



Universidad de San Carlos de Guatemala  
Facultad de Ingeniería  
Escuela de Ingeniería en Ciencias y Sistemas

**GESTIÓN DEL CONOCIMIENTO MEDIANTE EL USO DE ONTOLOGÍAS EN EL  
SERVICIO DE ATENCIÓN MÉDICA HOSPITALARIA DE GUATEMALA**

**Raúl Estuardo Cuyún González**

Asesorado por el Ingeniero Freiry Javier Gramajo López.

Guatemala, octubre de 2009

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA.



FACULTAD DE INGENIERÍA

**GESTIÓN DEL CONOCIMIENTO MEDIANTE EL USO DE ONTOLOGÍAS EN  
EL SERVICIO DE ATENCIÓN MÉDICA HOSPITALARIA DE GUATEMALA**

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA

FACULTAD DE INGENIERÍA

POR:

**RAÚL ESTUARDO CUYÚN GONZÁLEZ**

ASESORADO POR EL INGENIERO FREIRY JAVIER GRAMAJO LÓPEZ

AL CONFERÍRSELE EL TÍTULO DE

**INGENIERO EN CIENCIAS Y SISTEMAS**

GUATEMALA, OCTUBRE 2009.

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

FACULTAD DE INGENIERÍA



**NÓMINA DE JUNTA DIRECTIVA**

DECANO	Ing. Murphy Olympo Paiz Recinos
VOCAL I	Inga. Glenda Patricia García Soria
VOCAL II	Inga. Alba Maritza Guerrero Spínola de López
VOCAL III	Ing. Miguel Ángel Dávila Calderón
VOCAL IV	Br. Milton De León Bran
VOCAL V	Br. Isaac Sultán Mejía
SECRETARIA	Inga. Marcia Ivónne Véliz Vargas

**TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO**

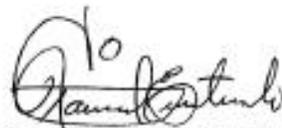
DECANO	Ing. Murphy Olympo Paiz Recinos
EXAMINADOR	Ing. Edgar Estuardo Santos Sutuj
EXAMINADOR	Inga. Virginia Victoria Tala Ayerdi
EXAMINADOR	Ing. Ludwin Federico Lantan Sac
SECRETARIA	Inga. Marcia Ivónne Véliz Vargas

## HONORABLE TRIBUNAL EXAMINADOR

Cumpliendo con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

### **GESTIÓN DEL CONOCIMIENTO MEDIANTE EL USO DE ONTOLOGÍAS EN EL SERVICIO DE ATENCIÓN MÉDICA HOSPITALARIA DE GUATEMALA,**

tema que me fuera asignado por la Dirección de la Escuela de Ingeniería en Ciencias y Sistemas, en Febrero 2007.



Raúl Estuardo Cuyún González

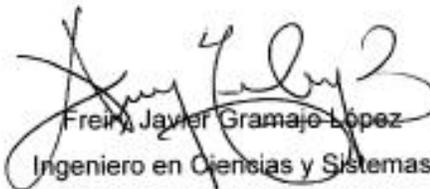
Guatemala, 13 de mayo de 2009

**Ing. Marlon Perez Turk**  
**Escuela de Ciencias y Sistemas**  
**Facultad de Ingeniería**

Por este medio hago de su conocimiento que he revisado el trabajo de graduación de el estudiante RAUL ESTUARDO CUYUN GONZALEZ con cané 200010423 titulado: "**GESTION DEL CONOCIMIENTO MEDIANTE EL USO DE ONTOLOGÍAS EN EL SERVICIO DE ATENCIÓN MÉDICA HOSPITALARIA DE GUATEMALA.**", y a mi criterio el mismo cumple con los objetivos propuestos para su desarrollo, según el protocolo propuesto, por lo que firmo la presente para que proceda con los trámites correspondientes.

Sin otro particular, me suscribo de usted.

Atentamente,



Freiry Javier Gramajo López  
Ingeniero en Ciencias y Sistemas

Colegiado No. 4036

Asesor y Revisor de Trabajo de Graduación

*Freiry Javier Gramajo López*  
Ing. Ciencias y Sistemas  
Col. 4036  
DEA, Inteligencia Artificial



Universidad San Carlos de Guatemala  
Facultad de Ingeniería  
Escuela de Ingeniería en Ciencias y Sistemas

Guatemala, 29 de Julio de 2009

Ingeniero  
**Marlon Antonio Pérez Turk**  
Director de la Escuela de Ingeniería  
En Ciencias y Sistemas

Respetable Ingeniero Pérez:

Por este medio hago de su conocimiento que he revisado el trabajo de graduación del estudiante **RAUL ESTUARDO CUYUN GONZALEZ**, titulado: **"GESTIÓN DEL CONOCIMIENTO MEDIANTE EL USO DE ONTOLOGIAS EN EL SERVICIO DE ATENCION MEDICA HOSPITALARIA DE GUATEMALA"**, y a mi criterio el mismo cumple con los objetivos propuestos para su desarrollo, según el protocolo.

Al agradecer su atención a la presente, aprovecho la oportunidad para suscribirme,

Atentamente,

  
**Ing. Carlos Alfredo Azuela**  
Coordinador de Privados  
y Revisión de Trabajos de Graduación



E  
S  
C  
U  
E  
L  
A  
  
D  
E  
  
C  
I  
E  
N  
C  
I  
A  
S  
  
Y  
  
S  
I  
S  
T  
E  
M  
A  
S

UNIVERSIDAD DE SAN CARLOS  
DE GUATEMALA



FACULTAD DE INGENIERÍA  
ESCUELA DE CIENCIAS Y SISTEMAS  
TEL.: 24767644

*El Director de la Escuela de Ingeniería en Ciencias y Sistemas de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del asesor con el visto bueno del revisor y del Licenciado en Letras, de trabajo de graduación titulado "GESTIÓN DEL CONOCIMIENTO MEDIANTE EL USO DE ONTOLOGÍAS EN EL SERVICIO DE ATENCIÓN MÉDICA HOSPITALARIA DE GUATEMALA", presentado por el estudiante RAÚL ESTUARDO CUYÚN GONZÁLEZ, aprueba el presente trabajo y solicita la autorización del mismo.*

**"ID Y ENSEÑAD A TODOS"**

  
Ing. Marlon Antonio Pérez Turk  
Director, Escuela de Ingeniería Ciencias y Sistemas



Guatemala, 14 de octubre 2009



Facultad de Ingeniería  
Decanato

Ref. DTG.381.09

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería en Ciencias y Sistemas, al trabajo de graduación titulado: **GESTIÓN DEL CONOCIMIENTO MEDIANTE EL USO DE ONTOLOGÍAS EN EL SERVICIO DE ATENCIÓN MÉDICA HOSPITALARIA DE GUATEMALA**, presentado por el estudiante universitario Raúl Estuardo Cuyún González, procede a la autorización para la impresión del mismo.

IMPRÍMASE.

Ing. Murphy Olimpo Paiz Recinos  
DECANO

Guatemala, octubre de 2009



/cc

---

## **ACTO QUE DEDICO A:**

"El Señor dirige los pasos del hombre y lo pone en el camino que a él le agrada" (Salmo 37:23), gracias Dios por guiarme en el sendero de tu luz.

Mis padres Raúl Cuyún y María Guadalupe González, por ser mi ejemplo, mi fuerza interna, gracias papi, mami por estar conmigo a lo largo del camino y brindarme amor incondicional.

Hermanas Ana Lucia, Julia María y Ana Fabiola, por ser esos rayitos de luz que me alegran cada mañana al despertar, gracias por hacer de mi vida, una vida llena de alegría.

Mi familia, a todas estas hermosas personas que han reservado un pedacito de su corazón para mi, gracias por querer a este niño que siempre he sido para ustedes.

Mis amigos, por ser parte de mi aprendizaje diario, gracias por dejarme ser un poquito de cada uno de ustedes, y a todas aquellas personas que sin ser mis amigos han dejado una huella imborrable en mi vida.

Mi amigo y asesor Javier Gramajo, por compartir su conocimiento y guiarme durante todo el trabajo de graduación.

# ÍNDICE GENERAL

<b>ÍNDICE DE ILUSTRACIONES</b>	<b>V</b>
<b>GLOSARIO</b>	<b>VII</b>
<b>RESUMEN</b>	<b>XIII</b>
<b>OBJETIVOS</b>	<b>XV</b>
<b>INTRODUCCIÓN</b>	<b>XVII</b>

<b>1. GESTIÓN DEL CONOCIMIENTO</b>	<b>1</b>
------------------------------------	----------

Gestión del conocimiento	1
Ontología	2
1.1. ¿Qué es una ontología?	2
1.2. Aspectos generales de una ontología	4
1.2.1. Componentes de una ontología	4
1.2.2. Ontologías existentes	6
1.2.3. Principios y criterios para construir ontologías	7
1.3. El uso de ontología en inteligencia artificial	8
1.4. Tipos de ontologías	11
1.4.1. Ontologías para representación del conocimiento	11
1.4.2. Ontologías generales	11
1.4.3. Meta - ontologías	11
1.4.4. Ontologías de dominio	12

1.4.5. Ontologías lingüísticas	12
1.4.6. Ontologías de tareas	12
1.4.7. Ontologías de tareas de dominios	13
1.4.8. Ontologías de métodos	13
1.4.9. Ontologías de aplicación	13
<b>2. LENGUAJES PARA ESTANDARIZAR EL CONOCIMIENTO</b>	<b>15</b>
2.1. RDF	15
2.2. RDF Schema	15
2.3. OIL	15
2.4. DAM + OIL	15
2.5. OWL	16
<b>3. HERRAMIENTAS PARA GESTIÓN DE ONTOLOGÍAS</b>	<b>17</b>
3.1. Ontolingua server	17
3.2. LinkFactory	18
3.3. OILed	18
3.4. Protege	19
3.5. OntoEdit	20
3.6. WebOde	20
3.7. Tabla comparativa de herramientas	22
3.8. API para acceder ontologías	23

<b>4. SERVICIO DE ATENCIÓN MÉDICA HOSPITALARIA EN GUATEMALA</b>	<b>27</b>
4.1. Actualidad en el servicio médico hospitalario de Guatemala	27
4.1.1. Tecnología utilizada en la actualidad	28
4.2. Clasificación de servicios hospitalarios que se prestan	29
4.3. Clasificación internacional de enfermedades (CIE10)	31
4.4. Las 10 primeras causas de morbilidad y mortalidad	36
<b>5. GESTIÓN DEL CONOCIMIENTO MEDIANTE EL USO DE ONTOLOGIAS EN EL SERVICIO DE ATENCIÓN MÉDICA HOSPITALARIA DE GUATEMALA</b>	<b>39</b>
5.1. Identificación de requerimientos para la clasificación de enfermedades CIE10	41
5.1.1. Descripción de los campos de capítulo	41
5.1.2. Descripción de los campos para categorías de tres caracteres	41
5.1.3. Descripción de los campos para categorías de cuatro caracteres	42
5.2. Identificación de requerimientos del protocolo para CIE10	42
5.3. Herramienta a utilizar para crear ontología	43
5.3.1. Protégé	43
5.4. Construcción de la ontología en protégé	44
5.4.1. Creación de clases para la CIE10	45

5.4.1.1. Clase capítulo	45
5.4.1.2. Clase categoría de tres caracteres	47
5.4.1.3. Clase categoría de cuatro caracteres	48
5.4.2. Creación de clases para protocolo CIE10	49
5.4.2.1. Síntoma	49
5.4.2.2. Sintoma_CIE10	50
5.4.2.3. Inclusión	50
5.4.2.4. Exclusión	51
5.4.2.5. Criterio de aceptación	52
<b>CONCLUSIONES</b>	<b>55</b>
<b>RECOMENDACIONES</b>	<b>57</b>
<b>REFERENCIAS ELECTRÓNICAS</b>	<b>59</b>
<b>BIBLIOGRAFÍA</b>	<b>61</b>
<b>APÉNDICE</b>	<b>65</b>

## ÍNDICE DE ILUSTRACIONES

### FIGURAS

1. Ejemplo de instrumentos musicales	20
2. Estructura de los servicios básicos médicos hospitalarios	44
3. Modelo mental de la estructura de la CIE10	45
4. Modelo mental de la solución básica propuesta	54
5. Sub clase CIE_capítulo	60
6. Sub clase CIE_3CHR	61
7. Sub clase CIE_4CHR	62
8. Sub clase síntoma	63
9. Sub clase síntoma_CIE10	64
10. Sub clase inclusión	65
11. Sub clase exclusión	66
12. Sub clase criterio_aceptación	67
13. Ontología de CIE10 y protocolo asociado propuesto.	68
14. Administrador del sistema OntoGuate - Salud.	78
15. Cargar ontología, Sistema OntoGuate - Salud.	79
16. Mostrar ontología, Sistema OntoGuate - Salud.	80
17. Borrar ontología, Sistema OntoGuate - Salud.	81
18. Búsqueda de capítulos, Sistema OntoGuate – Salud.	82
19. Búsqueda de tres caracteres, Sistema OntoGuate - Salud.	83
20. Búsqueda de cuatro caracteres, Sistema OntoGuate - Salud.	84

## TABLAS

I. Herramientas para ontologías	22
II. Las 10 principales casusas de la morbilidad	36
III. Las 10 principales causas de la mortalidad	37

## GLOSARIO

- Aplicación** En informática, una aplicación es un tipo de programa informático diseñado para facilitar al usuario la realización de un determinado tipo de trabajo. Esto lo diferencia principalmente de otros tipos de programas como los sistemas operativos (que hace funcionar al ordenador), las utilidades (que realiza tareas de mantenimiento o de uso general), y los lenguajes de programación (con el cual se crean los programas informáticos), que realizan tareas más avanzadas y no pertinentes al usuario común.
- API** Una API (del inglés Application Programming Interface - Interfaz de Programación de Aplicaciones) es el conjunto de funciones y procedimientos (o métodos si se refiere a programación orientada a objetos) que ofrece cierta biblioteca para ser utilizado por otro software como una capa de abstracción.
- Axioma** Es una proposición o evidencia que no es susceptible de demostración. Son verdades obvias a la razón humana que sirven de base a la ciencia. Los axiomas se refieren a los principios que se aceptan en todas las ciencias.

<b>Atributo</b>	Es una función de la sintaxis adicional. Es un sintagma que acompaña a los verbos copulativos y que se refiere al mismo tiempo al sujeto, con el que concuerda en género y número.
<b>Base de datos</b>	Una base de datos o banco de datos es un conjunto de datos pertenecientes a un mismo contexto y almacenados sistemáticamente para su posterior uso. En este sentido, una biblioteca puede considerarse una base de datos compuesta en su mayoría por documentos y textos impresos en papel e indexados para su consulta. En la actualidad, y debido al desarrollo tecnológico de campos como la informática y la electrónica, la mayoría de las bases de datos están en formato digital (electrónico), que ofrece un amplio rango de soluciones al problema de almacenar datos.
<b>Categoría</b>	Se denomina Categoría al proceso en el cual las ideas y los objetos son reconocidos, diferenciados y entendidos. Mediante las categorías, se pretende una ubicación jerárquica de elementos. Elementos muy parecidos y con características comunes formarán un grupo (categoría), y a su vez varias categorías con características afines formarán una categoría superior.
<b>Clase</b>	Estructura que define como son los objetos, indicando sus atributos y sus acciones.

<b>Conocimiento</b>	Es un conjunto de datos sobre hechos, verdades o de información ganada a través de la experiencia o del aprendizaje (a posteriori), o a través de introspección (a priori). El conocimiento es una apreciación de la posesión de múltiples datos interrelacionados que por si solos poseen menor valor cualitativo. Significa, en definitiva, la posesión consciente de un modelo de la realidad en el alma.
<b>Diagnóstico</b>	En medicina, el diagnóstico es el procedimiento por el cual se identifica una enfermedad, entidad nosológica, síndrome, o cualquier condición de salud-enfermedad. El diagnóstico médico establece a partir de síntomas, signos y los hallazgos de exploraciones complementarias, qué enfermedad padece una persona. Generalmente una enfermedad no está relacionada de una forma biunívoca con un síntoma, es decir, un síntoma no es exclusivo de una enfermedad. Cada síntoma o hallazgo en una exploración presenta una probabilidad de aparición en cada enfermedad.
<b>Embebido</b>	Que está contenido en uno.
<b>Estándar</b>	Especificación que regula la realización de ciertos procesos o la fabricación de componentes para garantizar la interoperabilidad.
<b>Inteligencia artificial</b>	Es aquella inteligencia exhibida por artefactos creados por humanos (es decir, artificial). A menudo se aplica

hipotéticamente a las computadoras. El nombre también se usa para referirse al campo de la investigación científica que intenta acercarse a la creación de tales sistemas.

**Internet** Red de redes a escala mundial de millones de computadoras interconectadas con el conjunto de protocolos TCP/IP. También se usa este nombre como sustantivo común y por tanto en minúsculas para designar a cualquier red de redes que use las mismas tecnologías que Internet, independientemente de su extensión o de que sea pública o privada.

**Jerarquía** Orden de los elementos de una serie según su de valor. Principio de orden que dispone, con un rango, a unos elementos de un conjunto por encima de otros.

**Lenguaje** Conjunto de símbolos que en conjunto dejan transmitir un mensaje, y es una capacidad exclusiva del ser humano (los animales tienen sistemas de comunicación) que lo capacita para abstraer, conceptualizar y comunicarse.

**Metodología** Se refiere a los métodos de investigación en una ciencia.

**Modelo mental** Un modelo mental es un mecanismo del pensamiento mediante el cual un ser humano u otro animal, intenta explicar la manera de cómo funciona el mundo real. Es un tipo de símbolo interno o representación de la realidad externa, hipotética que juega un papel importante en la cognición.

<b>Parser</b>	Un parser es un programa o la parte de un programa que analiza un string alfanumérico para poder reconocer unidades gramaticales.
<b>Protocolo</b>	Para este trabajo llamaremos protocolo a las reglas y actividades relacionadas a la búsqueda de un diagnóstico certero.
<b>Protégé</b>	Es un editor libre de código abierto y un sistema de adquisición de conocimiento.
<b>Síntoma</b>	En medicina, la referencia subjetiva que da un enfermo por la percepción o cambio que puede reconocer como anómalo o causado por un estado patológico o enfermedad. El síntoma es un aviso útil de que la salud puede estar amenazada sea por algo psíquico, físico, social o combinación de las mismas.
<b>Subclase</b>	Estructura que hereda todas las variables miembro de su superclase que puedan ser accesibles desde la subclase.
<b>WEB</b>	World Wide Web (del inglés, Telaraña Mundial), la Web o WWW, es un sistema de hipertexto que funciona sobre Internet.
<b>W3C</b>	El World Wide Web Consortium, abreviado W3C, es un consorcio internacional que produce estándares para la World Wide Web. Está dirigida por Tim Berners-Lee, el

creador original de URL (Uniform Resource Locator, Localizador Uniforme de Recursos), HTTP (HyperText Transfer Protocol, Protocolo de Transferencia de HiperTexto) y HTML que son las principales tecnologías sobre las que se basa la Web.

## RESUMEN

El conocimiento es inherente a las personas, y es el resultado de procesar la información que recibe, es por eso que la Gestión del Conocimiento busca gestionar espacios para crear el conocimiento de los individuos y convertirlo en conocimiento social. Hoy en día éste concepto se ha convertido en un medio fundamental para la integración de lo que hasta ahora se encontraba en procesos aislados, tal es el caso de la información, capital humano y la tecnología.

Desde hace algún tiempo la palabra Ontología se ha convertido en un término relevante y popular en la comunidad científica relacionada al estudio del conocimiento y otras áreas relacionadas. Es por ello que este trabajo de graduación titulado “Gestión del Conocimiento mediante el uso de Ontologías en el servicio de atención médica hospitalaria de Guatemala” pretende brindar un aporte a la sociedad médica guatemalteca, específicamente al personal asociado con la realización de diagnósticos y estadísticas, mediante la construcción e implementación de una Ontología basada en la Clasificación Internacional de Enfermedades en su décima versión (CIE 10); propuesta en Ginebra en el año 1992 por la Organización Mundial de la Salud, ésta a su vez será cargada a una base datos en MySQL mediante RAP y una herramienta programada en un lenguaje PHP. Esta Ontología pretende ser el comienzo para la construcción de sistemas orientados al área médica, utilizando

tecnologías innovadoras para el almacenamiento de conocimiento como ya lo mencionamos, específicamente para la Clasificación Internacional de Enfermedades en su decima versión (CIE10). Y por último, éste trabajo concluye con la construcción de un software sencillo que permite visualizar el alcance de la propuesta planteada acerca de las búsquedas inteligentes en el contexto de la CIE10

## OBJETIVOS

1. Realizar la comparativa de las herramientas mas populares para la gestión de ontologías dentro del capítulo tres, de ésta manera poder elegir una herramienta que mas se adecue a las necesidades para la creación de la ontología de la CIE10 propuesta.
2. Establecer los requerimientos dentro del contexto de la CIE10 para la creación de la ontología propuesta en el quinto capítulo de este trabajo de graduación.
3. Mediante un método de programación tradicional se pretende cargar a la ontología CIE10 creada todos los diagnósticos asociados a la misma, a través de la escritura de las instancias con su estructura al archivo plano rdf.
4. Realización de un protocolo asociado a la CIE10 para que esta pueda llevar a cabo de una mejor manera los procesos de elección del diagnóstico mediante la especificación de conceptos requeridos para dicho propósito en el contexto del quinto capítulo.
5. Construcción de un módulo de consulta, para búsquedas inteligentes de diagnósticos dentro de la CIE10, programado en lenguaje PHP y

razonado por RAP (RDF Api for PHP) para personas relacionadas con el área medica interesadas en utilizarlo.

## INTRODUCCIÓN

Actualmente Internet es un lugar diseñado para el intercambio de información, el cual a su vez es amplio y complejo debido a su crecimiento rápido en volumen, siendo este de grandes magnitudes por lo que no puede ser clasificado de una manera inteligente para búsquedas más precisas; es por ello que es importante conocer las llamadas Ontologías. Desde hace algún tiempo la palabra Ontología se ha convertido en un término relevante y popular en la comunidad científica relacionada al estudio del conocimiento (en inglés Knowledge Engineering) y otras áreas relacionadas. Uno de los objetivos de una Ontología como lo veremos en las líneas posteriores, básicamente es la de brindar ayuda en cuanto a clasificación de dominios se refiere, convirtiendo el conocimiento tácito a explícito. También es importante conocer que gracias a la utilización de Ontologías se comparte y reutiliza mas eficaz y eficientemente la información almacenada.

En los primeros capítulos de este trabajo se describirán mediante un estado del arte muchos aspectos puramente teóricos, tales como los distintos lenguajes de estandarización de ontologías, así como las herramientas utilizadas para éste fin, ventajas y desventajas de cada una de ellas mediante algunas métricas comparativas. Para finalizar dentro del quinto capítulo

abordaremos la construcción de la ontología como tal, dentro del marco contextual médico.

Es por ello que este trabajo de graduación titulado “Gestión del Conocimiento mediante el uso de Ontologías en el servicio de atención medica hospitalaria de Guatemala” pretende brindar un aporte a la sociedad médica guatemalteca, específicamente al personal asociado con la realización de diagnósticos y estadísticas mediante la construcción e implementación de una Ontología basada en la Clasificación Internacional de Enfermedades (CIE 10) proporcionada por la Organización Mundial de la Salud en su sitio de internet, ésta a su vez será cargada a una base datos en MySQL mediante RAP y una herramienta programada en un lenguaje PHP.

Esta Ontología pretende ser el comienzo para la construcción de sistemas orientados al área médica, utilizando tecnologías innovadoras para el almacenamiento de conocimiento como ya lo mencionamos específicamente para la Clasificación Internacional de Enfermedades en su decima versión (CIE10). Al finalizar el trabajo se podrá comprobar que una ontología médica como la propuesta ayudará a tener mejores resultados dentro de un marco estadístico de clasificación de pacientes por diagnóstico.



# 1. GESTIÓN DEL CONOCIMIENTO

Para abordar el primer capítulo de éste trabajo es importante hablar antes de un término denominado “Gestión del Conocimiento”, éste termino es la razón de ser de muchas de las ideas aportadas y desarrolladas en este trabajo, las cuales se tratarán en su momento.

Para Javier Gramajo López, la gestión del conocimiento constituye un enfoque gerencial que ha tomado fuerza en las últimas décadas, su tratamiento teórico ha sido amplio pero en menor medida se ha implementado en la práctica. Pocos son los que han corrido el riesgo de incorporarse a una nueva forma de integrar lo que hasta ahora se encontraba en procesos aislados; la información, el capital humano y la tecnología. En las últimas dos décadas se han desarrollado muchas actividades alrededor de lo que se ha dado en llamar "Gestión del Conocimiento". A partir de éste proceso se ha iniciado una nueva teoría de la empresa basada en sus recursos, como alternativa a los enfoques tradicionales, que basan la empresa en sus productos y servicios.<sup>1</sup>

Para Kemly Camacho de la organización Sulá Batsú, el conocimiento no se gestiona, lo que se gestiona son los espacios para crear conocimiento; por lo que gestión del conocimiento es la disciplina de gestión que se responsabiliza de crear las condiciones adecuadas para la creación colectiva de nuevos conocimientos.<sup>2</sup>

---

<sup>1</sup> <http://cmc.ihmc.us/papers/cmc2004-016.pdf>, Agosto 2008.

<sup>2</sup> <http://www.slideshare.net/em3marquez/gestion-del-conocimiento-presentation-852277> Enero 2009.

Dejando en claro ésta idea es posible ahora comprender la razón que ha motivado a realizar este trabajo; trabajo que permite gestionar un espacio para la construcción del conocimiento en el área de salud dentro del contexto de la Clasificación Internacional de Enfermedades CIE10 de manera que pueda ser utilizado por personas asociadas al área médica.

## **Ontología**

Hoy en día las ventajas de alcance e información global que ofrece Internet, hace que exista cada vez más información a nuestra disposición, sin que esta a su vez pueda ser clasificada de una manera inteligente para búsquedas más precisas, por ello es importante tener toda esta información clasificada a manera de generar un conocimiento dentro de la Web, mejor aún, dentro de nuestros sistemas computacionales en un dominio específico y que ésta pueda ser reutilizada de una manera fácil y sencilla por otros sistemas. Se pretende básicamente tener información semántica rica en contenido, es decir clasificada respecto de un dominio específico, lo cual puede lograrse a través del uso de Ontologías. En las líneas que se presentan a continuación se detallaran algunas generalidades dentro del marco contextual de lo que estas representan.

### **1.1 ¿Qué es una ontología?**

Para conocer acerca de éste concepto se darán varias definiciones que han sido propuestas por varios investigadores.

- Según Gruber, define que ontología es una especificación de una conceptualización, dice que es una descripción de conceptos y relaciones que pueden existir para un agente a una comunidad de

agentes<sup>3</sup>. Un cuerpo de conocimiento presentado formalmente está basado en una conceptualización: los objetos, los conceptos y otras entidades en un área de interés determinada y las relaciones que la contienen. La conceptualización es una forma abstracta y simplificada de ver el mundo que se quiere representar con algún propósito.

- Para Macelo Tallarico, Ontología es una descripción formal y explícita de conceptos de un dominio, propiedades de cada concepto que describen varias características y atributos del concepto (roles o propiedades) y restricciones (facetas o restricciones de rol.)<sup>4</sup>

Para fines de estudio y poder brindar una definición formal a las Ontologías nos dirigimos a la proporcionada por la W3C [3] la cual dice:

“Una ontología define los términos a utilizar para describir y representar un área de conocimiento. Las ontologías son utilizadas por las personas, las bases de datos, y las aplicaciones que necesitan compartir un dominio de información (un dominio es simplemente un área de temática específica o un área de conocimiento, tales como medicina, fabricación de herramientas, bienes inmuebles, reparación automovilística, gestión financiera, etc.). Las ontologías incluyen definiciones de conceptos básicos del dominio, y las relaciones entre ellos, que son útiles para los ordenadores. Codifican el conocimiento de

---

<sup>3</sup> <http://www.di.uniovi.es/~dani/research/trabajo-investigacion.pdf>, Agosto 2008

<sup>4</sup> <http://usuarios.fceia.unr.edu.ar/~tallaric/tesina/presentacion.pdf> Agosto 2008

un dominio y también el conocimiento que extiende los dominios. En este sentido, hacen el conocimiento reutilizable”<sup>5</sup>.

Cabe mencionar que en este trabajo se realizará basado en ontologías formales, es decir son todas aquellas que pueden ser interpretadas por las maquinas. Solamente se hará referencia siempre a semánticas que describen el significado del conocimiento de una forma precisa, dónde la palabra *precisa* significa que la semántica no se refiere a opiniones ni intuiciones, y que las máquinas y las personas deben interpretarla de una misma forma. Esto quiere decir que las ontologías deberán ser clasificadas mediante argumentos válidos para el dominio y no al azar.

## 1.2 Aspectos generales de una Ontología

### 1.2.1 Componentes de una Ontología

Una ontología puede estar compuesta por los siguientes tipos de componentes:

- **Conceptos (Clase):** son las ideas básicas que se intentan formalizar. Los conceptos pueden ser clases de objetos, métodos, planes, estrategias, procesos de razonamiento, etc.

---

<sup>5</sup> <http://www.alonsoperez.com/diicc2006.pdf> Noviembre 2008

- **Sub clase:** es en si misma una clase, pero hija de alguna otra clase, es decir representan conceptos mas específicos<sup>6</sup>.
- **Clase jerárquica:** es una colección de clases conectadas por relaciones.
- **Atributos:** son propiedades de cada concepto que describen varias características y atributos del concepto.
- **Axiomas:** son teoremas sobre relaciones que deben cumplir los elementos de la ontología<sup>7</sup>.
- **Valor:** Los axiomas o valores describen atributos que se aplican a alguna clase o caso.<sup>8</sup>
- **Valor por defecto:** Es posible declarar un valor por defecto, es decir son valores que se instancias desde su creación.
- **Tipo:** con este se define el tipo de valor que se ingresará en un atributo específico por ejemplo: entero, carácter, flotante.
- **Relaciones:** representan la interacción y enlace entre los conceptos del dominio. Ejemplos de relaciones pueden ser subclase-de, conectado-a, parte-de, etc.<sup>9</sup>

---

<sup>6</sup> <http://www.hipertexto.info/documentos/ontologias.htm> Agosto 2008.

<sup>7</sup> [http://bvs.sld.cu/revistas/aci/vol13\\_6\\_05/aci030605.htm](http://bvs.sld.cu/revistas/aci/vol13_6_05/aci030605.htm) Noviembre 2008.

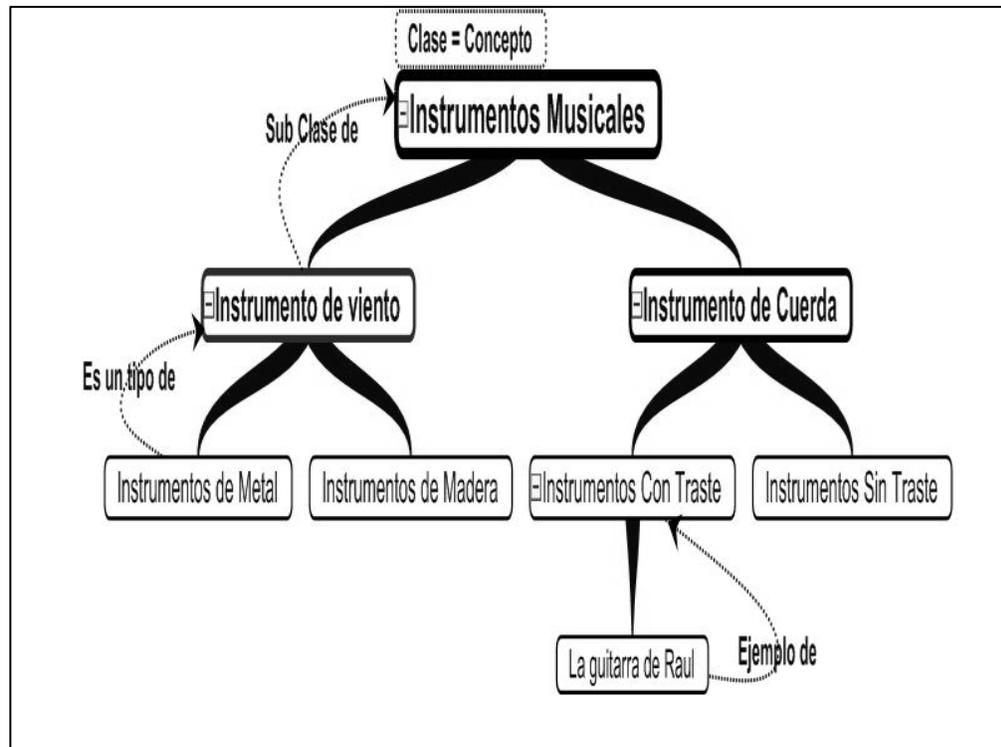
<sup>8</sup> <http://www.iaa.upf.es/~jblat/material/doctorat/students/jccbis/Ontologias2.htm> Noviembre 2008.

<sup>9</sup> <http://www.informandote.com/jornadasIngWEB/articulos/jiw02.pdf> Noviembre 2008.

- **Funciones:** son un tipo especial de relación donde se identifica un elemento mediante el cálculo de una función.
- **Instancias:** son usadas para representar objetos determinados de un concepto.

Para poder ejemplificar algunos de los conceptos antes mencionados veremos el siguiente caso como lo muestra la figura 1:

**Figura 1. Ejemplo de Instrumentos Musicales**



Referencia: figura realizada en Mind Genius 2005 por Raúl Estuardo Cuyún González, Noviembre 2008.

### **1.2.2 Ontologías existentes.**

Hoy en día existe ya un gran número de ontologías disponibles en la Web, prueba de ello es la biblioteca de ontologías DAML, la cual contiene unos 250 ejemplos hechos con OWL o DAML+OIL. Además, varias ontologías extensas han sido publicadas en OWL entre ellas se encuentran una ontología del cáncer en OWL, desarrollada por el Centro de Bioinformática del Instituto Nacional del Cáncer de los EEUU, y que contiene cerca de 17,000 términos relacionados con el cáncer y sus definiciones. Actualmente en Guatemala se han desarrollado varias ontologías muy interesantes, tanto en lenguajes owl como rdf, entre las que cabe mencionar trabajos relacionados con el desarrollo de Pymes, ontologías relacionadas con el sector de turismo guatemalteco entre otras.

### **1.2.3 Principios y criterios para construir Ontologías**

En esta parte del trabajo se exponen una serie de principios y criterios de diseño que suelen ser útiles en el desarrollo de ontologías [1]. Es importante tener muy en cuenta que la construcción de ontologías no responde a una única aproximación lógica sino que depende, en gran manera, del contexto en el que se construyen, debido a que siempre existirán muchos puntos de vista asociados a la ontología objeto de estudio. Para este trabajo se ha logrado identificar un conjunto de principios que son básicos, los cuales se deberán de ser la base en la construcción de una Ontología.

Estos principios básicos son:

- La representación de objetos del mundo real siempre depende del contexto en el que los objetos son usados, así como también de las necesidades que se tengan. Este contexto puede verse como un punto de vista tomado del objeto. Es generalmente, imposible enumerar de antemano todos los posibles puntos de vista útiles de objetos.<sup>10</sup>
- La reutilización de alguna parte de conocimiento requiere una descripción explícita de los puntos de vista que están inherentemente presentes en el conocimiento, de esta manera estaremos completamente seguros que esto es lo que realmente se necesita. De acuerdo a estos principios, una ontología especifica una conceptualización, una forma de ver el mundo.<sup>11</sup>

Criterios para la construcción de una Ontología.

- Claridad y objetividad: la ontología debe proveer el significado de los términos definidos proveyendo definiciones objetivas y documentación en lenguaje natural sencillo.
- Completitud: una definición expresada en términos de condición necesaria y suficiente es preferible sobre una definición parcial.

---

<sup>10</sup> <http://www.mor.itesm.mx/~rdec/node196.html> Noviembre 2008.

<sup>11</sup> <http://usuarios.fceia.unr.edu.ar/~tallaric/tesina/presentacion.pdf> Noviembre 2008.

- Coherencia: permite hacer inferencias válidas consistentes con las definiciones.
- Máxima extensibilidad monofónica: los nuevos términos (generales o especializados) debieran ser incluidos en la ontología de manera que no requiera la revisión de definiciones existentes.

### **1.3 El uso de Ontologías en Inteligencia Artificial**

En su trabajo realizado Miguel Ángel Abián [6] señala mecanismos para mejorar la búsqueda de información en la Web y en las intranets de las organizaciones, así como para navegar por ellas. Si se definieran una o más ontologías para cada dominio, los contenidos Web podrían describirse en función de los términos ontológicos. De este modo, se podrían expandir las búsquedas mediante términos de las categorías más específicas de la ontología.

Asimismo, podrían usarse axiomas que definieran los términos en función de otros términos, por ejemplo, *persona* sería lo mismo que ser *humano*, lo que solucionaría el problema de los sinónimos. La facilidad con que las ontologías se integran en sistemas de razonamiento automático permitirá hacer búsquedas más rápidas y eficientes.

También las ontologías pueden usarse para comprobar la validez de los datos. Para una ontología dentro de cualquier contexto válido podría usarse para comprobar si ciertas afirmaciones son válidas o no dentro de ella. En general las ontologías se usarán para programar

agentes inteligentes, que entenderán e integrarán las informaciones procedentes de distintas fuentes en distintos lugares. En el futuro, los servicios Web se describirán mediante ontologías. Los agentes las usarán para buscar los servicios Web que les interesen y utilizarlos automáticamente, sin intervención humana.

Además, las ontologías dotan de información semántica comprensible para las máquinas y, por tanto, permiten la automatización de muchos procesos. Así, los agentes que usen ontologías podrán buscar productos, negociar compras, localizar servicios de interés para los usuarios. En el caso de que varias empresas usen una misma ontología, sus sistemas de información podrán procesar automáticamente los datos que intercambien, pues compartirán una misma semántica. Incluso en el caso de empresas que usen distintas ontologías, existen herramientas que permiten rápidamente traducir los términos de una ontología a otra o mezclar ontologías, incluso de manera automática o semiautomática. En suma, el uso de ontologías en las empresas permitirá hacer negocios electrónicos con otras empresas, de modo automático y sin que sea necesario ponerse de acuerdo sobre la semántica de los datos. Dicho con otras palabras: las ontologías y las tecnologías que las usan permitirán, por un lado, que los compradores localicen productos y servicios con características específicas y, por otro, que los vendedores localicen a posibles compradores que correspondan a ciertos perfiles.

Para la W3C [3] el uso de ontologías puede verse clasificada de la siguiente manera.

- Portales Web

- Reglas de categorización utilizadas para mejorar la búsqueda.
- Colecciones Multimedia
  - Búsquedas basadas en contenido para medios no textuales.
- Administración de Sitios Web Corporativos
  - Organización taxonómica automatizada de datos y documentos.
  - Asignación entre Sectores Corporativos.
- Documentación de Diseño
  - Explicación de partes "derivadas".
  - Administración explícita de Restricciones.
- Agentes Inteligentes
  - Expresión de las Preferencias y/o Intereses de los usuarios.
  - Mapeo de contenidos entre sitios Web.
- Servicios Web y Computación Ubicua.
  - Composición y descubrimiento de Servicios Web.
  - Administración de Derechos y Control de Acceso.

#### **1.4 Tipos de ontologías**

En función del objetivo y dominio de las ontologías podemos tipificarlas de la siguiente manera:

**1.4.1 Ontologías para representación de conocimiento:** capturan las primitivas de representación usadas para formalizar conocimiento en paradigmas de representación de conocimiento.

Ejemplo de este tipo es Frame-Ontology, que captura la representación de primitivas usadas en lenguajes basados en marcos (*frame-based*).

**1.4.2 Ontologías generales (o comunes):** incluyen vocabulario relacionado a cosas, eventos, tiempo, espacio, comportamiento, etc.

Ejemplo de este tipo es CYC que posee un monto considerable de conocimiento humano fundamental.

**1.4.3 Meta-ontologías:** también llamadas Generic Ontologies o Core Ontologies son reusables a través de distintos dominios.

Ejemplo de este tipo es Mereology. Se define la relación *parte-de*, estudia sus propiedades, y permite expresar, por ejemplo, que un dispositivo está ensamblado por componentes, cada uno de los cuales puede ser descompuesto en subcomponentes.

**1.4.4 Ontologías de dominio:** son reusables en un dominio dado. Proveen vocabularios sobre los conceptos de un dominio y sus relaciones, sobre las actividades que toman lugar en el dominio y sobre las teorías y principios elementales que gobiernan tal dominio.

Ejemplos de este tipo son Eng-Math y PhysSys. Se formalizan modelos matemáticos usados en ingeniería. PhysSys es una librería de ontologías y describe el dominio de conocimiento requerido para hacer modelos de simulaciones de dispositivos, como sistemas de calefacción, sistemas automotores y máquinas.

**1.4.5 Ontologías lingüísticas:** tienen el objetivo de modelar el lenguaje natural.

Ejemplo de este tipo de ontología es Sensus, que provee una amplia estructura conceptual para trabajar en traducciones. Se desarrolló para mezclar y extraer información de recursos electrónicos existentes.

**1.4.6 Ontologías de tareas:** proveen vocabulario sistemático de términos usados para resolver problemas asociados con tareas que pueden o no pertenecer al mismo dominio. Incluyen nombres genéricos, verbos genéricos y adjetivos genéricos.

**1.4.7 Ontologías de tareas de dominios:** son ontologías reusables en un dominio dado pero no a través de otros dominios.

**1.4.8 Ontologías de métodos:** proveen definiciones de conceptos relevantes y relaciones usadas para especificar un proceso de razonamiento para lograr una tarea particular.

**1.4.9 Ontologías de aplicación:** contienen el conocimiento necesario para modelar una aplicación particular.



## 2. LENGUAJES PARA ESTANDARIZAR EL CONOCIMIENTO

Una vez que los componentes de la ontología están definidos, la ontología puede ser representada en varios lenguajes. Los mismos comenzaron a surgir a comienzos de 1990 y se basan principalmente en lógicas de primer orden, en marcos combinados con lógicas de primer orden y lógicas descriptivas. A continuación se listan algunos de los más conocidos:

**2.1 RDF (Resource Description Framework):** fue desarrollado con base a lenguajes sobre redes semánticas con el fin de describir recursos de la web. Podemos decir también que RDF es un vocabulario basado en XML y que como tal representa ésta sintaxis.

**2.2 RDF Schema:** se desarrolló como una extensión de RDF con primitivas basadas en marcos. La combinación de ambas suele denotarse como RDF. No es muy expresivo, sólo permite la representación de conceptos, taxonomías de conceptos y relaciones binarias. Algunos motores de inferencia fueron creados para este lenguaje, en especial para verificación de restricciones.

**2.3 OIL:** agrega al RDF(S) primitivas basadas en marcos y su semántica formal está basada en lógicas descriptivas. Se provee una clasificación automática de conceptos a través de FaCT.

**2.4 DAML + OIL:** agrega primitivas de representación de conocimiento basado en lógica descriptiva a RDF(S). Permite representar conceptos,

taxonomías, relaciones binarias, funciones e instancias. Se está invirtiendo un considerable esfuerzo en proveer mecanismos de razonamiento para este lenguaje.

**2.5 OWL (Ontology Web Language):** recientemente surgido. Está basado en las principales características de DAML + OIL.

Para el presente trabajo será importante poseer una noción básica al implementar una ontología las necesidades en términos de expresividad y servicios de inferencia, es decir el nivel semántico al que llevaremos la ontología, porque no todos los lenguajes permiten representar los mismos componentes de la misma forma. La representación y razonamiento con información básica, como conceptos, taxonomías, y relaciones binarias, no siempre es suficiente si se quiere crear una ontología pesada y hacer razonamientos complejos. Es frecuente que las traducciones entre lenguajes no sea lo suficientemente precisa por lo que puede perderse información en el proceso de traducción. Por eso, la decisión del uso de un lenguaje específico para representar una ontología, juega un papel importante.

### 3. HERRAMIENTAS PARA LA GESTIÓN DE ONTOLOGÍAS

En los últimos años, el número de entornos de desarrollo para ontologías y herramientas relacionadas ha crecido considerablemente. Las herramientas ayudan a proveer soporte para el proceso de desarrollo y uso de las ontologías.

A continuación se presentan las características de algunas de las herramientas más conocidas.

**3.1 Ontolingua Server:** esta fue la primera herramienta desarrollada. Surgió a principios de 1990 en el Knowledge Systems Laboratory (KSL) de la Universidad de Standford, su última versión data del 12 de octubre del 2002. Es una aplicación basada en formularios web para que distintos editores remotos puedan ver y editar ontologías.

También esta herramienta ofrece un entorno de colaboración distribuido para crear, editar, modificar y usar ontologías, basado en el principio de la jerarquía múltiple. La creación de la ontología se facilita mediante la posibilidad de incluir partes de otras ya existentes en un *registro de ontologías* y la posibilidad de incluir términos básicos incluidos en una *frameontology*.

La licencia de este herramienta es gratuita, solo bastará con registrarse en el sitio de descarga. Ontolingua es una solución flexible y escalable para la gestión de ontologías. Se está utilizando para modelar

distintos tipos de términos como la gestión de documentos, vocabulario físico, químico, matemático.

**3.2 LinkFactory:** es un sistema de gestión de ontologías que se utiliza para construir y gestionar el conocimiento lingüístico médico LinkBase (representación formal del mundo de la medicina). LinkFactory es una herramienta que puede ser utilizada para construir grandes y complejas ontologías independientes del lenguaje formal. No se limita a pequeñas ontologías como la mayoría de editores de ontologías. Cabe mencionar que la licencia no es gratuita, esta herramienta flexible y escalable, la única deficiencia es que no tiene gestión de Backup. El acceso a dicha herramienta es vía Web a través de un navegador.

**3.3 OilEd:** es un editor gráfico de ontologías que permite al usuario construir ontologías usando DAML+OIL. Está basado en Protégé solo que no proporciona un entorno de desarrollo de ontologías completo. Es compatible con RDF y RDFs. El modelo del conocimiento es el mismo que el usado por DAML+OIL, aunque mejorado por el uso de un marco para representar el modelado. Las clases están definidas en función de sus súper clases y las restricciones de sus propiedades, con axiomas adicionales captando más relaciones tales como las desuniones. Una de las tareas principales de OilEd está centrada en editar ontologías o schemas, en oposición a la adquisición de conocimiento o la construcción de grandes bases de conocimientos formados por instancias. Un aspecto clave del comportamiento de OilEd es el uso del “razonador” llamado FaCT para clasificar ontologías y chequear la consistencia a través de una traducción desde DAML+OIL hasta la descripción lógica de SHIQ. Esto permite al usuario describir las clases

de su ontología y que su razonador determine el sitio apropiado en la jerarquía.

Tiene varias desventajas, entre las que se encuentran las más importantes: no tiene gestión de backup. No tiene extensibilidad, es decir no es flexible ni escalable.

En los últimos años se desarrollaron entornos de desarrollo un tanto más ambiciosos, con el objetivo de integrar ontologías a sistemas de información existentes. Por lo general son extensibles, basados en arquitecturas de componentes y lo más importante: por lo general son independientes del lenguaje de implementación de la ontología. Entre los más populares pueden citarse los siguientes:

**3.4 Protégé:** esta herramienta fue desarrollada por el Stanford Medical Informatics (SMI) de la Universidad de Stanford. Es un proyecto Java que provee una arquitectura extensible para la creación de herramientas de bases de conocimiento. Es una herramienta porque permite al usuario construir el dominio de una ontología, configurar formularios para reemplazo de datos e ingresar el dominio de conocimiento. Es una plataforma que puede ser extendida con gráficos, diagramas, componentes animados para acceder a aplicaciones embebidas en sistemas de bases de conocimientos. Es también una librería que otras aplicaciones pueden acceder.

Las aplicaciones desarrolladas con Protégé-2000 se usan para la resolución de problemas y la toma de decisiones en un dominio determinado. El usuario puede realizar las siguientes operaciones:

- Creación de una ontología
- Optimización de datos de entrada de formularios
- Inserción de datos o instancias.

En cuanto al licenciamiento se puede mencionar que el código fuente es gratis bajo la licencia de código abierto. La última versión estable es la 3.3.1.

**3.5 OntoEdit:** es una herramienta de edición de ontologías que apoya el desarrollo y mantenimiento de las mismas utilizando medios gráficos en un entorno web. Permite la representación semántica de lenguajes conceptuales y estructuras mediante conceptos, jerarquías de conceptos, relaciones y axiomas. La herramienta ha sido desarrollada por la empresa Ontoprise GmbH. Ontoprise es una empresa fundada en 1999 como una spin off de la Universidad de Karlsruhe (Alemania). La mayoría de los productos comercializados por Ontoprise han sido creados en modo experimental por el AIFB. Estos productos son transferidos a la empresa para que desarrolle una versión comercial de los mismos.

En cuanto a la licencia ésta se vende aunque existe una versión limitada al manejo de 50 conceptos, 50 relaciones y 50 instancias que es gratuita. Posee extensibilidad a través de Plug – ins, cabe mencionar que OntoEdit no proporciona soporte directo para acceder a la base de conocimiento a través de Internet.

**3.6 WebOde:** WebODE es una herramienta para modelar el conocimiento usando ontologías. Está basado en el ampliamente

probado y testeado Methontology de la Escuela Técnica de Ciencias Informáticas de Madrid. Es la parte Web de ODE (Ontology Design Environment) e incluye muchas características nuevas e interesantes que la convierten en una herramienta simple y a la vez potente para la creación de ontologías. Las ontologías disponen de los siguientes componentes principales: conceptos, grupos de conceptos, relaciones predefinidas, taxonomías (herencias simples y múltiples), relaciones ad-hoc, constantes, fórmulas, ejemplos (de conceptos y relaciones), propiedades, términos importados y referencias. Otras características de interés son:

- Colección de ejemplos, que permite vislumbrar el mismo modelo conceptual en distintos escenarios.
- Distintas vistas del mismo modelo conceptual, que permiten crear y almacenar distintas partes de la ontología, personalizando la visualización de la misma según los requerimientos de uso.
- Posibilidad de añadir a las *relaciones ad-hoc*, propiedades matemáticas como: reflexión, simetría, etc. y otras propiedades definidas por el usuario.

El editor de ontologías de la plataforma disponible es una aplicación Web, la cual es de libre acceso, basta registrarse. Si posteriormente se desea integrar la plataforma en algún otro sistema, se realizaría un estudio de los módulos y servicios que serían necesarios y las licencias se harían en función a las necesidades, por lo que dichas licencias tendrían un costo individual.

### 3.7 Tabla comparativa, herramientas de creación de Ontologías

Para este trabajo se hace necesario realizar una tabla comparativa entre las herramientas de creación de Ontologías, con base a la descripción básica que se realizó de cada una de ellas. Cabe mencionar que dicha tabla comparativa pretende ser de ayuda en la toma de decisión para elegir la herramienta a utilizar para la creación de nuestra ontología, que como ya lo mencionamos se encuentra en el contexto del perfil del paciente.

**Tabla I. Herramientas para Ontologías.**

Nombre	Licencia	Arquitectura Software	Extensibilidad	Gestión de Backup	Acceso	Almacenamiento de la Ontología
Ontolingua Server	Uso libre tras registrarse	Aplicación Cliente/servidor	Flexible y escalable	No	No proporciona acceso vía Web, se necesita una aplicación para acceder al conocimiento	En el Servidor
Link Factory	No es gratuita	Aplicación Cliente/servidor, 100% Java	Flexible, escalable y probado en dominios Médicos y otros	No	Vía Web a través de Navegador	En BBDD compatible con Java
OilEd	Open Source bajo GPL	Standalone	No	No	RDF URI's, nombres de dominio	En archivos de extensión *.daml

Continúa

					limitados	
Protegee	Es Gratis, Bajo licencia MPL	Standalone	A través del uso de Plug-ins	No	A través de Servlets o Applets	En Archivos de texto, RDFS, o cualquier BBDD compatible con JDBC
OntoEdit	Bajo licencia, Versión express limitada a 50 conceptos, relaciones e instancias	Standalone	A través del uso de Plug-ins	No	Para acceder al conocimiento se necesita una aplicación externa	OXML
WebOde	Uso libre tras registrarse	Tecnologías usadas: Java, Corba, XML, Minerva	Escalable, capaz de editar ontologías de diferentes dominios	Si	Vía Web a través de Navegador	En BBDD compatible con Java

Referencia: herramientas para la construcción de ontologías, Noviembre 2008.

### 3.8 API para acceder ontologías

Cuando ya se tiene una herramienta para la creación de la Ontología se hace necesario poder tener acceso a esa información. Para ello necesitamos tener alguna herramienta que nos permita parsear los archivos generados. En este trabajo se mencionarán tres de esas herramientas: Jena, JDom y RAP.

#### JENA

Jena es una API que permite crear y manipular modelos RDF desde una aplicación Java. Adicionalmente proporciona clases java para representar:

- Modelos.
- Recursos.
- Propiedades.
- Literales.
- Statements.

Jena por medio de sus clases permite consultar y navegar por la Ontología creada, si desea mayor información de ésta API la podrá encontrar en <http://jena.sourceforge.net>.

## **JDOM**

Para [8] JDOM son las siglas en inglés de Java Document Object Model (Documento de Modelado de Objetos en Java) es una librería de código fuente para manipulaciones de datos XML optimizados para Java. JDOM se creó específicamente para usarse con Java y por lo tanto beneficiarse de las características de Java, incluyendo sobrecarga de métodos, colecciones, etc. Para los programadores de Java, JDOM es una extensión más natural y correcta. Se asemeja a como la librería RMI optimizada para Java (invocación remota de métodos) se amolda mejor que el CORBA (arquitectura de intermediario solicitador de objetos comunes) que es más neutral respecto a los lenguajes. <sup>12</sup>

---

<sup>12</sup> <http://es.wikipedia.org/wiki/JDOM> Enero 2009.

Si desea mayor información acerca de esta API lo podrá hacer en <http://www.idom.org>.

### **RAP (RDF Api for PHP)**

Es un paquete que brinda mecanismos para parsear, consultar, manipular, serializar y diseñar modelos RDF, éste fue diseñado para programar soluciones en PHP. RAP tiene sus inicios mediante un proyecto de código abierto en la universidad "Freie Universität Berlin" en el año de 2002. Para poder obtener más información acerca de esta API puede acceder a la web a la dirección <http://www4.wiwiwiss.fu-berlin.de/bizer/rdfapi/>.

### **RACER (Robust Server for Scalable Ontology Reasoning)**

Racer es un miembro más de la familia de razonadores, proporciona una inferencia altamente optimizada para aplicaciones relacionadas con ontologías. Además proporciona un lenguaje de consulta para OWL más sofisticado llamado nRQL, mucho más que el lenguaje conocido SPARQL. Uno de los inconvenientes de éste razonador es la necesidad de comprar la licencia debido a que no es libre, pero puede solicitarse en su versión académica para poder probarlo, éste a su vez como todos sabemos son versiones que no son completas.

Si se desea mayor información acerca de esta razonador puede acceder a la pagina oficial de la herramienta a la siguiente dirección: <http://www.racer-systems.com/>



## **4. SERVICIO DE ATENCIÓN MÉDICA HOSPITALARIA EN GUATEMALA**

### **4.1 Actualidad en el servicio médico hospitalario**

Hoy en día el sistema hospitalario de Guatemala posee grandes problemas en la prestación de los servicios básicos que se brindan a la población (pacientes potenciales), esto básicamente se debe a la falta de normativas que regulen los servicios de admisión y documentación médica. Por ello las funciones derivadas de la gestión de pacientes han adquirido una relevancia muy importante tanto en términos clínicos como económicos, esta gestión es la actividad en torno a la cual giran todo el resto de actividades de la Institución médica, por lo tanto las mismas pueden ser automatizadas si se establece un control adecuado sobre la fuente primordial que es el paciente a través de un sistema flexible, estable y seguro.

Para la conveniencia de este trabajo se mencionará en gran parte al Hospital Roosevelt de Guatemala, debido a que se posee información relevante de dicha institución, y a su vez es uno de los dos hospitales más grandes de Guatemala. Se presta servicio a personas que están vecindadas en un área geográfica específica, pero regularmente se atiende a pacientes de todo el país, debido a que son remitidos para este centro asistencial por los hospitales departamentales debido a la seriedad de sus enfermedades o sencillamente porque en esos centros no cuentan con el equipo

necesario para tratar dicha enfermedad. Hasta el momento se han realizado varios intentos para poder automatizar el sistema de salud de Guatemala, pero esto solo ha quedado en intentos, sin que haya prosperado alguna propuesta seria y con bases sólidas, dónde el paciente sea su principal objetivo.

#### **4.1.1 Tecnología utilizada en la actualidad**

En la actualidad en el Hospital Roosevelt de Guatemala se utiliza un sistema para la administración de pacientes llamado SIADPA, dicho sistema tiene las siguientes características:

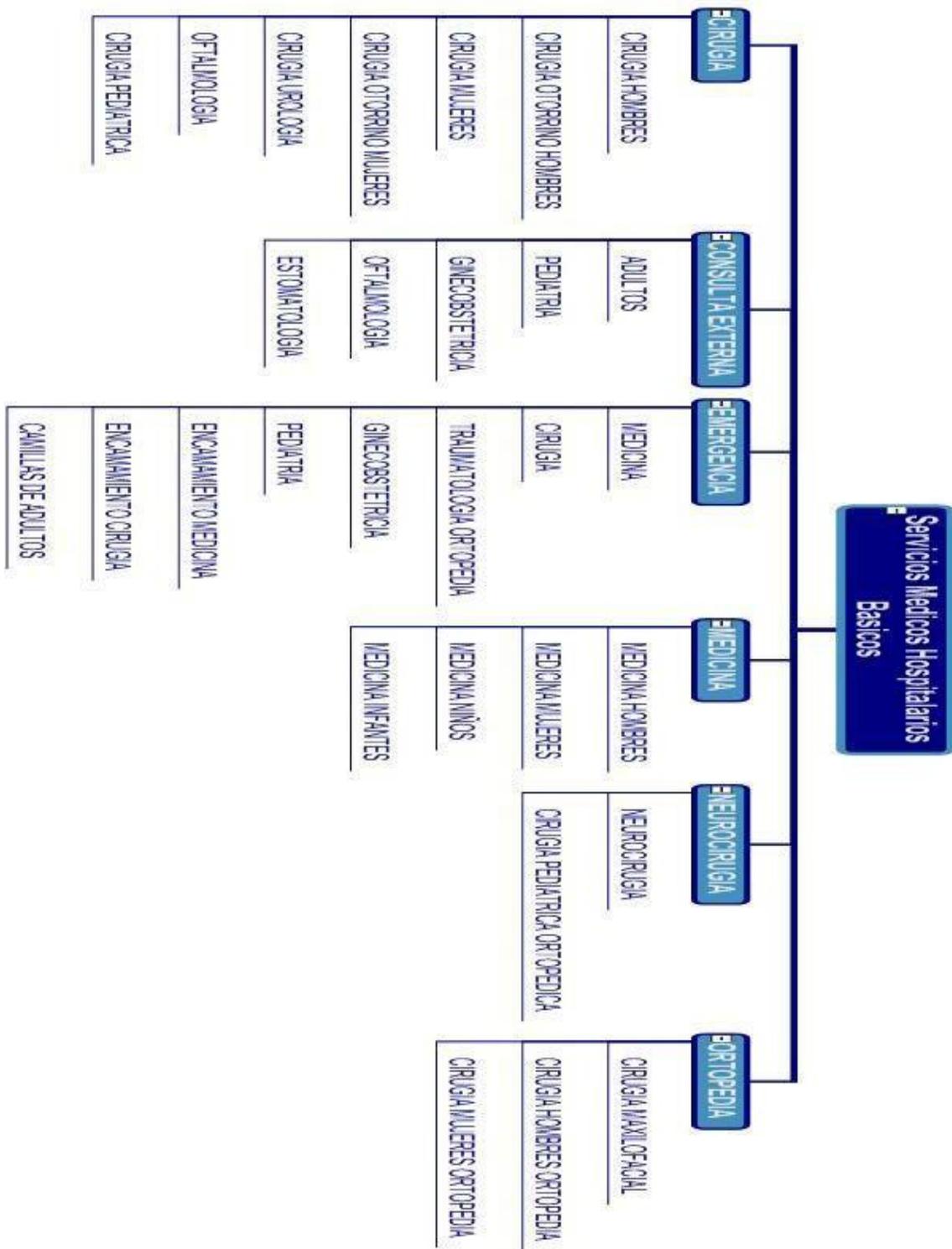
- Esta programado en ambiente Windows.
- Fue realizado en Visual Studio 5.0.
- Tiene opción de registro médico nuevo.
- Este sistema está implantado en cada una de las computadoras pertenecientes a las admisiones de tres áreas importantes: Adultos, Pediatría y Maternidad, en dichas terminales son tomados los datos de las personas para luego imprimir el reporte que lleva el registro medico.
- Este sistema no brinda información del historial clínico del paciente, es decir, las enfermedades por las cuales ha sido atendido en el hospital.

- El sistema no brinda reportes que contengan estadísticas acerca de los diagnósticos de pacientes, por lo que las proyecciones acerca de las enfermedades son totalmente nulas

#### **4.2 Clasificación de servicios hospitalarios que se prestan**

En el Hospital Roosevelt de Guatemala se prestan varios servicios en los que puede ser diagnosticadas enfermedades para los pacientes, los cuales están categorizados como se muestra en la siguiente figura 2:

Figura 2. Estructura de los Servicios Básicos Medico Hospitalarios.

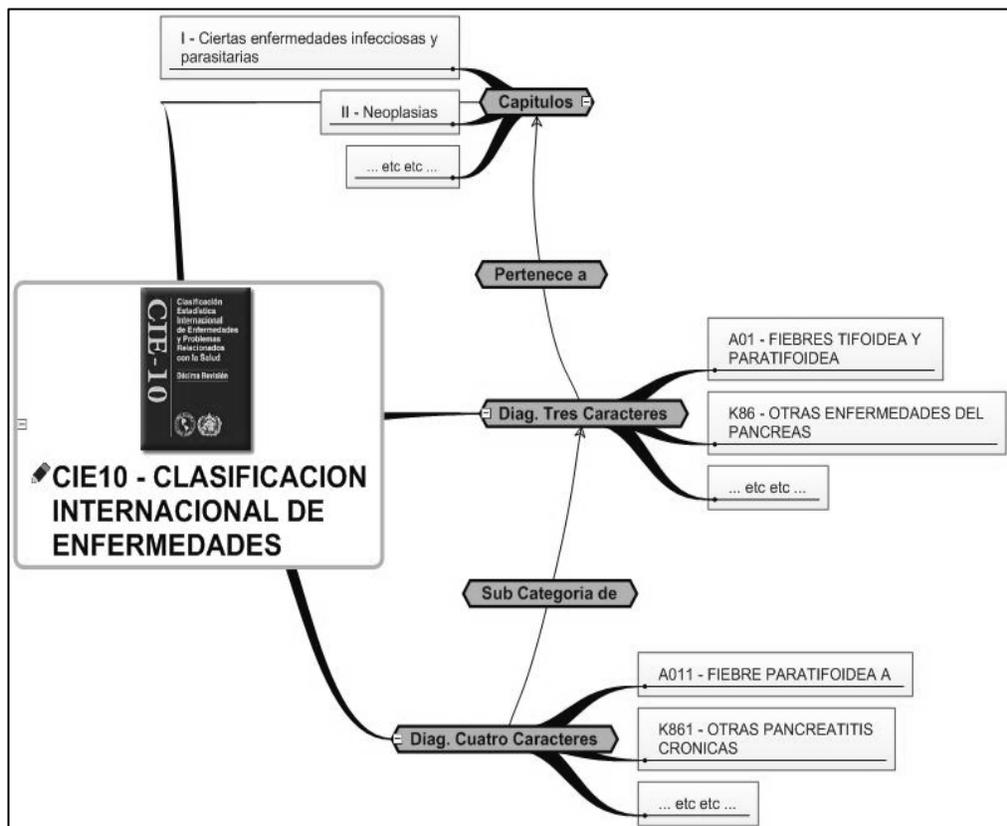


Referencia: figura realizada en Mind Genius 2005 por Raúl Estuardo Cuyún González, Noviembre 2008.

### 4.3 Clasificación Internacional de Enfermedades (CIE10)

La clasificación de enfermedades es usada mundialmente en áreas de morbilidad y mortalidad, éste sistema está diseñado para promover la comparación internacional de la recolección, procesamiento, clasificación y presentación de estadísticas alrededor del mundo.<sup>13</sup> Esta clasificación puede girar en torno a varios ejes posibles y la elección de uno en particular estará determinada por el uso que se hará de las estadísticas. La CIE10 básicamente está compuesta como muestra la figura 3.

**Figura 3. Modelo mental de la estructura de la CIE10.**



Referencia: figura realizada en Mind Genius 2005 por Raúl Estuardo Cuyún González, Noviembre 2008.

<sup>13</sup> [http://es.wikipedia.org/wiki/Lista\\_de\\_códigos\\_CIE-10](http://es.wikipedia.org/wiki/Lista_de_códigos_CIE-10) Enero 2009.

La revisión número diez es la más reciente que se ha realizado para los problemas relacionados con la salud. La categorización de enfermedades nació a partir de las reuniones de los Centros Colaboradores de la OMS, es por ello que esta estructuración fue regida por una buena parte de médicos. Cabe mencionar que el “Manual de Diagnóstico y estadístico de trastornos mentales” en su cuarta versión (DMS IV), favoreció en gran medida en la construcción de la CIE10 para estructurarla de una mejor manera en su quinto Capítulo, el cual trata los Trastornos mentales y de Comportamiento, así como también parte de la protocolarización que intenta adoptar la CIE10 de la que más adelante hablaremos.

La CIE10 provee los códigos para clasificar las enfermedades y una amplia variedad de signos, síntomas, hallazgos anormales, denuncias, circunstancias sociales y causas externas de daños y/o enfermedad<sup>14</sup>. Para poder elegir una afección dentro del sistema de categorización, ésta se deberá de realizar al final del proceso de atención como la causante primaria de la necesidad de tratamiento o investigación que tuvo el paciente. Si hay más de una afección así caracterizada, debe seleccionarse la que se considera causante del mayor uso de recursos.<sup>15</sup>

Esta CIE10 utiliza un código alfanumérico, con una letra en la primera posición y números en la segunda y tercera. Estas categorías suelen contener un grupo de enfermedades similares. Para poder interpretar el conocimiento de la CIE10 se ha identificado la siguiente

---

<sup>14</sup> [http://es.wikipedia.org/wiki/Lista\\_de\\_códigos\\_CIE-10](http://es.wikipedia.org/wiki/Lista_de_códigos_CIE-10) Enero 2009.

<sup>15</sup> <http://www.codamedver.gob.mx/p040ssa2.htm> Enero 2009.

estructura de la misma, que a su vez es sencilla pero funcional para propósitos de almacenamiento digital:

**Capítulos:** la CIE10 es una nomenclatura muy extensa por eso se ha identificado que se divide en alrededor veintidós capítulos, los cuales está compuestos por un numeral romano y una descripción, así como también tienen información de los rangos de códigos de tres caracteres que están contenidos en cada uno de ellos.

**Categoría de tres caracteres:** actualmente la CIE10 cuenta con alrededor de más de 2000 diagnósticos básicos los cuales son clasificados en ésta nomenclatura, éstos a su vez pertenecen a alguno de los capítulos antes mencionados.

**Categoría de cuatro caracteres:** estos diagnósticos son clasificados en una nomenclatura más específica, es decir, cada uno de los diagnósticos de tres caracteres a menudo contienen varias especificaciones distintas para el mismo tipo de diagnóstico, ésta clasificación tiene aproximadamente 1,2400 afecciones diferentes.

Para poder brindar una mejor explicación acerca de la descripción de la CIE10 se muestra un ejemplo 1 como sigue:

### **Ejemplo 1**

#### **Capítulo**

Código: I.

Descripción: Ciertas enfermedades infecciosas y parasitarias.

### **Categoría de tres caracteres**

Capítulo de CIE donde se encuentra contenido: I.

Código: A01.

Descripción: Fiebres tifoidea y paratifoidea.

### **Categoría de cuatro caracteres**

Categoría de tres caracteres de CIE donde se encuentra contenida:  
A01.

Código: A014.

Descripción: Fiebre Paratifoidea, No especificada.

Es muy importante mencionar que, para que la CIE10 funcione de manera adecuada dentro de los procesos para diagnosticar, es necesario que funcione en paralelo con algún tipo de protocolo asociado que deberá ser supervisado de alguna manera por una organización relacionada con el estudio epidemiológico de la región dónde se requiera implementar la CIE10, debo dejar en claro que éste protocolo no es parte de, sino un complemento de la CIE10 como ya se mencionó.

Se ha investigado y discutido mediante entrevistas con médicos como Leonel González y Odilia Magzul entre otros, actualmente residentes en el Hospital Nacional Pedro de Betancourt de Guatemala, en la localidad de Antigua Guatemala así como de la propia CIE10 acerca de lo que podría ser un protocolo como el que se ha señalado anteriormente, protocolo que de una manera consensuada debe de completar información como:

- Síntomas asociados a la enfermedad, esto con la finalidad de poder realizar algún censo con síntomas comunes entre enfermedades.
- Criterios de aceptación, se refiere a la cantidad de síntomas que se requiere que estén presentes para que se cumpla el diagnóstico.
- Exclusiones, es decir la revisión de otros diagnósticos antes de diagnosticar el presente.
- Inclusiones, muchas veces hay ciertas enfermedades que están incluidas dentro de otras enfermedades un tanto más complejas, para este caso el presente diagnóstico tiene varios diagnósticos incluidos en sí mismo. Es importante mencionar que en caso de las inclusiones, la CIE10 se encuentra poco madura hasta el momento, debido a que no está codificada como las exclusiones.

Dicho protocolo mencionado anteriormente es implementado por la CIE10 (pero no del todo) de la siguiente manera como muestra el ejemplo 2, diagnóstico obtenido de la página 106 de la Clasificación Internacional de Enfermedades en su décima versión.

## **Ejemplo 2**

### **Capítulo.**

Código: I.

Descripción: Ciertas enfermedades infecciosas y parasitarias.

### **Categoría de tres caracteres.**

Capítulo de CIE donde se encuentra contenido: I.

Código: A06.

Descripción: Amebiasis.

**Incluye:** infección debida a Entamoeba histolytica.

**Excluye:** Otras enfermedades intestinales debidas a protozoarios (A07).<sup>16</sup>

#### **4.4 Las 10 Primeras causas de morbilidad y mortalidad**

Habiendo ya comentado acerca de la composición estructural de la CIE10 hablaremos acerca de las enfermedades más frecuentes que son atendidas, actualmente en el Hospital Roosevelt de Guatemala se tienen estadísticas de las primeras 10 causas de morbilidad y mortalidad de los pacientes que son diariamente atendidos en dicho centro asistencial. La siguientes dos tablas muestran el resumen de estas estadísticas durante el período de enero a junio de 2008.

---

<sup>16</sup> <http://www.dne.sld.cu/librocie10/capitulo%20I.htm> Noviembre 2008.

**Tabla II. Las 10 principales causas de morbilidad.**

<b>CAUSAS DE MORBILIDAD</b>	<b>Hombres</b>	<b>Mujeres</b>	<b>Total</b>	<b>Capítulo</b>
DIABETES MELLITUS	67	96	163	IV
VIH/SIDA	102	55	157	I
INSUFICIENCIA RENAL CRONICA	60	69	129	XIV
HEMORRAGIA INTESTINAL SUPERIOR	53	48	101	
NEUMONIA	36	51	87	X
ENFERMEDAD CEREBRO VASCULAR	28	52	80	IX
PANCREATITIS	30	48	78	XI
PANCREATITIS AGUDA	27	44	71	XI
LEUCEMIA LINFOCITICA AGUDA	26	29	55	II
SINDROME CORONARIO AGUDO	11	4	15	XI
RESTO DE CAUSAS	600	619	1219	Otros
<b>TOTAL</b>	<b>440</b>	<b>496</b>	<b>936</b>	

Referencia: datos proporcionados por el Departamento de Estadística del Hospital Roosevelt de Guatemala, de enero a junio del 2008.

Como podemos verificar, la tabla tiene una columna llamada capítulo, en dicha columna se especifica el número de capítulo perteneciente a la CIE10 al cual pertenece el diagnóstico.

**Tabla III. Las 10 principales causas de mortalidad.**

<b>CAUSAS DE MORTALIDAD</b>	<b>Mujeres</b>	<b>Hombres</b>	<b>Total</b>	<b>Capitulo</b>
VIH/SIDA	41	11	52	I
ENFERMEDAD CEREBRO VASCULAR	15	24	39	IX
PARO CARDIO RESPIRATORIO	19	15	34	IX
NEUMONIA	15	14	29	X
DIABETES MELLITUS	13	11	24	IV
INSUFICIENCIA RENAL CRONICA	13	9	22	XIV
HEMORRAGIA GASTRO INTESTINAL	9	7	16	
FALLECIDO VIH/SIDA	11	3	14	I
FALLECIDO PARO CARDIORESPIRATORI	8	5	13	IX
FALLO VENTILATORIO	7	5	12	
RESTO DE CAUSAS	165	135	300	Otros
<b>TOTAL</b>	<b>316</b>	<b>239</b>	<b>555</b>	

Referencia: datos proporcionados por el Departamento de Estadística del Hospital Roosevelt de Guatemala de enero a junio del 2008.

## **5. GESTIÓN DEL CONOCIMIENTO MEDIANTE EL USO DE ONTOLOGIAS EN EL SERVICIO DE ATENCIÓN MÉDICA HOSPITALARIA DE GUATEMALA**

El quinto capítulo es sin duda la parte central de este trabajo de graduación por lo que es muy importante, esto no quiere decir que los anteriores no lo hayan sido, sino que es aquí donde se aborda el análisis de requerimientos, la elección de la tecnología a utilizar para la construcción de la solución básica propuesta, todo esto basado en los antecedentes mencionados en los capítulos anteriores. También en la llegada de éste capítulo se hace necesario aclarar que éste trabajo es solamente una parte de una solución integral para la atención médica hospitalaria de Guatemala, ya que ésta va mas allá de diagnósticos y consultas a pacientes.

Se hace necesario comentar que éste trabajo forma parte de un proyecto denominado OntoGuate–SGC propuesto por Javier Gramajo, que está destinado en gestionar un espacio para poder crear conocimiento básicamente en un área geográfica delimitada por Sur de México, Centroamérica y Colombia.

Esta solución básica pretende ser el comienzo para la construcción de sistemas orientados al área médica, utilizando tecnologías innovadoras para el almacenamiento de conocimiento como lo son las ontologías para búsquedas más precisas, específicamente para la Clasificación Internacional de Enfermedades en su decima versión (CIE10), todo esto en su mayoría para fines estadísticos y de colaboración para personas asociadas al área médica.

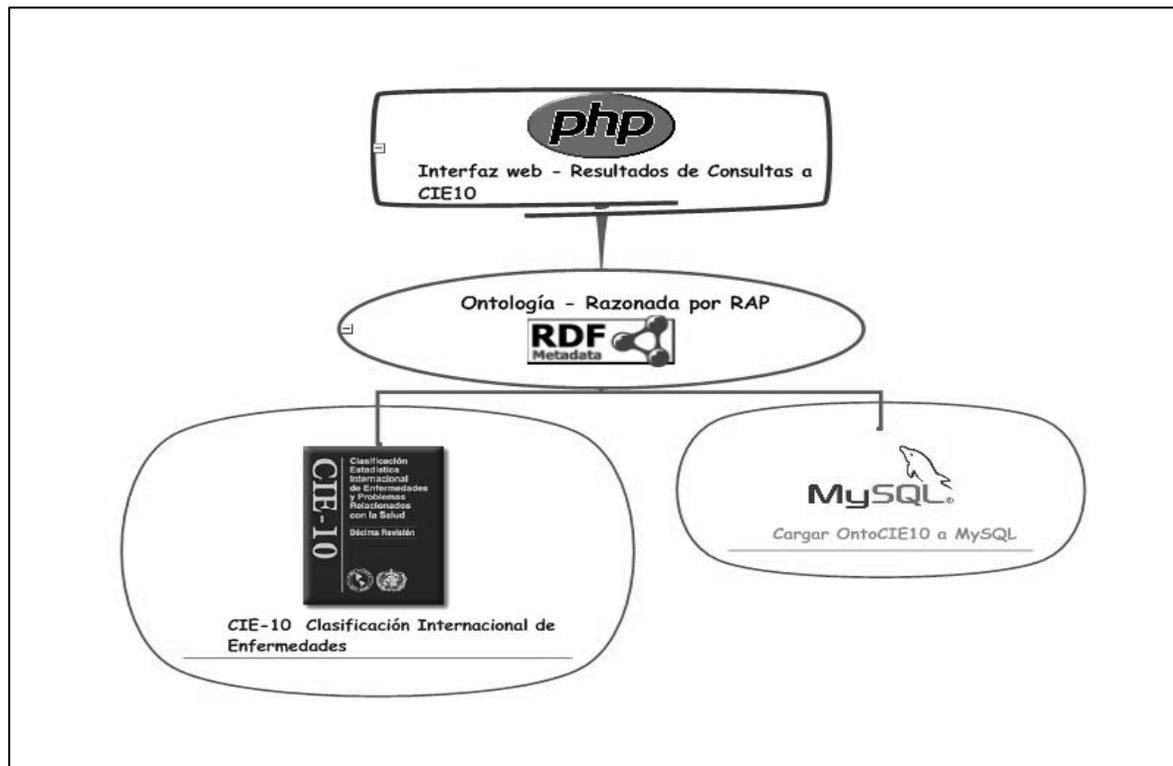
La solución básica propuesta para éste trabajo consta de las siguientes tres partes:

**Primera parte:** creación de una ontología para la Clasificación Internacional de Enfermedades CIE10.

**Segunda parte:** cargar la ontología mediante una API a una base de datos de MySQL.

**Tercera parte:** aplicación dónde las búsquedas serán realizadas mediante un razonador a la ontología y a la base de datos por el motor de búsqueda predeterminado. La figura 4 muestra modelo mental de la propuesta básica como sigue:

**Figura 4. Modelo mental de la solución básica propuesta.**



Referencia: figura realizada en Mind Genius 2005, por Raúl Estuardo Cuyún González, Noviembre 2008.

Entrando ya en materia, es importante empezar por la identificación de requerimientos, para ello se realizará el análisis de requerimientos de la Clasificación Internacional de Enfermedades.

## **5.1 Identificación de requerimientos para la Clasificación Internacional de Enfermedades (CIE10)**

Es de mucha importancia poder identificar la información necesaria para el almacenamiento de la CIE10 en un entorno digital. En el siguiente orden haremos ésta identificación de requerimientos: capítulos, categoría de tres caracteres, categoría de cuatro caracteres.

### **5.1.1 Descripción de los campos de capítulos**

Como ya se ha mencionado, ésta la parte donde se encuentran contenidas muchas categorías de tres caracteres, para ello necesitamos conocer información acerca de:

- **Código:** campo que representa el código único de cada uno de los capítulos.
- **Descripción:** Representa el nombre o descripción del capítulo a crear.
- **Rango de códigos:** éste campo contiene un rango de códigos de tres caracteres que pertenecen al capítulo.

### **5.1.2 Descripción de los campos para categorías de tres caracteres**

Los siguientes campos son muy importantes para almacenar la información de categorías de tres caracteres:

- **Capítulo:** debido a su estructura categorizada, es imprescindible que ésta categoría contenga al capítulo de la CIE10 a la que pertenece.
- **Código:** es la representación única de la categoría, no se podrá repetir, este campo tendrá que ser alfanumérico.
- **Descripción:** almacena el nombre o descripción completa para este tipo de diagnóstico.

### **5.1.3 Descripción de los campos para categorías de cuatro caracteres**

- **Código de tres caracteres:** la categoría de cuatro caracteres es una clasificación más específica de la categoría de tres caracteres por lo que éste campo hace referencia a ésta categoría superior a la cual pertenece.
- **Código:** es la representación única de la categoría, no se podrá repetir, este campo tendrá que ser alfanumérico.
- **Descripción:** almacena el nombre o descripción completa para este tipo de diagnóstico.

## **5.2 Identificación de requerimientos del protocolo para CIE10**

Como ya se ha mencionado en las líneas finales del capítulo anterior, se necesita que de un protocolo que funcione de manera paralela a la CIE10 para que esta pueda brindar mejores resultados, este protocolo se resume en los siguientes conceptos:

- Síntomas asociados a la enfermedad.
- Criterios de aceptación.
- Exclusiones.
- Inclusiones.

Debo aclarar que este protocolo no es el único, puede que halla alguna otra solución, pero para éste trabajo se propone éste, por la sencillez y facilidad en la manipulación de la ontología.

### **5.3 Herramienta a utilizar para crear la ontología**

La ontología del proyecto se construirá en la herramienta Protégé en la versión 3.3.1 que es la versión más reciente y estable. Se ha elegido esta herramienta por varias razones ventajosas sobre algunas otras como las descritas en la tabla 1 del capítulo 3. Una de ellas es que se cuenta con las librerías RAP que hacen que el enlace entre PHP y RDF sea posible. Como también se ha puntualizado en el capítulo 3, RAP es un conjunto de librerías que interactúan exclusivamente para el entorno PHP, éste aspecto es fundamental debido a que es importante darle impulso a la programación con Software libre como lo es éste lenguaje. Además los sitios que venden alojamiento de PHP por lo regular son más económicos que el de otros lenguajes como aspx, jsp etc.

#### **5.3.1 Protégé**

Protégé es una herramienta de edición de ontologías que apoya el desarrollo y mantenimiento de las mismas utilizando medios gráficos en un entorno Web. Permite la representación

semántica de lenguajes conceptuales y estructuras mediante conceptos, jerarquías de conceptos, relaciones y axiomas.

La interfaz es abierta y permite a los usuarios ajustar Protege a sus necesidades. Además tiene disponibles una gran variedad de funcionalidades fáciles de usar y adaptables a los diferentes escenarios.

Lo que Protégé pretende, es ofrecer una herramienta para representar gráficamente ontologías y que además, pueda almacenarlas y posteriormente manipularlas en una base de datos relacional.

Otros editores de ontologías se centran en ofrecer interfaces gráficas ricas en funcionalidades para editar, integrar y traducir ontologías; pero estas herramientas no poseen integración en una base de datos, las ontologías no pueden ser guardadas, gestionadas y manipuladas directamente por el editor.

#### **5.4 Construcción de la ontología en Protégé**

Para comenzar se quisiera dejar en claro que ésta sección no pretende ser una introducción a Protégé ni mucho menos un manual del mismo, simplemente es una guía para poder construir la ontología planteada para este trabajo de graduación.

Inicialmente se crea el proyecto con el nombre de OntoCIE10 de tipo RDF, debemos asegurarnos que en el lugar donde almacenamos el proyecto se crean dos archivos, con extensiones RDF y RDFS

respectivamente. Dentro de la ontología se deberán de crear dos clases maestras que representarán los dominios a las cuales llamaremos CIE10 y Protocolo, ya dentro de ellas se crearan las sub clases que mencionaremos posteriormente para su realización. Es importante mencionar también que se debe de crear una nomenclatura para identificar los nombres de cada una de las clases y/o propiedades ya que no se sabe que tanto crecerá la ontología en su esquema, por lo que el nombre de cada una de las clases y/o propiedades deben de comenzar con un sub fijo el cual lo identificará a la clase que pertenece, o bien nombres que realmente sea significativos a la comprensión humana.

#### **5.4.1 Creación de clases para la CIE10**

Dentro de la clase maestra CIE10 se crearán las siguientes sub clases que componen ésta parte de la ontología y que servirá para identificar la composición de la CIE10.

##### **5.4.1.1 Clase capítulo**

Esta clase será utilizada para poder realizar el almacenamiento de instancias de los capítulos en los que se compone la CIE10. Las siguientes propiedades como ya se identificaron en los requerimientos son las siguientes:

- **Capítulo:** se necesita para ésta clase una identificación única para representar cada una de las instancias, éste campo tiene que ser "String" debido a que la información a

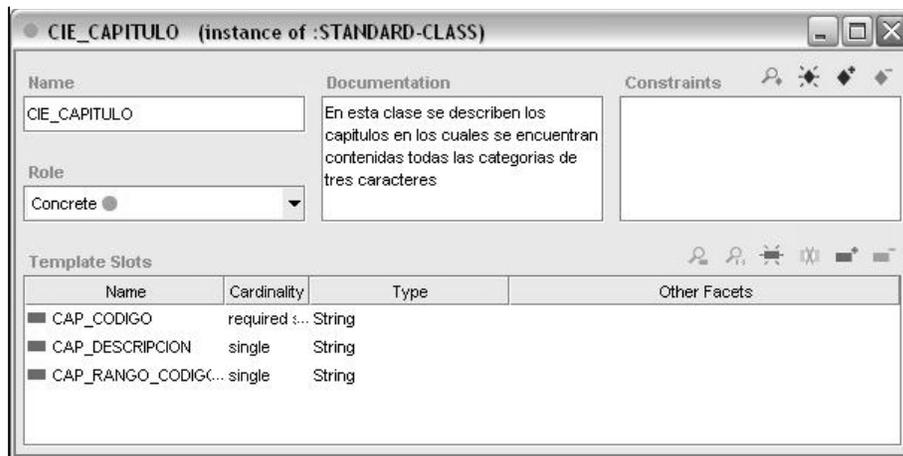
almacenar será de tipo alfanumérico, por lo que éste tipo de dato se adapta perfectamente.

- **Descripción:** cada capítulo cuenta con un nombre o descripción, este será un campo de tipo “String” debido a que cuenta con información alfanumérica.

- **Rango de códigos:** también como las anteriores, esta propiedad debe ser alfanumérica y el tipo que se le adapta es “String”.

En la figura 5 se muestra la construcción de la clase capítulo con las propiedades descritas adoptando la nomenclatura propuesta en el apartado anterior.

**Figura 5. Sub Clase CIE\_Capitulo.**



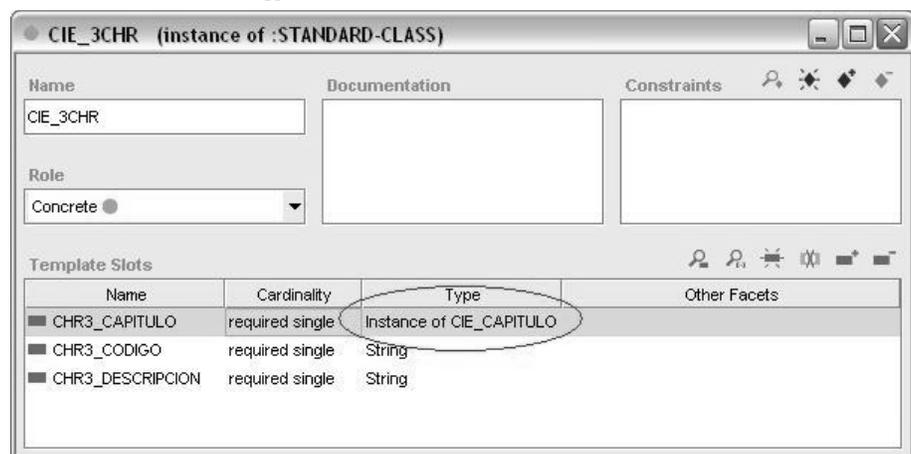
Referencia: figura obtenida de Protégé 3.3.1, Noviembre 2008.

### 5.4.1.2 Clase categoría de tres caracteres

- **Capítulo:** como se puede comprobar en la figura 3, ésta sub clase forma parte de la sub clase capitulo, esto debido a su estructura categorizada.
- **Código:** cada uno de los diagnósticos de tres caracteres necesita un código único, la información contenida en esta propiedad posee datos alfanuméricos por lo que el tipo de dato nuevamente debe ser “String”.
- **Descripción:** así mismo esta propiedad es de tipo de dato “String”.

La figura 6 describe de una manera grafica las propiedades creadas en protege utilizando como ya se menciona la nomenclatura para nombrarla. Como se puede observar en dicha figura, la relación entre la propiedad de capitulo de ésta clase y la clase capitulo se logra a través de asociar mediante una instancia a “Instance of CIE\_Capitulo”.

**Figura 6. Sub Clase CIE 3CHR.**



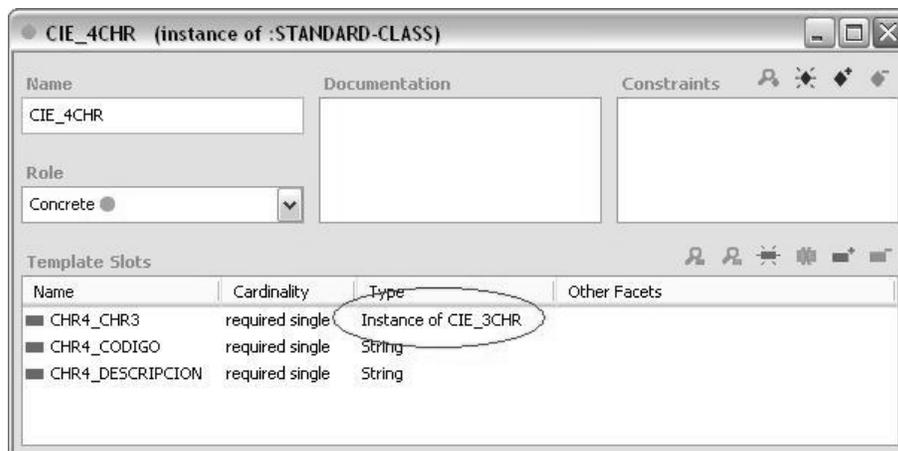
Referencia: figura obtenida de Protégé 3.3.1, Noviembre 2008.

### 5.4.1.3 Clase categoría de cuatro caracteres

- **Código de tres caracteres:** ésta propiedad se relaciona mediante una instancia de la clase “CIE\_3CHR” como se muestra mas adelante en la figura 7, porque como se ha dicho, la categoría de cuatro caracteres es una clasificación más específica de la categoría de tres caracteres.
- **Código:** los datos contenidos para ésta propiedad son alfanuméricos por lo que nuevamente el tipo de dato debe ser “string”.
- **Descripción:** almacena el nombre o descripción completa para este tipo de diagnóstico.

Como podemos ver la CIE10 es una estructura muy pequeña, pero que puede crecer si la necesidad lo amerita, por lo que una ontología como ésta se encuentra lista para este cambio. La figura 7 muestra gráficamente la manera en que se encuentran creadas las propiedades descritas.

**Figura 7. Sub Clase CIE\_4CHR.**



Referencia: figura obtenida de Protégé 3.3.1, Noviembre 2008.

## 5.4.2 Creación de clases para protocolo CIE10

Continuando con la ontología, pero esta vez en la parte de dominio del protocolo se procede a crear las sub clases asociadas, clases en las cuales enlazaremos información de la CIE10 a información puramente del protocolo.

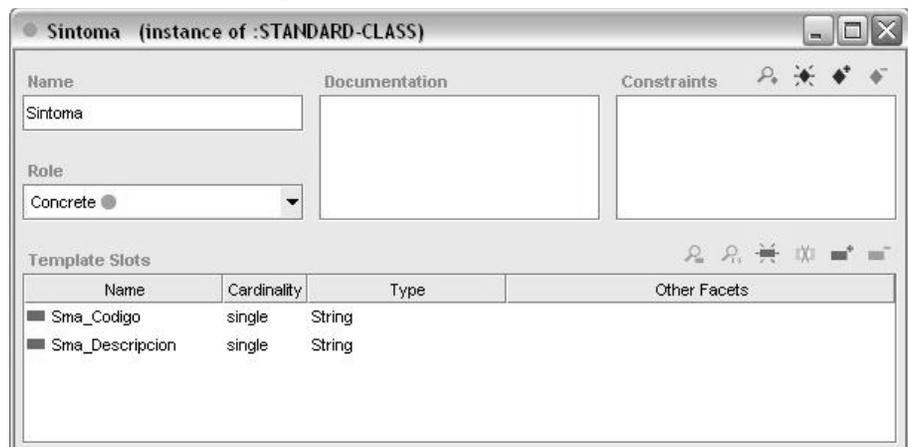
### 5.4.2.1 Síntoma

Es una clase que representa el listado de padecimientos que pueda estar contenida en el paciente y que hace que pueda asociarse al diagnóstico.

**Sma\_Codigo:** es un atributo que representa el código único de cada uno de los síntomas asociados al protocolo.

**Sma\_Descripcion:** almacena la descripción o cualquier otra particularidad del síntoma.

Figura 8. Sub clase sintoma.



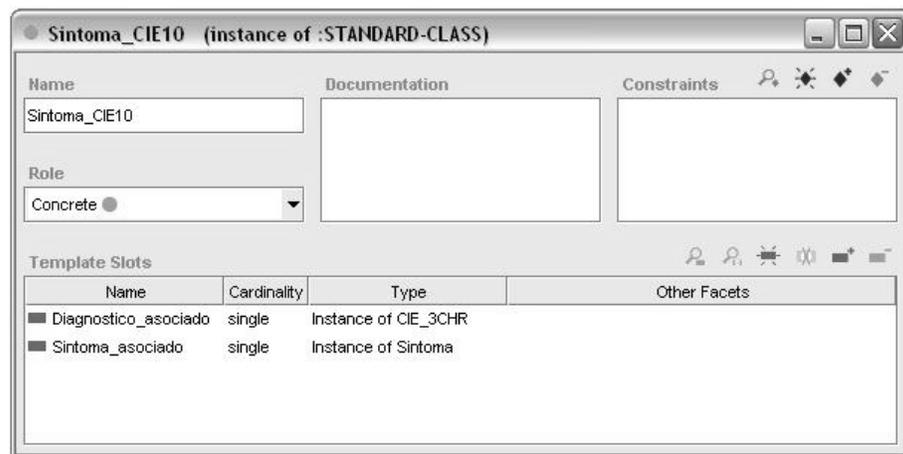
Referencia: figura obtenida de Protégé 3.3.1, Noviembre 2008.

#### 5.4.2.2 Síntoma\_CIE10

**Diagnóstico\_asociado:** que no es mas que un dato de una instancia del diagnostico, de la clase CIE3\_CHR.

**Síntoma\_asociado:** es la instancia de alguno de los síntomas, asociados de una manera coherente con el diagnóstico.

**Figura 9. Sub Clase Sintoma\_CIE10.**



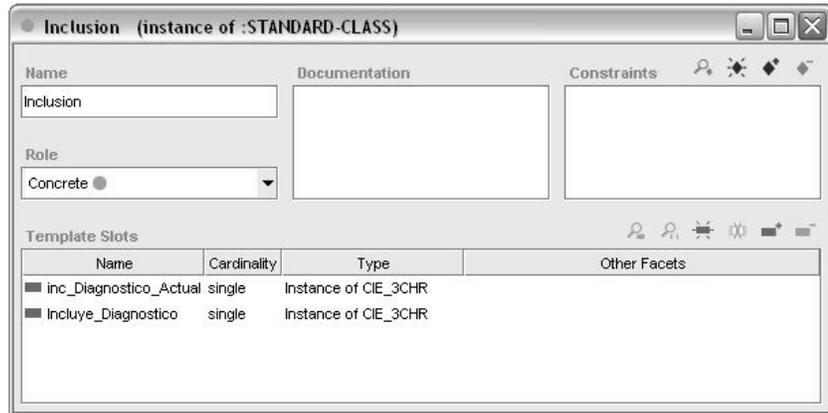
Referencia: figura obtenida de Protégé 3.3.1, Noviembre 2008.

#### 5.4.2.3 Inclusión

**Inc\_diagnóstico\_actual:** se refiere a una instancia del diagnóstico de la clase CIE\_3CHR, se asocia a ésta clase como parte del protocolo y también se refiere al diagnostico que se está consultando actualmente.

**Incluye\_diagnóstico:** también es una instancia de la CIE\_3CHR solamente que esta representa el diagnostico que se encuentra incluido dentro del diagnóstico actual.

**Figura 10. Sub clase inclusión**



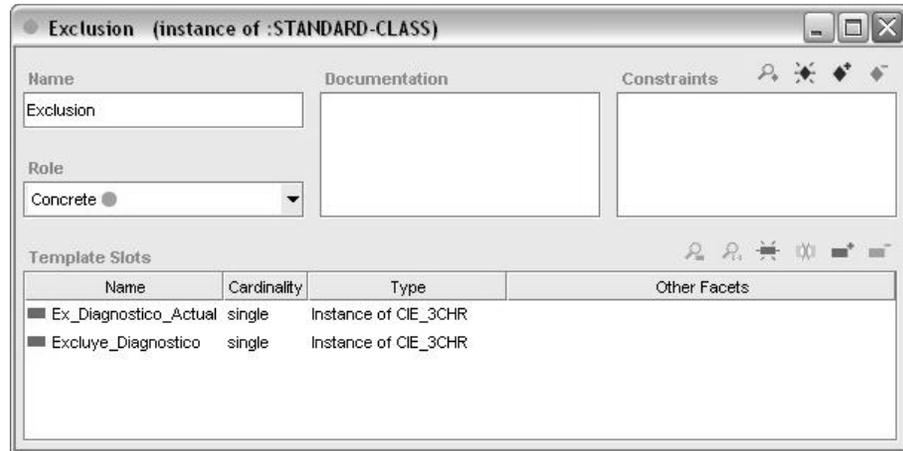
Referencia: figura obtenida de Protégé 3.3.1, Noviembre 2008.

#### 5.4.2.4 Exclusión

**Ex\_diagnóstico\_actual:** se refiere a una instancia del diagnóstico de la clase CIE\_3CHR, se asocia a ésta clase como parte del protocolo, y se refiere al diagnóstico que se está consultando actualmente.

**Excluye\_diagnóstico:** es un atributo que almacena el diagnóstico de tres caracteres mediante una instancia de la clase CIE\_3CHR, y como se ha mencionado este tipo de diagnósticos son los que hay que consultar antes de poder proporcionar uno definitivo.

**Figura 11. Sub clase exclusión**



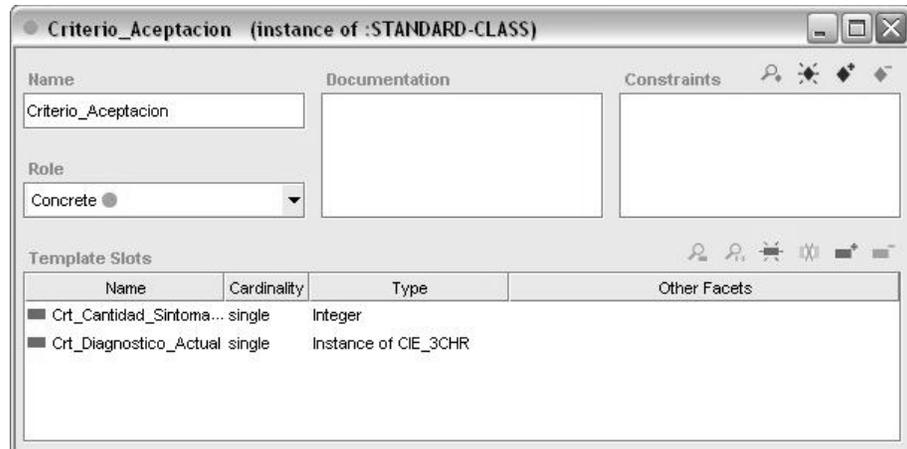
Referencia: figura obtenida de Protégé 3.3.1, Noviembre 2008.

#### **5.4.2.5 Criterio de aceptación**

**Crt\_cantidad\_síntoma:** éste atributo almacena la cantidad de síntomas que se requiere para que este diagnostico sea realmente el indicado, o en su defecto mayormente probable, es lógico que cada uno de los diagnósticos tenga un número diferente ya que todos son distintos.

**Crt\_diagnostico\_actual:** se refiere a una instancia del diagnóstico de la clase CIE\_3CHR, se asocia a ésta clase como parte del protocolo, y se refiere al diagnostico que se está consultando actualmente.

**Figura 12. Sub clase criterio\_aceptación.**



Referencia: figura obtenida de Protégé 3.3.1, Noviembre 2008.

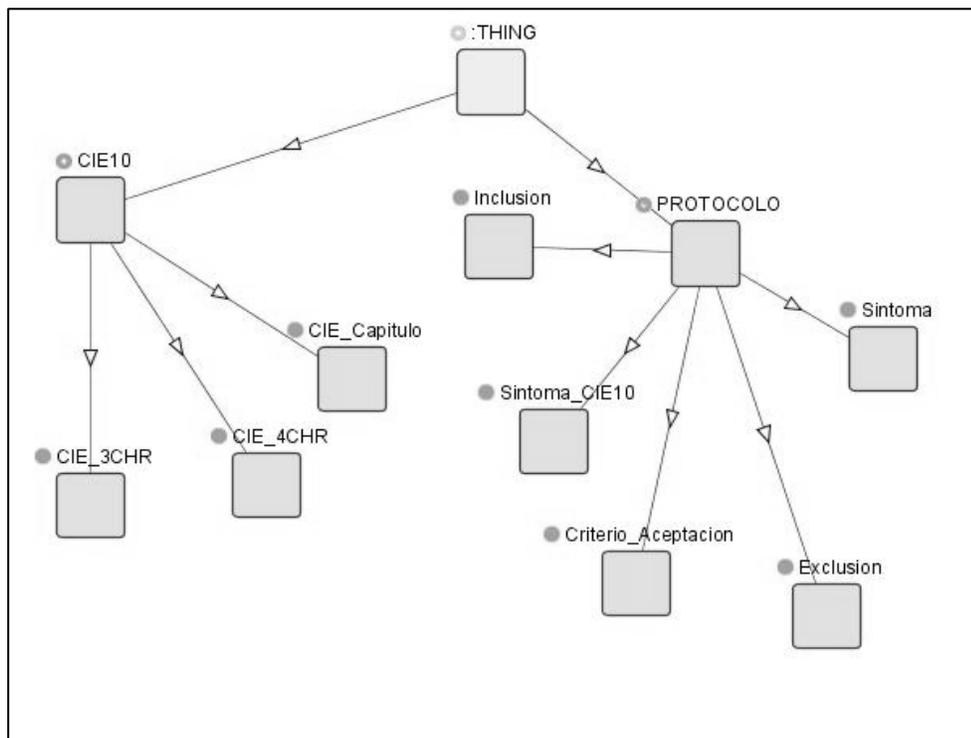
De esta manera queda finalizada la construcción de nuestra ontología llamada OntoCIE10 para la que luego manipularemos mediante el módulo propuesto, donde se consultarán diagnósticos en una interfaz construida en PHP mediante librerías RAP.

Quisiera hacer un paréntesis para comentar acerca de un plug-in que posee Protégé 3.3.1 llamada Jambalaya, que es un visualizador para poder graficar cualquier ontología dentro de la herramienta de Protégé 3.3.1, éste plug-in tiene varios tipos de visualizaciones como:

- Vista de anillo.
- Arbol, horizontal y vertical.
- Orden, este puede ser por tipo, por orden alfabético.

En la figura 13 se muestra un ejemplo de visualización que se logra obtener mediante el plug-in Jambalaya, la figura es como sigue:

**Figura 13. Ontología de CIE10 y protocolo asociado**



Referencia: figura obtenida del Plug-in Jambalaya de Protégé 3.3.1, Noviembre 2008.

## CONCLUSIONES

1. Se estableció mediante algunas métricas sencillas la comparativa de las herramientas utilizadas en la creación de ontologías, para luego tomar una decisión en la herramienta a utilizar dentro de este trabajo.
2. Se ha logrado satisfactoriamente cubrir las expectativas en los requerimientos encontrados para la solución planteada en este trabajo, mediante la construcción de la ontología CIE10 en la herramienta Protégé versión 3.3.1.
3. Se realizó la carga de datos de todos los diagnósticos asociados hasta el momento de la Clasificación Internacional de Enfermedades en su décima versión (CIE10) publicada por la Organización Panamericana de la Salud en 1992, a través de la escritura de las instancias con su estructura a un archivo plano RDF.
4. Realización de un protocolo médico de colaboración, en paralelo a la CIE10 para un funcionamiento eficaz de los procesos asociados en encontrar el diagnóstico más certero, mediante búsquedas, utilizando parámetros concretos como síntomas, diagnósticos incluyentes, diagnósticos excluyentes.
5. Basado en los requerimientos identificados en el quinto capítulo se ha realizado satisfactoriamente la programación de un módulo de consulta en PHP para las búsquedas inteligentes, razonado mediante la API RAP (RDF API for PHP).



## RECOMENDACIONES

Dentro de las inquietudes que se han despertado a lo largo de este trabajo se han podido formular las siguientes recomendaciones:

1. Hasta el momento los diagnósticos cargados a la CIE10 en este trabajo son obtenidas de la última versión (versión 10), la cual fue publicada en 1992, por la Organización Mundial de la Salud, por lo que es recomendable que las autoridades relacionadas con el área médica (de la región propuesta por OntoGuate: Centroamérica, sur de México, Colombia) organicen un comité para validar las enfermedades con los nuevos descubrimientos realizados en dicha área.
2. Actualmente no hay una política para que los diagnósticos se estandaricen a nivel médico, es decir la CIE10 es solamente una guía, de seguir así los sistemas médicos que se construyan no podrán funcionar a la perfección como se desea, por lo que se recomienda dejar de improvisar en la marcha con sistemas de alta tecnología y enfocar esfuerzos en establecer una cultura en la que los involucrados en el área médica utilicen la CIE10 de una manera más responsable y activa.
3. Es necesario que dependencias gubernamentales brinden apoyo a los proyectos concernientes con el desarrollo de la sociedad, y para este

caso a los relacionados con la salud, creando un clima adecuado para que estos se realicen de una manera correcta.

En cuanto a la implementación técnica del proyecto propuesto en este trabajo se hace necesario hacer las siguientes recomendaciones:

4. Para poder cargar una ontología de magnitud considerable (por ejemplo 500 KB) a una base de datos mediante las librerías del razonador RAP, es necesario realizar algunos cambios en el servidor web, específicamente en el archivo de configuración php.ini en el área de "Resource Limits" debido a que aquí se configura el tiempo de espera para las transacciones realizadas, el cual deberá de aumentar para que la ontología pueda cargar a la base de datos en MySql.
5. Cuando la cantidad de los datos en la ontología crezca de manera considerable se recomienda utilizar algún API para implementar AJAX en PHP, esto para agilizar la carga pesada de datos en las páginas dinámicas de búsquedas.

## REFERENCIAS ELECTRÓNICAS

1. [www.w3c.org](http://www.w3c.org) (Agosto 2008)
2. [http://personales.upv.es/ccarrasc/doc/2003-2004/OntoEdit/presentacion\\_html/definiciononto.html](http://personales.upv.es/ccarrasc/doc/2003-2004/OntoEdit/presentacion_html/definiciononto.html) (Agosto 2008)
3. <http://jena.sourceforge.net> (Agosto 2008)
4. <http://www.jdom.org> (Agosto 2008)
5. <http://www.who.int/es/>. (Septiembre 2008)
6. <http://devserver.paho.org/> (Septiembre 2008).
7. <http://es.wikipedia.org/wiki/CIE-10>. (Septiembre 2008)
8. <http://es.wikipedia.org/wiki/DSM-IV> (Septiembre2008)
9. <http://www.hospitalroosevelt.gob.gt>. (Noviembre)
10. [http://es.wikipedia.org/wiki/Administraci%C3%B3n del Conocimiento](http://es.wikipedia.org/wiki/Administraci%C3%B3n_del_Conocimiento) (Agosto 2008).
11. <http://es.wikipedia.org/wiki/Ontolog%C3%ADa> (Septiembre 2008)

12. <http://portal.mspas.gob.gt/> (Septiembre 2008).
13. <http://protege.stanford.edu/> (Octubre 2008).
14. [http://es.wikipedia.org/wiki/Inteligencia\\_artificial](http://es.wikipedia.org/wiki/Inteligencia_artificial) (Junio 2008).
15. <http://www.aifb.uni-karlsruhe.de/WBS/ysu/publications/SEKT-D3.1.1.b.pdf> (Octubre 2008).
16. [http://es.wikipedia.org/wiki/Capital\\_humano](http://es.wikipedia.org/wiki/Capital_humano) (Enero 2009).
17. [http://es.wikipedia.org/wiki/Propiedad\\_intelectual](http://es.wikipedia.org/wiki/Propiedad_intelectual) (Enero 2009).
18. <http://www.gnu.org/licenses/licenses.es.html#GPL> (Febrero 2009)
19. <http://www.gnu.org/copyleft/gpl.html> (Febrero 2009)
20. <http://www.youtube.com/watch?v=jC2H4vIRdFs> (Noviembre 2008)
21. <http://www.sulabatsu.com/>. (Enero 2009)
22. <http://www.ontologia.net/> (Enero 2009).
23. <http://www.ontoguate-salud.com/> (JUNIO 2009)

## BIBLIOGRAFÍA

1. Abián, Miguel Angel. El futuro de la Web. Octubre 2005.
2. Ariza, Cesar. Uso de RDF y Base de Datos de Metadatos nativas dentro del proyecto Omnipaper. Universidade do Minho. Portugal 2004.
3. Blázquez Entonado, Florentino. Sociedad de la información y Educación. Consejería de Educación, Ciencia y Tecnología. Mérida 2001.
4. Camacho, Kemly Jiménez. Gestión del Conocimiento: aportes para una discusión latinoamericana. Costa Rica 2008.
5. CIE10 Glosario de Síntomas para los trastornos mentales, Unidad de investigación de Psiquiatría Social de Cantabria, Hospital Universitario Marques de Valdecilla, Santander España, 1994.
6. Clasificación Internacional de Enfermedades en su decima versión CIE10, Organización Panamericana de Salud, Organización Mundial de la Salud, Ginebra, 1992.
7. Declaración de principios para construir la sociedad de la información: un desafío mundial para el nuevo milenio. Ginebra 2003 – Tunes 2005, Diciembre 2003.

8. Diaz Ortuño, Pedro. Problemática y Tendencias en la arquitectura de metadatos Web, Facultad de Ciencias de la Documentación, Universidad de Murcia. España 2003.
9. Fernández Breis, Jesualdo. Tecnologías semánticas para facilitar la integración de datos e interoperabilidad de sistemas de información médicos. Departamento de informática y sistemas, Facultad de Informática, Universidad de Murcia. España.
10. Gramajo López, Javier. OntoGuate-SGC Sistemas de gestión de conocimiento basados en ontologías aplicados a sectores de desarrollo social y económico. Universidad de San Carlos de Guatemala. Guatemala 2008.
11. Gramajo López, Javier. Ontologies State of the Art. Ingeniería en Ciencias y Sistemas, Facultad de Ingeniería, Universidad de San Carlos de Guatemala. Guatemala Abril 2005.
12. Labra, José. Web Semántica. Universidad de Oviedo, Departamento de informática. España
13. Lozano Tello, Adolfo. Ontologías en la Web Semántica, Area de Lenguajes y sistemas informáticos. Departamento de Informática, Escuela Politécnica, Universidad de Extremadura. España.
14. Perfil de los sistemas de Salud en Guatemala. Monitoreo y Análisis de los Procesos de Cambio y Reforma. USAID. Guatemala. Febrero 2007.

15. Sabino, Carlos. Cómo hacer una tesis, segunda edición. Caracas Venezuela 1994, 240pp.
16. Senso, José. Herramientas para trabajar con RDF. El profesional de la información, volumen 12. Marzo – Abril 2003.
17. Tallarico, Marcelo. Uso de Ontologías en tareas de recupero de información. Trabajo de graduación. Universidad Nacional de Rosario.
18. T. R. Gruber. A translation approach to portable ontology specifications. Knowledge Acquisition 5 (1993) 199-220.
19. Tramullas Saz, Jesús. Agentes y Ontologías para el tratamiento de la información: Clasificación y recuperación en internet, Universidad de Zaragoza, España 2006.
20. Una estrategia nacional en el desarrollo de las TIC's y el comercio electrónico: la experiencia de Costa Rica. Curazao, 25 – 27 de Junio del 2002.



## **APÉNDICE**

### **A.1 Sistema OntoGuate-Salud.**

Como se ha comentado dentro de la introducción y en algunas líneas a lo largo de éste trabajo, este concluye con la construcción de un software sencillo que permita visualizar el alcance de la propuesta planteada acerca de las búsquedas inteligentes en el contexto de la CIE10. Como todo proyecto se necesita de recursos de algunas áreas específicas como las que sigue.

#### **A.1.1 Análisis de Recursos.**

##### **Recurso Humano.**

En el desarrollo del proyecto presentado en este trabajo de graduación, se contó con la participación de Javier Gramajo, autor intelectual de OntoGuate – SGC, Raúl Cuyún González, autor del módulo OntoGuate-Salud. Para darle continuidad es necesaria la participación de un grupo de médicos, preferentemente de un hospital donde sea utilizada la CIE10.

##### **Herramientas Informáticas.**

Para poder construir este módulo de búsquedas, se utilizaron las herramientas que SE listan a continuación.

Protege versión 3.3.1  
Apache versión 2.2.8  
MySQL versión 5.0.51b  
RAP (RDF Api for PHP) version 0.9.6  
Macromedia Dreamweaver 8.

## A.2 Administración de OntoGuate - Salud.

Para fines educativos, en éste sistema se han cargado solamente las instancias del capítulo V de la CIE10, correspondientes a los “Trastornos Mentales y del Comportamiento”. En éste sistema la ontología podrá ser administrada y consultada de una manera sencilla.

**Figura 14. Administrador del Sistema OntoGuate - Salud.**



Referencia: figura obtenida de OntoGuate – Salud creado por Raúl Cuyún González, Febrero 2009.

Este módulo de administración brinda la posibilidad de Cargar, Mostrar y Borrar la ontología cuando el administrador del sistema lo crea oportuno.

### A.2.1 Cargar Ontología.

Dentro del sistema OntoGuate – Salud deberá cargarse la ontología a la base de datos en MySQL, para lo cual se tiene la siguiente opción dentro de la administración del sistema.

**Figura 15. Cargar Ontología, Sistema OntoGuate - Salud.**



Referencia: figura obtenida de OntoGuate – Salud creado por Raúl Cuyún González, Febrero 2009.

Este proceso tardará algunos minutos, dependiendo de la rapidez del servidor que procese la información. Al final de la carga el sistema muestra un resumen de las operaciones realizadas con éxito. De ésta

manera se tiene almacenada la ontología en la una base de datos MySQL.

## A.2.2 Mostrar Ontología.

Es importante verificar que la ontología se encuentre almacenada en la base de datos MySQL, para lo cual el sistema OntoGuate – Salud brinda la opción de “Mostrar”, proceso que despliega en pantalla la ontología de manera conceptual.

**Figura 16. Mostrar Ontología, Sistema OntoGuate - Salud.**

The screenshot shows the 'OntoGuate - Salud' web application. The header includes the title and a description: 'Gestion del conocimiento para la Clasificación Internacional de Enfermedades CIE10, desarrollado en Protege, RAP y PHP.' The navigation menu contains 'Principal', 'Busqueda CIE10', 'Admin', 'Nosotros', and 'Contactos'. A central text block states: 'Este trabajo es realizado como parte de un proyecto denominado OntoGuate-SGC que pretende conectar una Ontología de un sector específico a una base de datos, permitiendo la realización de búsquedas inteligentes, enfocado a la región Centroamericana, sur de México y Colombia.' Below this, it says: 'En este apartado podrá realizar algunas de las gestiones respecto de la ontología asociada, acciones como:' followed by buttons for 'Cargar', 'Mostrar' (highlighted with a magnifying glass), and 'Borrar'. Status messages indicate: 'Petición para conectar a BD...', 'Éxito al conectar a la BD', and 'Modelo cargado a memoria satisfactorio Desplegando a Pantalla =>'. A table displays RDF data with columns for 'Base URI', 'Prefix', 'Namespace', 'No.', 'Subject', 'Predicate', and 'Object'. The table shows one row with 'kb.CIE10' as the subject and 'RDF Node: rdfs:Class' as the object.

No.	Subject	Predicate	Object
1.	kb.CIE10	RDF Node: rdfs:Class	RDF Node: rdfs:Class

Referencia: figura obtenida de OntoGuate – Salud creado por Raúl Cuyún González, Febrero 2009.

En éste despliegue en pantalla de la ontología se puede diferenciar el sujeto, predicado y objeto.

### A.2.3 Borrar Ontología.

Cuando sea necesario realizar una nueva carga de datos, se deberá de borrar la versión que se tiene actualmente y cargar una nueva. Para esto es necesario borrar la ontología de la base de datos, el sistema OntoGuate – Salud brinda una opción de borrado de datos como los muestra la figura 17.

**Figura 17. Borrar Ontología, Sistema OntoGuate - Salud.**



Referencia: figura obtenida de OntoGuate – Salud creado por Raúl Cuyún González, Febrero 2009.

Estas son las tres opciones con las que el administrador cuenta para gestionar la ontología.

### A.3 Búsquedas inteligentes en OntoGuate - Salud.

Actualmente, para fines educativos, dentro de este trabajo de graduación se ha cargado a la ontología solamente los diagnósticos del capítulo cinco de la CIE10, dicho capítulo clasifica a todos los diagnósticos relacionados con los “TRASTORNOS MENTALES Y DEL COMPORTAMIENTO”.

#### A.3.1 Búsquedas de capítulos.

Los prefijos asociados para esta búsqueda son “capitulo” o “capitulos” los cuales deberán colocarse antes del texto a buscar por ejemplo: “capítulos enfermedades del sistema respiratorio”, el resultado se muestra en la figura 18.

**Figura 18. Búsqueda de capítulos, Sistema OntoGuate - Salud.**



Referencia: figura obtenida de OntoGuate – Salud creado por Raúl Cuyún González, Febrero 2009.

La búsqueda retorna dos tipos de resultados, en primer lugar se muestran los resultados “concretos” es decir, son las instancias que son iguales a el texto a buscar, mientras que las cercanas son una aproximación a las palabras ingresadas en el texto.

### A.3.2 Búsquedas de tres caracteres.

Para poder realizar búsquedas de los diagnósticos dentro de ésta clasificación es necesario anteponer la clase de búsqueda que será “tres caracteres” y luego el texto a buscar en las instancias para ésta clase. La figura 19 visualiza de una mejor manera el ejemplo “tres caracteres demencia vascular”

**Figura 19. Búsqueda de tres caracteres, Sistema OntoGuate - Salud.**



The screenshot displays the 'OntoGuate - Salud' web application interface. At the top, the title 'OntoGuate - Salud' is followed by the subtitle 'Gestion del conocimiento para la Clasificacion Internacional de Enfermedades CIE10, desarrollado en Protege, RAP y PHP.' Below this is a navigation menu with options: 'Principal', 'Busqueda CIE10', 'Admin', 'Nosotros', and 'Contactos'. A descriptive paragraph explains the project's goal: 'Este trabajo es realizado como parte de un proyecto denominado OntoGuate-SGC que pretende conectar una Ontologia de un sector especifico a una base de datos, permitiendo la realizacion de busquedas inteligentes, enfocado a la region Centroamericana, sur de Mexico y Colombia.' The main content area is titled 'Busqueda inteligentes a CIE10' and features a search input field containing 'tres caracteres demencia vascular' and a 'Buscar' button. Below the search results, there are two sections: 'Se han encontrado los siguientes resultados para la busqueda concreta de 'DEMENCIA VASCULAR' ..' and 'Se han encontrado los siguientes resultados para la busqueda cercana de 'DEMENCIA VASCULAR' ..'. The first section lists: CODIGO: F01, DIAGNOSTICO: DEMENCIA VASCULAR, CAPITULO: TRASTORNOS MENTALES Y DEL COMPORTAMIENTO. The second section lists: CODIGO: F00\*, DIAGNOSTICO: DEMENCIA EN LA ENFERMEDAD DE ALZHEIMER, CAPITULO: TRASTORNOS MENTALES Y DEL COMPORTAMIENTO. The third section lists: CODIGO: F02\*, DIAGNOSTICO: DEMENCIA EN OTRAS ENFERMEDADES CLASIFICADAS OTRA PARTE, CAPITULO: TRASTORNOS MENTALES Y DEL COMPORTAMIENTO. The fourth section lists: CODIGO: F03, DIAGNOSTICO: DEMENCIA NO ESPECIFICADA, CAPITULO: TRASTORNOS MENTALES Y DEL COMPORTAMIENTO. A small image of a person is visible on the right side of the page. The footer shows the Internet Explorer logo.

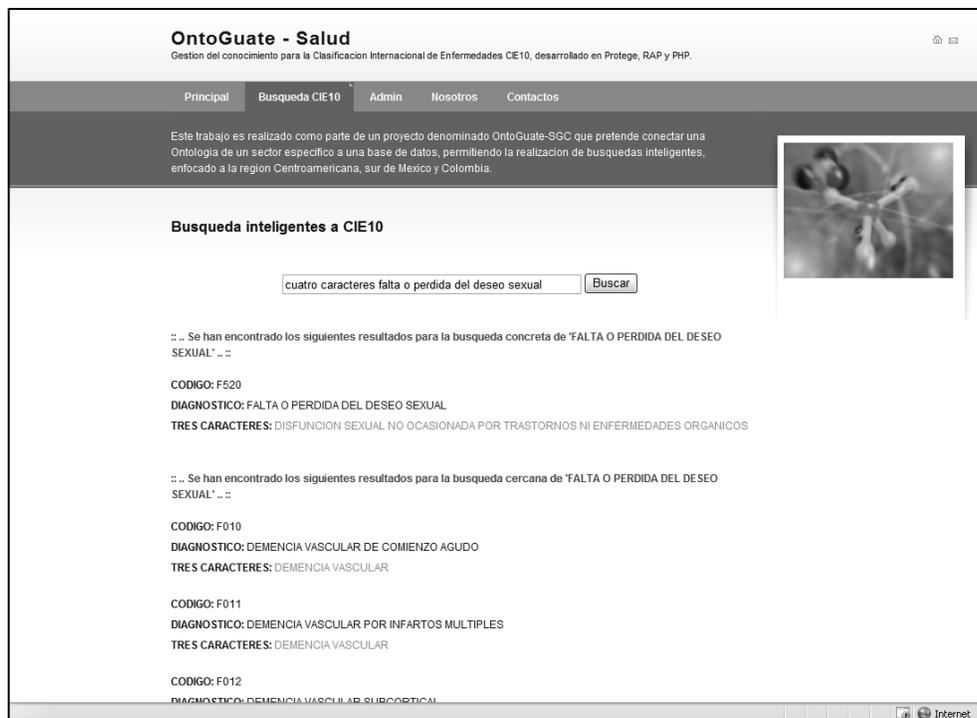
Referencia: figura obtenida de OntoGuate – Salud creado por Raúl Cuyún González, Febrero 2009.

De igual manera como en capítulos, el sistema realiza búsquedas concretas y luego las búsquedas cercanas, para éste caso ha encontrado un resultado concreto, y luego muchos resultados cercanos a cada una de las palabras ingresadas en la búsqueda.

### A.3.3 Búsquedas de cuatro caracteres.

Repetimos el proceso de las otras búsquedas, es decir ahora antepone las palabras clave “cuatro caracteres” para poder buscar dentro de las instancias de la clase “cuatro caracteres”. La figura 20 muestra la pantalla del resultado de la búsqueda “cuatro caracteres falta o perdida del deseo sexual”.

**Figura 19. Búsqueda de cuatro caracteres, Sistema OntoGuate - Salud.**



Referencia: figura obtenida de OntoGuate – Salud creado por Raúl Cuyún González, Febrero 2009.