. Código en XML		112
Figura 36. Lenguaje OWL.....		113
Figura 37. Ontología de la Investigación		114
Figura 38. Clases/Recursos de la Ontología de la Investigación		115
Figura 39. Configuración de Test Settings de Protégé L.....		115
Figura 40. Resultados de la Ontología		118
Figura 41 Arquitectura Propuesta L.....		127
Figura 42 Pagina Principal de Descarga de Protégé		128
Figura 43 Página de selección de Protégé según la Plataforma.....		129
Figura 44 Plugins		130
Figura 46 Ventana Options en Protégé.....		130

RESUMEN

UNIVERSIDAD CENTROCCIDENTAL "LISANDRO ALVARADO"

DECANATO DE CIENCIAS Y TECNOLOGIA

MAESTRIA EN CIENCIAS DE LA COMPUTACION

VOCABULARIO BASADO EN XML Y ONTOLOGÍA PARA LA COORDINACIÓN DE POSTGRADO DEL DECANATO DE CIENCIAS Y TECNOLOGÍA DE LA UNIVERSIDAD CENTROCCIDENTAL LISANDRO ALVARADO.

Autor (a): Clorelbys Escobar Querales

Tutor (a): Edgar González Muñoz

La presente investigación se encuentra enmarcada dentro de los estudios de proyectos factibles, apoyada en la investigación documental. En la actualidad existe gran cantidad de información que ocasiona que cada día sea más difícil buscar dicha información de manera precisa, clara y oportuna, es por esto que los usuarios deben dedicar más tiempo a filtrar los enlaces que ofrecen los buscadores disponibles, para ello el presente trabajo se basó en una Ontología Web que permite definir, estructurar y trabajar sobre un vocabulario común para referirse a los diferentes conceptos del dominio, en este caso la Coordinación de Postgrado es decir, asociarle un significado inequívoco y preciso a cada concepto. Los datos por si solos, sin tener una semántica asociada, no son de utilidad por resultar ambiguos. Por lo tanto, hay que introducir esa semántica añadiendo datos que describen a su vez los propios datos. Sin embargo, para dar una descripción formal de contenidos, hace falta una Ciencia eficiente: las ontologías, apoyándose de la metodología Methontology, la herramienta de desarrollo Protégé y los frameworks: Jena y Pellet. En consecuencia, el propósito de la presente investigación es Desarrollar un Vocabulario basado en XML y Ontología para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

Palabras Claves: Ontologías, Web Semántica, XML, Vocabulario, Methontology, Protégé, Jena, Pellet, OWL, RDF

INTRODUCCION

Hoy en día, buscar información que nos pueda resultar útil sobre algún tema en la web puede resultar complicado y tedioso. El motivo es la cantidad de información mal estructurada que puebla la red debido al uso de HTML, un lenguaje más de maquetación que de definición de contenido. Por esto actualmente se está apoyando el uso de XML que nos permitiría definir ontologías del tema o temas que trata nuestro sitio web. Esto es una aproximación a la llamada Web Semántica, en la que tendrán más importancia en la búsqueda de información los significados de los conceptos que las meras palabras o etiquetas que suelen usar los buscadores.

Al considerar la aplicación de las ontologías en el entorno Web, puede decirse que Internet es primero un medio de comunicación y después un sistema de almacenamiento y recuperación de la información. La productividad de las búsquedas y la relevancia de los documentos recuperados no siempre resulta la esperada y mucho menos la deseada. El problema no radica en cómo saber si hay algo determinado sino dónde está, si la información es veraz, es confiable, de que fuente viene, si la información se encuentra estructurada de forma correcta, es por esto que tanto los defensores como los detractores del Web, lo califican como una biblioteca con información desordenada, donde abunda el material inservible.

La información que se encuentra en Internet, a diferencia de la que se halla en una biblioteca física, por ejemplo, no se concentra en un catálogo ni en forma estructurada. Si bien en estas bibliotecas, las fuentes de información se seleccionan y se comportan de manera confiable, en el ambiente de Internet, no son frecuentes los procesos de validación antes de su adición y colocación en el sitio Web. Los documentos existentes en instituciones como las bibliotecas, los archivos y centros de información, no presentan tampoco la dinámica sobre la que se mueve la información en Internet. Por ejemplo, el URL de un sitio Web puede cambiar e incluso desaparecer. Por tanto, adquirir la información relevante se convierte en una tarea de primer orden.

Las ontologías en este sentido, pueden venir a minimizar esta situación, si se parte del hecho que mientras más ordenada y mejor procesada se encuentre la información, será más fácil desarrollar la selección y posterior adquisición de la misma, de esta manera se garantiza el procesamiento de la información a partir de una descripción semántica más profunda y propician un almacenamiento que sienta nuevos espacios para la búsqueda y recuperación efectiva de la información. Basado en esta premisa se plantea la siguiente investigación que se orienta al desarrollo de un vocabulario basado en XML y Ontologías para la búsqueda efectiva de información para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

La Investigación se encuentra estructurada de la siguiente manera:

Capítulo I El Problema. Se inicia con el planteamiento de la situación que genera el trabajo investigativo, se establece el objetivo general, los objetivos específicos, los cuales responderán al problema planteado, se expone la justificación de la investigación, sus alcances y limitaciones.

Capítulo II Marco Teórico. Presenta los antecedentes, bases teóricas sobre XML, Ontologías, Archivos RDF y prácticas que soportan el conocimiento del tema en estudio.

Capítulo III Marco Metodológico. Esta conformado por los lineamientos metodológicos utilizados, la naturaleza de la investigación, factibilidad y diseño de la propuesta.

Capítulo IV: Desarrollo de la Propuesta. Esta conformado por la aplicación de la metodología Methontology, herramienta Protégé y los Frameworks Jena y Pellet.

Capítulo V: Conclusiones y Recomendaciones por parte del autor de la investigación.

CAPITULO I

EL PROBLEMA

Planteamiento del Problema

Uno de los mayores problemas a los que se enfrentan hoy día los usuarios de Internet es la sobrecarga de información, problemas que se incrementan por el constante crecimiento de recursos disponibles en la Web. Este hecho repercute de forma directa en la navegación a través de la red para la búsqueda de una determinada información útil o necesaria, incrementando la cantidad de tiempo invertido a pesar de la potencia y sagacidad creciente de los buscadores disponibles. Adicionalmente a esto, los buscadores sólo acceden a lo que se denomina contenido estático de la Web e ignoran la parte dinámica de la misma, es decir, aquellas páginas generadas a partir de bases de datos. Por lo anteriormente expuesto Blazquez (2008) afirma que se produce el surgimiento de la Web Semántica, cuyo principal instrumento para la mejora de la organización, búsqueda e interrelación de los recursos presentes en la Web son las ontologías.

Actualmente estas búsquedas son sintácticas es decir los ordenadores realizan la presentación visual (tarea fácil) y las personas navegan, seleccionan e interpretan el contenido (tarea difícil). Esta situación se debe a que los lenguajes utilizados desarrollo de páginas web (HTML - HyperText Markup Language), únicamente incluye información sobre características de visualización de la información (tipo de letra, color, etc.), por lo que el significado sólo es accesible a las personas y no para los ordenadores. Esto pone de manifiesto la falta de capacidad de las representaciones en que se basa la Web actual para expresar significados, ya que la presentación de los datos en el navegador es inmediatamente comprendida por una persona. Para el ordenador es una tarea imposible poder entender cuál es la semántica del documento, ya que aparece entremezclada con las etiquetas del formato de representación elegido. Esta misma

limitación expresiva, hace que la noción de semántica que manejan los buscadores se limite a palabras claves, lo que no permite reconocer ni solicitar significados más elaborados, algunos problemas surgen cuando se trata con grandes cantidades de información semiestructuradas. Los actuales buscadores basados en palabras clave suelen devolver información irrelevante que usa una cierta palabra con un significado diferente del que se pretende en la búsqueda, y pierden información cuando no reconocen palabras diferentes pero con el mismo significado que la buscada, por ejemplo una búsqueda por palabra clave en un buscador puede ser por tres tipos definidos a continuación:

- Por palabras clave en el texto del cuerpo. Emplee el operador "allinbody" para buscar las palabras clave sólo en el cuerpo de un artículo. Por ejemplo, para encontrar todos los artículos con la palabra "oro" en el texto del cuerpo, escriba allinbody:oro.
- Por palabras clave en el titular. Use el operador "allintitle" para buscar un artículo con un titular específico. Por ejemplo, introduzca allintitle:lémur para encontrar artículos con el término "lémur" en el titular.
- Por palabras clave en la URL. Use el operador "allinurl" para buscar palabras clave en la URL de un artículo. Por ejemplo, para encontrar todos los artículos con la palabra "política" en la URL, indique allinurl:política.

Castells (2003) destaca que como respuesta ante la problemática explicada anteriormente derivada de las dificultades que conllevan el acceso masivo a los datos, la complejidad de su interpretación posterior y la dificultad que supone la búsqueda de una determinada información de forma precisa y rápida. Como una opción para atacar esta situación, Castells plantea la estrategia de búsqueda semántica como vía para ofrecer soluciones más efectivas a los usuarios de cualquier sistema basado en la Informática.

Por otro lado Torres y Pomares (2005) en su Modelo de Evolución de Ontologías para la Gestión del Conocimiento Empresarial afirman que los cambios que han sufrido las empresas en cuanto a administración, estructura y ejecución de procesos, han modificado el rol de la información dejando de ser un simple recurso para

convertirse en un activo, que constituye el capital intelectual de la empresa. Esta nueva perspectiva ha hecho que las empresas se vean avocadas a adoptar nuevas técnicas y estrategias que les permita gestionar de una manera eficiente la información y el conocimiento, haciendo de ellos generadores de ventajas competitivas sostenibles. Del estudio de Torre y Pomares, también se visualiza que una de las estrategias que ha evolucionado desde los años 90's es la preocupación por el manejo formal de las actividades en torno al conocimiento, entendiéndose por él, una colección apropiada de información que es asimilada por un individuo dándole utilidad y posibilidad de aplicación. La formalización de estas actividades se concreta en lo que actualmente se conoce como gestión del conocimiento (GK) que abarca los procesos de planeación, generación, representación, almacenamiento, integración, transferencia, transformación, aplicación, protección y mantenimiento del conocimiento organizacional incluyendo documentos y textos manejados por las coordinaciones empresariales en la que se basó el estudio de los autores antes mencionados.

En cuanto al proceso de Gestión del conocimiento basado en Web, Corcoles (2007) señala que La Web y la gestión del conocimiento ha cambiado profundamente la forma en la que nos comunicamos, hacemos negocios y realizamos nuestro trabajo. La comunicación prácticamente con todo el mundo en cualquier momento y a bajo costo es posible hoy en día. Podemos realizar transacciones económicas a través de Internet. Tenemos acceso a millones de recursos, independientemente de nuestra situación geográfica e idioma. Todos estos factores han contribuido al éxito de la Web, sin embargo, al mismo tiempo, estos factores que han propiciado el éxito de la Web, también han originado sus principales problemas: sobrecarga de información y heterogeneidad de fuentes de información con el consiguiente problema de interoperabilidad. El mismo autor señala que La Web Semántica basada en Ontologías ayuda a resolver estos dos importantes problemas permitiendo a los usuarios delegar tareas en software. La mayor ventaja que representa la aplicación del concepto de semántica en la Web, es que el software es capaz de procesar su contenido, razonar con este, combinarlo y realizar deducciones lógicas para resolver problemas cotidianos automáticamente.

En Venezuela, en el área de educación universitaria, apreciamos que las universidades son un tipo de organización que manejan mucha información, específicamente en La Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA, universidad de la ciudad de Barquisimeto estado Lara no escapa a esta realidad ya que posee gran cantidad de información como son: Textos, tesis, documentos de Inscripción, documentos de admisión, textos científicos, tesis, memorias de organizaciones, proyectos de grado resúmenes curriculares de profesores adscritos a dicha coordinación que no se encuentra ordenada ni estructurada de manera correcta, además de presentar una creciente demanda de acceder a dicha información por el personal administrativo, y población estudiantil que no se encuentran disponibles de manera rápida que pueda ser aprovechada de una manera directa, sin tener que involucrarse en el proceso de organizar, ya que esto incrementa los costos de almacenamiento, ordenamiento y estructuración. Actualmente la Coordinación de Postgrado maneja una gran cantidad de términos, palabras, y documentos que cada uno tiene un fin pero esta información no se encuentran estructurada de manera semántica es decir que cada palabra y/o término tenga su significado asociado, ya que a su vez la terminología aparezca de manera estructurada y clasificada que exprese entendimiento y clasificación.

Basados en los objetivos de la Universidad Centrooccidental Lisandro Alvarado UCLA donde se persigue: *“Producir nuevos conocimientos en las ciencias puras y aplicadas, generar innovaciones tecnológicas para acrecentar el saber, enriquecer las funciones de la docencia, extensión, y satisfacer las necesidades del entorno, y Lograr una Gerencia Universitaria dinámica y flexible que permita adecuar permanentemente la Institución a las necesidades y cambios que demanda el entorno.”*, La UCLA es una institución educativa de la Región Centrooccidental comprometida en la búsqueda de conocimiento, aprendizaje y difusión de saberes que debe caracterizarse por la actualización en Tecnología y así adaptarse a los cambios permanentes de nuestro entorno.

En la actualidad en la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA se necesita información importante que no es accesible ya sea por falta de una clasificación, estructuración o por contener datos ambiguos que pudiesen solaparse, es importante recalcar que dicha información necesita ser ordenada y clasificada para ser utilizada en el momento preciso y evitar la duplicidad de la misma, esto impide que se obtengan los datos requeridos para una toma de decisión ágil evitando así el cumplir uno de los objetivos de la Universidad como es el generar innovaciones tecnológicas adaptándose a los cambios que demanda el entorno. En consecuencia, se crean problemas de comunicación por falta de entendimiento y consenso compartido entre las partes involucradas que limitan la interoperabilidad y, por tanto, el potencial de reutilizar y compartir documentos ocasionando problemas de ambigüedad en los datos generados por el software.

García Hernández (2005) señala que para evitar estos problemas de ambigüedad las teorías ontológicas ayudan a asegurarse de que seleccionan una gramática de modelado conceptual necesaria para producir modelos de alta calidad e indican cómo usar la gramática para generar descripciones claras y completas del dominio, y acota que el uso de ontologías para la recuperación de la información tiene ventajas sobre los métodos de acceso simples basados en palabras clave, al suministrar un vocabulario compartido común para expresar información sobre el contenido de los documentos.

Gruber, (2006) afirma que una ontología constituye una especificación explícita y formal de una conceptualización compartida. Esta definición tiene una serie de conceptos claves que ayudan a comprenderla, así por “conceptualización” se entiende un modelo abstracto de la realidad, de tal manera, que mediante esta se identifican los conceptos relevantes de un área. Por “explícita” se entiende que todos sus componentes deben estar definidos explícitamente. Respecto a “formal”, se refiere al hecho de que la ontología debe ser entendible por las máquinas. Por último, “compartida” refleja el hecho de que una ontología debe capturar conocimiento consensuado/aceptado por un grupo o comunidad de expertos, esto es, no debe ser privado o algo individual. De allí que la Ingeniería Ontológica según Gómez-Pérez (2003), surgida de la Web Semántica,

proporciona soluciones a los problemas actuales relacionados con la accesibilidad y las búsquedas distribuidas de información. Fundamentalmente, esta mejora se basa en el cambio de perspectiva en los procesos de búsqueda, ya que se pasa de utilizar palabras claves, es decir, de considerar únicamente los aspectos sintácticos, a centrarse en los significantes de los conceptos, es decir, en la semántica de la información. De esta manera, es obvia la asunción de que los datos deben ser entendidos, exclusivamente, por los usuarios y se pasa a un proceso de entendimiento recíproco entre hombre y máquina, en el que las máquinas pasan a “comprender” los datos que procesan.

Es necesario plantearse las siguientes interrogantes para esta investigación:

- ¿Qué Requisitos se necesitan para la Creación de un vocabulario de búsqueda de información como documentos, términos, para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA?
- ¿Qué terminología es usada por la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA?
- ¿Qué Tipo de Información es manejada por la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA?
- ¿La creación de un vocabulario basado en XML y Ontología reduciría de forma significativa la ambigüedad terminológica al considerar sinónimos (Mas de una palabra tiene el mismo significado) y polisemias (Una palabra tiene mas de un significado), repercutiendo sobre la comunicación y gestión de la información de la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA?
- ¿Qué tipo de términos y/o palabras de la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA necesita ser estructurada de manera que la búsqueda sea flexible y oportuna?
- ¿Qué principios, tecnologías y herramientas de representación del conocimiento y de definición de ontologías son las mas apropiadas para abordar este problema?

- ¿Qué mecanismos pueden utilizarse para permitir que cualquier usuario acceda de forma transparente o a través de interacciones sencillas a la información deseada por él, expresada según sus requerimientos?

En Consecuencia, el propósito de la presente investigación es Desarrollar un Vocabulario basado en XML y Ontología para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la Universidad Centro Occidental Lisandro Alvarado, destinado al manejo de términos y documentos con su respectiva semántica.

OBJETIVOS DE LA INVESTIGACION

Objetivo General

Diseñar un Vocabulario basado en XML y Ontología para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la Universidad Centro Occidental Lisandro Alvarado.

Objetivos Específicos

(a) Modelar los Requisitos para la creación del vocabulario para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA

(b) Construir la terminología no estructurada existente en la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

(c) Diseñar el vocabulario para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA

(d) Desarrollar un prototipo basado en Web que use el vocabulario para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA de tal manera que la búsqueda sea flexible y oportuna.

.

JUSTIFICACIÓN DE LA INVESTIGACIÓN.

Como toda investigación que reduce su campo de trabajo, se cree que sus posibles contribuciones se pueden resumir en los siguientes aspectos:

La aplicación basada en ontologías terminológicas y vocabulario basado en XML unifica la terminología de cada concepto y las relaciones entre ellos mejorando así la búsqueda de documentos utilizados por el personal administrativo adscrito a la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

La representación de los documentos utilizados en la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA mediante un conjunto de términos que especifican los temas de que trata ese documento, con el fin de ayudar a su posterior recuperación: La indexación de un sistema de información con apoyo de una ontología terminológica comienza con la extracción de los términos más relevantes de cada documento utilizado, y después de asociarlos a estos términos y/o conceptos candidatos, se evalúa la capacidad de representación del documento, logrando que las consultas se procesen a un nivel conceptual, lo que reportará un mayor grado de acierto evitando las ambigüedades en la búsqueda de documentos en la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA..

Con una ontología semántica, se pueden organizar los términos que sean sinónimos y polisémicos evitando multiplicidad de información en la búsqueda de documentos en la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

Según Guarino, 1998 Una ontología debe ser capaz de proveer un buen entendimiento del dominio que representa. Esto incluye relaciones de términos y conceptos, sus definiciones o significados, sus relaciones con cada uno de ellos y las características del dominio. Esto ayuda en la solución de las diferencias entre los modelos y el dominio usado por las fuentes disponibles, lo que se hace mapeando los

datos disponibles de todos los recursos para todos los dominios desde el modelo local usado por la fuente para el modelo especificado por medio de la ontología, el autor afirma que las ontologías proveen una comprensión compartida y concensuada del conocimiento de un dominio que puede ser comunicada entre personas y sistemas heterogéneos y fueron desarrolladas en el área de Inteligencia Artificial (IA) para facilitar el intercambio y reuso del conocimiento.

Al utilizar semántica las consultas ya no sólo se efectuaran por palabras claves proporcionando una calidad superior en los resultados de una búsqueda basado en ontologías y web semántica, las consultas serán tratadas desde un punto de vista conceptual, de este modo se redujo el margen de error en los resultados de una búsqueda, lo que permitió que no se omitieran aquellos resultados, que aún siendo conceptualmente sinónimos al de la consulta, no se encuentran por ser distintos terminológicamente. Uno de los aportes de la presente investigación es el incentivo a las líneas de investigación del Decanato de Ciencias y Tecnología a través de aportes originales en el campo de la Semántica, vocabularios Bases de conocimientos para su enriquecimiento y desarrollo tecnológico y científico.

Con objeto de poder usar y enlazar la gran cantidad de información disponible en la Web, los agentes software deben ser capaces de comprender la información, es decir, los datos deben estar escritos haciendo uso de una semántica. Por tanto, en los documentos XML o HTML, deberá de añadirse semántica adicional o metadatos para que los programas y/o software puedan establecer el significado de las etiquetas de dichos documentos. Para conseguir lo anteriormente descrito, la Web Semántica extiende la Web actual con conocimiento formalizado y datos que son procesados por ordenadores. Para ser capaz de buscar y procesar información relativa a alguna materia de interés, los programas necesitan información que haya sido modelada de una forma coherente. Una ontología modela todas las entidades y relaciones en un dominio. La ontología es necesaria para la representación del conocimiento. La clave de las ontologías es que pueden ser compartidas y por lo tanto incrementan en eficiencia e interoperabilidad.

ALCANCES Y LIMITACIONES

Alcances

A través de esta investigación se pretendió realizar un vocabulario basado en XML y Ontología para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología para la búsqueda, estructuración, ordenamiento de cada unos de los términos de los documentos utilizados en dicha coordinación por el personal administrativo adscrito a él. Aunque el alcance de esta investigación se suscribe a Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA existe la posibilidad que su aplicación se extienda a otras coordinaciones de otros decanatos de la misma universidad y otras Universidades del país. Dentro de este orden de ideas como el proyecto desarrollado fue basado en software libre tendrá libertad de uso y redistribución ya que las licencias de software libre existentes permiten la instalación del software tantas veces y en tantas máquinas como el usuario desee, por otra parte se contó con independencia tecnológica ya que el acceso al código fuente permite el desarrollo de nuevos productos sin la necesidad de desarrollar todo el proceso partiendo de cero.

A través de esta investigación también se desea lograr que la ontología sea usada para capturar el conocimiento del dominio de la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA, dicho conocimiento se debe expresar por medio de relaciones entre conceptos, y de esta manera proveer herramientas que facilitan la inferencia entre las clases permitiendo mantener correctamente la jerarquía entre las clases.

A través de la presente investigación se pretendió desarrollar una opción para la migración y/o representación a una Ontología de la Web Semántica de un vocabulario controlado y dinámico, en la cual estos términos tienen relaciones semánticas que se aplica a un dominio particular del conocimiento, con el propósito de permitir su uso en la recuperación de información concerniente a términos, vocablos y documentos

tratados por de la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

La Ontología de la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA se modeló con Methontology y se utilizó la herramienta Protégé para desarrollar y modelar dicha ontología, dicha herramienta permitió generar el vocabulario basado en XML tomando en cuenta el framework RDF y el lenguaje que se utilizó fue el Lenguaje de Ontologías Web OWL, posteriormente se procedió a utilizar el test de Protégé y realizar un código en Jena y Pellet que validara que la Ontología era correcta y consistente.

Limitaciones

- ❖ El trabajo de investigación principalmente esta limitado a los documentos y terminología utilizados en la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de La UCLA, específicamente estos términos y palabras claves.
- ❖ Para realizar el estudio del dominio, en este caso la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA, la información fue extraída de la Pagina Web del Postgrado de la UCLA
- ❖ Se limitó a crear el vocabulario basado en XML y Ontología, con la utilización de los términos semántica, gramática y Web semántica.
- ❖ Las Definiciones que se utilizó para la creación del Vocabulario basado en XML fueron extraídos de los reglamentos de la UCLA y los reglamentos internos del Decanato de Ciencias y Tecnología de la UCLA.

CAPITULO II

MARCO TEORICO

Antecedentes

Los Antecedentes de la presente investigación vienen enmarcados por proyectos desarrollados por expertos en el área de búsqueda de datos e información basados en términos y/o palabras claves, como es el caso del siguiente proyecto que fue diseñado por la Cooperación GALEN-CORBAMed, en respuesta a la necesidad de una Gestión Terminográfica de calidad, ya que según miembros del Equipo GALEN, el presente y el futuro de centran en el desarrollo de Sistemas informáticos que ofrezcan servicios terminográficos Basados en Ontologías.

1.- Proyecto GALEN (2006). Unión Europea. Cooperación GALEN-CORBAMed

GALEN, Arquitectura General para Enciclopedias y Nomenclaturas en Medicina, es un proyecto desarrollado y financiado por la Unión Europea que tiene como objetivo el desarrollo de herramientas y métodos para una construcción y mantenimiento de clasificaciones de procedimientos quirúrgicos basados en Ontologías para mejorar la búsqueda de Información de términos que sean sinónimos. En la actualidad, para afrontar el problema de la terminología en Ciencias de la Salud, GALEN está construyendo un modelo semántico para la gestión de terminología clínica: el modelo CORE, *Coding Reference*. Este modelo consiste en:

- **relaciones** tales como, fracturas que pueden ocurrir en los huesos, que eflajan posibles combinaciones de términos,
- **conceptos** complejos tales como *fractura de Húmero Izquierdo* que son combinaciones de conceptos más simples.

Dentro del proyecto GALEN, las normas formales que manipulan estos conceptos y sus relaciones se encuentran en el formalismo GRAIL, Lenguaje de Representación e Integración. GRAIL funciona como un programa *Concepto-Modelo* (CM) con herramientas de modulación. El CM se puede usar bien con herramientas de búsqueda o con un Servidor de Terminología. El CM, junto con sus herramientas hace que los terminógrafos puedan crear modelos que contengan conceptos y relaciones, y al mismo tiempo, se puedan derivar nuevos conceptos que sean composiciones plausibles de los ya existentes. GRAIL hace que el sistema utilice los conceptos y las relaciones basados en Ontologías para:

- determinar si una composición en particular tiene sentido o no
- generar todos los conceptos posibles, basándose en el conocimiento
- derivar automáticamente otras relaciones, tales como jerarquías clasificadoras.

Este potencial de GALEN hace que los integrantes del grupo no tengan que enumerar de forma explícita todos los términos posibles. Por ejemplo, si se ha establecido que los huesos se pueden romper, el sistema puede generar y clasificar a lo largo de múltiples ejes aquellos conceptos que representen fracturas de todos los huesos. Estos fundamentos formales nos conducen a modelos robustos que se pueden expandir de forma bastante segura, consistente y predecible.

Los módulos principales de GALEN (conceptual, multilingüe y el de conversión de códigos) están integrados en un único sistema de software en red: el Servidor de Terminología (TeS) de GALEN. El TeS combina la funcionalidad de estos tres módulos para garantizar un servicio terminológico sofisticado a la vez que uniforme para las distintas aplicaciones del cliente. En esta estructura se refleja la visión de la terminología de los integrantes de GALEN como un compendio de sistemas funcionales y dinámicos más que unos archivos que sirven para el almacenamiento estático de datos. Los usuarios pueden hacer preguntas de alto nivel tales como "cuáles son los tipos de" o "qué se puede decir sobre".

El aporte fundamental de este proyecto se enumera a continuación:

- garantiza descripciones clínicas detalladas basadas en un modelo semántico terminológico;
- los conceptos complejos se almacenan con una representación fija, para que su extracción sea mucho más fácil;
- contiene herramientas lingüísticas que hacen posible el desarrollo de sistemas multilingües;
- facilita el intercambio de datos clínicos entre sistemas con diferentes estructuras clasificadoras;
- aumenta y expande los esquemas de codificación y clasificación ya existentes. Y todo esto gracias a la utilización del motor de búsqueda basada en Ontologías.

Este proyecto fue de valor para la investigación aportando un modelo semántico para la gestión de terminología, dicho modelo es el CORE, *Coding Reference*. Este modelo sirvió de ayuda para aumentar y expandir los esquemas de codificación y clasificación de términos utilizados para la creación del vocabulario en esta investigación.

Entre los antecedentes también se puede mencionar el Framework basado en Ontología desarrollado Hun, H. y Du, X, a continuación se explicará el objetivo principal de este Framework.

2.- Hun, H. y Du, X. Framework Adoptando Reglas y Ontologías en los servicios de Búsqueda Web. Enero 2005.

Este trabajo tuvo como finalidad proponer un Framework de búsqueda Web usando Ontologías y Reglas. El Framework usa Ontologías para tratar con dos problemas principales como son los sinonimia: más de una palabra tienen el mismo significado y la polisemia: Una palabra tiene más de un significado, por lo tanto la búsqueda tiende a ampliarse debido a la polisemia y a encogerse debido a los sinónimos. Este Framework resolvería estos problemas a través la aplicación de un

Enfoque de Ontologías con Semántica Web y Sistemas basados en Reglas. En ese Framework las Ontologías:

(a) Proporcionan un vocabulario común y la definición de reglas para usar independientemente los servicios desarrollados.

(b) Logra Acuerdos entre compañías, organizaciones que comparten servicios comunes que pueden ser hechos con respecto a su uso y el significado de los conceptos relevantes que puedan ser expresados de forma no ambigua.

(c) Facilitan conversaciones entre agentes para coleccionar, procesar, fusionar e intercambiar información.

(d) Mejoran la búsqueda precisa a través de la búsqueda contextual usando conceptos de definiciones y relaciones entre ellos en lugar de palabras claves de relevancia estadísticas. A continuación se especificara la arquitectura y el sistema utilizado basado en Reglas.

Arquitectura del Framework:

Este Framework coloca las tareas de selección, creación y las reglas de filtraje en el sitio de los usuarios. Los Autores argumentan que esta arquitectura trae mucha carga al sitio de los usuarios, sin embargo esta arquitectura puede garantizar una máxima flexibilidad para toda la aplicación de búsqueda Web. Cuidadosamente el diseño de Interfases de Usuario y Software de Configuración pueden aliviar la carga de los Usuarios Web. A continuación el Framework de Búsqueda Web. Ver Grafico 1.

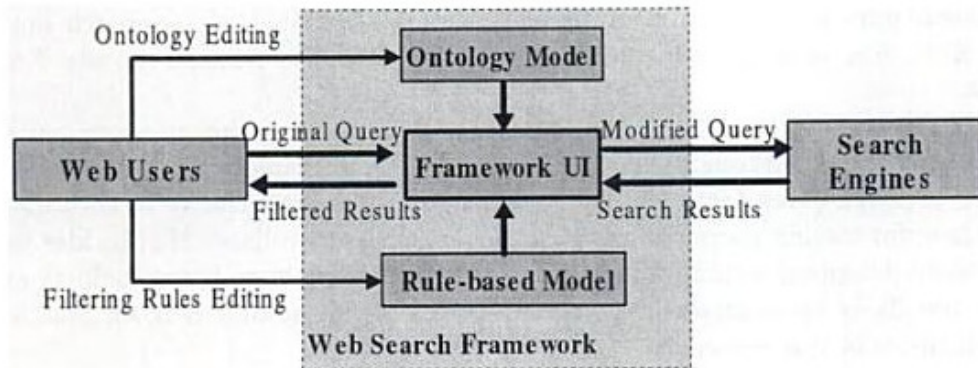


Figura 1. Framework de Búsqueda Web.

Fuente: <http://www.springerlink.com/content/4x4k154hm3ck9rpp/>

Autor: Hun, H. y Du, X 2005.

SISTEMAS BASADOS EN REGLAS DEL FRAMEWORK

Estos sistemas comenzaron en los años 60. Y este modelo fue propuesto por Newell y Simon y es llamado Modelo de Sistema de Producción y está basado en la idea que los Humanos resuelven problemas aplicando sus bases de conocimiento expresado como un modelo de reglas para un problema específico dado que a su vez esta representado por una información específica dada. Las reglas de producción son almacenadas en la memoria de término ancho y la información del problema son almacenadas en la memoria de término corto. El termino Regla puede ser definida como una estructura IF – THEN – ELSE. Las reglas son relativamente fáciles de crear y entender. Un sistema basado en regla usualmente consiste de un conjunto de reglas a ser evaluadas, una memoria de trabajo y un esquema de ejecución que decide cuales reglas serán aplicadas para resolver un conflicto y/o problema dado y a su vez si más de una regla es aplicable se decide como proceder. Principalmente se utilizo el siguiente modelo de Ontología:

Modelo de Ontología

El modelo de ontología principalmente se concentra en resolver el problema de los sinónimos. Por ejemplo los términos: “Plano”, “Aeroplano”, “AirPlano” tienen el

significado de maquina voladora; Cuando un usuario Web incluye el termino aeroplano en su consulta el modelo de Ontología debería tomar automáticamente los otros dos términos en su contador de búsqueda. El modelo de Ontología soportado por el proceso de consulta se muestra a continuación:

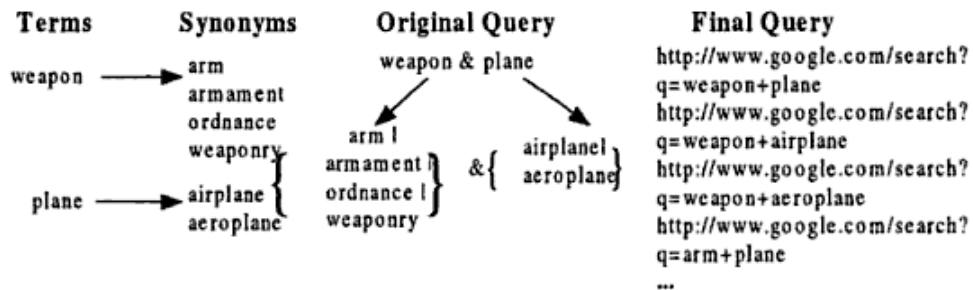


Figura 2:. Query de Procesamiento soportado por el Modelo de Ontología.

Fuente: <http://www.springerlink.com/content/4x4k154hm3ck9rpp/>

Autor: Hun, H. y Du, X 2005.

Este Framework significa un gran aporte a la presente investigación, ya que contribuyo en la creación del modelo de arquitectura de la propuesta al incluir un MVC (Modelo-Vista-Controlador), Ver Grafico 1, que encapsula el Modelo de Ontología, el Framework de Interfaz y el modelo basado en reglas utilizados para el motor de Búsqueda.

Nicola Guarino, Formal Ontology in Information Systems, in Proceedings of FOIS'98, 2002. Ha particionado la ontología en tres diferentes tipos de acuerdo al nivel de generalidad:

- (a) Nivel de Aplicación: Describe los conceptos muy generales como son: Espacio, tiempo, Asunto, Objeto, Evento, Acción, los cuales son independientes de un problema o dominio en particular.
- (b) Nivel Dominio y Tareas: escribe el vocabulario relacionado a un dominio genérico como por ejemplo: Medicina, Automóviles o Tarea Y/o actividades Genéricas como por ejemplo: ventas.

(c) Nivel SUMO: Viene siendo el repositorio de Datos e Información.

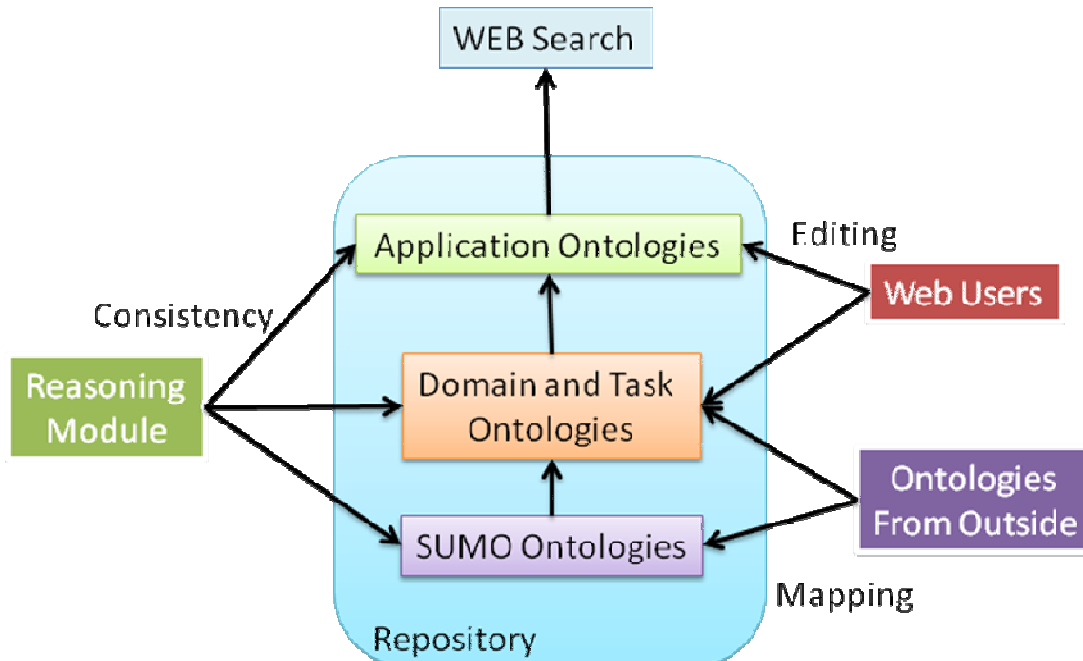


Figura 3. Modelo de Ontología.

Fuente: <http://www.springerlink.com/content/4x4k154hm3ck9rpp/>

Autor: Hun, H. y Du, X 2005.

Otro de los proyectos que ha sido desarrollado recientemente es el Proyecto SAW que será explicado a Continuación:

3.- PROYECTO SAW. Lozano Tello (2006):

SAW, Sistema de Accesibilidad a la Web, cuenta con diversas herramientas integradas que cooperan entre si y trabajan a distintos niveles para conseguir una accesibilidad real, por un lado ofrecen una herramienta para el diseñador de paginas web, EDITSAW, un editor que permite realizar anotaciones semánticas enfocadas a mejorar la accesibilidad de los contenidos de una pagina web. En cuanto al usuario con discapacidad visual se ofrece NAVISAW, un navegador multimodal que incorpora síntesis y reconocimiento de voz e interacción a través de MOUSESAW, un gestor de

correo accesible. Todo ello construido sobre ONTOSAW, una ontología que permite definir los componentes de una pagina web añadiendo atributos de accesibilidad basada en estándares W3C. Ver Grafico 4.

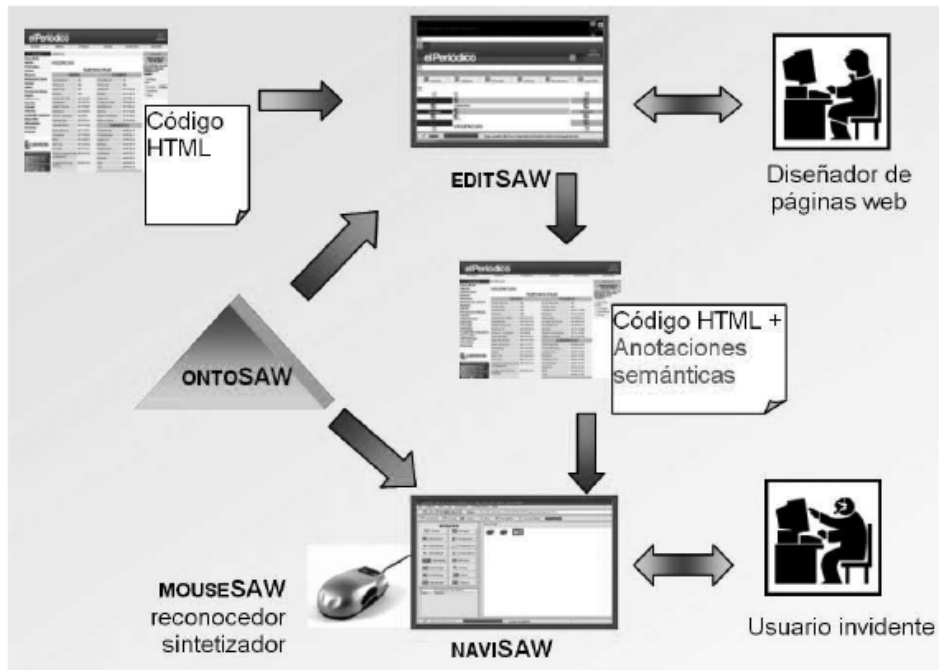


Figura 4: Componentes de SAW y su Interacción.

Fuente: Herramientas de Software basados en Ontología para mejorar la accesibilidad en la Web

Autor: Lozano Tello (2006).

Una de las aportaciones de SAW consiste en identificar los atributos de los elementos que componen las páginas web para hacerlas accesibles y representarlos en una ontología denominada ONTOSAW. El conocimiento que se representa en Ontologías trata de recoger de manera perfecta y precisa los elementos de un dominio y esta disponible para ser utilizable por varias aplicaciones en múltiples propósitos y en los últimos años su uso está muy extendido debido a la iniciativa de la web semántica.

Lozano Tello (2006) afirma que una de las ventajas del proyecto SAW es que la implementación de las herramientas software (el editor y el navegador) no debe modificarse en absoluto para adaptarse a las nuevas especificaciones de la W3C. Las herramientas usan la representación de la Ontología para identificar dinámicamente los elementos, atributos, requisitos y relaciones que serán utilizados para identificar los elementos web con sus marcas de accesibilidad y para navegar por las páginas marcadas conforme se especifica en la Ontología.

Uno de los aportes de este proyecto a la presente investigación fue que al representar los elementos web y sus requisitos de accesibilidad en una ontología se logró que esta información este abierta y disponible a cualquier usuario en una dirección web (<http://quercusseg.unex.es/SAW/ONTOSAW.OWL>); de esta forma, otros grupos de investigación que deseen desarrollar sistemas relacionados con accesibilidad web podrán usar esta ontología y no volver a realizar el esfuerzo de adquisición de conocimiento.

**4.- Vocabulario en XML para la Gestión Catastral de Inmuebles Urbanos.
Ing. Maribel Mendonça Da Mata. 2004. UCLA.**

Su objetivo principal fue el diseño e Implementación de un vocabulario en XML para la gestión catastral de inmuebles. Para ello se identificaron los elementos involucrados en la gestión catastral mediante el análisis de la información manejada por las oficinas de catastro de 3 municipios, tomados como muestra de todos los municipios del país, y principalmente considerando los requerimientos del Instituto Geográfico de Venezuela Simon Bolívar, con respecto a la gestión catastral. Luego se definieron los elementos relacionados con la gestión catastral en los aspectos físicos, jurídicos y valorativos, con sus atributos de manera estándar. Seguidamente mediante el UML se realizó el análisis y diseño del modelo para la gestión catastral, con los elementos, atributos y restricciones propias. Posteriormente se generó a partir del modelo, el esquema UML para la estructuración e Implementación del vocabulario.

A través de esta propuesta se logró una estandarización de la información y así la Gestión Catastral pueda realizar el intercambio de información entre las entidades involucradas y también se pueda crear y mantener actualizada una base de datos Central con la información catastral nacional. Al adoptar este vocabulario como estándar en las entidades municipales e institucionales involucradas con la gestión catastral, se puede tener grandes beneficios en cuanto al mejor y rápido procesamiento de información, emisión de estadísticas, reportes mas exactos y con información precisa y real a nivel nacional.

Este trabajo contribuyó a la presente investigación ya que en el desarrollo de este vocabulario surgió un elemento con una gran potencialidad de desarrollo futuro, como es la ampliación del vocabulario, la creación de nuevas aplicaciones en base a él, servicios web, entre otros, que podrían también ser de interés para otras futuras investigaciones. A su vez que se demostró que al utilizar el vocabulario exista homogeneidad en la información reduciendo la duplicidad y se pueda intercambiar información entre las entidades involucradas.

5.- Ontologías Para Servicios Web Semánticos de Información de Tráfico: Descripción y Herramientas de Explotación José J. Samper Z. Universidad de Valencia. 2005.

Esta tesis integra dos áreas científicas: por una parte la Ingeniería de Tráfico y por otra las Ciencias de Computación. Dentro de la Ingeniería de Tráfico, forma parte de los denominados Sistemas Inteligentes de Tráfico (ITS) los cuales se basan en la aplicación de tecnologías de Telecomunicaciones e Informática (Telemática) a los sistemas de ayuda al tráfico vial, perteneciendo al dominio “Servicios de Información al Viajero” (TIS). Como se pone de manifiesto en esta tesis, existen algunos problemas derivados del uso de este tipo de sistemas, desde el punto de vista del usuario, que dan lugar a que éste sea incapaz de obtener la información clara y precisa correspondiente a sus requerimientos. Los principales problemas identificados han sido la falta de un vocabulario común de términos que haga uso de una semántica bien definida que

permita obtener el significado (incluso aquél no explícito) de cada concepto de tráfico vial, así como la recuperación de información desde distintas fuentes heterogéneas, de una forma intuitiva y sencilla. Varios aspectos describen la problemática actual relativos al acceso y distribución de la información:

- Un gran volumen de información de tráfico se distribuye entre varios sitios web, el principal problema para un usuario que necesita información de este tipo es encontrar estos sitios web y además tratar con los diferentes accesos a ésta, así como sus diferentes formas de presentación.
- Por otra parte, un usuario puede necesitar información de diferente naturaleza o tipo, y por lo tanto el almacenaje de toda esta información en un solo sitio web no es viable a nivel de costes de almacenamiento ni incluso a nivel de mantenimiento.
- El tratamiento de la información no permite realizar inferencias sobre ella de manera que pueda ser obtenido como resultado, información que a priori no estuviera explícitamente detallada.

Tras el descubrimiento de la problemática existente y la identificación de elementos que pudieran solventar esta situación, el presente trabajo se centra en el estudio de los sistemas ITS orientados a facilitar información de tráfico al usuario, y que a su vez permitan la gestión de la información, su tratamiento e intercambio, de manera eficaz entre los diferentes elementos que componen la arquitectura de servicios de información de tráfico, como los usuarios, aplicaciones, y proveedores de la información. Para la consecución de los objetivos marcados han sido desarrollados diversos elementos como parte integrante de una arquitectura, que supusieron determinados aportes y/o resultados:

1.- Construcción de una infraestructura ontológica cuyo dominio queda enmarcado en la información sobre tráfico vial. Las ontologías construidas forman el núcleo principal del trabajo, ya que por una parte, constituyen la ontología de dominio utilizada como soporte en el emparejamiento de servicios, y a su vez, conforman un mecanismo de descripción de los diferentes servicios de información de tráfico, los cuales son especificados mediante las referencias a conceptos descritos en estas ontologías. La infraestructura ontológica creada puede considerarse una base sólida en

el desarrollo de un futuro y completo vocabulario semántico a utilizar por las Administraciones de Tráfico.

2.- Se ha expuesto un **marco de trabajo para la conversión de portales Web convencionales de información en Servicios Web Semánticos (SWS)** de manera que la información aportada pueda ser almacenada con la adición de significado consiguiendo a su vez, potenciar nuevas capacidades que en un principio no existían.

3.- Extensión a las metodologías de construcción de ontologías, previamente existentes, para plantear el proceso a seguir en la obtención de un modelo semántico formal a partir de un modelo de Entidad-Relación (ER). En la presente aproximación, se han integrado diferentes aspectos de las metodologías actuales así como aquéllos surgidos de lecciones aprendidas de la experiencia. Este método fue aplicado en la construcción de la infraestructura ontológica creada, y puede servir como guía de desarrollo en construcciones ontológicas similares.

Esta tesis fue de gran aporte al presente trabajo ya que se han mostrado una serie de propuestas que permiten crear y manejar servicios web semánticos de información de tráfico vial: nuevo parámetro no funcional, ontología de categorización de servicios de tráfico y elección de factores de medida en la Calidad de Servicio.

6.- Ontología del Instituto Nacional de Cáncer en OWL. 2007.

Esta Ontología fue desarrollada por el Centro de Bioinformática del Instituto Nacional del Cáncer de los EEUU, y contiene cerca de 17.000 términos relacionados con el cáncer y sus definiciones. La Ontología NCI es una descripción de la lógica de dominio público basado en la terminología elaborada por el Instituto Nacional del Cáncer distribuido como in Componente del centro de Bioinformática DEL NCI. Esta Ontología es muy amplia y compleja con los vocabularios clínicos más grandes existentes hoy en día, tiene una gran variedad de interrelaciones entre los términos y sus taxonomías. Esta ontología fue implementada para facilitar la búsqueda, traducción y soporte de la Infraestructura de Bioinformática del Instituto. Los tópicos descritos incluye: enfermedades, medicinas, quimioterapias, diagnósticos, genes, tratamientos,

anatomía, organismos y proteínas. La Ontología NCI esta basado en la librería Nacional de Medicina y en el Sistema de Lenguaje Medico Unificado (UMLS).

Características de la Ontología NCI:

- Mapas de 3.600.000 términos de 76 fuentes en 1.400.000 conceptos biomédicos que representan su significado.
- Muestra términos preferidos, sinónimos, definiciones.
- Enlaces a Reporte NCI y otras fuentes de información relacionada.
- Contiene 20.000.000 enlaces cruzados entre los elementos de contenido.
- Se actualiza con frecuencia por un equipo de terminología biomédica y expertos en la materia.
- NCI Ontologies contiene la mayoría de las terminologías de dominio público de la Biblioteca Nacional de Medicina Metathesaurus sinónimos, así como muchas otras terminologías biomédica creados o de interés para el NCI y sus socios.

Class Hierarchy for *nciOncology* Project

- [NCI Kind](#)
- [Findings and Disorders Kind](#)
- [Properties or Attributes Kind](#)
- [Anatomy Kind](#)
- [Chemicals and Drugs Kind](#)
- [Biological Process Kind](#)
- [EO Anatomy Kind](#)
- [Organism Kind](#)
- [Clinical or Research Activity Kind](#)
- [Gene Kind](#)
- [EO Findings and Disorders Kind](#)
- [Protein Kind](#)
- [Diagnostic and Prognostic Factors Kind](#)
- [Technique Kind](#)
- [Equipment Kind](#)
- [Retired Kind](#)

Generated on Fri Aug 01 21:16:00 EDT 2003

Figura 5 : Jerarquía de Clases para el Proyecto NCI

Fuente: <http://www.mindswap.org/2003/CancerOntology/htmls/>

Autor: NCI,2007

Este proyecto fue de mucha ayuda para la presente investigación ya que aportó el uso de herramientas como es el Protege 2000 y el OWL Plugin y OWL lenguaje que fueron

utilizados para el diseño y desarrollo de la Ontología para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

Project: nciOncology

Class Thorax

Concrete Class Extends

[Body Region](#)

Direct Instances:

None

Direct Subclasses:

None

Template Slots				
Slot name	Documentation	Type	Allowed Values/Classes	Cardinality
<i>rAnatomic_Structure_Has_Location</i>		Instance	Anatomy Kind	0:*
<i>rAnatomic_Structure_is_Physical_Part_of</i>		Instance	Anatomy Kind	0:*

Own Slots				
Slot name	Documentation	Type	Value	Ca
<i>:OWL-DISJOINT-CLASSES</i>		Instance		

[Return to class hierarchy](#)

Figura 6: Propiedades de la Clase Torax de NCI.

Fuente: <http://www.mindswap.org/2003/CancerOntology/htmls/Thorax.html>

Autor: NCI,2007

A través de la revisión de los antecedentes se logró relacionar cada concepto utilizado en la creación de la ontologías de cada estudio presentado en los antecedentes y así estructurar de manera correcta y entendible los pasos a seguir para el modelaje, desarrollo e implementación de los término del vocabulario de la presente investigación. Todo esto se logró gracias a la introducción del concepto de Web Semántica que es una extensión de la Web para dotarla de mayor significado, de manera que cualquier usuario pueda encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida.

BASES TEORICAS

Según Guarino (2002) la perspectiva de la construcción de ontologías, en un entorno tecnológico, permite abordar la problemática de la indización, recuperación y divulgación de la Información depositada en internet, facilitando una gestión rápida, eficaz, pertinente y permanentemente actualizada de las necesidades a cubrir. La aplicación de ontologías tecnológicas unifica la terminología de cada concepto y las relaciones entre ellos; y las ontologías de información unifican las estructuras de almacenamiento, de forma que pueden ser reutilizadas por varias aplicaciones informáticas con la misma fuente de información. Los modos de almacenamiento y recuperación de información han jugado un papel fundamental en el uso generalizado y estandarizado de sistemas documentales, permitiendo a los documentalistas tener unos modelos comunes y aceptados en la descripción y almacenamiento de información. Los conceptos se describen explícitamente para entender su significado, mediante acuerdos ontológicos. De este modo, un usuario que desee reutilizar una ontología desarrollada por otros puede conseguir la información de todos los conceptos que soporta, su taxonomía y los axiomas.

WEB SEMANTICA

Para la W3C En los últimos años, se ha comenzado a experimentar el uso de técnicas de representación del conocimiento para extender la web actual con metadatos, en la que los ordenadores no sólo serán capaces de presentar toda la información contenida en ella, sino que además podrán entenderla y gestionarla de forma «inteligente». Esta web mejorada se conoce actualmente como web semántica. Así, la web se convertirá en un espacio autonavegable y autocomprendible, siendo posible la relación entre los términos independientes de una misma búsqueda, ofreciendo como resultado sólo aquellas páginas web que realmente se adaptan a los requisitos especificados. De este modo, la información se convierte actualmente en conocimiento codificando los datos con metadatos mediante las ontologías. Cada una de estas ontologías se crea atendiendo a los principios de reutilización y diseño de forma consensuada. Los datos por si solos, sin tener una semántica asociada, no son de utilidad

por resultar ambiguos. Por lo tanto, hay que introducir esa semántica añadiendo datos que describen a su vez los propios datos o metadatos. Sin embargo, para dar una descripción formal de contenidos, hace falta una herramienta más potente que los metadatos: las ontologías. Mientras los metadatos estructuran contenidos, las ontologías estructuran la semántica de un recurso.

Pérez Valdez(2007) afirma que la Web Semántica de forma conceptual es una web extendida, la misma está dotada de un mayor significado. La Web semántica tiene como objetivo crear un medio universal para el intercambio de información basado en representaciones del significado de los recursos de la Web, de una manera inteligible para las máquinas. Con ello se pretende ampliar la interoperabilidad entre los sistemas informáticos y reducir la mediación de operadores humanos en los procesos inteligentes de flujo de información. Se desarrolla con lenguajes universales que permitirán a los usuarios encontrar respuestas a sus preguntas de una forma más rápida y sencilla gracias a la mejor estructuración de la información. La misma permitirá a los usuarios delegar tareas a las herramientas de la Web semántica, las cuales podrán ser capaces de procesar la información. La Web Semántica es una extensión de la World Wide Web en la que los contenidos de la Web pueden ser expresado mucho mas que en un lenguaje natural, y también en un formato que pueda ser entendido, interpretado y usado por diferentes aplicaciones, permitiéndoles buscar, compartir e integrar información más fácil. Ver las Capas de la Web Semántica Grafico 5.

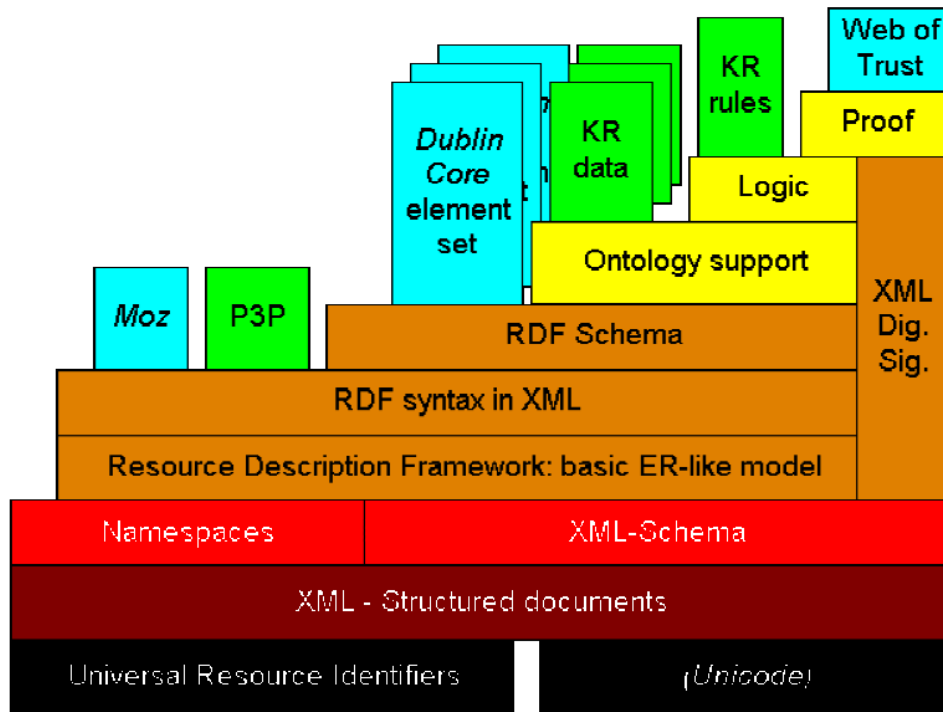


Figura 7: Capas de la Web Semántica.

Fuente: The Semantic Web: A new Form of Web Content that is Meaningful to Computers will Unleash a Revolution of new Possibilities.

Autor: Berners-Lee, 2007.

Funcionamiento de la Web Semántica:

Díaz Torner (2005), supone que la Web tiene la capacidad de construir una base de conocimiento sobre las preferencias de los usuarios y que, a través de una combinación entre su capacidad de conocimiento y la información disponible en Internet, sea capaz de atender de forma exacta las demandas de información por parte de los usuarios. Si esto ocurriese así en la vida real, el usuario, en su intento, por ejemplo, por encontrar todos los vuelos a Praga para mañana por la mañana, obtendría unos resultados exactos sobre su búsqueda. Sin embargo la realidad es otra. La información se procesara en términos de su **SEMÁNTICA**. La Web Semántica como infraestructura basada en metadatos aporta un camino para *razonar* en la Web, extendiendo así sus capacidades.

W3C dice que para obtener esa adecuada definición de los datos, la Web Semántica utiliza **RDF** y **OWL**, dos estándares que ayudan a convertir la Web en una infraestructura global en la que es posible compartir, y reutilizar datos y documentos entre diferentes tipos de usuarios.

- **RDF** proporciona **información descriptiva** simple sobre los **recursos** que se encuentran en la Web y que se utiliza,
- **OWL** es un **mecanismo para desarrollar** temas o **vocabularios específicos** en los que asociar esos recursos. Proporcionar un lenguaje para definir ontologías estructuradas que pueden ser utilizadas a través de diferentes sistemas.

RDF FRAMEWORK PARA LA DESCRIPCIÓN DE RECURSOS

Según W3C, 2004 Es un Framework para la descripción de recursos, ya que en un principio la web se construyó para el uso humano y, a pesar de que todo en ella era legible por máquina, los datos todavía no lo son del todo. Una de las soluciones es el uso de metadatos para describir los datos contenidos en ella mediante empaquetadores genéricos que permitan la integración y comunicación entre distintos esquemas de metadatos. RDF se creó en agosto de 1997, bajo el auspicio del W3C con la finalidad de crear una infraestructura para la descripción de recursos que proporcione una base para procesar meta-y posibilitar la interoperabilidad entre aplicaciones que intercambian información legible por máquina en el web.

RDF es una base para procesar metadatos; proporciona interoperabilidad entre aplicaciones que intercambian información legible por máquina en la Web. RDF destaca por la facilidad para habilitar el procesamiento automatizado de los recursos Web. RDF puede utilizarse en distintas áreas de aplicación; por ejemplo: en *recuperación de recursos* para proporcionar mejores prestaciones a los motores de búsqueda, en *catalogación* para describir el contenido y las relaciones de contenido disponibles en un sitio Web, una página Web, o una biblioteca digital particular, por los *agentes de software inteligentes* para facilitar el intercambio y para compartir conocimiento; en la

calificación de contenido, en la descripción de *colecciones de páginas* que representan un "documento" lógico individual, para describir los *derechos de propiedad intelectual* de las páginas web, y para expresar las *preferencias de privacidad* de un usuario, así como las *políticas de privacidad* de un sitio Web. RDF junto con las *firmas digitales* será la clave para construir el "Web Semántica" para el comercio electrónico, la cooperación y otras aplicaciones.

XML provee información semántica como un subproducto de definir la estructura del documento, porque define una estructura en forma de árbol para un documento, de manera que sus hojas contienen la información. Se puede observar entonces que la estructura y la semántica de un documento XML están entrelazadas, porque RDF provee medios para agregar semántica a un documento sin referirse a su estructura, como la DTD recomendada por el W3C. Su objetivo general es definir un mecanismo para describir recursos de forma que no cree dependencia alguna sobre un dominio de aplicación particular, ni defina, a priori, la semántica de algún dominio de aplicación.

El fundamento o base de RDF es un modelo para representar propiedades designadas y valores de propiedades. Las propiedades RDF pueden recordar atributos de los recursos y, en este sentido, corresponden con los tradicionales pares de atributo-valor. Las propiedades RDF representan también la relación entre recursos y por lo tanto, un modelo RDF puede parecer un diagrama entidad-relación. De forma más precisa, los esquemas RDF que son objetos específicos de la categoría del modelo de datos RDF son diagramas entidad-relación. En la terminología del diseño orientado a objetos, los recursos corresponden con objetos y las propiedades corresponden con objetos específicos y variables de una categoría.

OWL

Según W3C 2004, El Lenguaje de Ontología Web esta diseñado para el uso de aplicaciones que necesitan procesar el contenido de la información en vez de presentar la información a los seres humanos. Facilita en gran medida a la maquina que intérprete el contenido Web soportado por XML, RDF y RDF Schema (RDF-S) proporcionando un vocabulario adicional con una semántica formal. OWL tiene tres sublenguajes expresivos: OWL Lite, OWL DL y OWL Full.

OWL es un lenguaje de Ontologías *Web*. Lenguajes anteriores han sido utilizados para desarrollar herramientas y ontologías destinadas a comunidades específicas (especialmente para ciencias y aplicaciones específicas de comercio electrónico), no fueron definidos para ser compatibles con la arquitectura de la World Wide Web en general, y la Web Semántica en particular.

OWL rectifica esto proporcionando un lenguaje que utiliza la conexión proporcionada por RDF para añadir las siguientes capacidades a las ontologías:

- Capacidad de ser distribuidas a través de varios sistemas
- Escalable a las necesidades de la Web
- Compatible con los estándares Web de accesibilidad e internacionalización
- Abierto y extensible

OWL extiende RDFS para permitir la expresión de relaciones complejas entre diferentes clases RDFS, y mayor precisión en las restricciones de clases y propiedades específicas. Esto incluye por ejemplo: - los recursos para limitar las propiedades de clases con respecto a número y tipo, - los recursos para inferir qué elementos que tienen varias propiedades son miembros de una clase en particular - los recursos para determinar si todos los miembros de una clase tendrán una propiedad en particular, o si puede ser que sólo algunos la tengan - los recursos para distinguir entre relaciones uno-a-uno, varios-a-uno o uno-a-varios, permitiendo que las "claves externas" de las bases de datos puedan representarse en una ontología - los recursos para expresar relaciones entre clases definidas en diferentes documentos en la Web - los recursos para construir

nuevas clases a partir de uniones, intersecciones y complementos de otras, y - los recursos para restringir rangos y dominios para especificar combinaciones de clases y propiedades.

Los tres sublenguajes de OWL Según W3C 2006.

OWL proporciona tres lenguajes, cada uno con nivel de expresividad mayor que el anterior, diseñados para ser usados por comunidades específicas de desarrolladores y usuarios.

- *OWL Lite* está diseñado para aquellos usuarios que necesitan principalmente una clasificación jerárquica y restricciones simples. Por ejemplo, a la vez que admite restricciones de cardinalidad, sólo permite establecer valores cardinales de 0 ó 1. Debería ser más sencillo proporcionar herramientas de soporte a OWL Lite que a sus parientes con mayor nivel de expresividad, y OWL Lite proporciona una ruta rápida de migración para tesauros y otras taxonomías. OWL Lite tiene también una menor complejidad formal que OWL DL. *Nota del traductor:* como aclaración del término "tesauro":
- OWL DL está diseñado para aquellos usuarios que quieren la máxima expresividad conservando completitud computacional (se garantiza que todas las conclusiones sean computables), y reusable (todos los cálculos se resolverán en un tiempo finito). OWL DL incluye todas las construcciones del lenguaje de OWL, pero sólo pueden ser usados bajo ciertas restricciones (por ejemplo, mientras una clase puede ser una subclase de otras muchas clases, una clase no puede ser una instancia de otra). OWL DL es denominado de esta forma debido a su correspondencia con la Lógica de Descripción, un campo de investigación que estudia la lógica que compone la base formal de OWL.
- OWL Full está dirigido a usuarios que quieren máxima expresividad y libertad sintáctica de RDF sin garantías computacionales. Por ejemplo, en OWL Full una clase puede ser considerada simultáneamente como una colección de clases individuales y como una clase individual propiamente dicha. OWL Full permite una ontología para aumentar el significado del vocabulario preestablecido (RDF o OWL). Es poco probable que cualquier software de razonamiento sea capaz de obtener un razonamiento completo para cada característica de OWL Full.

ONTOLOGIA

Una de las primeras definiciones en el área de la ciencia de la información la hizo Neches, 1991 y su equipo de trabajo: “una ontología define los términos básicos y relaciones incluyendo el vocabulario de un área así como las reglas para la combinación de términos y relaciones para definir ampliaciones de un vocabulario”. Se puede decir que esta definición aporta unas líneas a seguir para crear una ontología: identificar términos básicos y relaciones entre los términos, identificar reglas para combinar las relaciones, aportar definiciones de los términos y las relaciones. Así que según esta definición, una ontología no incluye solo los términos que son explícitamente definidos en ella, sino que también los términos que pueden ser deducidos usando reglas.

En 1993, Gruber dio una de las definiciones más empleadas: “las ontologías se definen como una especificación explícita de una conceptualización”. Una conceptualización es una abstracción, una vista simplificada del mundo que queremos representar”. Una ontología no es una base de datos ni un programa (porque tienen sus propios formatos internos), no es una conceptualización (porque no es una especificación, es sólo una vista) ni una tabla de contenidos (aunque una taxonomía sí es una ontología).

En 1997, Borst modificó ligeramente la definición de Gruber diciendo que, “las ontologías se definen como una especificación formal de una conceptualización compartida”. Estas dos definiciones fueron ampliamente explicadas por Studer (1998) y su equipo de trabajo. Para Guarino, 1998 “una ontología es una fuerte estructura semántica que codifica reglas implícitas restringiendo la estructura de una porción de la realidad”. Según Weigand 1997, una ontología puede tomar una gran variedad de formas, pero necesariamente incluirá un *vocabulario de términos*, y alguna *especificación de su significado*. Esto incluye definiciones y una indicación de cómo los conceptos se interrelacionan lo que colectivamente impone una estructura sobre el dominio y restringe las posibles interpretaciones de los términos. Con las definiciones dadas, se puede ver que una ontología puede ser, una teoría lógica, una descripción

formal de semántica, el vocabulario de una teoría lógica y una especificación de una conceptualización.

Alexander Maedche, Boris Motik. Centro de Búsqueda para Tecnología de Información. Universidad de Karlsruhe.2009 Afirman que la mayoría de las empresas trabajan hoy en día con la creación de sistemas pequeños individuales para cada área, no se tiene uno central que se enfoque en la gestión de conocimiento, esto puede traer ciertas dificultades porque cada uno tal vez tenga su propio modelo conceptual, la idea es integrarlos para resolver problemas de semántica heterogénea entre los modelos. Ellos plantean que el uso de ontologías ayuda en gran medida a la representación del conocimiento y a su vez, esto hace más fácil la comprensión de los datos plasmados en ella es decir, una ontología bien diseñada, va permitir que el conocimiento representado en ella sea de fácil comprensión para los usuarios y por ende les permitirá tomar mejores decisiones.

Las ontologías son acuerdos, en un contexto social, para cubrir una serie de objetivos. Se crean, entre otras cosas, para:

- (a) Permitir el intercambio de datos entre programas.
- (b) Simplificar la unificación (o traducción) de distintas representaciones.
- (c) Facilitar la comunicación entre personas.

Gaitanou (2005) afirma que las Ontologías son especialmente útiles, dado que ellas tienen la habilidad de facilitar la interoperabilidad entre sistemas de computadoras, esto se logra con la definición de un vocabulario común, los errores entre conceptos son drásticamente reducidos y la disponibilidad y la situación existente de los diversos recursos se describen con precisión y claridad, nos señala un ejemplo de Ontología descrito en el Grafico 6. También realiza una enumeración de las aplicaciones de la Ontologías descritas a continuación:

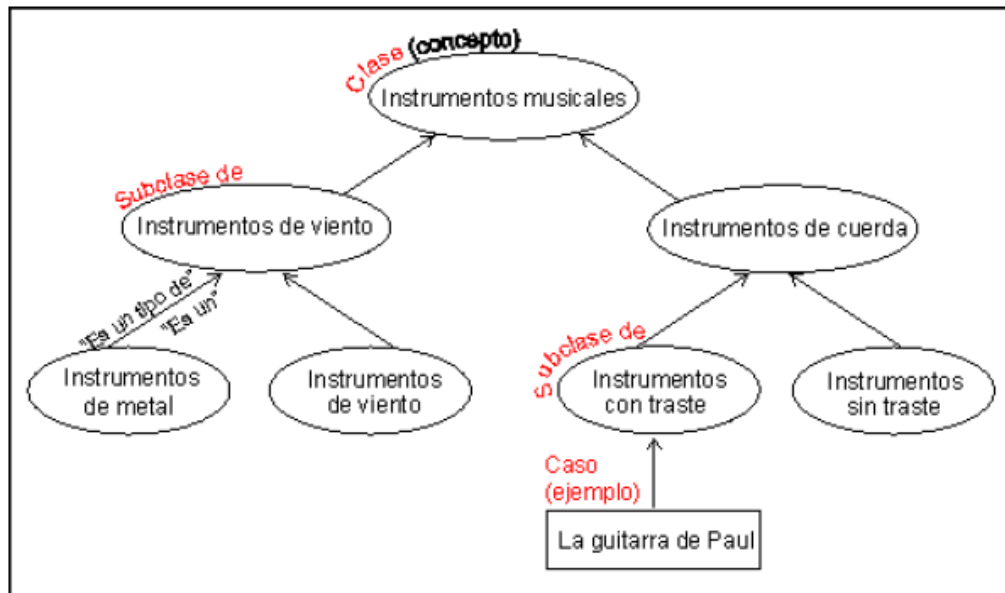


Figura 8. Ejemplo de Ontología.
 Fuente: Ontology and Applications
 Fuente: Gaitanou, 2004.

TIPOS DE ONTOLOGIAS

Van Heijst, 1996 propone una clasificación de las ontologías de acuerdo con la cantidad y tipo de la conceptualización. Así se pueden diferenciar los siguientes tipos de ontologías:

- **Terminológicas:** especifican los términos que son usados para representar conocimiento en el universo de discurso. Suelen ser usadas para unificar vocabulario en un dominio determinado.
- **De Información:** especifican la estructura de almacenamiento de bases de datos. Ofrecen un marco para el almacenamiento estandarizado de información.
- **De modelado del conocimiento:** Especifican conceptualizaciones del conocimiento. Contienen una rica estructura interna y suelen estar ajustadas al uso particular del conocimiento que describen.

Stevens, 1998 señala que Hay otras posibles clasificaciones de ontologías atendiendo a diversos criterios, por ejemplo si se atiende al asunto que conceptualizan, se distinguen tres tipos fundamentales de ontologías:

- **Ontologías de un dominio**, en las que se representa el conocimiento especializado pertinente de un dominio o subdominio, como la medicina, las aplicaciones militares, tráfico etc.
- **Ontologías genéricas**, en las que se representan conceptos generales y fundacionales del conocimiento como las estructuras parte/todo, la cuantificación, los procesos o los tipos de objetos, independientes de un dominio en particular.
- **Ontologías representacionales**, en las que se especifican las conceptualizaciones que subyacen a los formalismos de representación del conocimiento, por lo que también se denominan *meta-ontologías (meta-level o top-level ontologies)*.

A estos tres tipos, Guarino [Gua98] añade las ontologías que han sido creadas para una actividad o tarea específica (denominadas *task ontologies*), como por ejemplo la venta de productos o el diagnóstico de una enfermedad y las ontologías creadas para una aplicación específica. En dicho artículo, se pone de manifiesto la posibilidad de desarrollar diferentes tipos de ontologías teniendo en cuenta el nivel de generalidad.

APLICACIONES DE LAS ONTOLOGIAS

Aplicación de las ontologías en la Web Semántica

(a) *Indización de documentos*: La indexación de un sitio web con apoyo de una ontología terminológica comienza con la extracción de los términos más relevantes de cada página, y después de asociar a estos términos conceptos candidatos, se evalúa la capacidad de representación de la página de cada uno de estos conceptos, que determina su nivel de representatividad, y finalmente se construye el índice. De esta manera, las consultas se procesan a un nivel conceptual, lo que reportará un mayor grado de acierto.

- (b) *Agrupamiento*: Las técnicas de *clustering* permiten el crecimiento de un sistema mediante la adición de procesadores o CPU (*central processing units*) a la unidad primitiva. Las ontologías aportan las herramientas para que los distintos equipos puedan entenderse entre sí y funcionar como si fuera uno sólo.
- (c) *Servicios web*: Las ontologías representarán los datos en la red de tal forma que puedan ser utilizados y comprendidos por las máquinas sin necesidad de la intervención humana.

Aplicación de Ontologías en ingeniería del Conocimiento

- (a) **Ingeniería del Conocimiento**: Se usa la Ontología como un conjunto de Esquemas o Contenedores de Conocimiento creando un glosario de la terminología del dominio.
- (b) **Procesamiento del Lenguaje Natural**: Una Ontología tiene la capacidad de mantener la definición de elementos gramaticales del lenguaje y sus relaciones permitiendo el análisis sintáctico de un texto.
- (c) **Reutilización**: Permite reutilizar el conocimiento para sistemas futuros. Se pueden integrar ontologías para la constitución de una nueva que mejore la conceptualización que aportaban las demás por separado.

Aplicación de Ontologías en Sistemas de Información

- (a) **Interoperatividad entre Sistemas Heterogéneos**: Las Ontologías se presentan como una solución para lograr una integración inteligente.
- (b) **Sistemas de Información Cooperativa**: La meta es lograr que múltiples sistemas de información sean capaces de trabajar de forma cooperativa combinando datos y funcionalidades con ayuda de las ontologías.
- (c) **Mejorar la Formulación de Consultas**: De este modo, se reducirá el ruido y el silencio en los resultados de una búsqueda, lo que permitirá que no se omitan aquellos resultados, que aún siendo conceptualmente sinónimos al de la consulta, no se encuentran por ser distintos terminológicamente.

Según Gruber T 2003, las ontologías tienen los siguientes componentes que servirán para representar el conocimiento de algún dominio

- (a) **Conceptos:** son las ideas básicas que se intentan formalizar. Los conceptos pueden ser clases de objetos, métodos, planes, estrategias, procesos de razonamiento, entre otros.
- (b) **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.
- (c) **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 parecer funciones como categorizar-clase, asignar fecha.
- (d) **Instancias:** se utilizan para representar objetos determinados de un concepto.
- (e) **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”.

Guarino (2002) clasifica las ontologías de acuerdo con su dependencia y relación con una tarea específica desde un punto de vista:

- (a) **Ontologías de Alto Nivel o Genéricas:** Describen conceptos más generales. En relación con los Sistemas de Información, estas ontologías describirían conceptos básicos.
- (b) **Ontologías de Dominio:** Describen un vocabulario relacionado con un dominio genérico. Por ejemplo, podría ser una descripción de datos y entidades relacionados con la sensorización remota con un ambiente urbano.
- (c) **Ontologías de Tareas o de Técnicas básicas:** Describen una tarea, actividad o artefacto. Por ejemplo la evaluación de la contaminación sonora en ambientes urbanos o la descripción de características generales de componentes, procesos o funciones.

(d) **Ontologías de Aplicación:** Describen conceptos que dependen tanto de un dominio específico como de una tarea específica y, generalmente son una especialización de ambas. Federico Fonseca et al. (2000) proponen que este tipo de ontologías nazcan a partir de una combinación de ontologías de niveles superiores. Ellas representan las necesidades de los usuarios relacionados con una aplicación específica como, por ejemplo, una evaluación de disponibilidad de camarones en la costa de Espírito Santo (Brasil). Otra clasificación es la propuesta por Van Heist (2005) descrita a continuación:

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

(b) **Ontologías de Información:** Especifican la estructura de almacenamiento de bases de datos. Ofrecen un marco para el almacenamiento estandarizado de información.

(c) **Ontologías de Modelado de Conocimiento:** Especifican conceptualizaciones del conocimiento. Contienen una rica estructura y suelen estar ajustadas al uso particular del conocimiento que describen.

Herrera, D.2009 afirma El uso de ontologías para la *recuperación de la información* tiene ciertas ventajas sobre los métodos de acceso simples basados en palabras clave.

Estas ventajas son:

- Una ontología suministra un vocabulario compartido común para expresar información sobre el contenido de los documentos.
- Las ontologías incluyen axiomas para especificar relaciones entre conceptos.
- Facilidad de uso que ofrece, frente a otros sistemas que requieren un mayor entrenamiento por parte de los compiladores de la ontología.
- Posibilidad de crear un recurso independiente y autónomo.
- Riqueza de la información que permite representar.
- Trazar una estructura conceptual tan detallada y completa como sea necesario.

XML

Berners Lee (2006) señala que el lenguaje XML, siglas en inglés de *Extensible Markup Language* (lenguaje de marcas extensible), es un metalenguaje extensible de etiquetas desarrollado por el World Wide Web Consortium (W3C). Es una simplificación y adaptación del SGML y permite definir la gramática de lenguajes específicos (de la misma manera que HTML es a su vez un lenguaje definido por SGML). Por lo tanto XML no es realmente un lenguaje en particular, sino una manera de definir lenguajes para diferentes necesidades. Algunos de estos lenguajes que usan XML para su definición son XHTML, SVG, MathML. XML no ha nacido sólo para su aplicación en Internet, sino que se propone como un estándar para el intercambio de información estructurada entre diferentes plataformas. Se puede usar en bases de datos, editores de texto, hojas de cálculo y casi cualquier cosa imaginable. XML es una tecnología sencilla que tiene a su alrededor otras que la complementan y la hacen mucho más grande y con unas posibilidades mucho mayores. Tiene un papel muy importante en la actualidad ya que permite la compatibilidad entre sistemas para compartir la información de una manera segura, fiable y fácil.

IBM (2003) señala las principales ventajas del XML enumeradas a continuación:

- Es extensible: Después de diseñado y puesto en producción, es posible extender XML con la adición de nuevas etiquetas, de modo que se pueda continuar utilizando sin complicación alguna.
- El analizador es un componente estándar, no es necesario crear un analizador específico para cada versión de lenguaje XML. Esto posibilita el empleo de cualquiera de los analizadores disponibles. De esta manera se evitan *bugs* y se acelera el desarrollo de aplicaciones.
- Si un tercero decide usar un documento creado en XML, es sencillo entender su estructura y procesarla. Mejora la compatibilidad entre aplicaciones.

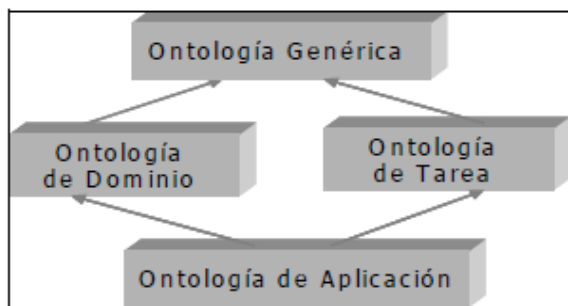


Figura 9. Tipos de ontologías según el nivel de dependencia de una tarea particular o punto de vista.

Fuente: Formal Ontology and Information Systems. Formal Ontology in Information Systems.

Autor: Guarino, 1998.

Alfredo Reino en la INTRODUCCIÓN A XML EN CASTELLANO (2007) dice que los DTD definen los tipos de elementos, atributos y entidades permitidas, y puede expresar algunas limitaciones para combinarlos. Los documentos que se ajustan a su DTD, se denominan "válidos". El concepto de "validez" no tiene nada que ver con el de estar "bien formado". Un documento "bien formado" simplemente respeta la estructura y sintaxis definida por la especificación de XML. Un documento "bien formado" puede además ser "válido" si cumple las reglas de una DTD determinada. También existen documentos XML sin una DTD asociada, en ese caso no son "válido", pero tampoco "inválido"... simplemente "bien formados"... o no. Una DTD puede residir en un fichero externo, y quizás compartido por varios (puede que miles) de documentos. O bien, puede estar contenido en el propio documento XML, como parte de su declaración de tipo de documento.

En el mismo documento Alfredo Reino afirma que Un "schema XML" es algo similar a un DTD, es decir, que define qué elementos puede contener un documento XML, cómo están organizados, y que atributos y de qué tipo pueden tener sus elementos.

Según Reino las ventajas de los schemas con respecto a los DTDs son:

- Usan sintaxis de XML, al contrario que los DTDs.
- Permiten especificar los tipos de datos.
- Son extensibles.

Por ejemplo, un schema nos permite definir el tipo del contenido de un elemento o de un atributo, y especificar si debes ser un número entero, o una cadena de texto, o una fecha, etc. Los DTDs no nos permiten hacer estas cosas.

Veamos un ejemplo de un documento XML, y su schema correspondiente:

```
<documento xmlns="x-schema:personaSchema.xml">
<persona id="fulano">
<nombre>Fulano Menganez</nombre>
</persona>
</documento>
```

Martín (2004). Universitat de Valencia señala Diez principios del diseño de XML

- 1.- Ser utilizable directamente sobre Internet
- 2.- Facilidad para escribir programas que procesen sus documentos.
- 3.- Soportan una amplia variedad de aplicaciones
- 4.- Compatible con SGML
5. Reducir las características opcionales, idealmente ninguna.
- 6.- Ser legible por un humano y razonablemente claro
- 7.- Facilidades para la creación de documentos.
- 8.- Diseño fácil, formal y conciso
- 9.- La brevedad del marcado no es importante.
- 10.- Definir vocabularios compartidos.

VOCABULARIOS EN XML

Galan Herrero (2007) afirma que RDF (Resource Description Framework) es uno de los vocabularios de XML que mas se ha hecho sentir últimamente, RDF propone ser el gran habilitador de la web semántica y convertirse en el standard universal para los metadatos en la web. Con el fin de lograr sus objetivos, RDF utiliza el lenguaje XML (eXtensible Markup Language) como método para representar y "transportar" la información. Hay que tener muy presente que XML no es un lenguaje de etiquetado. Se

trata de un lenguaje que establece un conjunto de reglas que permiten la creación de lenguajes de etiquetado. Para informar sobre el contenido de cada conjunto de datos, su interpretación, y establecer la forma más correcta de trabajar con ellos debe crearse una DTD (Document Type Definition) donde se plasme el esquema específico de trabajo (Hjelm, Johan , 2001). Por ese motivo RDF utiliza una DTD de XML para desarrollar sus etiquetas. RDF se beneficia de XML por su flexibilidad a la hora de generar nuevos conjuntos de etiquetas, su orientación multiplataforma y por que proporciona el mecanismo semántico perfecto para expresar la descripción de cualquier tipo de recurso.

El Framework de Descripción de Recurso RDF provee un medio de agregar semántica a un documento sin referirse a su estructura. RDF es una aplicación XML recomendada como estándar por la W3C

Galan Herrero (2007) realiza una Comparación de RDF con otros lenguajes:

- RDF = XML + restricciones estructurales para asegurar consistencia
- El Orden es relevante en XML
- XML permite mezclar texto y tags
- RDF tiene una interpretación (semántica) propia (graph data model)

También enumera Las Ventajas de RDF:

- Definición de clases (también XML Schema).
- Jerarquía de clases, se puede examinar y manipular en tiempo de ejecución.
- Definición dispersa de ontología y de instancias (y metadatos separado de recursos).
- Gran facilidad de extensión: la estructura de los schemas es abierta, se pueden añadir relaciones a posteriori.

El modelo de datos de RDF provee tres tipos de objetos: recursos, propiedades y sentencias, explicadas a continuación:

- Un recurso es una entidad que puede ser referenciada por un Identificador Único de Recursos (URI).
- Una propiedad define una relación binaria entre recursos y/o valores atómicos de los tipos de datos primitivos provistos por XML.
- Una sentencia especifica un valor en una propiedad para un determinado recurso.

Todo lo expresable en RDF, es expresable en sintaxis lineal de XML, podría surgir la pregunta entonces de porque es necesario RDF si todo metadato representado en RDF puede también ser representado en XML. La razón es que RDF provee un modo estándar de representar metadatos en XML usando directamente XML para representar metadatos, podrían obtenerse varias representaciones diferentes. Para procesar modelos RDF, éstos deben ser serializados en XML previamente, de esta forma RDF explota la denotación de XML para permitir que diferentes orígenes de información puedan intercambiar el conocimiento que expresa.

Para la presente investigación se utilizará como motor de Base de datos MySQL y a continuación se describe sus principales características:

MySQL es un gestor de base de datos sencillo de usar y increíblemente rápido. También es uno de los motores de base de datos más usados en Internet, la principal razón de esto es que es gratis para aplicaciones no comerciales. Las características principales de MySQL son:

- **Es un gestor de base de datos.** Una base de datos es un conjunto de datos y un gestor de base de datos es una aplicación capaz de manejar este conjunto de datos de manera eficiente y cómoda.
- **Es una base de datos relacional.** Una base de datos relacional es un conjunto de datos que están almacenados en tablas entre las cuales se establecen unas relaciones para manejar los datos de una forma eficiente y segura. Para usar y gestionar una base de datos relacional se usa el lenguaje estándar de programación SQL.
- **Es Open Source.** El código fuente de MySQL se puede descargar y está accesible a cualquiera, por otra parte, usa la licencia GPL para aplicaciones no comerciales.
- **Es una base de datos muy rápida,** segura y fácil de usar. Gracias a la colaboración de muchos usuarios, la base de datos se ha ido mejorando

optimizándose en velocidad. Por eso es una de las bases de datos más usadas en Internet.

- **Existe una gran cantidad de software que la usa.**

METHONTOLOGY

Corcho, 2007 en su documento Metodologías, herramientas y lenguajes para la construcción de de Ontologías Methontology es una metodología creada en el Laboratorio de Inteligencia Artificial de la Universidad Técnica de Madrid. La creación de la ontología puede empezar desde cero o en base a la reutilización de otras existentes. Methontology incluye la identificación del proceso de desarrollo de la ontología (calendario, control, aseguramiento de calidad, adquisición de conocimiento), un ciclo de vida basado en la evolución de prototipos, para lo cual se sigue -los pasos definidos en el estándar IEEE 1074 de desarrollo de software (Corcho, 2001): que son:

- **Especificación:** Consiste en delimitar los objetivos de su creación (compartir información entre personas o por agentes software; permitir la reutilización del conocimiento de un dominio; hacer explícitas las suposiciones que se efectúan en un dominio; separar el conocimiento del dominio del conocimiento sobre su fabricación o forma de operación; o simplemente analizar el conocimiento del dominio), decidir el dominio de actuación de la ontología (para no modelizar objetos poco relevantes en perjuicio de otros más importantes), quién la usará y para qué, las preguntas a las que deberá responder (ayuda al establecimiento de las dos anteriores), y quién se encargará de su mantenimiento (decidiendo si se limitará a introducir nuevas instancias, se permitirá la modificación de conceptos o atributos, etc.)
- **Conceptualización:** Consiste en crear un glosario de términos que pertenecen al dominio, definirlos y crear una taxonomía (estableciendo una clasificación o jerarquía entre los conceptos, sus niveles, las relaciones entre ellos, sus instancias, sus propiedades o atributos, e igualmente los axiomas o reglas).

- **Formalización:** Proceso consistente en convertir el modelo anterior en un modelo formal o semi computable. Se puede emplear en este paso una herramienta como Protégé.
- **Implementación:** Convierte el modelo formalizado en un modelo computable mediante un lenguaje para construcción de ontologías.
- **Mantenimiento:** Labor que puede acarrear desde el borrado de instancias ya inútiles o la incorporación de nuevas instancias que se han ido produciendo con el tiempo, hasta las tareas de introducción de cambios en el contenido de la información, ya sea redefiniendo atributos, relaciones o incluso conceptos.

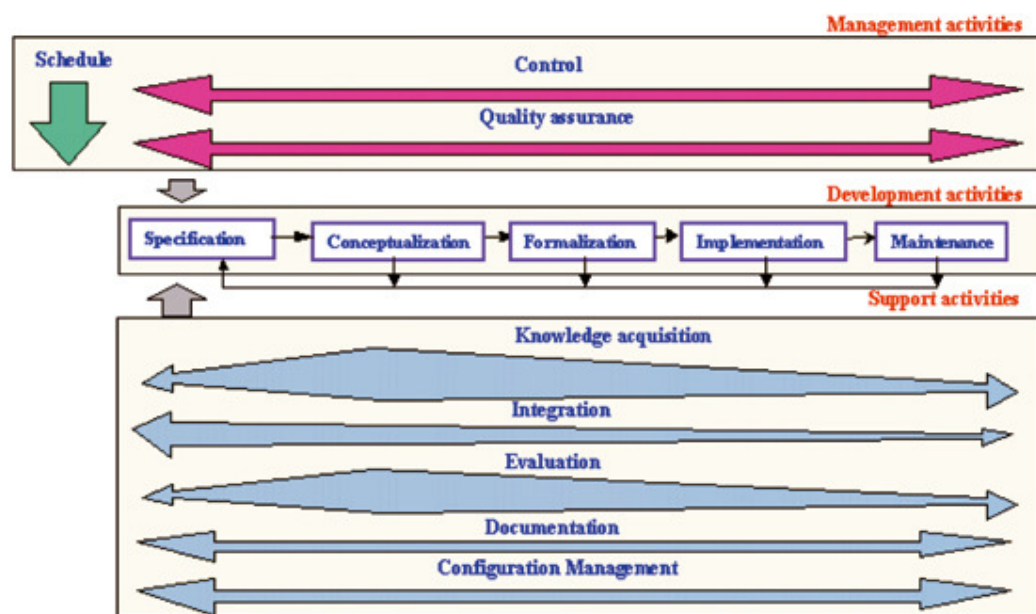


Figura 10. Arquitectura Creación de Ontología.

Fuente: Metodologías, herramientas y lenguajes para la construcción de de Ontologías.

Autores: Oscar Corcho, Mariano Fernández López, Asunción Gómez Pérez. 2007.

Methontology propone conceptualizar las ontologías con un juego de representaciones intermedias (RI) gráfica y tabular. Dichas RI permiten modelar los componentes que se describen a continuación:

- Conceptos. Se refieren en un sentido amplio. Son normalmente organizados en taxonomías a través de mecanismos de herencia (subclases).
- Relaciones. Se refiere al tipo de asociación entre los conceptos de un dominio. Si la relación es entre dos conceptos, se denomina relación binaria. Una relación binaria importante es Subclase de (Subclass-Of), que se usa para construir la taxonomía de clases. Cada relación binaria puede tener una relación inversa, que une los conceptos en la dirección opuesta.

A Continuación se muestra las tareas de la Metodología Methoontology y el ciclo de Vida de una Ontología:

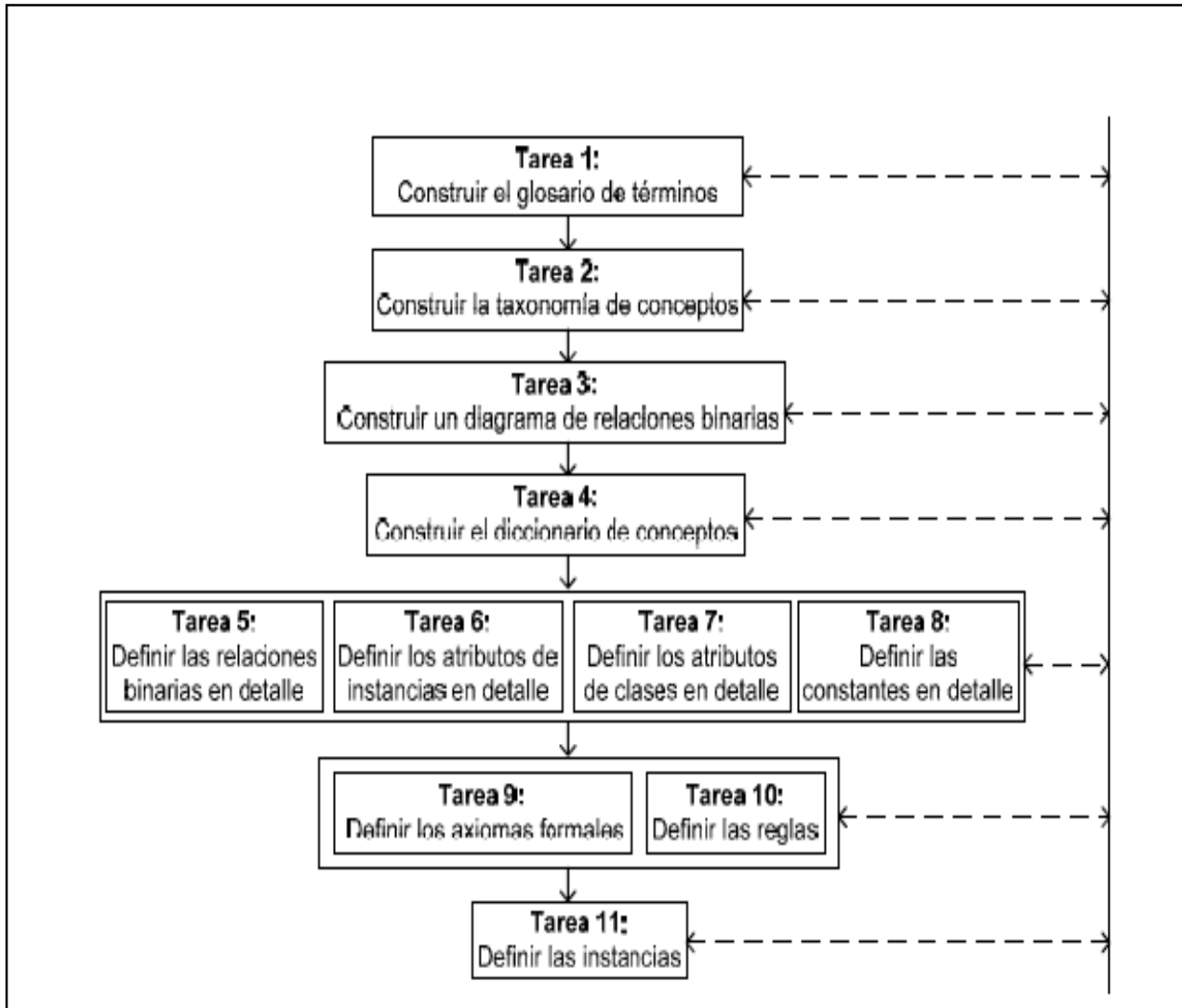


Figura 11: Tareas de Conceptualización de la Ontología.

Fuente: Metodologías, herramientas y lenguajes para la construcción de de Ontologías.

Autores: Oscar Corcho, Mariano Fernández López, Asunción Gómez Pérez. 2007.

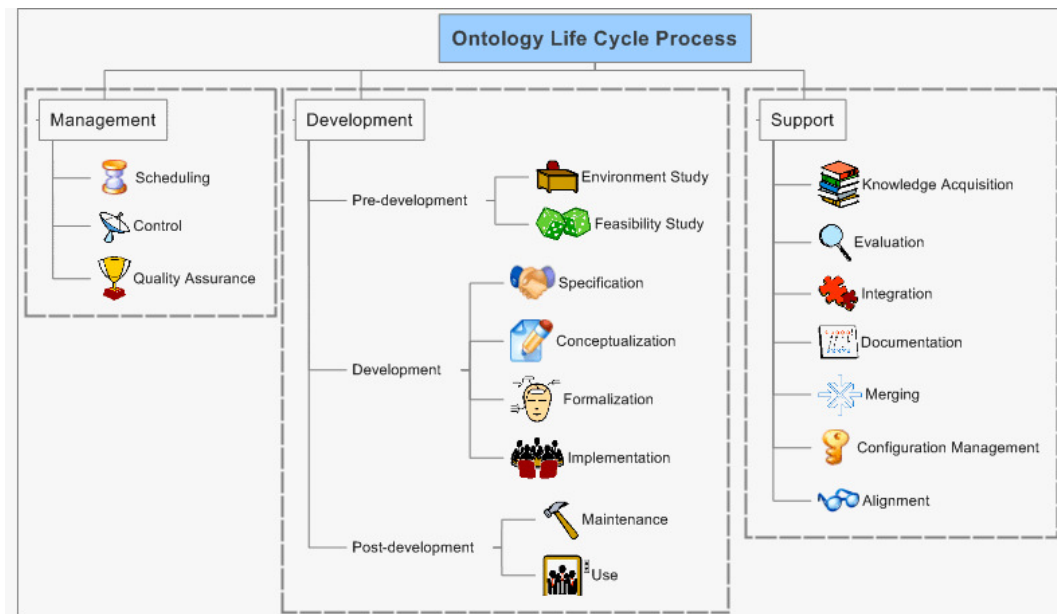


Figura 12. Ciclo de Vida de una Ontología.

Fuente: Metodologías, Herramientas, y Lenguajes para la Construcción de Ontologías.

Autor: Corcho, O., Fernández, M, Gómez, A. 2007

HERRAMIENTA PROTÉGÉ 2000

Según W3C,2005 Protege es un editor de ontologías open source desarrollado principalmente por la Universidad de Standford, con el apoyo de varias agencias gubernamentales y privadas, éste implementa un amplio entorno para el diseño, modelamiento, implementación, manipulación y visualización de ontologías en varios formatos. Es importante tener en cuenta que para poder hacer uso de toda la potencias de Protege se deben tener instaladas algunos pluggin y aplicaciones adicionales, como:

- RACER – Razonador basado en lógicas descriptivas.
- OWLViz – Pluggin de Protege para visualizar grafos
- OWL Wizard - Pluggin de Protege.
- GraphViz.

Como muchas otras herramientas de modelamiento, Protégé se encuentra separada entre una parte de modelo y una parte de vista. El modelo de Protégé es el mecanismo de representación interna de ontologías y bases de conocimiento. Los

componentes de la vista de Protégé proveen al usuario una interfase para visualizar y manipular el modelo, teniendo en cuenta todas las características del dominio. Por medio de Protégé la elaboración de propiedades y restricciones permite generar inferencia sobre el dominio, generando conceptualización y conocimiento del mismo.

Las aplicaciones desarrolladas con Protégé son empleadas en resolución de problemas y toma de decisiones en dominios particulares. La herramienta Protégé emplea una interfaz de usuario que facilita la creación de una estructura de *frames* con clases, slots e instancias de una forma integrada. Con la ayuda de este software se pueden construir aplicaciones basadas en el conocimiento que representen la información de forma ontológica. La característica de este tipo de herramientas es que utilizan un lenguaje muy flexible para ser utilizado en la Web, por eso se ha acuñado el término de Web Semántica, es decir, que facilitan la tarea no sólo de ser manejable a través de la Web, sino que hacen una descripción semántica de la información, con lo que todo el sistema es menos rígido y por tanto se hace más flexible y potente. Protégé reconoce por tanto, Frames, XML Schema, RDF Schema y OWL, que son lenguajes semánticos utilizados en la Web, en contraposición a la rigidez del HTML. El lenguaje con más perspectiva de futuro en la representación del conocimiento usando ontologías es el OWL. Protégé permite realizar programas en OWL con una gran facilidad, debido a su entorno gráfico, tan sólo con el uso de otro programa asociado a él que se denomina Protégé-OWL Plugin.

Base de Conocimiento

Una base de conocimiento es una forma avanzada de base de datos que no solo pretende almacenar, recuperar y modificar grandes cantidades de información, sino también, plasmar elementos de conocimiento (generalmente en forma de hechos y reglas de inferencia), así como la forma en que éste ha de ser utilizado.

Este recurso lingüístico es un modelo de un mundo/empresa/sección de la realidad, como declara Mylopoulos (1982) citado por Pérez Hernández (2002), en donde se considera el mundo/universo como una colección de individuos o entidades y

el conjunto de relaciones que existen entre esos individuos. La colección de individuos, más las relaciones entre éstos, constituye un estado, cuyos cambios causan la creación o modificación de individuos o de las relaciones entre ellos. La base de conocimiento utiliza al *esquema de representación* como notación precisa para representar el conocimiento que contiene. El tipo de *esquema de representación* de mayor difusión son las redes semánticas que estructuran sus datos en nodos que simbolizan a los conceptos unidos por arcos que representan las relaciones conceptuales. La red semántica también utiliza la herencia como método de inferencia, donde los nodos inferiores heredan las características de los nodos superiores permitiendo así una economía de codificación. Para realizar dicha investigación se utilizará el lenguaje de programación descrito a continuación:

LENGUAJE DE PROGRAMACION JAVA

Según Álvarez,2001 Java es un lenguaje de programación con el que podemos realizar cualquier tipo de programa. En la actualidad es un lenguaje muy extendido y cada vez cobra más importancia tanto en el ámbito de Internet como en la informática en general. Está desarrollado por la compañía Sun Microsystems con gran dedicación y siempre enfocado a cubrir las necesidades tecnológicas más punteras. El mismo autor afirma que una de las principales características por las que Java se ha hecho muy famoso es que es un lenguaje independiente de la plataforma. Eso quiere decir que si hacemos un programa en Java podrá funcionar en cualquier ordenador del mercado. Es una ventaja significativa para los desarrolladores de software, pues antes tenían que hacer un programa para cada sistema operativo, por ejemplo Windows, Linux, Apple, etc. Esto lo consigue porque se ha creado una Máquina de Java para cada sistema que hace de puente entre el sistema operativo y el programa de Java y posibilita que este último se entienda perfectamente. La independencia de plataforma es una de las razones por las que Java es interesante para Internet, ya que muchas personas deben tener acceso con ordenadores distintos. Pero no se queda ahí, Java está desarrollándose incluso para distintos tipos de dispositivos además del ordenador como móviles, agendas y en general para cualquier cosa que se le ocurra a la industria.

Jena Semantic Web Framework

Jena es un framework open-source para la construcción de aplicaciones Java relacionadas con la Web Semántica. Incluye:

API para RDF (*Resource Description Framework*)

API para OWL (*Ontology Web Language*)

Lectura y escritura RDF en formato RDF/XML, N3 y N-Triples

Motor de consultas SPARQL.

Almacenamiento en memoria y almacenamiento persistente.

Pellet OWL Razonador para JAVA

Pellet es un razonador para OWL. Proporciona un estándar para servicios de razonamiento de Ontologías basados en OWL. Las principales características incluye:

- Esta desarrollado para aplicaciones de código abierto.
- Destinado para aplicaciones que tengas propiedades, clases y objetos.
- Incorpora optimizaciones para consultas nominales y conjuntivas.
- Incrementa el razonamiento para ontología dando respuesta acerca de si la ontología es consistente y correcta.

NETBEANS 6.8

Según Alvarez, 2001, NetBeans es un proyecto exitoso de código abierto con una gran base de usuarios, una comunidad en constante crecimiento, y con cerca de 100 socios (¡y creciendo!) en todo el mundo. Sun Microsystems fundó el proyecto de código abierto NetBeans en junio 2000 y continúa siendo el patrocinador principal de los proyectos. Al día de hoy hay disponibles dos productos: el NetBeans IDE y NetBeans Platform.

NetBeans IDE es un entorno de desarrollo – una herramienta para que los

programadores puedan escribir, compilar, depurar y ejecutar programas. Está escrito en Java – pero puede servir para cualquier otro lenguaje de programación. Existe además un número importante de módulos para extender el NetBeans IDE. NetBeans IDE es un producto libre y gratuito sin restricciones de uso.

También está disponible NetBeans Platform; una base modular y extensible usada como estructura de integración para crear grandes aplicaciones de escritorio. Empresas independientes asociadas, especializadas en desarrollo de software, proporcionan extensiones adicionales que se integran fácilmente en la plataforma y que pueden también utilizarse para desarrollar sus propias herramientas y soluciones.

BASES LEGALES

Según **Córdova Aponte** (2007) En las Bases Legales, tal como la denominación de la sección lo indica, se incluyen todas las referencias legales que soportan el tema o problema de investigación. Para ello, se pueden consultar: (a) la constitución nacional; (b) las leyes orgánicas; (c) las gacetas gubernamentales. Dentro de este marco para la presente investigación se consultó Gobierno Electrónico en Venezuela, *ISSN 1316-6239.2007*, el cual enfatiza que uno de los aspectos importantes es el de lograr una mayor inclusión de los sectores académicos y empresariales, las políticas que se orientan a hacer del gobierno electrónico una realidad, trascendiendo al medio académico y educativo que es la mayor vertiente que debe involucrarse en ese esfuerzo y en ello tiene alta responsabilidad las instituciones universitarias privadas y públicas.

En el artículo 110 de la Constitución General de la República. Diciembre 1999, se destaca la importancia de la Innovación en los sistemas desarrollados por empresas públicas y privadas para el desarrollo económico del país, debe señalarse que la presente investigación pretende alcanzar esta innovación. Dicho artículo expresa lo siguiente: *“El Estado reconocerá el interés público de la ciencia, la tecnología, el conocimiento, la innovación y sus aplicaciones y los servicios de información necesarios por ser*

instrumentos fundamentales para el desarrollo económico, social y político del país, así como para la seguridad y soberanía nacional. Para el fomento y desarrollo de esas actividades, el Estado destinará recursos suficientes y creará el sistema nacional de ciencia y tecnología de acuerdo con la ley. El sector privado deberá aportar recursos para los mismos. El Estado garantizará el cumplimiento de los principios éticos y legales que deben regir las actividades de investigación científica, humanística y tecnológica. La ley determinará los modos y medios para dar cumplimiento a esta garantía.”

Dentro de este orden de ideas, la presente investigación se desarrolló con la Herramienta Protege 2000 que es un entorno de desarrollo web de código abierto que está optimizado para satisfacción de los programadores y de la productividad. Permite escribir un buen código favoreciendo la convención antes que la configuración. y se basa en el Decreto N° 3.390, que señala el Uso Prioritario de Software Libre. En atención a lo expuesto dicho artículo señala que *“el uso del Software Libre desarrollado con Estándares Abiertos fortalecerá la industria del software nacional, aumentando y fortaleciendo sus capacidades reducirá la brecha social y tecnológica en el menor tiempo y costo posibles, con calidad de servicio, gracias al uso de Software Libre desarrollado con Estándares Abiertos, y Que el Software Libre desarrollado con Estándares Abiertos, permite mayor participación de los usuarios en el mantenimiento de los niveles de seguridad e interoperatividad, contribuyendo a dar respuestas rápidas y oportunas a los usuarios.”*

CAPITULO III

MARCO METODOLÓGICO

TIPO DE INVESTIGACIÓN

Según Tamayo, Mario, El proceso de la Investigación científica. Limusa Editores, 2004. La Investigación es un proceso formal, sistemático, estructurado e interdisciplinario que va encaminado al descubrimiento de nuevos conocimientos en el campo de las ciencias, las artes o las letras, basado en hechos verificables conectados de manera lógica con el fin de concluir en conocimiento científico no conocido o comprobado, es decir, nuevo conocimiento.

En el mismo orden de ideas la investigación es un proceso, que, como lo define Zina O'Leary (2004), debe ser creativo y estratégico, en el que constantemente se evalúa, se reevalúa, y se toman decisiones sobre los medios para obtener información confiable, haciendo análisis apropiados y concluyendo ideas simples. En este proceso lo primero es establecer una meta, un lugar a donde llegar, es necesario enfocar la idea inicial de tal manera que guíe el resto del proceso; luego se deben identificar las fuentes de información para después empezar el proceso de recolección, en esta etapa es mejor no intentar enfocarse en un asunto específico sino por el contrario establecer un marco general, siguiendo lo que John Sawatsky (2004) denomina investigación pura, que permitirá identificar cuales son las preguntas que se deben hacer en el proceso de investigación; de acuerdo con Sawastsky centrarse en temas específicos al iniciar la recolección de información solo logrará que nos perdamos detalles que podrían ser importantes.

Balestrini (2001), hace referencia de que el fin esencial del marco metodológico es el de situar en el lenguaje de investigación los métodos e instrumentos que se emplearan en el trabajo planteado, desde la ubicación acerca del tipo de estudio y el diseño de investigación, su universo o población, su muestra, los instrumentos y

técnicas de recolección de datos, la medición hasta la codificación, análisis y presentación de los datos.

Naturaleza de la Investigación

La naturaleza de esta investigación es inductiva de carácter descriptiva bajo un diseño documental donde para recabar la información se emplearan herramientas para la ubicación y búsqueda de fuentes bibliográficas, documentales, textos, informes, Proyectos, Foros, Seminarios, Artículos de revistas Digitales, Paginas Web que permiten obtener la información pertinente al tema en estudio el cual esta definido por la Universidad Pedagógica Experimental Liberador UPEL (2003) *“como la investigación, elaboración y desarrollo de una propuesta de un modelo operativo viable para solucionar problemas, requerimientos, o necesidades de organizaciones o grupos sociales; puede referirse a la formulación de políticas, programas, tecnologías, métodos o procesos “.*

Diseño de la Investigación

Para el diseño de la presente investigación se utilizará la investigación documental donde se emplearan procedimientos para la ubicación y búsqueda de fuentes bibliográficas y documentales, textos, informes, Proyectos, Foros, Seminarios, Artículos de revistas Digitales, Paginas Web que permiten obtener la información pertinente al tema en estudio, ya que según señala Arias(2004) la investigación documental es aquella que se basa en la obtención y análisis de datos provenientes de materiales impresos u otros tipos de documentos y en el mismo orden de ideas el manual de la UPEL (2003) conceptualiza la investigación documental como el estudio de problemas con el propósito de ampliar y profundizar el conocimiento de su naturaleza, con apoyo principalmente en trabajos previos, información y datos divulgados por medios impresos, audiovisuales o electrónicos.

En tal sentido la revisión bibliográfica estará enmarcada en la búsqueda de antecedentes y trabajos previos que servirán como base para el desarrollo del vocabulario basado en Ontologías, tomando como punto de partida los antecedentes de

sistemas diseñados con el objetivo de mejorar las búsquedas eliminando el problema de los sinónimos y la polisemia descrito en las bases teóricas.

Fases del Estudio

La investigación se enmarca dentro de la modalidad de Proyecto Factible. Sobre este concepto el Manual de Trabajos de Grado de Especialización y Maestría y Tesis Doctorales de la UPEL (ob.cit.), señala que esta modalidad debe ser desarrollada en tres (3) fases, con el fin de cumplir con los requisitos involucrados en este tipo de proyecto. En la primera de ellas, se desarrollara el conocimiento de la situación existente en la realidad objeto de estudio, a fin de describir la plataforma tecnológica, funcionalidades requeridas y marco legal que involucran el diseño de un Vocabulario basado en XML y Ontología y demostrar la factibilidad de conducir la presente investigación. En la segunda fase de la investigación y atendiendo a los resultados de la fase diagnóstica, se formulara un modelo arquitectónico, para dar respuestas a los problemas planteados inicialmente en la investigación. En la tercera fase se diseñara un software, alineado al modelo operativo propuesto.

Fase I Diagnóstico.

Balestrini (2001) define la fase de diagnóstico de una investigación, como la fase cuyo objetivo central es descubrir las características fundamentales de la realidad. No debería transformarse en un ejercicio extremadamente largo, minucioso y menos teórico, es decir, se trata de realizar un diagnóstico operativo y práctico. Al final, la idea es contar con un instrumento que realmente sirva a nuestro propósito de planificar. En la presente investigación la fase de diagnóstico dio inicio, con la documentación, bases teóricas y con el análisis exhaustivo de cada uno de los antecedentes presentados que se relacionaron para la determinación de los objetivos en función de los factores críticos que se descubran.

El objetivo básico de esta fase es encontrar las debilidades del modelo actual, con la finalidad de captar de una forma general las características para desarrollar el nuevo modelo de dominio, de esta manera ayudará el desarrollo en las etapas siguientes:

la factibilidad y diseño de la propuesta del nuevo modelo para la solución del problema planteado. En la presente investigación los datos serán recabados directamente en la Coordinación de Postgrado del Decanato de Ciencia y Tecnología de la Universidad Centrooccidental Lisandro Alvarado, con el propósito de determinar la necesidad de la investigación, de tal manera que se garantizará la obtención de un producto que sea un aporte al desarrollo científico de la Universidad.

Población o Universo de Estudio

Balestrini (2001) señala que se entiende por población cualquier conjunto de elementos de los que se quiere conocer o investigar alguna o algunas de sus características. En el caso de estudio la población esta constituida por el personal administrativo adscrito a la Coordinación de Postgrado del Decanato de Ciencia y Tecnología de la UCLA.

Instrumentos para la Recolección de Datos

Para Balestrini (2001) La recolección de datos se refiere al uso de una gran diversidad de técnicas y herramientas que pueden ser manejadas por el investigador para recabar la información que servirá de sustento a la investigación. Es importante seleccionar un instrumento de recolección de datos, el cual debe ser valido y confiable para poder aceptar los datos, ejemplos de estas técnicas son: La observación, la entrevista, el cuestionario, los test, la revisión bibliografía.

(a) Revisión Bibliográfica: La revisión bibliográfica es una actividad que se lleva a cabo como parte del desarrollo de una propuesta de tesis de investigación o trabajo de Grado., forma parte importante de la monografía y proporciona un contexto y una justificación de la investigación que se está llevando a cabo. Los principales objetivos de esta revisión Bibliográfica es:

- Distinguir lo que se ha hecho de lo que se necesita hacer
- Descubrir variables importantes, relevantes al tema

- Sintetizar y adquirir una nueva perspectiva
- Identificar la relación entre ideas y la práctica
- Establecer el contexto del tema o problema
- Racionalizar el significado del problema
- Mejorar y adquirir vocabulario sobre el tema
- Relacionar ideas y teoría con las aplicaciones
- Identificar la metodología principal y las técnicas de investigación que se han usado hasta el momento
- Colocar la investigación en un contexto histórico para demostrar familiaridad con los últimos adelantos

A través de la revisión bibliografía se logro investigar acerca de sistemas existentes que aplican la Web Semántica, y el uso de herramientas para el modelado y diseño de las mismas. Así como también de investigar acerca de sistemas que usan los Frameworks explicados en el capítulo II.

(b) Consulta de Términos en la Pagina Web de la UCLA y de Postgrado del Decanato de ciencias y Tecnología de la UCLA.

Por lo anteriormente expuesto se llevo a la conclusión de diseñar un Vocabulario basado en XML y Ontologías para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

Fase II: Factibilidad

Después de definir la problemática presente y establecer las causas que ameritan de un nuevo sistema, es pertinente realizar un estudio de factibilidad para determinar la infraestructura tecnológica y la capacidad técnica que implica la implantación del sistema en cuestión, los costos, beneficios y el grado de aceptación que la propuesta genera. Este análisis permitió determinar las posibilidades de diseñar el sistema propuesto. Los aspectos tomados en cuenta para este estudio fueron clasificados en tres áreas, los cuales se describen a continuación:

Factibilidad Técnica.

Factibilidad técnica y operativa para diseñar y desarrollar un modelo basado en Ontología y XML para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

Definición Conceptual

Definida conceptualmente como un proceso mediante el cual expertos en tecnología de la información e ingeniería de software determinan la viabilidad de diseñar y desarrollar un modelo basado en ontologías XML identificando y validando los recursos disponibles, capacidades, habilidades y conocimiento necesario para tal fin.

-Definición Procedimental

Determinar la factibilidad técnica y operativa para diseñar y desarrollar un modelo basado en Ontología y XML para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA. Se refiere a los recursos necesarios como herramientas, conocimientos, habilidades, experiencia, etc., que son necesarios para efectuar las actividades o procesos que requiere el proyecto. Generalmente nos referimos a elementos tangibles (medibles). El proyecto debe considerar si los recursos técnicos actuales son suficientes o deben complementarse.

Desde el punto de vista del hardware, el equipo mínimo a trabajar vendrá condicionado por los requisitos de hardware exigido por el sistema operativo, esto debido a que el vocabulario a desarrollar servirá de herramienta de apoyo y se espera que los niveles de requisitos cambiaran dependiendo de las nuevas investigaciones que surjan a partir de la presente

Desde el punto de vista del Software se basa en el Gobierno Electrónico de Venezuela. *ISSN 1316-6239 (2006)*, donde el Decreto N° 3.390, refleja el Uso Prioritario de Software Libre y el artículo 1 hace referencia que La Administración Pública Nacional empleará prioritariamente Software Libre desarrollado con Estándares Abiertos, en sus sistemas, proyectos y servicios informáticos. A tales fines, todos los órganos y entes de la Administración Pública Nacional iniciarán los procesos de migración gradual y progresiva de éstos hacia el Software Libre desarrollado con Estándares Abiertos.

La herramienta de Desarrollo a utilizar será el basado en software libre Protege 2000, Protégé es un editor para construir ontologías y un marco general para representar el conocimiento. Protégé es una herramienta para el desarrollo de Ontologías y Sistemas basados en el conocimiento creada en la Universidad de Stanford. Y para el modelado de la Ontología se utilizó MethOntology Protégé está desarrollada en JAVA y puede funcionar perfectamente bajo WINDOWS.

Factibilidad Económica.

Se refiere a los recursos económicos y financieros necesarios para desarrollar o llevar a cabo las actividades o procesos y/o para obtener los recursos básicos que deben considerarse son el costo del tiempo, el costo de la realización y el costo de adquirir nuevos recursos. Con respecto a la presente investigación es factible económicamente ya que al utilizar Software libre con esquema de licenciamiento GNU/GPL se estará garantizando el acceso al código fuente sin pagar algún costo, esto debido al artículo 2 del Decreto N° 3.390 De Gobierno Electrónico de Venezuela. *ISSN 1316-6239 (2006)*, donde dice: *“A los efectos del presente Decreto se entenderá por: Software Libre:*

Programa de computación cuya licencia garantiza al usuario acceso al código fuente del programa y lo autoriza a ejecutarlo con cualquier propósito, modificarlo y redistribuir tanto el programa original como sus modificaciones en las mismas condiciones de licenciamiento acordadas al programa original, sin tener que pagar regalías a los desarrolladores previos, y por Estándares Abiertos: Especificaciones técnicas, publicadas y controladas por alguna organización que se encarga de su desarrollo, las cuales han sido aceptadas por la industria, estando a disposición de cualquier usuario para ser implementadas en un software libre u otro, promoviendo la competitividad, interoperatividad o flexibilidad.”

La GNU/GPL es una licencia creada por la Free Software Foundation a mediados de los 80, y está orientada principalmente a proteger la libre distribución, modificación y uso de software. Su propósito es declarar que el software cubierto por esta licencia es software libre y protegerlo de intentos de apropiación que restrinjan esas libertades a los usuarios.

Factibilidad Operativa.

La Factibilidad Operativa permite predecir, si se pondrá en marcha el sistema propuesto, aprovechando los beneficios que ofrece, a todos los usuarios involucrados con el mismo, ya sean los que interactúan en forma directa con este como también aquellos que reciben información producida por el sistema. Por otra parte el correcto funcionamiento del sistema en cuestión, siempre estará supeditado a la capacidad de los empleados encargados de dicha área.

Con la finalidad de garantizar el buen funcionamiento del sistema y que este impactara de forma positiva a los usuarios, el mismo será desarrollado en forma estándar a los sistemas existentes en las Coordinación de Postgrado del Decanato de Ciencias y tecnología de la UCLA, presentara una interfaz amigable a usuario, lo que se traduce en una herramienta de fácil manejo comprensión, contando con la opinión de lo mismos para cualquier modificación del sistema.

Fase III: Diseño de la Propuesta.

Se procedió a establecer los elementos que caracterizarán la arquitectura del diseño de un vocabulario basado en XML y Ontologías para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA, de acuerdo a lo anteriormente planteado y en función de sus objetivos, y en atención a la modalidad de investigación a fin de determinar los requerimientos para el desarrollo del Vocabulario se introducirá la utilización de una metodología llamada MethoOntology, se diseñara la Ontología con una herramienta llamada Protégé 2000 y se validó el vocabulario basado en XML y Ontología con el OntoTest de la herramienta Protégé y los Frameworks Jena y Pellet.

Fase IV: Fase Metodológica.

Para la presente investigación realice los siguientes pasos:

- 1) Investigaciones bibliográficas.
- 2) Investigaciones digitales.
- 3) Recopilación de la Información.
- 4) Clasificación de la Información.
- 5) Aplicación de la metodología methoontology, esto consistió en el desarrollo de la tareas que comprende esta metodología, consistió en dar respuesta a las siguientes interrogantes:
 - .- ¿Cual es el dominio de la Ontología?
 - .- ¿Para qué se usara la Ontología?
 - .- ¿Quién mantendrá la Ontología?
 - .- ¿Cual es el vocabulario que se utilizará?
- 6) Desarrollo de la Tareas que comprende Methoontology, esto será explicado en el capítulo IV.

- 7) Una vez desarrollada cada una de las tareas de la metodología, se procedió a instalar la herramienta Protégé.
- 8) Se instala la herramienta Protégé, Ver Guía de Instalación Anexo 1.
- 9) Crear un Proyecto en Protégé.

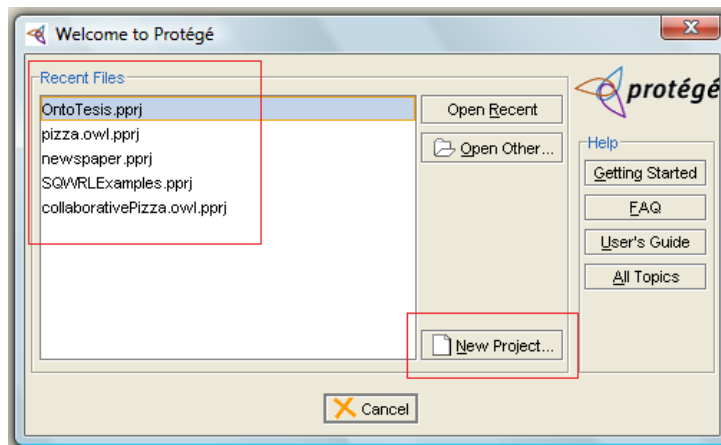


Figura 13. Bienvenido a Protégé.

Fuente: Autor.

- 10) Seleccionar Tipo de Proyecto.

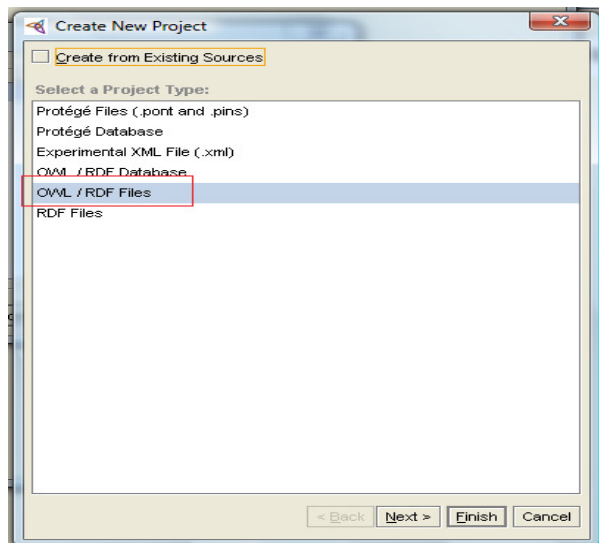


Figura 14. Crear un nuevo Proyecto.

Fuente: Autor.

11) Definir la URI de la Ontología.

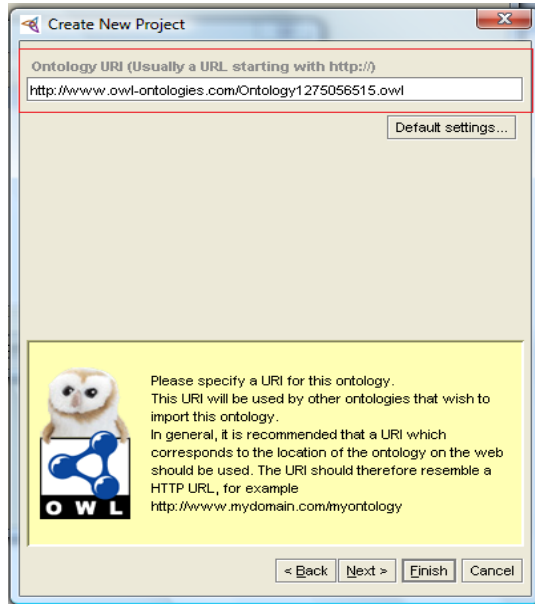


Figura 15. URI en Protégé.
Fuente: Autor.

12) Seleccionar el tipo de Lenguaje.

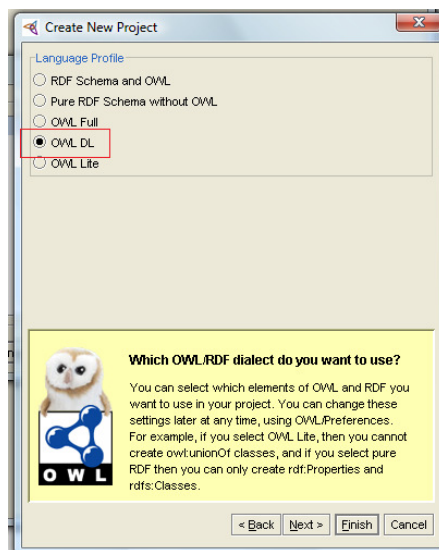


Figura 16. Tipo de Lenguaje en Protégé.
Fuente: Autor.

13) Seleccionar el tipo de Vista

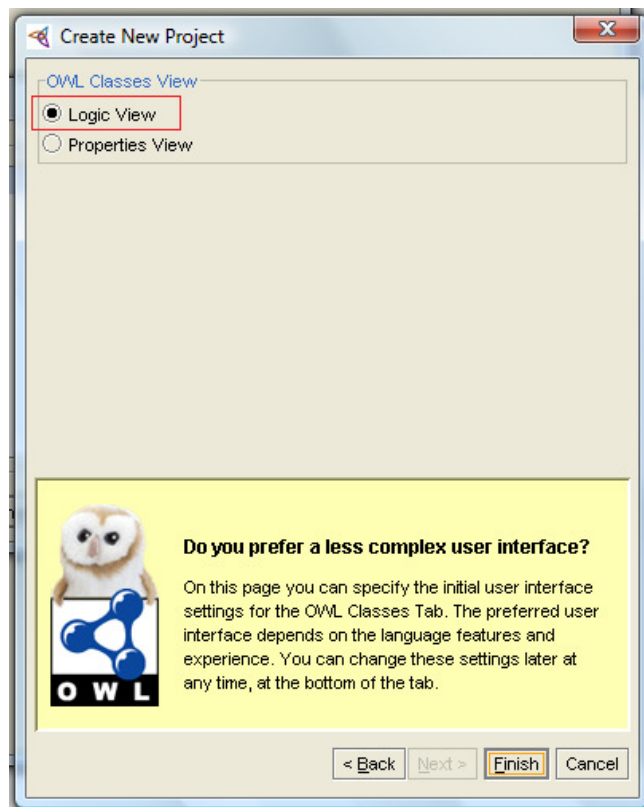


Figura 17. Vista en Protégé.

Fuente: Autor.

14) Pantalla Principal.

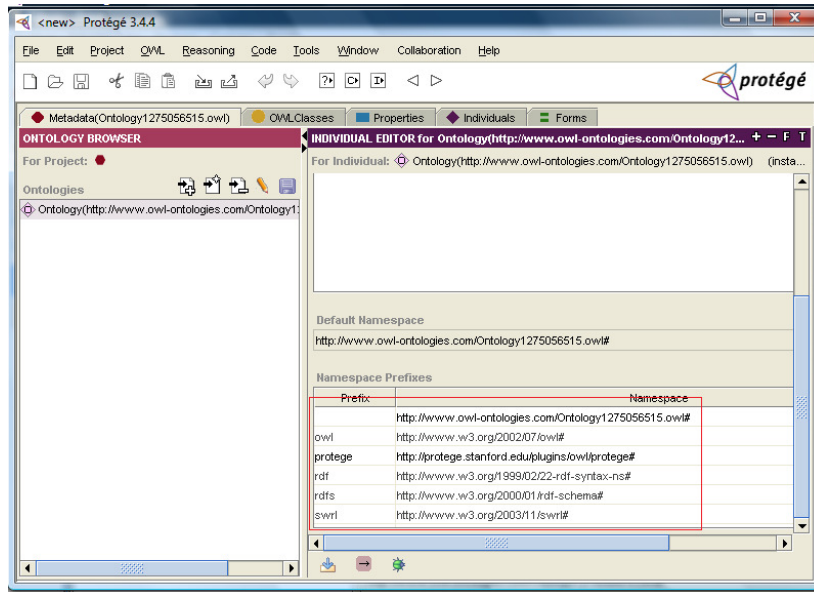


Figura 18. Pantalla Principal en Protégé.

Fuente: Autor.

- 15) Se procede a diseñar y desarrollar la Ontología, definiendo objetos, clases, propiedades, relaciones, consultas y utilizar cada Plugin destinado a cada actividad.
- 16) Una vez diseñada y desarrollada la ontología, se procede a validar que la Ontología no tenga errores.
- 17) Muestra de Código XML y RDF.
- 18) Una vez Generado el Vocabulario basado en XML y Ontología, se procede a utilizar los framework: Jena y Pellet.

CAPITULO IV

PROPUESTA DEL ESTUDIO

Introducción

La Ingeniería Ontológica es un conjunto de actividades concernientes a los procesos de desarrollo de una ontología, el ciclo de vida de la misma, los métodos y metodologías para su construcción y las herramientas y lenguajes que la soportan. Durante las últimas décadas se ha incrementado la atención en las ontologías y en la ingeniería Ontológica. Las Ontologías están siendo ampliamente usadas en la ingeniería del conocimiento, la inteligencia artificial y la ciencia de la computación en aplicaciones relacionadas con Gestión del Conocimiento, procesamiento de lenguaje natural, recuperación de información, integración de base de datos, y un nuevo campo que esta emergiendo como es la semántica Web. La construcción de sistemas basados en conocimiento usualmente implica la construcción de nuevas bases de conocimientos desde cero. En vez de eso debería construirse a partir de la unión de componentes reusables. Los desarrolladores de sistema deberían solamente preocuparse por la creación de conocimiento especializado y por especificar las tareas de sus sistemas, por ende este nuevo sistema podrá interoperar con sistemas existentes, en este sentido el conocimiento declarativo, las técnicas de resolución de problemas pueden ser compartidos entre sistemas, este enfoque facilitaría la construcción de sistemas mas grandes, mejores y mas baratos.

Según Guerrero y Lozano, 2006 La construcción de de Ontologías lleva implícito que cada termino y cada relación entre términos se defina formalmente. Los conceptos se describen explícitamente para entender su significado, mediante acuerdos ontológicos. Con ello, un usuario que desee reutilizar una ontología desarrollada por otros, puede conseguir la información de todos los conceptos que soporta, su taxonomía y los axiomas. EL mismo autor afirma que a diferencia de los Tesoros las Ontologías poseen una mayor variedad de relaciones entre conceptos, estas relaciones se crean a

propósito del modelo conceptual existente en el dominio que se está formalizando. El conocimiento almacenado en un sistema documental se puede organizar definiendo cada concepto (como; Revista, Libro, Artículo, autor, etc), y las relaciones que existen entre ellos de forma explícita (como: Escrito_por, Aparece_en , etc). Con ello se gana el establecimiento de un consenso tanto en el significado de cada término, como en una unificación formal del almacenamiento de esta información. Las Ontologías se construyen siguiendo esta filosofía, por lo que pueden ser reutilizadas en diferentes dominios y con diferentes fines. Poniendo axiomas del tipo: si A es coautor de B y sabes que B escribió el libro L1 y L2, se sabe que A escribió alguno de estos dos libros, es decir, se pueden deducir conocimientos que no están explícitos en la ontología.

El acceso en los documentos o información de interés en la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA y la calidad de la misma son afectadas por un amplio conjunto de variables, términos, vocabularios dinámicos. Cuando estos términos se evalúan en conjunto aparecen como interrelaciones que requieren de un lenguaje común para poder comunicarse e intercambiar información. Para representar estas interrelaciones y entidades como un conjunto implícito y explícito se ha planteado un modelo basado en la Ontología, para representar dicho modelo se usará la metodología Methontology y computacionalmente se utilizó una herramienta de la Tecnología semántica llamada: Protege, el producto pretende ser un Vocabulario basado en XML y Ontología para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la Universidad Centro Occidental Lisandro Alvarado.

Methontology incluye en la actividad de conceptualización un juego de tareas de estructuración del conocimiento, donde se enfatizan los componentes de la ontología (conceptos, atributos, relaciones, constantes, axiomas formales, reglas, y entidades) que se construyen dentro de cada tarea, y propone un orden de creación de dichos componentes. Este proceso de modelado no es secuencial, aunque debe seguirse algún orden que asegure la consistencia e integridad del conocimiento que se desea representar. En caso de cambios o actualizaciones es posible regresar en cualquiera de las tareas.

APLICACIÓN DE LA METODOLOGIA METHOONTOLOGY

Especificación:

El proceso de especificación consistió en responder a algunas preguntas: ¿Cuál es el dominio que la ontología cubrirá? ¿Para qué se usará la ontología? ¿Para qué tipos de preguntas la información en la ontología debería proveer respuestas? ¿Quién usará y mantendrá la ontología?. La Ontología deberá permitir obtener especificaciones de orientadas a la definición de Términos, vocabularios, glosarios, y además en la medida de lo posible estas especificaciones deben ser mapeadas en algún lenguaje ejecutable que facilite su implementación.

La creación de un Vocabulario basado en Ontología para la Coordinación de Postgrado del Decanato de Ciencia y Tecnología de la UCLA está fundamentado por la necesidad que existe de establecer un lenguaje común para la comprensión, compartir y recuperación de los datos a partir de la descripción de los conceptos y términos que intervienen en los procesos y actividades de dicha coordinación; de forma tal, que pueda ser entendible entre los distintos usuarios de la información, con independencia de las fuentes que generen los datos.

La fuente fundamental de información para la ejecución del trabajo lo constituyen por un lado las entidades Coordinación, Postgrado que intervienen en el proceso de ordenamiento y manejo sostenible de la información; que a su vez, tienen una importante influencia en los procesos y actividades vinculados a las principales dimensiones del desarrollo sostenible: económica, tecnológico y político-institucional.

Se han podido identificar una serie de requisitos, necesarios a tener en cuenta para la solución del problema. Ellos son:

Creación de una base de conocimiento de los términos utilizados en la Coordinación de Postgrado.

Realizar una descripción detallada de los términos utilizados.

Reflejar las propiedades de los objetos de forma tal, que se obtenga una correlación semántica entre la realidad y la propia representación. Esta descripción, llamada anotación semántica, permitirá hacer búsquedas más completas.

Tiene que ser entendible por los expertos.

Debe quedar formalizado para que, de cierto modo le permita apoyar el procesamiento automático de la información y permita la homogeneidad de los datos, para esto es indispensable ser capaz de extraer el conocimiento disponible en la gran variedad de recursos con los que cuenta la organización (información estructurada e información no estructurada). En este proceso, las ontologías ofrecen la posibilidad de definir la estructura que debe tener el conocimiento extraído de cada tipo de elemento.

Facilitar una rápida recuperación de la información.

DOMINIO DE LA ONTOLOGIA

Según Berners, 2001, El dominio o alcance, es similar al dominio que tomamos para el diseño de una base de datos. En el dominio nos dedicaremos a responder las preguntas planteadas en la especificación de la Ontología.

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

En la presente investigación, realizado el estudio, se obtiene las siguientes respuestas

1. ¿Cuál es el dominio que la ontología cubrirá?

R.- Conceptos relacionados para la coordinación de postgrado del Decanato de Ciencias y Tecnología, y estos conceptos serán extraídos de la Pagina Web de la UCLA. establecer las relaciones entre los conceptos y sus propiedades.

2. ¿Para qué usaremos la ontología?

Para la información, conceptos, vocabularios de la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

3. ¿Para qué tipos de preguntas la información en la ontología debería proveer respuestas?

R.- Las preguntas son las siguientes:

_ ¿Cómo se estructura la información de la Coordinación?

_ ¿Cómo se relacionan los conceptos de la Coordinación?

_ ¿Cuáles son las dependencias entre unidades?

_ ¿Cuáles son las autoridades posesionadas actualmente?

4. ¿Quien usará y mantendrá la ontología?

R.- Los administradores, y técnicos informáticos encargados del mantenimiento de la página Web del Decanato de Ciencias y Tecnología de la UCLA.

Modelado Conceptual de una Ontología:

Según Anaya,R, 2000 Los modelos presentan un marco conceptual donde reflejar las teorías, plasmar propiedades y establecer los principios del diseño de los sistemas. Su importancia radica en que permiten identificar, organizar y realizar razonamientos sobre los componentes y comportamiento del sistema, son la guía para el proceso de diseño del software y puede usarse posteriormente como una referencia para evaluar un diseño particular, razonar sobre la solución realizada y sobre el posible espacio de soluciones. Por su naturaleza debe ser expresivos, fáciles de usar y completos. Para Realizar el Modelo Conceptual para la presente investigación se utilizó MethoOntology y su aplicación se explicará a través de una serie de tareas que a continuación describiré.

TAREA 1: Construir un Glosario de Términos.

En este apartado se identifican los conceptos fundamentales del dominio en que trabajamos y que debemos tener en consideración. Nuestro dominio se refiere a los términos y/o vocabulario utilizado en la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA y principalmente para esta construcción se utilizó la información presente en la página Web de la UCLA y en la página Web de Postgrado del Decanato de Ciencias y Tecnología. Se establecen además las propiedades que permitan caracterizar y describir todos estos aspectos.

Tabla1: Glosario de Términos.

Nombre	Definición
Actividad	Conjunto de operaciones o tareas que son ejecutadas por una persona o unidad administrativa como parte de una función asignada. Conjunto de tareas u operaciones interrelacionadas entre sí, dirigidas a la obtención de un objetivo determinado. Conjunto de Tareas o asignaciones por el profesor a cada estudiante cursante de un Programa de Postgrado.
Admisión	Aceptación de una persona para iniciar estudios en una institución de educación superior. Permite la matriculación en el curso académico inicial de un programa.
Asignatura	Cada una de las materias en que se estructura un plan de estudios, aunque es posible que una materia comprenda varias asignaturas. Cada asignatura suele tener asignados unos créditos, de acuerdo con la dedicación de horas de docencia o de trabajo total de los estudiantes.
Calendario Académico.	Sistema que permite definir una forma de estructurar el tiempo, a través de patrones de medida uniformes, y así organizar días laborables, festivos de un periodo académico.

Coordinación de Postgrado	La Coordinación de Estudios de Postgrado es un órgano de carácter académico administrativo adscrito al Vice-Rectorado Académico, para supervisar y coordinar los Estudios de Postgrado de la Universidad Centro Occidental “Lisandro Alvarado”.
Comisión de Estudios de Postgrado	La Comisión de Estudios de Postgrado es la unidad asesora del Consejo de Decanato de Ciencias y Tecnología de la UCLA encargada de orientar, organizar, coordinar, promover y supervisar los estudios de Postgrado.
Curriculum Vitae.	Enfatiza la información sobre experiencia, habilidades y estudios, relevante para los objetivos que se tienen que cumplir en un determinado puesto de trabajo u oferta de empleo para el cual se opta o se está en proceso de selección. Curriculum Vitae: es una recopilación de todos los datos académicos y experiencia de una persona a lo largo de su vida (como bien indica el término latino vitae) independientemente del puesto de trabajo al cual se opta en el proceso de selección. La estructura suele ser datos personales, académicos, experiencia, idiomas, informática y otros datos, todo ello en orden cronológico de adquisición.
Decanato	Conjunto de dependencias en las cuales se divide la Universidad CentroOccidental Lisandro Alvarado, destinadas oficialmente al decano para el desempeño de sus funciones.
Dirección Técnica de Apoyo Académico	Es una unidad técnico administrativa, dependiente del Vicerrectorado Académico con la misión de promover y coordinar la integración de la tecnología en la educación, reforzando el uso de los métodos y medios de enseñanza innovadores, como apoyo a la docencia, investigación, extensión y difusión mediante la coordinación de los programas que en este sentido se han puesto en marcha en la Universidad.

Dirección de Desarrollo Estudiantil	<p>Unidad Técnica de apoyo académico que contribuye al desarrollo integral del estudiante a través de la atención bio-psicosocial y espiritual, para lograr una población sana, pro-activa, comprometida, con sentido de pertenencia, de alta calidad humana y tecno-científica, preparada y apta para desarrollar una vida profesional, preactiva y pertinente con los cambios de entorno social.</p>
Dirección de Extensión Universitaria	<p>La Extensión Universitaria es una función académica que tiene por finalidad mantener la interacción permanente con la comunidad, destinada a para lograr el fortalecimiento de la conciencia crítica y contribuir a elevar el nivel social, económico y cultural de las comunidades. En la Universidad Centroccidental “Lisandro Alvarado” esta función la coordina la Dirección de Extensión Universitaria (DEU), adscrita al Vice-Rectorado Académico.</p> <p>La esencia de la Extensión Universitaria es la interacción Sociedad – Universidad con compromiso social y vocación productiva que logra cambios favorables para el desarrollo sostenible de su entorno a través de un permanente intercambio de experiencias y conocimientos.</p> <p>Se sustenta en las experiencias y en la generación de conocimientos, brinda oportunidades de conocimiento, desarrollo sostenible y mejor calidad de vida a los productores, comunidades y al sector empresarial de la Región Centroccidental, a través de un proceso dialógico de transformación, de acuerdo a las pluralidades y capacidades de la institución con un sentido de responsabilidad, pertinencia y sensibilidad social.</p>

<p>Dirección de Planificación Universitaria</p>	<p>Es el ente motorizador y coordinador del proceso de creación de un sistema integrado de planificación y evaluación, como vía expedita para intervenir deliberadamente en la gestión de los cambios, para acelerarlos regularlos y orientarlos hacia la mejora continua.</p>
<p>Docencia</p>	<p>La docencia se inscribe dentro del campo educativo como actividad que promueve conocimientos, que sitúa al docente (Profesor) como factor especial, tanto con referencia a los conocimientos mismos, como con respecto a las condiciones específicas en que éstos son producidos.</p> <p>Se subraya con frecuencia que la relación pedagógica se establece alrededor <i>de</i> y <i>con</i> referencia a los saberes; saberes a adquirir tal como se presentan, como parte de un <i>currículum</i> y no como saberes a confrontar, a descifrar, en tanto el conocimiento aparece siempre en su carácter de relativo e inacabado, como algo siempre susceptible de ser comprendido, mejorado y completado.</p>
<p>Documentos</p>	<p>Es el conjunto de información referente o correspondiente a la Coordinación de Postgrado y estos pueden ser: Los expedientes, escritos, oficios, reportes, estudios, actas, resoluciones, acuerdos, directivas, directrices, circulares, contratos, convenios, instructivos, notas, memorandos, o bien, cualquier otro registro que documente información de los sujetos obligados y sus servidores públicos, sin importar su fuente o fecha de elaboración. Los documentos podrán estar en cualquier medio, sea escrito, impreso, fotográfico, sonoro, visual, digital, electrónico, magnético, informático, holográfico o cualquier otro elemento técnico que tenga ese carácter y que haya sido creado u obtenido por los sujetos obligados.</p>
<p>Egresado</p>	<p>Término que se aplica a la persona que ha completado sus</p>

	estudios universitarios. Denominación que recibe el estudiante universitario tras haber concluido sus estudios y abandonar la universidad. El término designa a los alumnos que han cubierto 100% de los créditos establecidos en un programa académico de pregrado o posgrado en el tiempo determinado por las normas institucionales.
Estudiante	Un estudiante es aquella persona dedicada a la lectura, puesta en práctica y aprehensión de conocimientos sobre alguna materia o arte. Persona que está formalmente matriculada en un programa de estudios. Hay distintos tipos de estudiantes, en función del modelo de enseñanza, de su dedicación temporal, del plan de estudios en el que se matricula o inscribe, por lo que las estadísticas universitarias pueden tener que atender a estas circunstancias.
Estudiantes Regulares	Los estudiantes de Postgrado pueden ser regulares y especiales. Son estudiantes regulares los que optan por título académico y reúnen las siguientes condiciones: a) Cursar, al menos, el mínimo de unidades crédito en cada período académico señaladas en este Reglamento, o estar inscrito en el trabajo o tesis de grado según corresponda. b) Cumplir con las normas referentes al índice de prosecución pautado en este Reglamento.
Estudiantes Especiales	Son estudiantes especiales de Postgrado, los que cursan una o más asignaturas o actividades en un Programa de Postgrado, sin la finalidad de obtener un grado. Los requisitos para el ingreso y permanencia como estudiante especial, así como el número máximo de créditos que pueda cursar como tal, serán fijados en cada Programa por la coordinación respectiva.
Estudios de Postgrado	Son estudios de Postgrado aquellos que se realizan después de haber obtenido en una Institución de Educación Superior

	Venezolana, o extranjera de reconocido prestigio, el Título de licenciado o equivalente como resultado de estudios cuya duración no ha sido inferior a cuatro años y que se efectúen conforme a las disposiciones del presente Reglamento y según las Normas para la Acreditación de Estudios para Graduados dictados por el Consejo Nacional de Universidades (C.N.U.).
Estudio de Postgrado No Conducente a Grado Académico	Los Estudios de Postgrado no conducentes a grado académico tienen como objetivo ofrecer a profesionales universitarios oportunidades de incorporar a su formación los avances humanísticos, científicos y tecnológicos en aspectos específicos del saber.
Estudio de Postgrado Conducente a Grado Académico	Los Estudios de Postgrado Conducentes a grado académico tiene como objetivo fundamental otorgar una formación académica pertinente, que a través de los conocimientos y competencias propios de su campo, permita el dominio, la comprensión y la aplicación del conocimiento.
Examen	Tipo de Actividad evaluada por el Profesor de una determinada materia de un programa de postgrado el cual es realizada o ejecutada por un estudiante, se refiere a la evaluación de ciertos tópicos o temas ya vistos en clase producto de exposiciones, temas de discusión, debates o investigaciones.
Extensión	La Extensión Universitaria es una función académica que tiene por finalidad mantener la interacción permanente con la comunidad, destinada a para lograr el fortalecimiento de la conciencia crítica y contribuir a elevar el nivel social, económico y cultural de las comunidades. En la Universidad Centroccidental “Lisandro Alvarado” esta función la coordina la Dirección de Extensión Universitaria (DEU), adscrita al Vice-Rectorado Académico. La esencia de la Extensión Universitaria es la interacción

	<p>Sociedad – Universidad con compromiso social y vocación productiva que logra cambios favorables para el desarrollo sostenible de su entorno a través de un permanente intercambio de experiencias y conocimientos.</p> <p>Se sustenta en las experiencias y en la generación de conocimientos, brinda oportunidades de conocimiento, desarrollo sostenible y mejor calidad de vida a los productores, comunidades y al sector empresarial de la Región Centroccidental, a través de un proceso dialógico de transformación, de acuerdo a las pluralidades y capacidades de la institución con un sentido de responsabilidad, pertinencia y sensibilidad social.</p>
Evaluación	Proceso sistemático y metódico, mediante el cual se recopila información cuantitativa o cualitativa a través de medios formales sobre un objeto determinado, con el fin de juzgar su mérito o valor y fundamentar decisiones específicas. Este proceso puede ser empleado en diferentes ámbitos del quehacer humano: social, económico, educativo o político.
Hoja de Referencia Académica.	Es un tipo de documento utilizados por la Coordinación de Postgrado del Decanato de Ciencias y Tecnología para recomendar a cierta persona para ser estudiante de una maestría en Particular.
Horario	Estipulación de los periodos del día en que se lleva a cabo las clases académicas. Y esto se refleja en un cuadro donde se indican las horas.
Investigación	Proceso sistemático que, por medio de la observación, búsqueda de información y metodologías especializadas, tiene por objeto incrementar el conocimiento en un campo disciplinario específico. La investigación forma parte esencial de los procesos de evaluación, adaptando sus finalidades a las

	propias de estos procesos. La que tiene por fin ampliar el conocimiento científico, sin perseguir, en principio, ninguna aplicación práctica.
Lapso Académico	Es el periodo establecido para el inicio, fin de un programa de Estudio de Postgrado establecido por la coordinación de Postgrado, hace uso de un Calendario académico previamente establecido.
Maestría	Los Estudios de Maestría tienen como objetivo la profundización en un campo del conocimiento para obtener en él elevados niveles científicos, humanísticos y la formación metodológica para la investigación. En ellos se otorgará el grado de Magíster Scientiarum con la mención correspondiente.
Mención	Es la especialización del profesional en un determinado subsector de aprendizaje o en un determinado nivel educativo, que puede ser reconocido como una formación profesional especial o adicional.
Programa	Se refiere a los diferentes estudios de Postgrado impartidos por la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA. Los programas deberán propender a la Utilización integral de los recursos de la Universidad y la región, prever la coordinación de esa utilización y corresponder con precisión a los objetivos del nivel de Postgrado al cual se refieren.
Postgrado	Los estudios de Postgrado del Decanato de Ciencias y Tecnología de la UCLA, tiene como objetivo fundamental formar personal altamente calificado en las áreas de las Ciencias Básicas, de la Información y afines, para atender las necesidades de la región y del país, en las labores de Investigación, Docencia y Extensión.
Plan de Estudios	Se refiere al instrumento que, con fundamento en una

	orientación educativa, precisa y articula objetivos, contenidos y estrategias de enseñanza-aprendizaje para sustentar la formación de una persona en el contexto de un ciclo completo de educación. Esta conformado por Asignatura, Cantidad de Asignatura, Unidades de Crédito.
Profesor	Un profesor es una persona que enseña una determinada ciencia o arte . Deben poseer habilidades pedagógicas para ser agentes efectivos del proceso de aprendizaje. El profesor, por tanto, parte de la base de que es la enseñanza su dedicación y profesión fundamental y que sus habilidades consisten en enseñar la materia de estudio de la mejor manera posible para el alumno.
Programa Académico	Documento institucional que describe la misión y los objetivos del programa, su organización académico-administrativa, el plan y los programas de estudio. Organización de los estudios de pregrado, grado o posgrado, dentro de una institución de educación superior.
Proyecto	Tipo de Actividad que es asignada por el profesor de una determinada materia de un programa de Postgrado que se refiere a realizar una investigación o desarrollo de un software referente a algunos tópicos discutidos en clase o para la generación de nuevas ideas.
Unidades de Crédito	Se entiende por Unidades de crédito una hora semanal de actividad teórica o de seminario durante dieciséis (16) semanas, o dos (2) o tres (3) de práctica u otras actividades durante el mismo período. Los créditos correspondientes a asignaturas o actividades realizadas en períodos de otra duración o modalidad se medirán proporcionalmente.
Unidad Administrativa	Es una dirección general o equivalente a la que se le confieren atribuciones específicas en el reglamento interno. Puede ser

	también un órgano que tiene funciones propias que lo distinguen de los demás en la UCLA.
Universidad	<p>La palabra “universidad” viene del latín <i>universitas</i> y se origina de los vocablos latinos <i>unus</i> (uno, que no admite división) y <i>versus (hacia)</i>, que en su conjunto significan “convertido en uno”. Desde la perspectiva filosófica, es la agremiación o unificación de maestros y estudiantes que estudian o realizan procesos de adquisición de conocimiento partiendo de la existencia de lo sobrenatural, el mundo y el hombre (Bautista, 2006). La definición dada por la Real Academia de la Lengua Española tiene un carácter más estructural. Una universidad es una “institución de enseñanza superior que comprende diversas facultades, y que confiere los grados académicos correspondientes., etc.”. La Universidad ha sido, desde la historia antigua, la fuente de las ideas. En la época moderna, las universidades han sido la fuente del desarrollo de la humanidad. El mayor número de grandes descubrimientos modernos se han gestado en las universidades. Ahora bien, no basta con generar conocimiento, si a éste no se le da un uso práctico.</p>

En las siguientes imágenes se aprecia la Ontología diseñada para la presente investigación con la herramienta Protégé 2000.

Metadata(Ontology1270352124.owl) | OWLClasses | Properties | Individuals | Forms

SUBCLASS EXPLORER | CLASS EDITOR for Calendario_Academico (instance of owl:Class)

For Project: **OntoTesis** | For Class: http://www.owl-ontologies.com/Ontology1270352124.owl#Calendario_Academico

Asserted Hierarchy

- owl:Thing
 - Actividad
 - Examen
 - Investigacion
 - Proyecto
 - Admisión
 - Asignatura
 - Basica
 - De_Area
 - Electivas
 - Generica
 - Topico_Especial
 - Trabajo_de_Grado
 - Calendario_Academico
 - Curriculum_Vitae
 - Documento
 - Arancel_de_Postgrado
 - Aviso_de_Prensa
 - Carta_de_Admission
 - Carta_de_Recomendacion

Property	Value
rdfs:comment	Sistema que permite definir una forma de estructurar el tiempo, a través de patrones de medida uniformes, y asi organizar días laborables, festivos de un periodo académico.

Asserted

owl:Thing

- Pertenece_a_un **only** Lapso_Academico

Metadata(Ontology1270352124.owl) | OWLClasses | Properties | Individuals | Forms

SUBCLASS EXPLORER | CLASS EDITOR for Calendario_Academico (instance of owl:Class)

For Project: **OntoTesis** | For Class: http://www.owl-ontologies.com/Ontology1270352124.owl#Calendario_Academico

Asserted Hierarchy

- Carta_de_Recomendacion
- Hoja_de_Referencia_Academica
- Listado_de_Admitidos
- Manuales
- Normas_para_Trabajo_de_Grado
- Normativas
- Recaudos_de_Admission
- Reglamento_General
- Solvencia_Administrativa
- Estudio_de_Postgrado
- Horario
- Lapso_Academico
- Persona
 - Adjunto_de_Investigación
 - Aspirante
 - Coordinador_de_Postgrado
 - Coordinador_de_Programa
 - Director
 - Estudiante
 - Especial
 - Regular

Property	Value
rdfs:comment	Sistema que permite definir una forma de estructurar el tiempo, a través de patrones de medida uniformes, y asi organizar días laborables, festivos de un periodo académico.

Asserted

owl:Thing

- Pertenece_a_un **only** Lapso_Academico

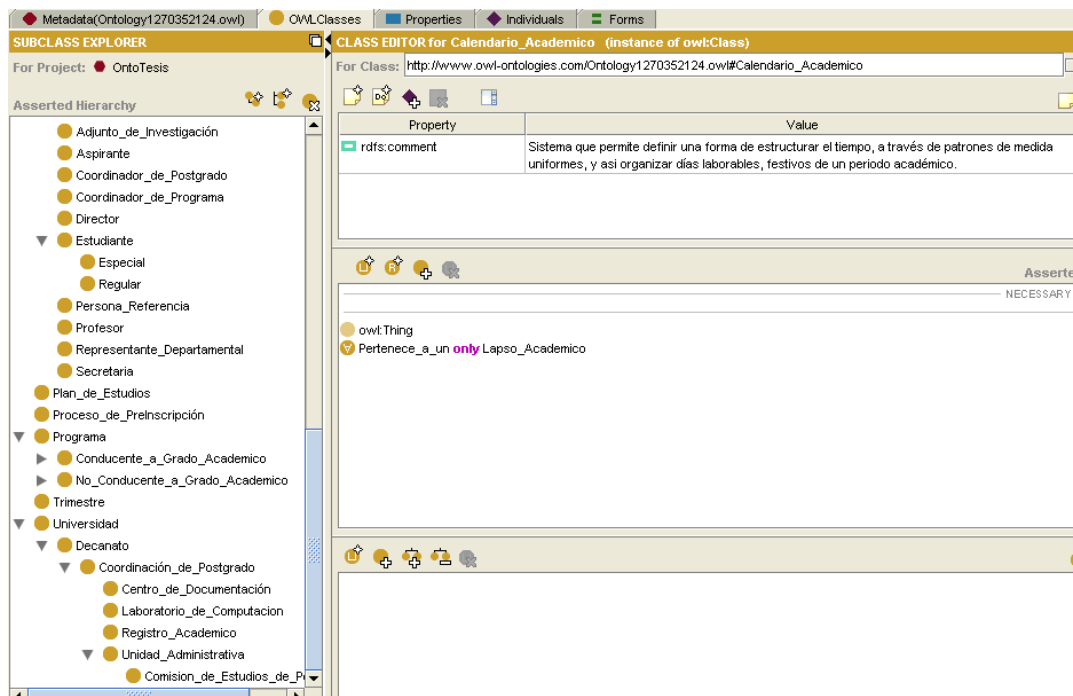


Figura 19: Vocabulario Basado en Ontología Diseñada con la Herramienta Protégé 2000.

Fuente: Autor.

TAREA 2: Construir una Taxonomía de Conceptos.

Etimológicamente hablando, *taxonomía* procede de los términos griegos “*taxis*”, ordenación, y “*nomos*”, norma. Aristóteles fue uno de los primeros en utilizar este término, en el 300 antes de Cristo, para designar esquemas jerárquicos orientados a la clasificación de objetos científicos. De acuerdo con Grove (2003), citado por Centelles, (2005), los principios que proporcionaban una guía rigurosa para la construcción de taxonomías eran la base lógica, la observación empírica, la estructura jerárquica basada en la herencia de propiedades, la historia evolutiva, y la utilidad pragmática. Las fuentes terminológicas de la lengua general todavía recogen el significado especialmente orientado al entorno de las ciencias experimentales.

Aplicando el principio etimológico, la *taxonomía* se ha definido como la ciencia que trata de los principios, métodos y fines de la clasificación y, procura la organización

jerarquizada y sistemática, dando nombre a grupos de elementos y a los elementos mismos.

Según Centelles, 2005 Una taxonomía es un tipo de vocabulario controlado en que todos los términos están conectados mediante algún modelo estructural (jerárquico, arbóreo, facetado, etc.) y especialmente orientado a los sistemas de navegación, organización y búsqueda de contenidos de los sitios web.

De acuerdo con Abed Gregio, Barbato, Duarte, et al., (2007): “Una taxonomía es un proceso científico (o un sistema particular) de categorizar entidades, es decir, de organizarlas en grupos. Un sistema taxonómico debe ser claro y consistente, flexible, exhaustivo y práctico”.

Los mismos autores, plantean además que la base para el desarrollo de una buena taxonomía son las características taxonómicas, que son las propiedades o atributos de los objetos que se categorizarán, y deben satisfacer requisitos. Estos son:

Objetividad: cuando la característica se identifica sobre la base de un conocimiento objetivo exclusivamente.

Determinación: cuando existe un proceso claro que pueda seguirse para identificar la característica.

Reproducibilidad: cuando varias personas describen de forma independiente las características de un mismo objeto y coinciden con el valor observado.

Mutualmente excluyente: cuando la inclusión de un grupo en una categoría lo excluye de la categorización en cualquier otra.

Exhaustivos: cuando los grupos incluyen todas las posibilidades.

Aceptable: cuando la característica es lógica e intuitiva de forma que la comunidad acepte las categorías.

Útil: cuando la característica puede utilizarse para obtener conocimiento. El cumplimiento de los atributos en una taxonomía aplicada a un sitio web favorece la navegabilidad y facilita además acceder al capital intelectual de las organizaciones.

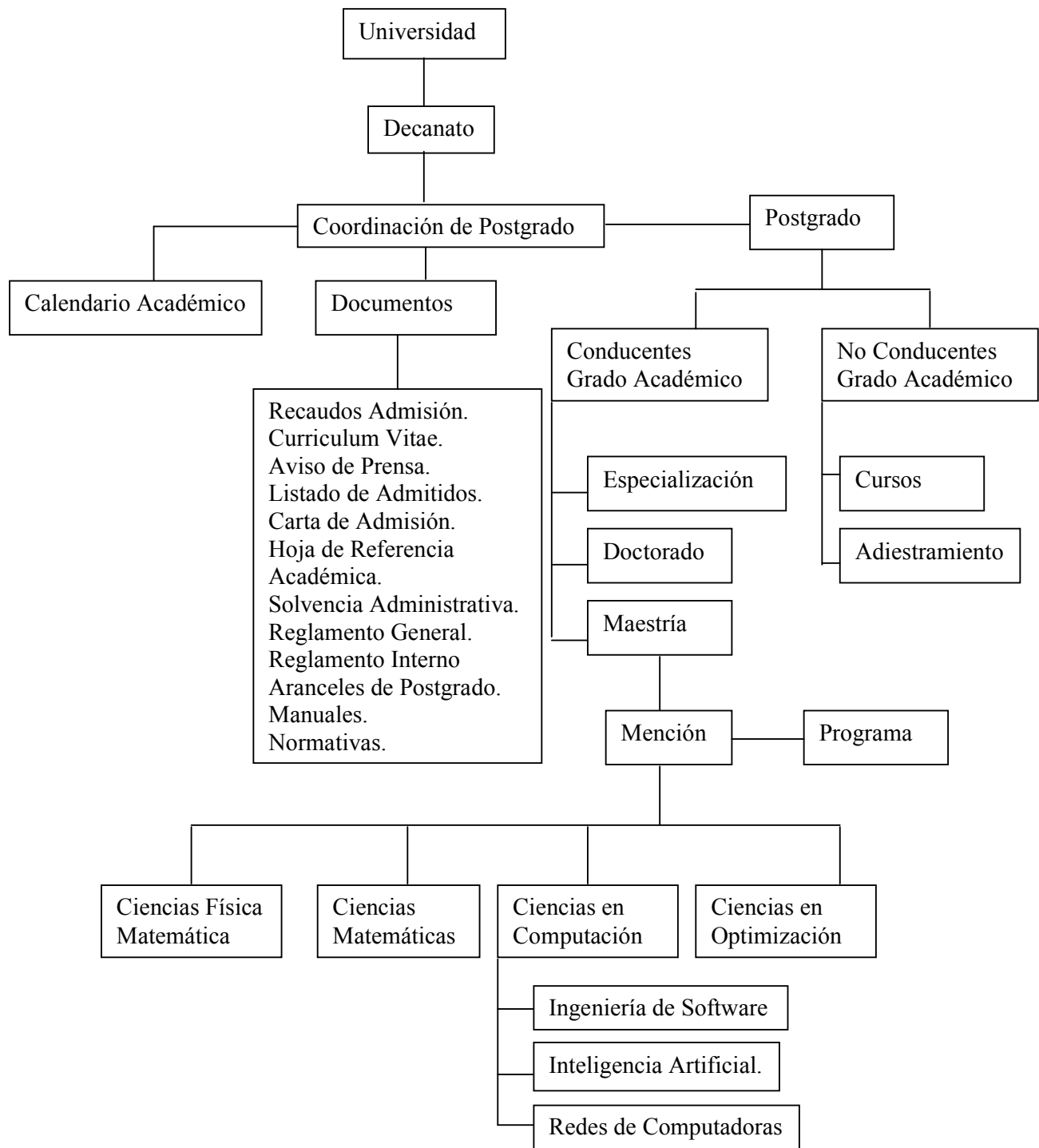


Figura 20: Diagrama de Taxonomía Nro. 1.
Fuente: Autor.

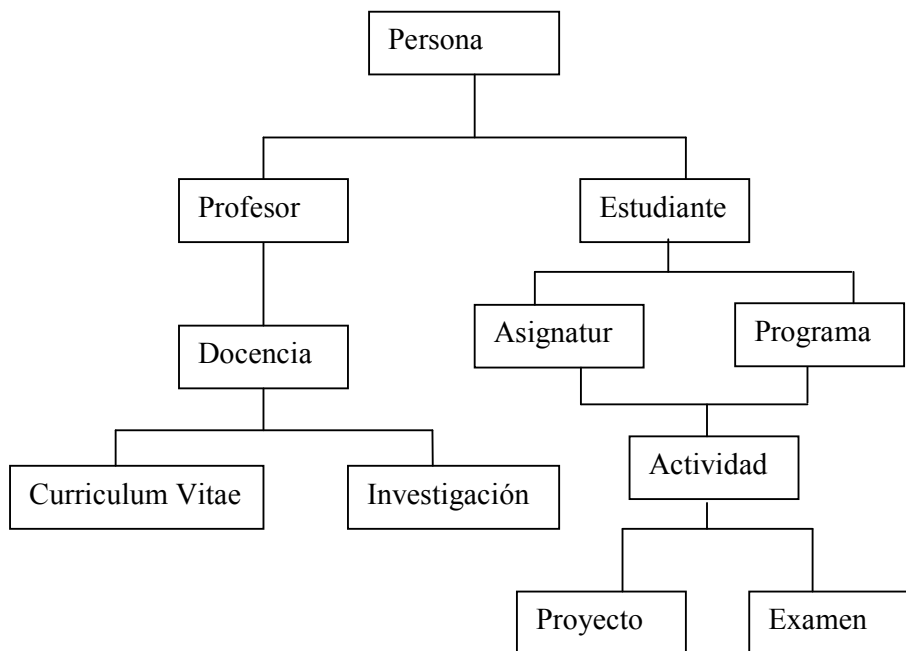


Figura 21: Diagrama de Taxonomía Nro. 2.

Fuente: Autor.

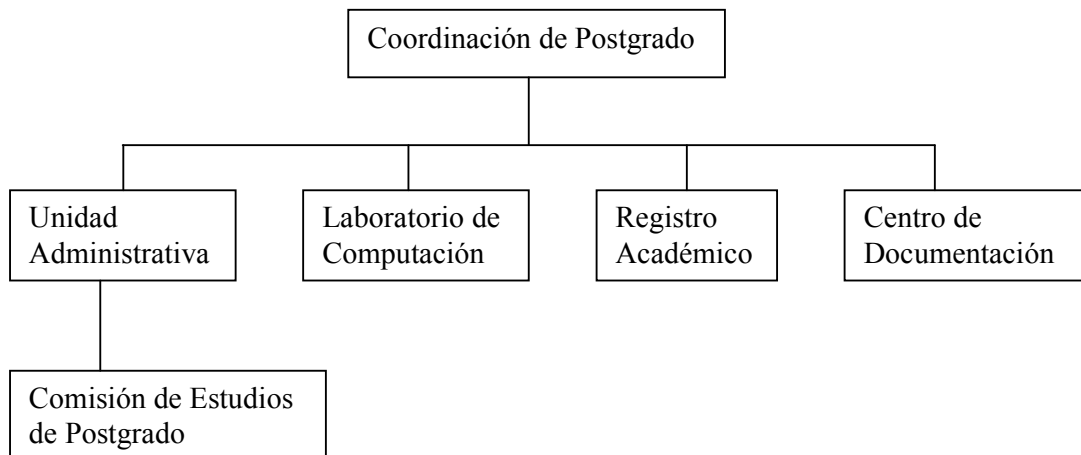


Figura 22. Diagrama de Taxonomía Nro. 3.

Fuente: Autor

A Continuación se presentan los Diagramas de Taxonomía realizados por el Plugin OWLViz de la Herramienta Protégé 2000.

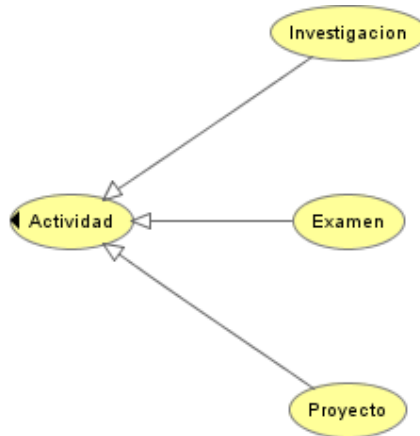


Figura 23. Taxonomía Nro 4.
Fuente: Autor



Figura 24. Taxonomía Nro 5
Fuente: Autor



Figura 25. Taxonomía Nro 6.
Fuente: Autor

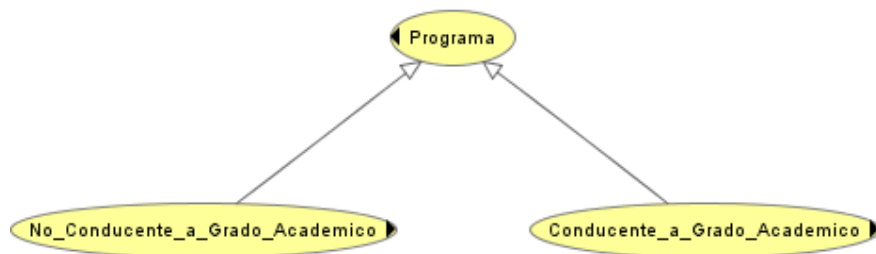


Figura 26. Taxonomia Nro 7.
Fuente: Autor

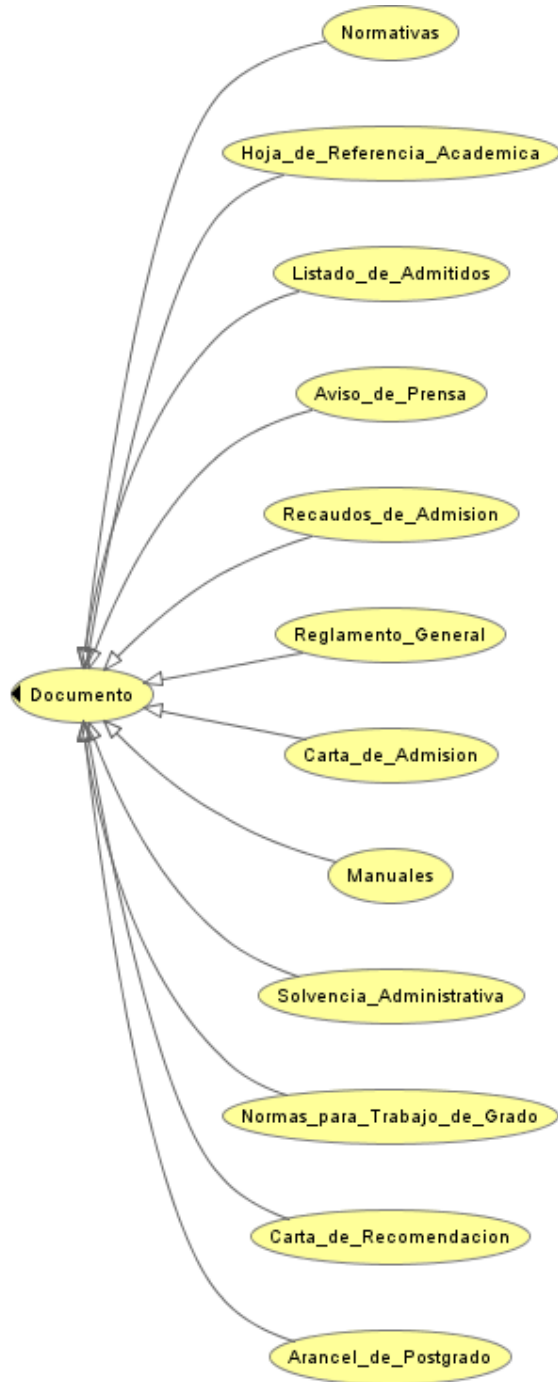


Figura 27. Taxonomia Nro 8.
Fuente: Autor

TAREA 3 Diagramas de Relación Binaria ad-hoc.

Ad hoc es una **locución latina** que significa literalmente “para esto”. Generalmente se refiere a una solución elaborada específicamente para un problema o fin preciso y, por tanto, no es generalizable ni utilizable para otros propósitos. Se usa pues para referirse a algo que es adecuado sólo para un determinado fin. En sentido amplio, *ad hoc* puede traducirse como “específico” o “específicamente”. Los Diagramas Ad-Hoc netamente especifican las relaciones entre los conceptos y clases del dominio de la investigación.

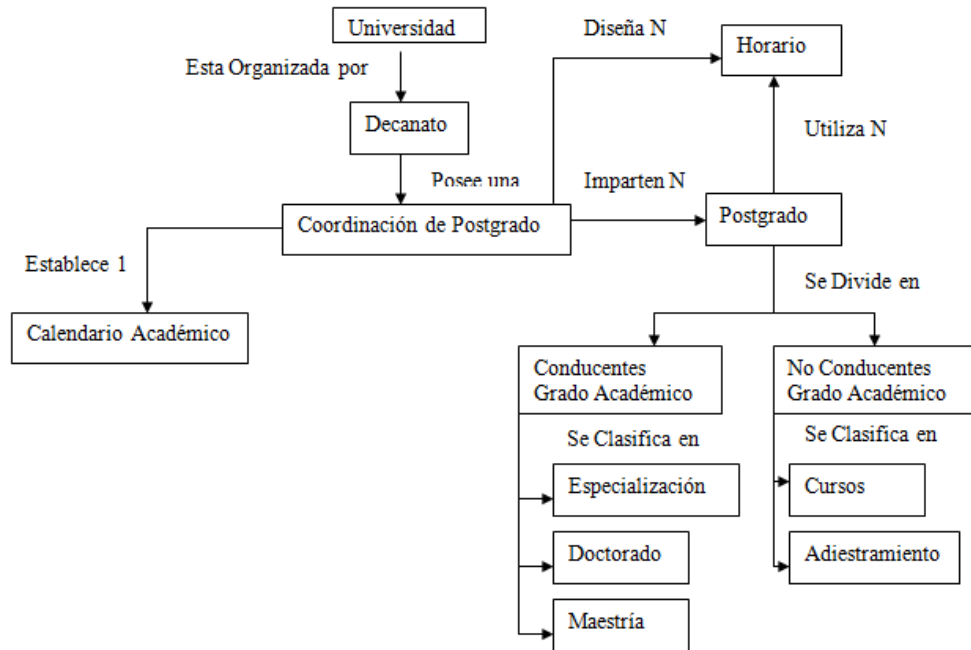


Figura 28. Diagrama de Relación Binaria Ad-Hoc Nro. 1.

Fuente: Autor.

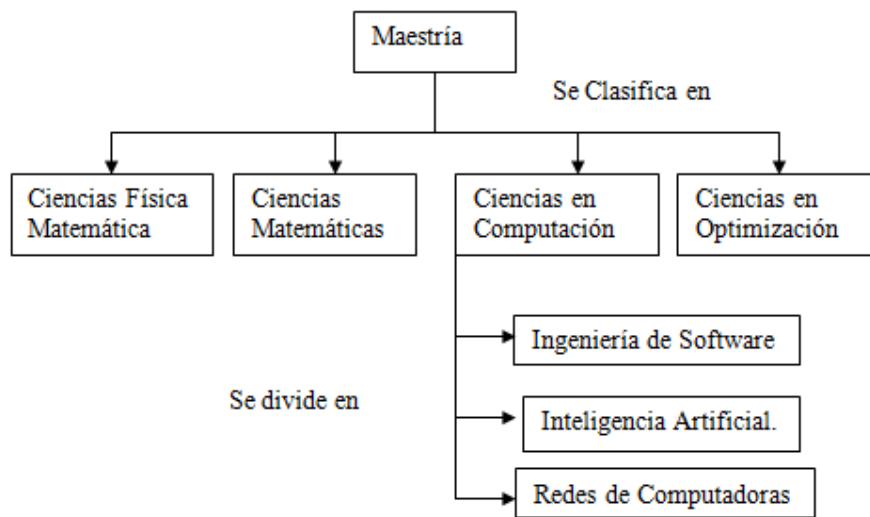


Figura 29. Diagrama de Relación Binaria Ad-Hoc Nro. 2.

Fuente: Autor.

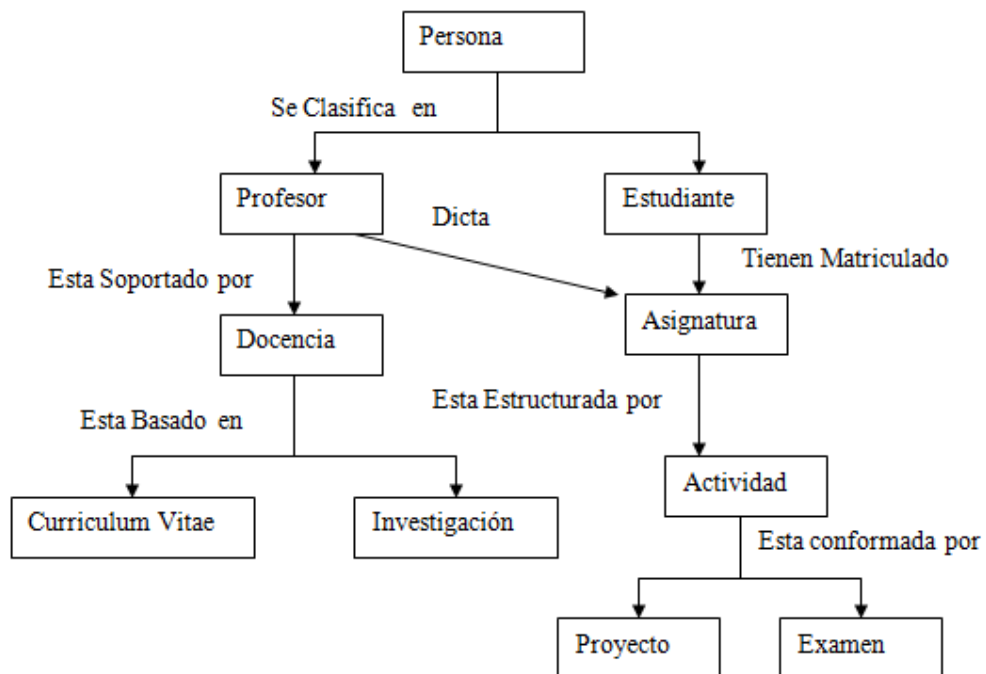


Figura 30. Diagrama de Relación Binaria Ad-Hoc Nro. 3.

Fuente: Autor.

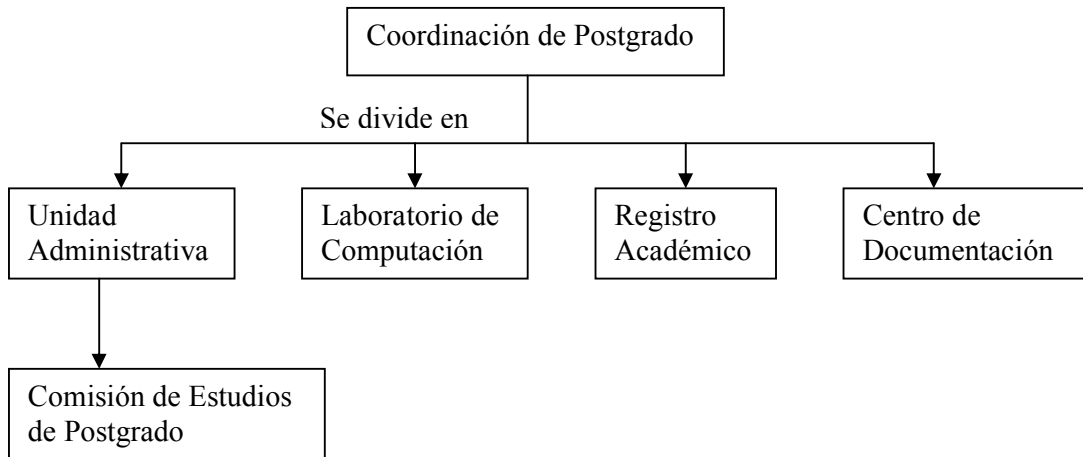


Figura 31. Diagrama de Relación Binaria Ad-Hoc Nro. 4.

Fuente: Autor.

A Continuación mostrare los diagramas de relaciones creados por la herramienta Protégé 2000.

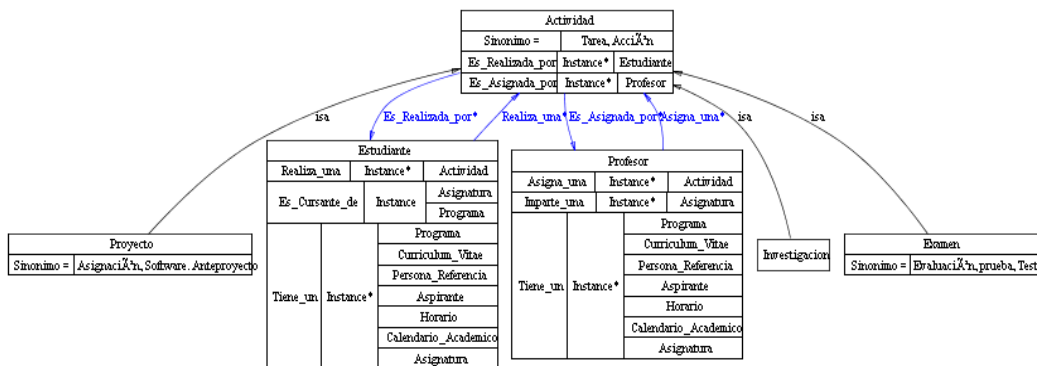


Figura 32. Diagrama de Relaciones con Protégé.

Fuente: Autor:

TAREA 4: Diccionario de Conceptos.

En este diccionario se incluyen todos los conceptos del Vocabulario basado en ontología, así como las propiedades a nivel de concepto y a nivel de instancia y las relaciones que los describen.

Tabla 2. Diccionario de Conceptos.

Concepto	Atributos Clase	Atributo Instancia	Relaciones
Actividad	Nombre Tipo de Actividad Puntuación Tema Fecha de Entrega	Proyecto Examen	
Admisión	Requisitos Programa		
Asignatura	Nombre Unidades de Crédito Nombre Profesor Horario		
Calendario Académico	Periodo Inicio Periodo Fin Periodo		
Coordinación de Postgrado	Código Decanato	-Unidad Administrativa - Laboratorio de Computación. -Registro Académico. -Centro de Documentación. -Comisión de Estudios de Postgrado.	
Decanato	Nombre Decanato Nombre Decano		
Docencia	Nombre Especialidad Curriculum de Requisitos.		
Documentos	Código Nombre Periodicidad	Recaudos Admisión. Curriculum Vitae.	

		Aviso de Prensa. Listado de Admitidos. Carta de Admisión. Hoja de Referencia Académica. Solvencia Administrativa. Reglamento General. Reglamento Interno Aranceles de Postgrado. Manuales. Normativas. Normas para Trabajo de Grado.	
Estudiante	-Nombre -Profesión -Empresa donde Labora -Programa -Nacionalidad -Fecha de Nacimiento -Es Profesor?		
Examen	Nombre Examen Puntaje Fecha de Examen		
Investigación	Código Titulo Autor(es) Tópicos/Temas Fecha de Elaboración		
Programa	Programa	-Ciencias de la computación -Ciencias Física Matemática. -Ciencias Matemática -Ciencias de Optimización	Se Clasifica en Se divide en
	Mención	Mención Ingeniería de Software	

		Mención Inteligencia Artificial Mención Redes de Computadoras.	
	Requisitos para Aprobar	-Aprobar las Unidades de Crédito de las Asignaturas Cursadas. -Tener el Promedio Mínimo ponderado de las Asignaturas Cursadas. -Tener el porcentaje mínimo de asistencias en las asignaturas cursadas. -Culminación de Trabajo de Grado.	
	Plan de Estudio	Nombre de las Asignaturas	
	Líneas de Investigación		
	Evaluación		
	Recursos Materiales		
Postgrado	Conducentes a Grado Académico	Especialización Doctorado Maestría	Se divide en Se Clasifica en
	No Conducentes a Grado Académico	Curso Adiestramiento	
Profesor	Nombre Cargo Fecha de Nacimiento Currículum Proyectos		
Proyecto	Código Nombre Proyecto Asignatura Fecha de Entrega Tópico/Tema Puntaje		
Universidad	Decanatos	Ciencias y Tecnología	Esta Organizado Por

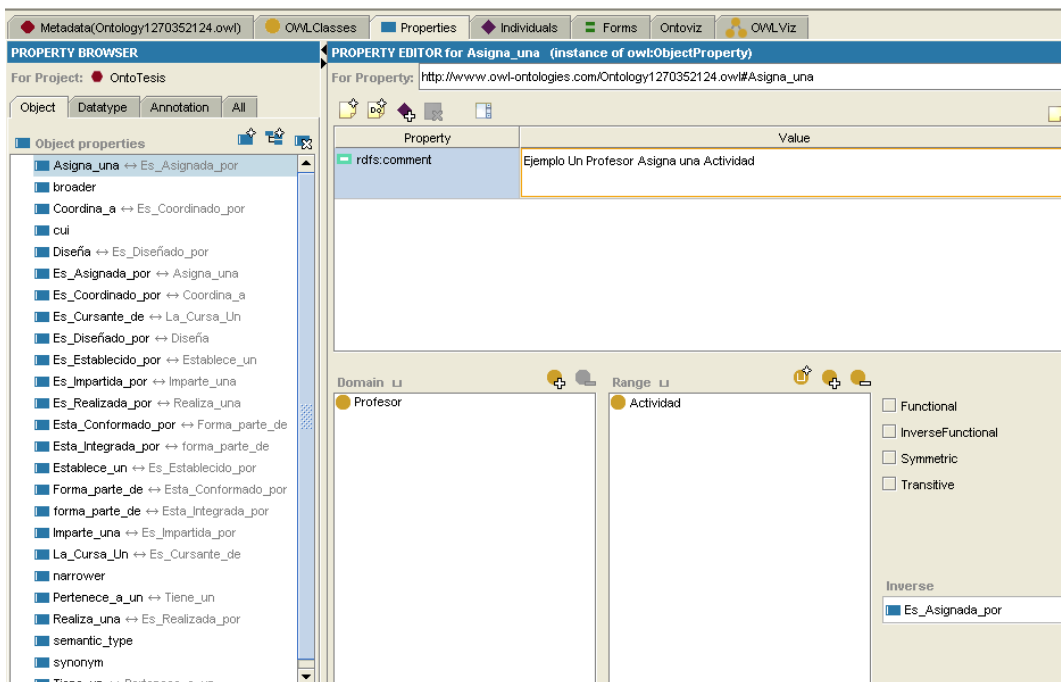
TAREA 5. Detalle de Relaciones.

En esta tarea se detalla las relaciones existentes en el Vocabulario Ontológico. Cuando nos referimos a slots o propiedades, inevitablemente existen características enlazadas con ellos, tal es el caso de: cardinalidad, tipo de dato. Estos refieren a restricciones del tipo de valores se aceptarían, en todo caso es una funcionalidad posible con el OWL. Entre las características más importantes están: Cardinalidad del slot., Tipo de valor de los spots, relación Inversa.

Tabla 3. Detalle de Relaciones.

Nombre Relación	Cardinalidad	Relación Inversa
Asigna_una	1...N	Es_Asignada_Por
Coordina_a	1...N	Es_Coordinado_Por
Diseña_a	1...N	Es_Diseñado_por
Es_Coordinando_Por	1...N	Coordina_a
Es_Cursante_de	1...N	La_Cursa_un
Es_diseñado_Por	1...N	Diseña
Es_Establecido_Por	1...N	Establece_Un
Es_Impartida_Por	1...N	Imparte_Una
Es_Realizada_Por	1...N	Realiza_Una
Esta_Conformada_Por	1...N	Forma_Parte_de
Esta_Integrada_Por	1...N	Forma_Parte_De
Establece_Un	1...N	Es_establecido_Por
Forma_Parte_de	1...N	Esta_Integrado_Por
Imparte_Una	1...N	Es_Impartida_Por
La_Cursa_Un	1...N	Es_Cursante_de
Pertenece_a_Un	1...N	Tiene_Un
Realiza_Una	1...N	Es_realizada_Por
Tiene_Un	1...N	Pertene_a_Un

A Continuación presentó como son configuradas las Propiedades en la Herramienta Protégé2000. Estas Propiedades son las que comúnmente se llaman relaciones, y en la Ontología de la presente investigación se refieren a tres tipos: Las de tipo Objeto que son las relaciones binarias, Tipo DataType son las de tipo de datos y de en las de Tipo Anotación configuraremos los sinónimos de la Ontologías del vocabulario de la presente investigación. Se usará el almacenamiento en fichero, la ontología se guardará en un fichero .owl. En dicho fichero estarán las uris de las ontologías que usen la que estamos definiendo. Si esos recursos no están accesible, por ejemplo porque los tengamos en archivos locales en el sistema de archivos, Protégé es capaz de cargarlos. Para ello usa un fichero con la extensión .repository. El tercer archivo que usa es el .pprj, en el que se almacena toda la información ya específica del nuestro proyecto con Protégé. Si importamos una ontología y esta a su vez depende de otra que no está accesible, bien porque no dispongamos de ella, no se pueda acceder online, o la ruta en el .repository esté mal, Protégé nos dará un error al cargarla.



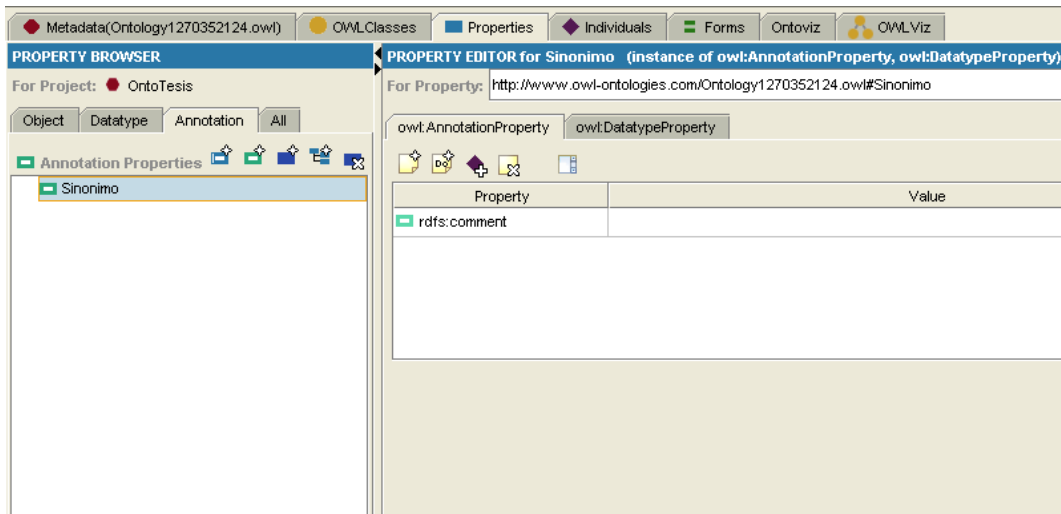
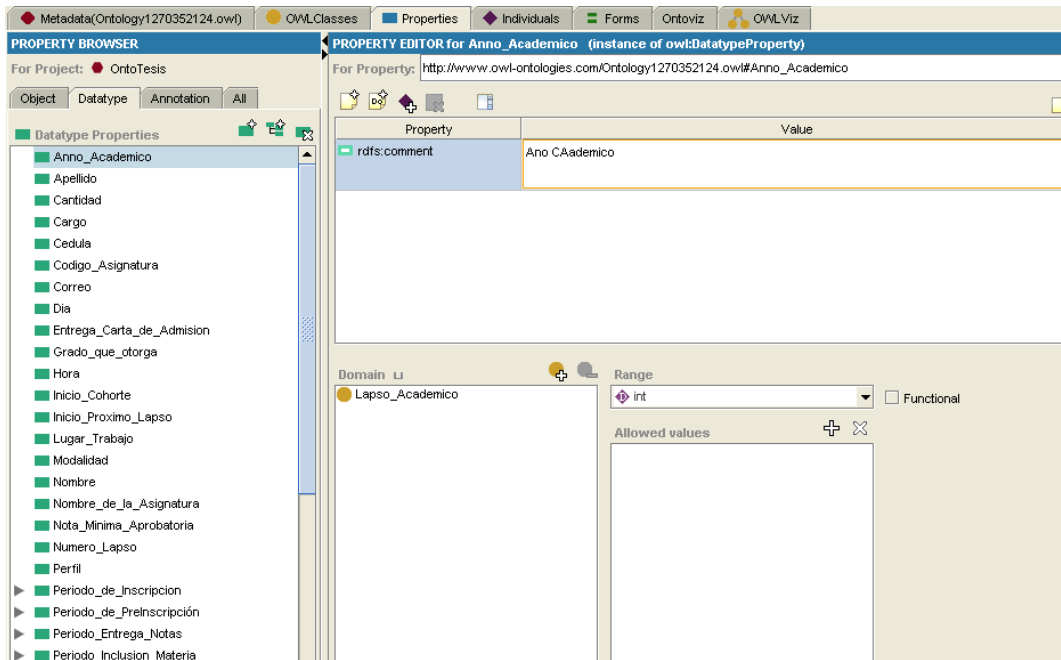


Figura 33. Diagrama de las Relaciones entre los Conceptos de la Ontología Con Protégé 2000.

Fuente: Autor:

TAREA 6. Definir las Instancia.

Este paso consiste en crear instancias individuales de clases en la jerarquía. La definición de una instancia individual de una clase requiere (1) elegir una clase, (2) crear una instancia individual de la clase y (3) rellenar los valores del slot.

TAREA 7. Definir los Atributos de Clases en detalle.

Tabla 4. Atributos de Clase.

Nombre del Atributo	Nombre del Concepto	Tipo	Descripción
Nombre de la Actividad Tipo de Actividad Puntuación Tema Fecha de Entrega	Actividad	Carácter (250)	Es el nombre de la Actividad.
Tipo de Actividad	Actividad	Carácter (1)	Es el tipo de Actividad. Puede tomar los siguientes valores: 'P' para proyecto y 'E' para examen.
Requisitos Programa	Admisión	Carácter(150)	Son todos los requisitos que exigen para inscribirse en u Programa de Grado
Nombre de Unidades de Crédito Nombre Profesor Horario	Asignatura	Numérico	Es la Cantidad de Unidades de Crédito que se le asignan a cada

			asignatura impartida en cada programa
Periodo Inicio Periodo Fin Periodo	Calendario Académico	Date	
Periodo Inicio Periodo Fin Periodo	Coordinación de Postgrado		
Código Decanato	Decanato		
Nombre Decanato Nombre Decano	Docencia		
Nombre Especialidad Curriculum de Requisitos.	Documentos		
Código Nombre Periodicidad	Estudiante		
-Nombre -Profesión -Empresa donde Labora -Programa -Nacionalidad -Fecha de Nacimiento -Es Profesor?	Examen		
Nombre Examen Puntaje Fecha de Examen	Investigación		
Código Titulo Autor(es) Tópicos/Temas Fecha de Elaboración	Maestría		
Programa			
Mención			
Requisitos para Aprobar			

Plan de Estudio			
Líneas de Investigación			
Evaluación			
Recursos Materiales	Postgrado		
Conducentes a Grado Académico			
No Conducentes a Grado Académico	Profesor		
Nombre Cargo Fecha de Nacimiento Curriculum Proyectos	Proyecto		
Código Nombre Proyecto Asignatura Fecha de Entrega Tópico/Tema Puntaje	Universidad		
Decanatos			

TAREA 8. Tabla de Constantes.

Tabla 5. Constantes.

Nombre	Tipo	Valor
Promedio Mínimo Ponderado en las Asignaturas Cursadas	Integer	15
Porcentaje Mínimo de Asistencia en las Asignaturas Cursadas	Integer	85
Cantidad Mínima de Dominio de Idiomas	Integer	2
Unidades de Crédito para Material de Información para estudiantes venezolanos	Integer	0.25
Unidades de Crédito para Material de Información para estudiantes extranjeros	Integer	0.5
Unidades de Crédito para la Inscripción Oficial en el Programa para estudiantes venezolanos	Integer	1.5

Unidades de Crédito para la Inscripción Oficial en el Programa para estudiantes extranjeros	Integer	3
Unidades de Crédito para el Manual de Elaboración de Trabajo de Grado para estudiantes venezolanos y extranjeros.	Integer	0.5
Unidades de Crédito Solicitud de Exámenes de Exención para estudiantes venezolanos.	Integer	1
Unidades de Crédito Solicitud de Exámenes de Exención para estudiantes extranjeros.	Integer	2
Unidades de Crédito para Solicitud de Inscripción del Proyecto del Trabajo de Grado para estudiantes venezolanos.	Integer	2
Unidades de Crédito para Solicitud de Inscripción del Proyecto del Trabajo de Grado para estudiantes extranjeros.	Integer	4
Unidades de Crédito para Solicitud de Retiro temporal del Programa de postgrado para estudiantes venezolanos.	Integer	2
Unidades de Crédito para Solicitud de Retiro temporal del Programa de postgrado para estudiantes extranjeros.	Integer	4
Unidades de Crédito para Solicitud de Retiro Definitivo del Programa de postgrado para estudiantes venezolanos.	Integer	2
Unidades de Crédito para Solicitud de Retiro Definitivo del Programa de postgrado para estudiantes extranjeros.	Integer	4
Unidades de Crédito para Solicitud de Tramites Administrativos del proceso de Equivalencia para estudiantes venezolanos.	Integer	2
Unidades de Crédito para Solicitud de Tramites Administrativos del proceso de Equivalencia para estudiantes extranjeros.	Integer	4
Unidades de Crédito para el reconocimiento de Créditos por equivalencia para estudiantes venezolanos.	Integer	3
Unidades de Crédito para el reconocimiento de Créditos por equivalencia para estudiantes extranjeros.		2
Unidades de Crédito para Expedición del Título de Postgrado para estudiantes venezolanos.	Integer	2

Unidades de Crédito para Expedición del Título de Postgrado para estudiantes extranjeros..	Integer	3
Unidades de Crédito para la solicitud de Acto de Grado y Gastos Administrativos para estudiantes venezolanos.	Integer	1
Unidades de Crédito para la solicitud de Acto de Grado y Gastos Administrativos para estudiantes extranjeros.	Integer	2
Unidades de Crédito para Medallas de Grado para estudiantes venezolanos.	Integer	1
Unidades de Crédito para Medallas de Grado para estudiantes Extranjeros.	Integer	1.5
Unidades de Crédito para la Certificación de Acta de Grado para estudiantes venezolanos.	Integer	1
Unidades de Crédito para la Certificación de Acta de Grado para estudiantes extranjeros.	Integer	1.5
Unidades de Crédito para Solicitud de Notas Certificadas por secretaria general para estudiantes venezolanos.	Integer	1.5
Unidades de Crédito para Solicitud de Notas Certificadas por secretaria general para estudiantes extranjeros.	Integer	1.75
Unidades de Crédito para Solicitud de Notas Certificadas refrendadas por el Rector para estudiantes venezolanos.	Integer	2
Unidades de Crédito para Solicitud de Notas Certificadas refrendadas por el Rector para estudiantes extranjeros.	Integer	2.25
Unidades de Crédito para Solicitud de Constancia de Inscripción otorgado por la Coordinación de Postgrado para estudiantes venezolanos.	Integer	0.25
Unidades de Crédito para Solicitud de Constancia de Inscripción otorgado por la Coordinación de Postgrado para estudiantes extranjeros	Integer	0.5
Unidades de Crédito para Solicitud de Constancia de Inscripción otorgado por la Secretaria General para estudiantes venezolanos.	Integer	0.3

Unidades de Crédito para Solicitud de Constancia de Inscripción otorgado por la Secretaria General para estudiantes extranjeros	Integer	0.5
Unidades de Crédito para Solicitud de Constancia de Inscripción refrendadas por el Rector para estudiantes venezolanos.	Integer	0.5
Unidades de Crédito para Solicitud de Constancia de Inscripción refrendadas por el Rector para estudiantes extranjeros.	Integer	1
Unidades de Crédito para Constancias Cursantes de postgrado para estudiantes venezolanos.	Integer	0.25
Unidades de Crédito para Constancias Cursantes de postgrado para estudiantes extranjeros.	Integer	0.5
Unidades de Crédito para Notas Simples otorgadas por la Coordinación de Postgrado para estudiantes venezolanos.	Integer	0.5
Unidades de Crédito para Notas Simples otorgadas por la Coordinación de Postgrado para estudiantes extranjeros.	Integer	1
Unidades de Crédito para Notas Simples otorgadas por la Secretaria General para estudiantes venezolanos.	Integer	1
Unidades de Crédito para Notas Simples otorgadas por la Secretaria General para estudiantes extranjeros.	Integer	1.5
Unidades de Crédito para Notas Simples Refrendadas por el rector para estudiantes venezolanos.	Integer	1.5
Unidades de Crédito para Notas Simples Refrendadas por el rector para estudiantes extranjeros.	Integer	2
Unidades de Crédito para solicitud de Rendimiento Académico para estudiantes venezolanos.	Integer	0.25
Unidades de Crédito para solicitud de Rendimiento Académico para estudiantes extranjeros.	Ineger	0.5
Unidades de Crédito para Solicitud de Solvencia Administrativa para estudiantes venezolanos.	Integer	0.5
Unidades de Crédito para Solicitud de Solvencia Administrativa para estudiantes extranjeros	Integer	0.5

Unidades de Crédito para Solicitud de Culminación de Escolaridad sin Trabajo de Grado para estudiantes venezolanos.	Integer	0.75
Unidades de Crédito para Solicitud de Culminación de Escolaridad sin Trabajo de Grado para estudiantes extranjeros	Integer	1
Unidades de Crédito para Solicitud de Culminación de Escolaridad con Trabajo de Grado para estudiantes venezolanos.	Integer	0.5
Unidades de Crédito para Solicitud de Culminación de Escolaridad con Trabajo de Grado para estudiantes extranjeros	Integer	1
Unidades de Crédito para la Aprobación del Trabajo de Grado para estudiantes venezolanos.	Integer	0.5
Unidades de Crédito para la Aprobación del Trabajo de Grado para estudiantes extranjeros.	Integer	1
Unidades de Crédito para solicitud de programas de cada asignatura otorgada por la Coordinación de Postgrado para estudiantes venezolanos.	Integer	0.25
Unidades de Crédito para solicitud de programas de cada asignatura otorgada por la Coordinación de Postgrado para estudiantes extranjeros.	Integer	0.5
Unidades de Crédito para solicitud de programas de cada asignatura certificadas por la Secretaria General para estudiantes venezolanos.	Integer	0.3
Unidades de Crédito para solicitud de programas de cada asignatura certificadas por la Secretaria General para estudiantes extranjeros.	Integer	0.6
Unidades de Crédito para solicitud de programas de cada asignatura certificadas por el Rector para estudiantes venezolanos.	Integer	0.5
Unidades de Crédito para solicitud de programas de cada asignatura certificadas por el Rector para estudiantes extranjeros.	Integer	1
Unidades de Crédito para solicitud de Plan de Estudio otorgada por la Coordinación de Postgrado para estudiantes venezolanos.	Integer	0.25
Unidades de Crédito para solicitud de Plan de Estudio otorgada por la Coordinación de Postgrado para estudiantes extranjeros.	Integer	0.5

Unidades de Crédito para solicitud de Plan de Estudio certificadas por la Secretaria General para estudiantes venezolanos.	Integer	0.3
Unidades de Crédito para solicitud de Plan de Estudio certificadas por la Secretaria General para estudiantes extranjeros.	Integer	0.5
Unidades de Crédito para solicitud de programas de Plan de Estudio certificadas por el Rector para estudiantes venezolanos.	Integer	0.5
Unidades de Crédito para solicitud de programas de Plan de Estudio certificadas por el Rector para estudiantes extranjeros.	Integer	0.5

DESARROLLO DE LA PROPUESTA

Según Berners, 2001, Las ontologías son un elemento importante en la Web Semántica, y es un término filosófico que trata sobre la naturaleza de la existencia y los tipos de objetos que existen. También, la ontología se puede definir como una especificación explícita de una conceptualización compartida. La conceptualización se refiere a un modelo abstracto de un fenómeno en el mundo, en donde se tienen que identificar los conceptos relevantes de este fenómeno. El término explícito significa qué tipos de conceptos son usados y las restricciones de estos conceptos. Y, por último, el término compartido se refiere a que las ontologías capturan conocimiento que es aceptado por todos y no es individual de un grupo.

Actualmente, se les está dando más uso a las ontologías en el campo de la Inteligencia Artificial. En este campo se define la ontología como una descripción formal y explícita de conceptos de un dominio (classes, a las que nos referiremos como conceptos), propiedades de cada concepto, las cuales describen las características y atributos de los conceptos (slots llamados también como propiedades o roles) y restricciones de las propiedades (facets conocidas como restricciones), es por esto que la herramienta Protégé 2000 es muy poderosa para el diseño y creación de la Ontología de la presente investigación ya que a través de ella se generan las clases, spots, propiedades y Relaciones existentes en la misma.

Una vez realizada las 8 tareas del Metodología explicada se procede a generar el Framework RDF que es el código de la Ontología basado en XML, este código será consumido por el Programa Java para la muestra de la Ontología.

```
<?xml version="1.0">
<rdf:RDF>
  <rdf:Description about="http://www.ece.uprm.edu/~juanr/tesis.pdf">
    <s:autor>Juan Rojas</s:autor>
  </rdf:Description>
</rdf>
```

Figura 34. Ejemplo de un archivo RDF.

Fuente: Autor.

The screenshot shows a web browser window titled "RDF/XML Source Code" displaying the source code of an OWL ontology. The code includes a DOCTYPE declaration for rdf:RDF, a list of XML entities (owl, swrl, swrlb, xsd, rdfs, rdf, protege, xsp), and the main RDF structure. The main structure includes namespace declarations, an ontology declaration, and class declarations for "Actividad" and "Adiestramiento".

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [
  <!ENTITY owl "http://www.w3.org/2002/07/owl#" >
  <!ENTITY swrl "http://www.w3.org/2003/11/swrl#" >
  <!ENTITY swrlb "http://www.w3.org/2003/11/swrlb#" >
  <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
  <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
  <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
  <!ENTITY protege "http://protege.stanford.edu/plugins/owl/protege#" >
  <!ENTITY xsp "http://www.owl-ontologies.com/2005/08/07/xsp.owl#" >
]>

<rdf:RDF xmlns="http://www.owl-ontologies.com/Ontology1270352124.owl#"
  xml:base="http://www.owl-ontologies.com/Ontology1270352124.owl"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:owl="http://www.w3.org/2002/07/owl#">
  <owl:Ontology rdf:about=""/>
  <owl:Class rdf:ID="Actividad">
    <rdfs:comment rdf:datatype="xsd:string"
      >Conjunto de Tareas o asignaciones por el profesor a cada estudiante cursante de un Programa de
    <Sinonimo rdf:datatype="xsd:string">Tarea, Acci#243;n</Sinonimo>
  </owl:Class>
  <owl:Class rdf:ID="Adiestramiento">
    <rdfs:comment rdf:datatype="xsd:string"
      >Conjunto de Tareas o asignaciones por el profesor a cada estudiante cursante de un Programa de
    <Sinonimo rdf:datatype="xsd:string">Tarea, Acci#243;n</Sinonimo>
  </owl:Class>
```

Figura 35. Código Basado en XML y Ontología para la Presente Investigación.

Fuente: Autor.

El Lenguaje OWL creado a través de un ambiente Web es:

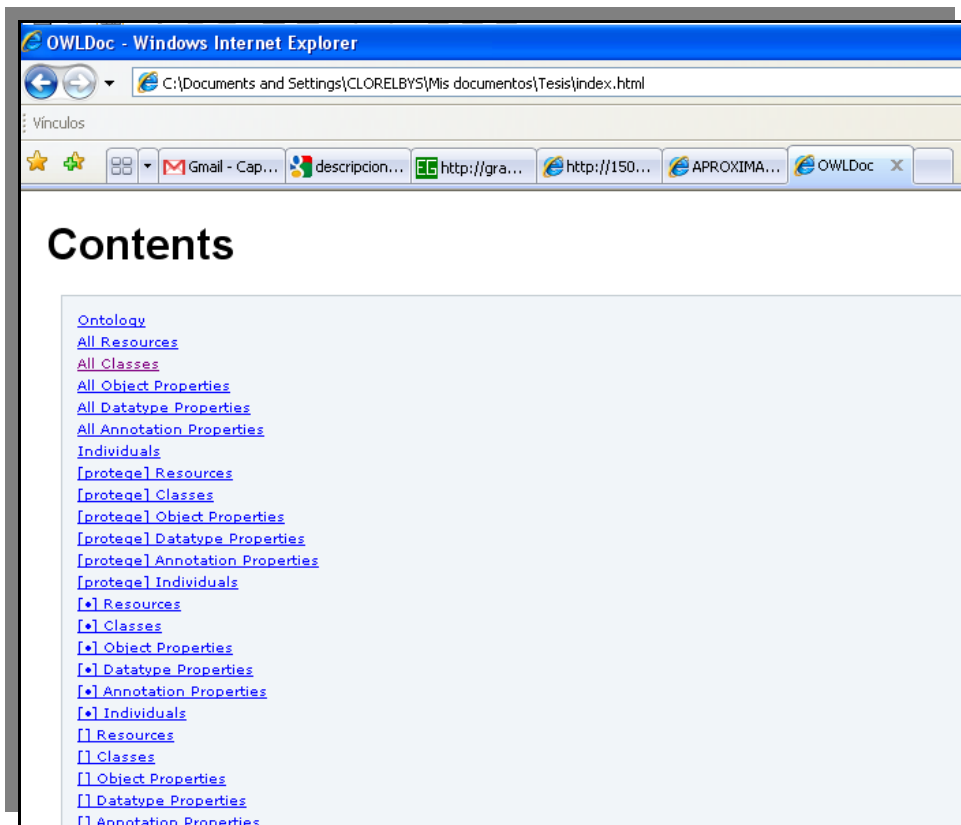


Figura 36. Lenguaje OWL creado a través de un ambiente Web.

Fuente: Autor:

A continuación Presento cada Framework usado por el Protege para generar la Ontología en ambiente Web.

Ontology

Namespaces

Default Namespace

<http://www.owl-ontologies.com/Ontology1270352124.owl#>

xsd	http://www.w3.org/2001/XMLSchema#
xsp	http://www.owl-ontologies.com/2005/08/07/xsp.owl#
swrl	http://www.w3.org/2003/11/swrl#
protege	http://protege.stanford.edu/plugins/owl/protege#
swrlb	http://www.w3.org/2003/11/swrlb#
rdfs	http://www.w3.org/2000/01/rdf-schema#
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
owl	http://www.w3.org/2002/07/owl#
	http://www.owl-ontologies.com/Ontology1270352124.owl#

Figura 37. Ontología Diseñada para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

Fuente: Autor:

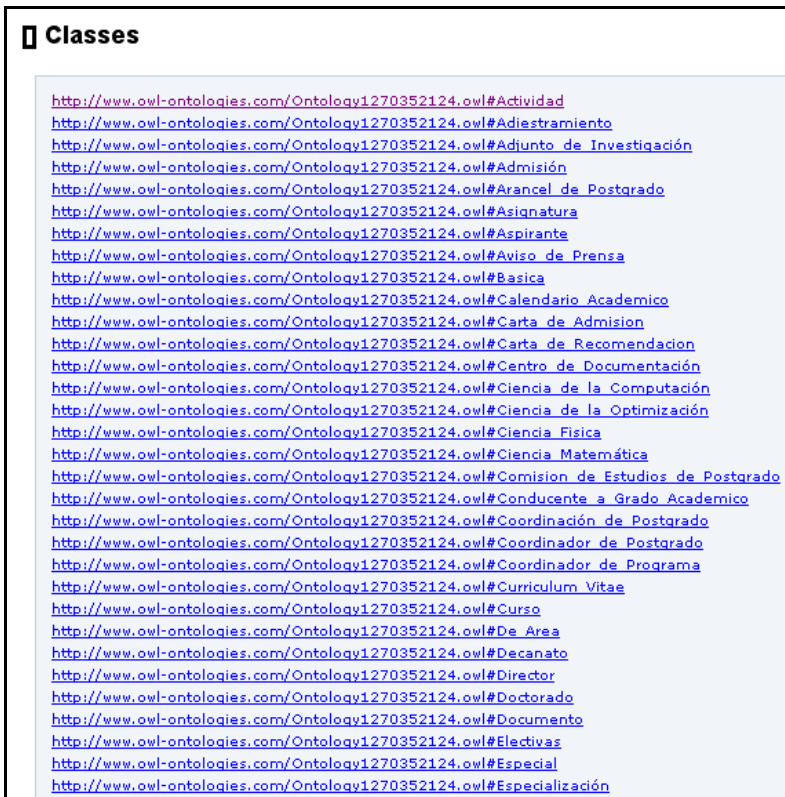


Figura 38. Clases/Recursos en Ambiente en Web de la Ontología en Estudio.

Autor: Clorelbys Escobar.

ONTOLOGIA TEST

Una vez desarrollado el Vocabulario basado en XML y Ontología para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología, se debe probar dicha Ontología para ello use el Framework Ontology Test que lo proporciona Protégé-OWL el cual contiene varias pruebas que pueden ser ejecutadas cuando la Ontología esta siendo editada, es decir a medida que vamos creando nuestra Ontología. Cuando es este test se comienza a ejecutar chequea que las características de las propiedades corresponden con la inversa de la misma. Este Framework esta basado en la Arquitectura de Plugins que permite que nuevas pruebas sean añadidas a la programación de la Ontología. Varios Tipos de Test pueden ser configurados a través

del Dialogo “Test Settings” mostrados en la Figura: , el cual es accesible a través del Menú OWL del Protégé.

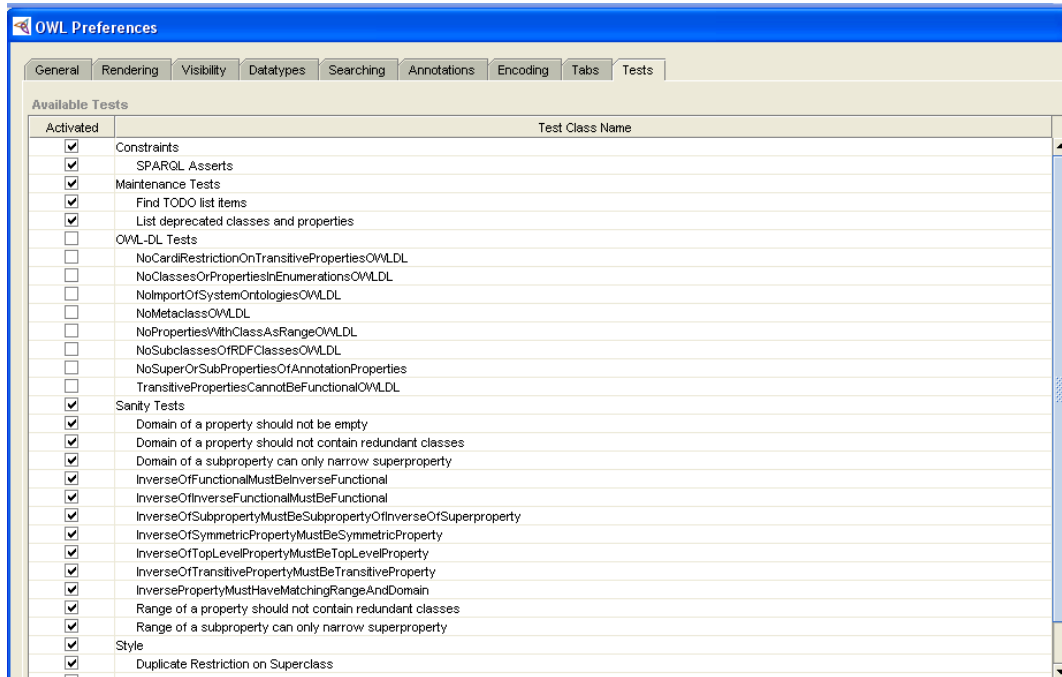


Figura 39. Configuración de Test Settings conm Protégé 200 .

Fuente: Autor.

Luego de que el Test Ontology ha sido ejecutado los resultados son mostrados en el panel Popup e la parte baja de la pantalla que sera mostrada a continuación:

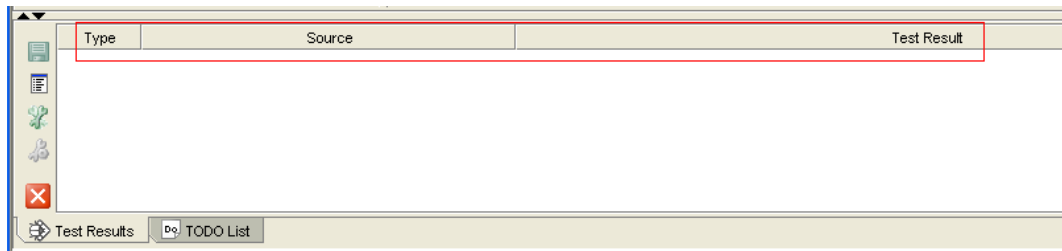


Figura 40. Resultados de la Ontología para la coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.

Fuente: Autor.

En la Columna Tipo muestra el tipo de test (error, Warning, etc.), en la columna fuente muestra si es tipo clase o propiedad y en la columna resultado muestra un mensaje que describe el resultado obtenido. Como se puede apreciar el resultado de la prueba de la ontología en estudio esta sin errores de ningún tipo.

ARQUITECTURA DESARROLLADA

Para la presente investigación la arquitectura propuesta es la siguiente: En la capa Principal se tiene las Interfases predeterminadas y las aplicaciones Externas que vienen siendo los Plug In: Slots, estas son las propiedades y los Tab que son los OWL Clases y las Relaciones. En la Capa de en medio se tiene el Módulo del Conocimiento que es la herramienta donde se diseño la Ontología en nuestro caso Protégé, Y como una ultima Capa tenemos la funciones de mapeado de lectura y escritura que son los ficheros: .ppr, .OWL, los archivos RDF/XML y los objetos de tipo clase en el lenguaje OWL. Esta Arquitectura se puede apreciar en la figura 35. Esta Arquitectura refleja toda el diseño de la Ontología.

En el tema en estudio nos planteamos almacenamiento en memoria es la forma clásica con la que suelen trabajar habitualmente las aplicaciones. En el caso de aplicaciones que involucran el manejo de ontologías, equivale a tener el modelo OWL (ontología) almacenado en memoria principal, como si se tratase de una variable de programa. Una forma de trabajar almacenamiento persistente con ontologías se logra con Jena y es declarando una variable OntModel e ir construyendo el modelo en la propia aplicación (creación de clases, subclases, propiedades, restricciones, etc). Con Jena también es posible cargar un modelo a partir de una fuente local, por ejemplo, un fichero que contenga una ontología definida en el lenguaje de marcado OWL.

El almacenamiento persistente surge ante la necesidad de guardar datos de programa de forma duradera para recuperarlos en otro momento. El término 'persistencia' es sinónimo de 'durabilidad' y 'permanencia'. Obviamente, el almacenamiento persistente en bases de datos supone grandes ventajas sobre el almacenamiento en memoria y el almacenamiento tradicional en el sistema de ficheros. Para ejecutar operaciones sobre bases de datos en una aplicación Java, es necesario JDBC (Java Database Connectivity). La API JDBC es una colección de interfaces Java y métodos de gestión de manejadores de conexión para cada modelo específico de base de datos. En este ámbito, Jena es capaz de trabajar con ontologías almacenadas de forma persistente en una base de datos utilizando el driver JDBC.

Para utilizar el Framework Jena se utilizó el editor NetBeans y de allí se exportó el archivo Onto.OWL que arroja la herramienta protege.

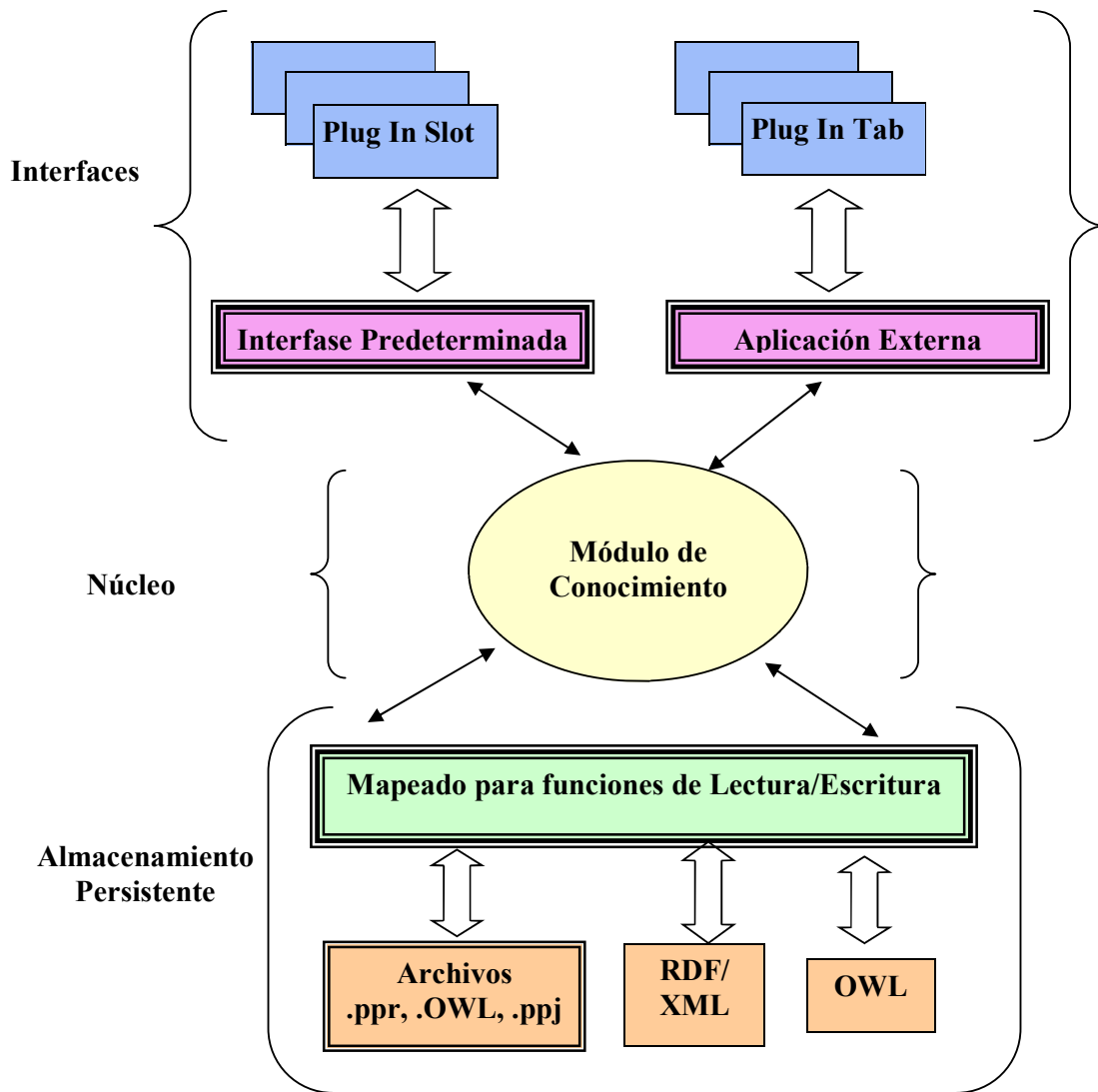


Figura 41. Arquitectura del Vocabulario basado en XML y Ontología para la Coordinación de Postgrado de la UCLA.

Fuente: Autor.

CAPITULO V

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

Al desarrollar la presente investigación, llegué a las siguientes conclusiones:

- El uso de la Metodología MethoOntology, en conjunto con las tareas que especifican la taxonomía y el diccionario de Conceptos permitió diseñar y estructurar la Ontología de Conceptos para el Vocabulario basado en XM para la Coordinación de Postgrado del Decanato de Ciencias y Tecnología.
- La revisión de la Pagina Web de la UCLA así como también de la página Web de Postgrado del Decanato de Ciencias y Tecnología de la UCLA sirvió para establecer la jerarquía de Clases y/o Conceptos, Relaciones, Slots y Propiedades del Vocabulario de la Presente Investigación.
- A través de la Herramienta Protégé se diseño la Ontología de forma amigable al usuario ya que las implementaciones son fáciles a nivel de interfase a través de menús, y las funciones adjuntas se separan en pestañas. Por otro lado estas librerías tienen la capacidad de insertar diferentes plugins, es decir que es un software adaptable a las necesidades del diseñador.
- Las Herramientas tecnológicas utilizadas para la construcción de la Ontología son accesibles para trabajar en diversas plataformas de Software y Hardware ya que son Código Abierto.
- El Framework de Diseño de Recursos RDF, un modelo de meta datos basado en XML, que promete ser un estándar flexible para la estructuración de la información en Clases jerárquica, y posee un potencial desarrollo para optimizar la recuperación de información en Internet. Se permitió apreciar que RDF es un

conjunto de informaciones relativas a las clases de recursos que sirve para explicitar las relaciones jerárquicas que establecen entre ellos, o bien para matizar el carácter obligatorio u opcional de las propiedades y otras restricciones como el número de ocurrencias.

- En la Presente Investigación se verifico que OWL es un mecanismo para desarrollar temas o vocabularios específicos. Lo que hace OWL es proporcionar un lenguaje para definir ontologías estructuradas que pueden ser utilizadas a través de diferentes sistemas. Las ontologías, que se encargan de definir los términos utilizados para describir y representar un área de conocimiento, son necesarias por las bases de datos y las aplicaciones que necesitan compartir información específica, es decir, en un campo determinado como puede ser el de las finanzas, medicina, deporte u educación.
- A través del Framework Test Ontology se logró realizar la prueba a la Ontología desarrollada para verificar que fue perfectamente realizada, es decir hubo congruencia entre clases, propiedades, relaciones y sus relaciones inversas.
- A través de los frameworks Jena y Pellet se logró probar que el vocabulario basado en XML y Ontología diseñado y desarrollado era correcto y consistente.
- El vocabulario basado en XML y Ontología para la coordinación de postgrado del Decanato de Ciencias y Tecnología.

Recomendaciones

Tomado en cuenta las conclusiones planteadas, el autor considera conveniente:

- Se recomienda implementar el vocabulario basado en XML y Ontología para la coordinación de Postgrado del Decanato de Ciencias y Tecnología de la UCLA.
- Se recomienda extender la aplicabilidad de dicha Ontología incluyendo términos que abarquen otras áreas y/o departamentos del Decanato de Ciencias y Tecnología de la UCLA.
- Se recomienda que a esta Ontología se aplique otros métodos de almacenamiento persistentes utilizando ficheros de lenguajes.
- En la presente investigación utilice Protégé para la creación de la ontología así como para las pruebas y la depuración, almacenando la información normalmente en un archivo de texto con la representación en OWL. Sin embargo recomiendo también el uso de un almacenamiento persistente basado en base de datos, para lo cual es necesario exportar la información OWL del archivo de texto a un modelo relacional, para ello recomiendo utilizar el plugin de Protégé: Protege2Jena, teniendo en cuenta que el objetivo posterior es acceder a la información desde el propio Protégé o mediante Jena.
- En la presente investigación se generó el vocabulario basado en XML y ontología por lo que se recomienda que al momento de desplegar los datos o metadatos, se utilice la codificación UNICODE en la creación de los metadatos contenidos en los documentos XML ya que esta codificación es parte de la capa más básica de la arquitectura de la Web Semántica.

REFERENCIAS BIBLIOGRAFICAS

Abed Gregio AR, Barbato LGC, Duarte LO, Montes A, Hoepers C, Stedding-Jessen K. 2007. Taxonomías de Vulnerabilidades: Situación Actual. U.R.L: <http://www.ppgia.pucpr.br/~maziero/pesquisa/ceseg/sbseg05/artigos/12463.pdf>

AHMED, I. 2008. Tag Clouds in Ruby on Rails. URL: <http://www.ibrahim-ahmed.com/2008/03/tag-clouds-in-ruby-and-rails.html>

ALVAREZ, M. 2001. Descripción y características de este potente y moderno lenguaje de programación. U.R.L: <http://www.desarrolloweb.com/articulos/497.php>.

ANAYA, R.; Ramos, I. 2000. AR₂CA:Un Herramienta para la Construcción de Componentes Reutilizables a Través de Niveles de Refinamiento. Memorias 3^{er} Workshop Iberoamericano de Ingeniería de Requisitos y Ambientes Software. Cancún, México.

ARTICULO N° 110. Constitución de la República Bolivariana de Venezuela. Diciembre 1999. URL: <http://www.venezuela-oas.org/Constitucion%20de%20Venezuela.htm>

BAADER, F., Calvanese, D., Nardi, D., Patel-Schneider, P. 2003. The Description Logic Handbook.

BECKETT, D. 2004. RDF/XML Syntax Specification (Revised). URL: <http://www.w3.org/TR/2004/REC-rdf-syntax-grammar-20040210/> . Latest version **available at** <http://www.w3.org/TR/rdf-syntax-grammar/>.

BERNERS,L., T., J. Hendler y O. Lassila. 2001 The Semantic Web: A new Form of Web Content that is Meaningful to Computers will Unleash a Revolution of new Possibilities. <http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21>.

Borst, W.N., “Construction of Engineering Ontologies”. University of Tweenty. Enschede, NL- Centre for Telemática and Information Technology. (1997).

BRICKLEY, D. 2004. RDF Vocabulary Description Language 1.0: RDF Schema. URL: <http://www.w3.org/TR/2004/REC-rdf-schema-20040210/> . Latest version available at <http://www.w3.org/TR/rdf-schema/>

Centelles, M. 2005. Taxonomías para la categorización y la organización de la información en sitios web . U.R.L: <http://www.hipertext.net>.

Corcho, O., Fernández, M, Gómez, A. 2007. Metodologías, Herramientas, y Lenguajes para la Construcción de Ontologías.

DEAN, M., Schreiber G. 2004. OWL Web Ontology Language Reference. URL: <http://www.w3.org/TR/2004/REC-owl-ref-20040210>

DECRETO N° 3.390, Uso Prioritario de Software Libre.
Diciembre 2004, URL:
<http://www.gobiernoenlinea.gob.ve/>

FERNANDEZ, O.2006. Cliver Title Pending. URL:
<http://seangeo.blogspot.com/2006/02/active-ontology-for-ruby-on-rails.html>

Fernández, M., Gómez-Pérez, A., Juristo, N. 1997. "METHONTOLOGY: From Ontological Art Towards Ontological Engineering", AAAI Symposium on Ontological Engineering, Stanford,

FUHR, N., Weikum, G. 2004. Classification and Intelligent Search on Information in XML.

GAITANOU, P. 2004. Ontology and Applications. URL:
<http://www.mtsr.ionio.gr/proceedings/gaitanou.pdf>

GOBIERNO electrónico de Venezuela. URL:
<http://dircompucv.ciens.ucv.ve/Documentos/RT-2006-09.pdf>

GONZÁLEZ, E 2004. Modelo de software basado en la integración de aplicaciones empresariales para el Decanato de Ciencias y Tecnología de la UCLA. Trabajo de Grado. Universidad Centrooccidental Lisandro Alvarado. Barquisimeto. 165 p.

Gruber, T.R., "A Translation Approach to Portable Ontology Specifications". Knowledge Acquisition Journal, Vol. 5 pp. 199-200,1993.

Guarino, N., "Formal Ontology and Information Systems. Formal Ontology in Information Systems", Proc. Of the 1st International Conference, Trento, Italy, 6-8 june 1998, Ed. N. Guarino, IOS Press.

GUERRERO BOTE, Vicente; LOZANO TELLO, Adolfo."Vínculos entre las Ontologías y la Biblioteconomía y Documentación" en La Representación y la Organización del Conocimiento en sus distintas perspectivas: su influencia en la Recuperación de la Información : Actas del IV Congreso ISKO-España EOCONSID'99 22-24 de abril de 1999 Granada. Granada: ISKO; Universidad de Granada. Facultad de Biblioteconomía y Documentación, 1999, p. 25-31.

HERNANDEZ Sampieri, Roberto; Fernández, Carlos y Batista, Pilar. 1999 Metodología de la Investigación. Segunda Edición. D.F.- México. McGraw H.

Herrera,D. 2009. U.R.L: <http://dvherrera.blogspot.com/>.

HU, H., Du, X. 2002. Adopting Ontologies and Rules in Web Searching Services. Information School, Renmin University of China, Beijing China. U.R.L: <http://www.springerlink.com/content/4x4k154hm3ck9rpp/>

KAPPA-PC OnLine Manual.

KIM, H. 2002. XML-hoo!: A Prototype Application for Intelligent Query of XML Documents using Domain-Specific Ontologies.

KLYNE, G., Carroll, J. 2004. Resource Description Framework (RDF): Concepts and Abstract Syntax. URL: <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/> . Latest version **available at** <http://www.w3.org/TR/rdf-concepts/>

MENDONÇA, M. 2004. Vocabulario en XML para la Gestión Catastral de Inmuebles Urbanos. Trabajo de Grado. Universidad Centrooccidental Lisandro Alvarado. Barquisimeto. 105 p.

MCGUINNESS, D., Harmelen, F. 2004. OWL Web Ontology Language Overview. URL: <http://www.w3.org/TR/owl-features/>

NCI. 2007. U.R.L: <http://ncimeta.nci.nih.gov/>

NECHES, R.; Fikes, R.E.; Finin T.; Gruber T.R.; Senator T., and Swartout W.R. "Enabling technology for knowledge sharing". AI Magazine. Vol.12 (3) p.p. 36-56, 1991

O'LEARY, Z. 2004. Essential Guide to Doing Research. GBR: Sage Publications, Incorporated. Londres.

Ontologías.2002. U.R.L: <http://www.ontology.org/> .

Protégé-2000. abril-mayo 2002. U.R.L: <http://protege.stanford.edu/index.html>.

Racer. U.R.L: : <http://www.racer-systems.com/products/tools/protege>.

Rowland, R. 2000. Creative Guide to Research : How to Find What You Need... Online or Offline. New Jersey, EUA.

Semantic Web. U.R.L: <http://www.semanticweb.org/>

SHEN Li, T., Han, T., Chen, G. 2004. XML Query Based on Ontology

Stevens, Robert; Wroe, Chris; Bechhofer, Sean; Lord, Phillip; Rector, Alan and Goble, Carole. Conference Review, "Building ontologies in DAML + OIL, Comparative and

Functional Genomics”, *Comp Funct Genom* 2003; 4: 133– 141. Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/cfg.233

Studer, R.; Benjamins, R. and Fensel, D. “Knowledge Engineering: Principles and Methods”. *DKE* 25(1-2).pp:161-197. 1998.

SURE, Y., Studer R. 2003. A Methodology for Ontology-based Knowledge Management.

TELLO, L. 2001. Herramientas de Software basados en Ontología para mejorar la accesibilidad en la Web. U.R.L: [tps://bdigital.ufp.pt/dspace/bitstream/10284/400/1/70-79.pdf](https://bdigital.ufp.pt/dspace/bitstream/10284/400/1/70-79.pdf)

TRAMUYAS, Jesús. "Agentes y ontologías para el tratamiento de la información: clasificación y recuperación en Internet" en *La Representación y la Organización del Conocimiento en sus distintas perspectivas: su influencia en la Recuperación de la Información : Actas del IV Congreso ISKO-España EOCONSID'99* 22-24 de abril de 1999 Granada. Granada: ISKO; Universidad de Granada. Facultad de Biblioteconomía y Documentación, 1999, p.247-252.

UCLA (2002). Manual para la Presentación del Trabajo Conducente al Grado Académico de: Especialización, Maestría, Doctorado. Barquisimeto, Venezuela

UNIVERSIDAD PEDAGÓGICA EXPERIMENTAL LIBERTADOR. (2006) Manual de Trabajos de Grado de Especialización y Maestría y Tesis Doctorales. Cuarta Edición. Barquisimeto, Venezuela.

VÉLIZ, A. 2005. Como hacer y defender una tesis. Editorial Texto. Caracas

Van Heijst, G.; Schreiber, A.T. and Wieling, B.J. “ Using Explicit Ontologies in KBS Development” *International Journal of Human and Computer Studies*, 1996.

W3C Recommendation February 2004. OWL Web Ontology Language Overview. U.R.L: <http://www.w3.org/TR/owl-features/>

Weigand, H., "Multilingual Ontology-Based Lexicon for News Filtering –The TREVI Project", en K. Mahesh (1997): 138-159.

WindsWap. 2002. U.R.L: <http://www.mindswap.org/2003/CancerOntology/>.

XML. U.R.L: http://www.xml.org/xml/resources_cover.shtml

ANEXOS

ANEXO 1

GUÍA DE INSTALACIÓN PROTÉGÉ 2000

Para la Instalación de la herramienta Protégé 2000 se debe:

1.- Ingresar a la página <http://protege.stanford.edu/>

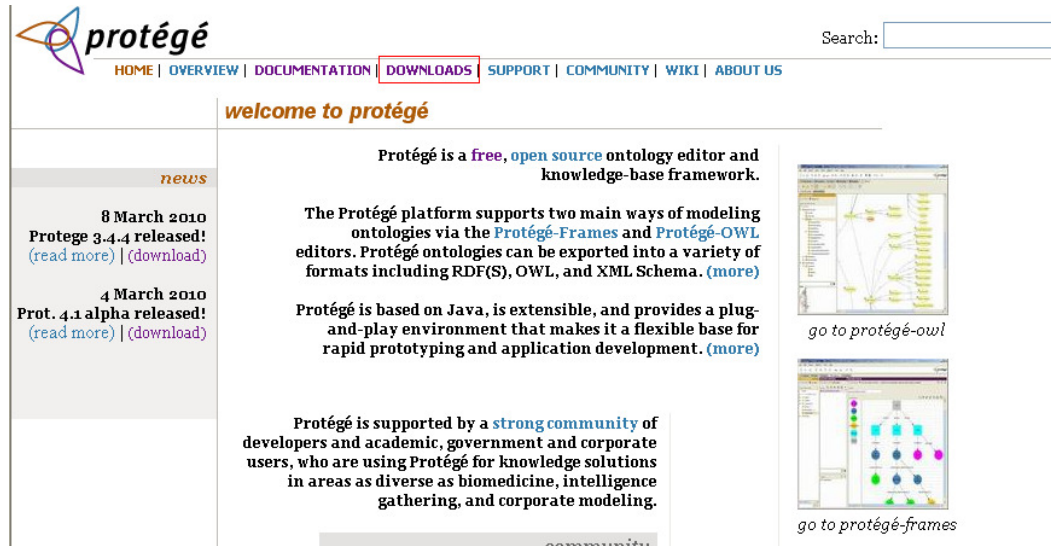


Figura 42: Pagina Principal de Descarga de Protégé.

Fuente: <http://protege.stanford.edu/>

2.- Ingresar al tab Downloads

3.- Para Usuarios Nuevos realizar el Registro.

4.- Una vez registrados se procede a hacer Click en Download

5.- Se despliega una ventana para escoger la versión del Protégé Software, para la presente investigación se utilizó el Protégé 3.4.4 ya que soporta OWL 1.0, RDF(S) y Frames.

6.- Al escoger la versión se despliega una ventana para seleccionar el tipo de descarga según la plataforma que se use y si se desea instala la Java Virtual Machine y si se desea usar una previamente instalada en la PC.

Protege 3.4.4 (build 579)

Recommended Installation for Your Platform:

[Download Installer for Windows...](#) Include a VM in download

Installer created with InstallAnywhere. Copyright © 1997-2008 Access Software Inc.
and/or InstallShield Co. Inc.

Available Installers



Platform	includes Java VM	without Java VM	Instructions
 MacOSX		Download (72.8M)	View
>  Windows 64bit	Download (94.4M)	Download (73.4M)	View
>  Windows	Download (98.6M)	Download (73.4M)	View
 Linux 64bit	Download (113.3M)	Download (73.4M)	View
 Linux	Download (122.2M)	Download (73.4M)	View
 Any Unix Platform		Download (73.4M)	View
 HPUX	Download (136.1M)	Download (73.4M)	View
 Solaris	Download (114.7M)	Download (73.4M)	View
 AIX	Download (132.1M)	Download (73.4M)	View
 Other Java-enabled Platforms		Download (73.1M)	View

Figura 43: Página de Selección de Protégé según la Plataforma.

Fuente: <http://protege.stanford.edu/>

7.- Se comienza a ejecutar la descarga para su instalación y se escoge una ruta de almacenamiento como por ejemplo: C:\Archivos de programa\Protege_3.4.4.

8.- En dicha carpeta se instalan una serie de carpetas las cuales son: examples, plugins, logs, jre, entre otros

9.- En la carpeta Plugins se deben extraer todos los plugins que se deseen de acuerdo la tipo de ontología que se dese desarrollar por ejemplo para la presente investigación se utilizo el plugin uk.ac.man.cs.mig.coode.owlviz para realizar los diagramas de Taxonomías.

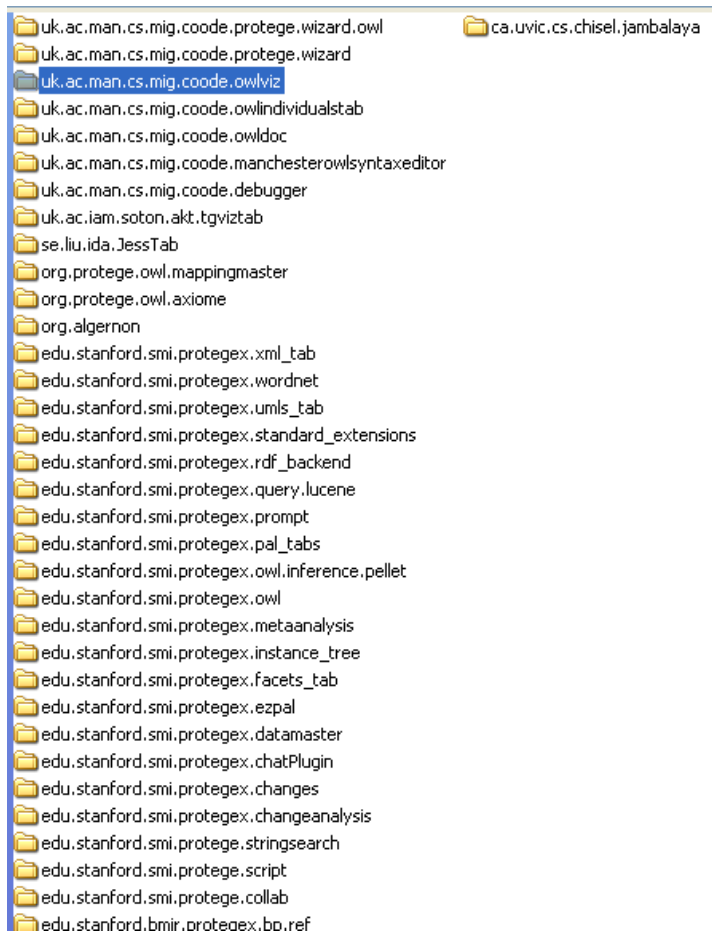


Figura 44: Plugins
Fuente: Protégé Instalado.

10.- Los plugins se descargan de la misma pagina en la sección Plugin y solo se tiene que descomprimir el plugin en la carpeta plugin del Protégé Instalado.

11.- Es importante recalcar que para el uso del Plugin antes mencionado se debe tener previamente Instalado el GraphViz y esta herramienta se puede descargar de la siguiente página: <http://graphviz.softonic.com/mac/descargar>.

12.- Una vez instalado el protege2000, el plugin OWLViz y la herramienta GraphViz se necesita decirle al Protégé donde encontrar dicha herramienta esto se hace por la pestaña OWLVIZ del protege, Botón Options

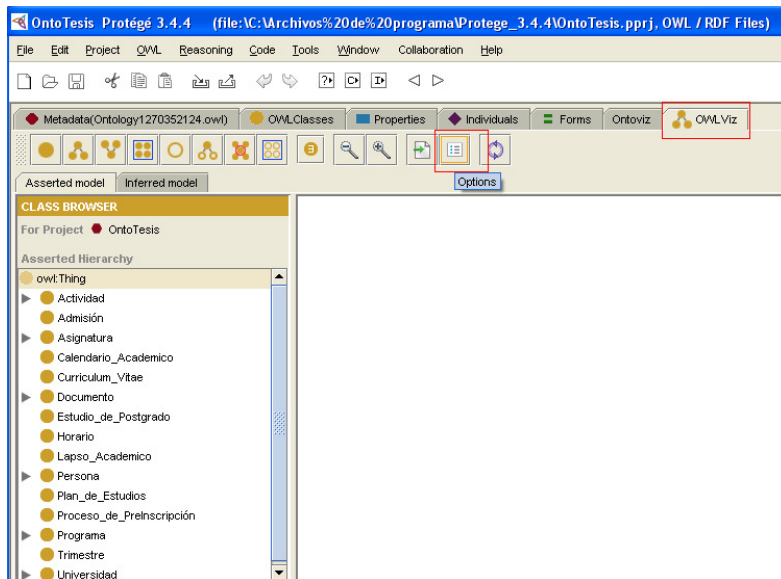


Figura 45: Ventana Clases en Protégé.

Fuente: Protégé Instalado.

13.- Allí se descarga la ventana Options y en Dot Application Path se coloca la ruta donde se encuentra el GraphViz instalado.

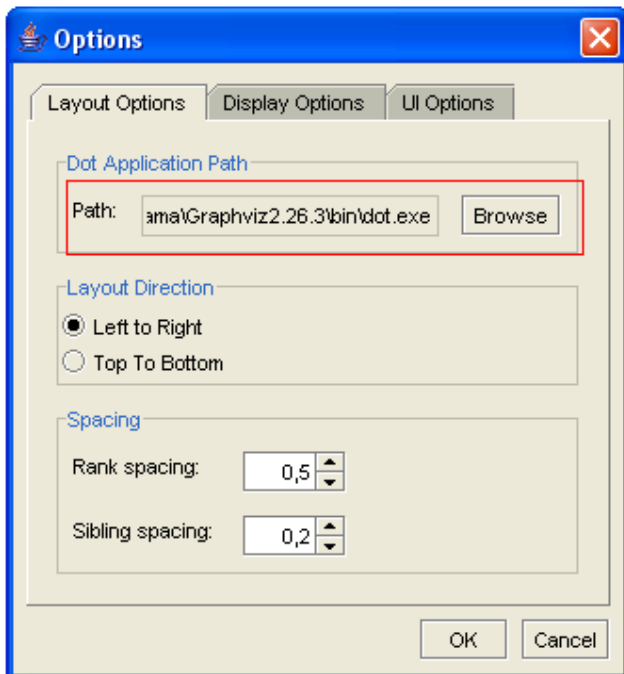


Figura 46: Ventana Options en Protégé.

Fuente: Protégé Instalado.

Currículum Vitae del Autor

Clorelbys Coromoto Escobar Querales, Portador de la C.I: N° V-15.004.241, nace en Barquisimeto el 26 de Junio de 1980. Realiza estudios de Primaria en la Escuela Nacional “José Gil Fortoul”. Obtiene el título de bachiller en el Liceo “Lisandro Alvarado” de Barquisimeto. Ingresa a la Universidad Centroccidental “Lisandro Alvarado” de Barquisimeto en el año 1998, para iniciar estudios de pregrado en la carrera de Ingeniería en Informática, egresando en el año 2003 con Mención CumLaude. En el año 2004 empieza a trabajar en la Empresa Orion Consultores C.A, desempeñando funciones de Programador y Administrador de Base de Datos en Sistemas Administrativos. En año 2006 comienza trabajar en Corporación BEL desempeñando funciones de Programador. En Mayo del 2007 comienza a trabajar en Inter adscrita a la Gerencia de Proyectos desempeñando funciones de Programador hasta la presente fecha.