

---

# ONTOLOGÍAS MÉDICAS: UNA REVISIÓN

## *MEDICAL ONTOLOGIES: A REVIEW*

**Harold Vásquez**

Universidad de Carabobo  
hvasquez@uc.edu.ve  
Carabobo-Venezuela

**Ana Aguilera**

Universidad de Carabobo  
aaguilef@uc.edu.ve  
Carabobo-Venezuela

**Leonid Tineo**

Universidad Simón Bolívar  
leonid@ldc.usb.ve  
Caracas-Venezuela

### RESUMEN

Existe un sin número de investigaciones destinadas a crear sistemas de información útiles al médico en su labores; orientadas a crear, reusar y consultar el conocimiento generado. Una de las herramientas ampliamente usadas en estos sistemas son las ontologías y de ahí que existen todo un número indefinido de estas orientas al campo medico. Debido a este gran número de repositorios de conocimiento y la vasta información que se ha generado al respecto, se hace necesario un compendio que muestre los últimos avances más significativos en esta área, de manera tal que futuras investigaciones no redunden en sus trabajos y más bien tomen como punto de partida estos avances. Esta revisión se elaboró como insumo para un trabajo de investigación que, entre otras cosas, pretende enriquecer las consultas difusas (SQLf) hechas a una base de datos médica del Hospital Ortopédico Infantil en Caracas, Venezuela; usando para ello Ontologías difusas.

**Palabras Clave:** Ontologías Médicas, Base de Datos Médicas, Bases de datos inteligentes, Sistemas expertos.

**Recibido:** 23/03/2010

### ABSTRACT

There are a lot of investigates to built useful information systems to help medico in his work, its oriented to create, reuse and consult the knowledge generated. One of the tools widely used in these systems is the ontologies and for this reason there are indefinite numbers of ontologies oriented to the medical field. Due to this great number of knowledge repositories and the vast information that has been generated at respect, becomes necessary a compendium to shows the last significant advances in this area, so future research do not repeat made works and they take initial point this advances instead. This summary was taken as input to one investigation where it is tried to enrich fuzzy queries (SQLf) done to a medical data base of Infantile Orthopedic Hospital at Caracas, Venezuela; using for it Fuzzy Ontologies.

**Keywords:** Medical Ontologies, Medical Database, Intelligent Databases, Expert systems.

**Aprobado:** 01/11/10

## INTRODUCCIÓN

En la actualidad existe un sin número de investigaciones que buscan: primero, recopilar y almacenar adecuadamente la información médica generada en consultas, tratamientos e investigaciones y; segundo, obtener conocimiento útil de esa data, para proveer el mayor bienestar posible al paciente. Estos esfuerzos se orientan en crear modelos que permitan formalizar de una manera aprovechable, todo el conocimiento inmerso en cada uno de los pasos del quehacer médico, de tal manera que pueda utilizarse posteriormente, sin trauma alguno, y hacer más efectivo la detección y tratamiento de las patologías ya conocidas. Uno de los mecanismos de formalización ampliamente utilizados últimamente es el de las Ontologías. Este ha demostrado ser una herramienta muy útil a la hora de reusar el conocimiento en ellas depositadas y cada vez existen más mecanismos que facilitan su construcción e incluso se acerca el momento en que se pueda automatizar completamente esta tarea.

En cada país o continente surgen grupos multidisciplinarios y mancomunados que buscan estandarizar el conocimiento y el lenguaje médico para lograr un mayor aprovechamiento de los recursos informativos y mayor bienestar a los pacientes indistintamente del país y la lengua hablada; es decir, las ontologías han logrado romper estas barreras idiomáticas y lingüísticas, logrando mayor cooperación entre los expertos médicos. Este estudio fue hecho en el marco de una investigación que busca extender las consultas difusas hechas con el lenguaje SQLf (Bosc & Pivert, 1995) a una base de datos médica del Hospital Ortopédico Infantil de la ciudad de Caracas en Venezuela, utilizando para ello Ontologías también difusas.

A continuación se presenta algunas definiciones previas que permitirán entender la forma como fueron agrupadas y estudiadas las distintas ontologías médicas aquí mostradas, seguidamente en la tercera parte se muestra el estudio comparativo por parte de las ontologías existentes en el campo de la Ingeniería Biomédica, se concluye en la cuarta parte mostrando un análisis de lo estudiado y se finaliza con las distintas referencias bibliográficas utilizadas en esta investigación.

## DEFINICIONES PREVIAS

### 1.Ontologías

Existe un sin número de definiciones al respecto, aunque parecidas, difieren en algunos puntos. Sin embargo, una de las más ampliamente aceptadas es dada en (Gruber, 1993): Una ontología es una especificación explícita de una conceptualización. El término proviene de la filosofía, donde una ontología es un recuento sistemático de la existencia. En sistemas

de Inteligencia Artificial, lo que existe es lo que puede ser representado. Cuando el conocimiento de un dominio se representa mediante un formalismo declarativo, el conjunto de objetos que puede ser representado se llama universo del discurso. Esos conjuntos de objetos, y las relaciones que se establecen entre ellos, son reflejados en un vocabulario con el cual representamos el conocimiento en un sistema basado en conocimiento.

Así, en el contexto de IA, podemos describir la ontología de un programa como un conjunto de términos. En tal ontología, las definiciones asocian nombres de entidades del universo del discurso con textos comprensibles por los humanos que describen el significado de los nombres, y axiomas formales que limitan la interpretación y buen uso de dichos términos. Formalmente, una ontología es una teoría lógica.

Sin embargo, y luego de varias críticas y sugerencias, este mismo autor redefine este concepto diciendo:

Bien, ... los componentes más importantes de esa definición de ontología son que la ontología es un artefacto de representación (una especificación), distinta del mundo que modela, y que es un artefacto diseñado, construido para un propósito. Creo que la mayoría de científicos en computación obtienen la diferencia entre una especificación del mundo, incluso para mundos sintéticos. Retrospectivamente, no cambiaría la definición pero intentaría enfatizar que nosotros diseñamos ontologías. La consecuencia de esta vista es que podemos aplicar una disciplina de ingeniería en su diseño y evaluación. Si las ontologías son cosas derivadas de una ingeniería, entonces no tenemos que preocuparnos tanto sobre si son correctas y favorecer el negocio de construir las para hacer algo útil. Podemos diseñarlas para conocer objetivos funcionales y restricciones. Podemos construir herramientas que nos ayuden a gestionarlas y validarlas. Y podemos tener múltiples ontologías que se coordinen o compitan basadas en un criterio objetivo más que en una marca de fábrica o una autoridad. (Gruber, 2004:4)

En todo caso, aun no existe un consenso absoluto sobre una única definición. Lo que si es bien sabido, es que estas herramientas permiten representar conocimiento y de ahí su utilidad en el campo médico, entre otras muchas áreas más.

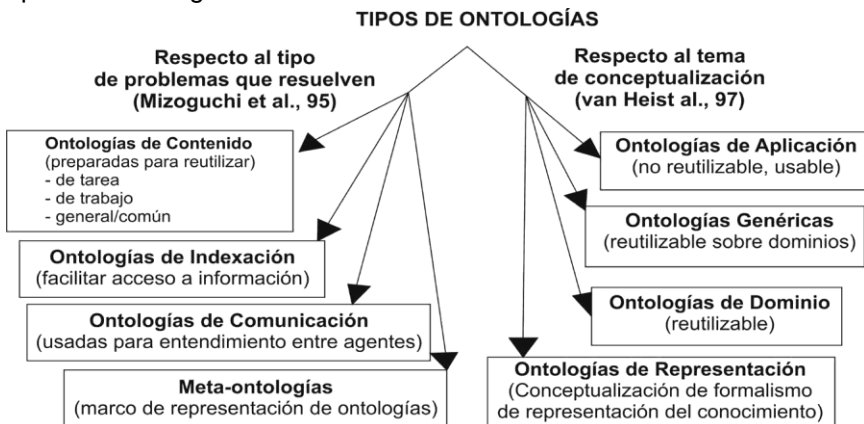
## **2. Tipos de ontologías**

Al igual que como en el caso anterior, no existe un consenso sobre como clasificar los distintos tipos de Ontologías. Sin embargo, en (Valencia, 2005) se muestra un compendio muy completo sobre las distintas clasificaciones que existen, a saber:

- 1) Clasificación por el conocimiento que contienen.

- a. Ontologías terminológicas, lingüísticas.
  - b. Ontologías de información.
  - c. Ontologías para modelar conocimiento.
  - o
  - a. Ontologías del dominio.
  - b. Ontologías de tarea.
  - c. Ontologías generales.
  - d. Clasificaciones por motivación
- 2) Ontologías para la representación de conocimiento.
- a. Ontologías genéricas.
  - b. Ontologías del dominio.
  - c. Ontologías de aplicación.
  - o
  - a. Ontologías generales.
  - b. Ontologías categóricas.
  - c. Ontologías del dominio.
  - d. Ontologías genéricas.
  - e. Ontología regional.
  - f. Ontología aplicada.
  - g. Clasificaciones por el grado de formalidad de la ontología.
  - h. Ontología descriptiva.
  - i. Ontología formal.

Por razones de espacio, no es posible mostrar las definiciones de cada tipo, pero puede ser consultado en (Valencia, 2005). Así como lo anterior descrito, en (Lozano, 2002) se muestra otra clasificación que se puede apreciar en la figura 1.



**Figura 1. Clasificaciones de Ontologías respecto al tipo de problemas que resuelven y respecto al tema de conceptualización.** Fuente: Lozano (2002)

Ahora bien, en Borgo (2004), se detalla otra clasificación pensada en las ontologías médicas, aunque se puede apreciar que es aplicable a cualquier ontología indistintamente de su utilidad final:

### **Ontologías Lingüísticas.**

- a. Glosarios.
- b. Vocabularios Controlados
- c. Taxonomías.
- d. Tesoros

### **Ontologías dirigidas a implementación.**

- a. Esquemas Conceptuales.
- b. Base de Conocimiento.

### **Ontologías Formales.**

- a. Ontologías de Dominio.
- b. Ontologías Bases (referencia).
- c. Ontologías Fundacionales.

Este mismo autor señala que los catálogos y los mapas de tópicos no deberían considerarse ontologías.

En definitiva, así como sucede en la definición y otros aspectos del mundo ontológico, aún falta mucho camino por recorrer para llegar a un consenso mayoritario que permita lograr un acuerdo en, al menos, una clasificación de las ontologías.

### **3) Elementos de una Ontología**

En razón de que algunos autores consideran como una Ontología a cualquier elemento que contenga conocimiento, es importante resaltar que para diferenciar estos de una verdadera Ontología, es deseable que la misma contenga los siguientes elementos:

1. Clases: También llamados conceptos. Suelen organizarse en taxonomías, por lo que no debe considerarse a estas solas como la ontología completa, si no más bien podrían tratarse como tesauros.

2. Atributos: Conforman la estructura interna de la clase y los mismos pueden ser propios o heredados cuando la clase funge como hijo de otra, en razón de la taxonomía antes descrita.

3. Relaciones: conocidos formalmente como un subconjunto del producto cartesiano de dos o más conjuntos, es decir,  $R \subseteq A_1 \times A_2 \times \dots \times A_n$ , tal que se cumple una propiedad en específico. En el caso de relaciones binarias, se suele utilizar propiedades como «subclase de» o «conectado a».

4. Funciones: es un tipo especial de relaciones, con la restricción de que al último elemento de la n-tupla, solo le corresponde una única (n-1)-tupla. Propiedades como «madre de» y «precio de un coche usado» generan funciones.

5. Axiomas: son expresiones asumidas como ciertas y no todas las ontologías las incluyen. Suelen usarse para definir el significado de los componentes ontológicos, definir restricciones complejas sobre los valores de los atributos, argumentos de relaciones, etc., verificando la corrección de la información especificada en la ontología o deduciendo nueva información. También suele usarse Lógica de Primer Orden (FOL) para describirlas.

6. Instancias: Al momento de tomar valores los atributos de una clase, se dice que la misma es instanciada, es decir, toma un valor concreto de la vida real.

#### 4) Metodologías de Construcción

La tarea de construir desde cero una Ontología no es fácil. Tampoco el hecho de crearla a partir de otra (reuso). De ahí que se necesiten mecanismos que faciliten estas tareas, es decir, deben existir métodos que indiquen los pasos más adecuados a seguir para lograr la creación en poco tiempo y con buena calidad de estos repositorios de conocimiento. A continuación se presentan algunas de las metodologías más destacadas en este campo:

1. La metodología Cyc (Lenat & Guha, 1990) consiste en varios pasos. En primer lugar, hay que extraer manualmente el conocimiento común que está implícito en diferentes fuentes. A continuación, una vez que tengamos suficiente conocimiento en nuestra ontología, podemos adquirir nuevo conocimiento común usando herramientas de procesamiento de lenguaje natural o aprendizaje computacional. Así se construyó la ontología Cyc.

2. La metodología de Uschold & King (1995) propone algunos pasos generales para desarrollar ontologías, a saber: (1) identificar el propósito; (2) capturar los conceptos y relaciones entre estos conceptos y los términos utilizados para referirse a estos conceptos y relaciones; (3) codificar la ontología. La ontología debe ser documentada y evaluada, y se pueden usar otras ontologías para crear la nueva. De esta forma se creó la Enterprise Ontology.

3. En la metodología presentada por Grüninger & Fox (1995), el primer paso es identificar intuitivamente las aplicaciones posibles en las que se usará la ontología. Posteriormente, se usa un conjunto de preguntas en lenguaje natural, llamadas cuestiones de competencia, para determinar el ámbito de la ontología. Se usan estas preguntas para extraer los conceptos principales, sus propiedades, relaciones y axiomas, los cuales se definen formalmente en Prolog. Por consiguiente, esta es una metodología muy formal que se aprovecha de la robustez de la lógica clásica y que puede ser usada como guía para transformar escenarios informales en modelos computables. Esta metodología se usó para construir la ontología TOVE.

4. Metodología KACTUS (Bernaras *et al.*, 1996). En esta metodología se construye la ontología sobre una base de conocimiento por medio de

un proceso de abstracción. Cuantas más aplicaciones se construyen, las ontologías se convierten en más generales y se alejan más de una base de conocimiento. En otras palabras, se propone comenzar por construir una base de conocimiento para una aplicación específica. A continuación, cuando se necesita una nueva base de conocimiento en un dominio parecido, se generaliza la primera base de conocimiento en una ontología y se adapta para las dos aplicaciones, y así sucesivamente. De esta forma, la ontología representaría el conocimiento consensuado necesario para todas las aplicaciones. Esta metodología ha sido utilizada para construir una ontología para diagnosticar fallos.

5. Methontology es una metodología para construir ontologías tanto partiendo desde cero como reusando otras ontologías, o a través de un proceso de reingeniería. Este entorno permite la construcción de ontologías a nivel de conocimiento, e incluye: (1) identificación del proceso de desarrollo de la ontología donde se incluyen las principales actividades (evaluación, gestión de configuración, conceptualización, integración, implementación, etc); (2) un ciclo de vida basado en prototipos evolucionados; y (3) la metodología propiamente dicha, que especifica los pasos a ejecutar en cada actividad, las técnicas usadas, los productos a obtener y cómo deben ser evaluados. Esta metodología está parcialmente soportada por el entorno de desarrollo ontológico WebODE y ha sido usada en la construcción de múltiples ontologías, como una ontología química, ontologías de hardware y software, entre otras

6. La metodología basada en Sensus (Swartout *et al.*, 1997) es un enfoque top-down para derivar ontologías específicas del dominio a partir de grandes ontologías. Los autores proponen identificar un conjunto de términos semilla que son relevantes en un dominio particular. Tales términos se enlazan manualmente a una ontología de amplia cobertura. Los usuarios seleccionan automáticamente los términos relevantes para describir el dominio y acotar la ontología Sensus. Consecuentemente, el algoritmo devuelve el conjunto de términos estructurados jerárquicamente para describir un dominio, que puede ser usado como esqueleto para la base de conocimiento. Esta metodología sirvió para construir la ontología Sensus.

7. Metodología On-To-Knowledge (Staab *et al.*, 1997). El proyecto OTK aplica ontologías a la información disponible electrónicamente para mejorar la calidad de la gestión de conocimiento en organizaciones grandes y distribuidas. La metodología proporciona guías para introducir conceptos y herramientas de gestión de conocimiento en empresas, ayudando a los proveedores y buscadores de conocimiento a presentar este de forma eficiente y efectiva. Esta metodología incluye la identificación de metas que deberían ser conseguidas por herramientas de gestión de conocimiento y está basada en el análisis de escenarios de uso y en los diferentes

papeles desempeñados por trabajadores de conocimiento y accionistas en las organizaciones. Cada una de las herramientas de la arquitectura de OKT se centra en el desarrollo de aplicaciones dirigidas por ontologías y, finalmente, describe el uso y la evaluación de la metodología mediante casos de estudio como, por ejemplo, la ontología Proper o AIFB.

8. Terminae (Aussenac-Gilles *et al.*, 2002) aporta tanto una metodología como una herramienta para la construcción de ontologías a partir de textos. Se basa en un análisis lingüístico de los textos, el cual se realiza mediante la aplicación de diferentes herramientas para el procesamiento del lenguaje natural. En particular, se usan dos herramientas: (1) Syntex para identificar términos y relaciones; y (2) Caméléon para identificar roles o relaciones.

La metodología funciona como sigue. Mediante la aplicación de Syntex obtenemos una lista de posibles palabras y frases del texto y algunas dependencias sintácticas y gramaticales entre ellas. Estos datos se usan como entrada para el proceso de modelado junto con el texto original. El siguiente paso es normalizar el conocimiento para obtener una ontología bien estructurada, donde cada concepto quede justificado por sus relaciones con otros conceptos. Esta metodología sugiere aplicar criterios diferenciadores para hacer explícitas las propiedades comunes y diferentes de un concepto con sus respectivos conceptos padre y hermanos debidas a sus roles. La última etapa es la formalización de la ontología en el lenguaje formal Terminae, que es un tipo de lógica descriptiva. Una función de clasificación sirve para comprobar la corrección de las definiciones de conceptos genéricos, ya que solo pueden ser definidos si tienen roles diferenciados.

9. ONIONS (Ontological Integration Of Naive Sources) (Steve *et al.*, 1997) es una metodología para la integración de varias ontologías. Viene siendo desarrollada desde 1990 y es una metodología para la integración del conocimiento terminológico de repositorios con conceptualizaciones heterogéneas. Esta metodología se funda en consideraciones filosóficas para una teoría semántica (es decir, nuestra teoría del significado) que conduzca a la estructura de la misma metodología. Crea un marco común para generalizar e integrar las definiciones que se utilizan para organizar un conjunto de fuentes terminológicas. Es decir, permite trabajar fuera de coherencia una ontología terminológica del dominio (una ontología terminológica se define generalmente como la conceptualización explícita de un vocabulario) para cada fuente, que se puede ser comparado con las otras y mapeado a una librería integrada antológica.

10. OntoClean (Guarino & Welty, 2002) es una metodología para validar la suficiencia ontológica de relaciones taxonómicas. Se basa en nociones ontológica altamente generales dirigidas de la filosofía, tal como esencia, identidad, y unidad, que se utilizan para caracterizar aspectos relevantes del significado previsto de las propiedades, clases, y relaciones que hacen



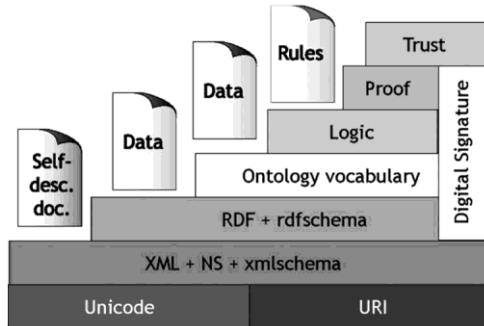
una ontología. Estos aspectos son representados por los metapropiedades formales, que imponen varias restricciones en la estructura taxonómica de una ontología. El análisis de estas restricciones ayuda en la evaluación y validación de las escogencias hechas.

Estas no son todas las metodologías existentes. Existen un gran número no mostrado aquí, pero no se tratan porque no es la finalidad de esta investigación. Aquí se detallaron las más significativas en razón de que son las más conocidas, por ser las más referenciadas y usadas para construir la mayoría de las ontologías.

## 5) Lenguajes de Representación

Han surgido varias propuestas concernientes a la forma más adecuada para representar las ontologías en un lenguaje entendible por la máquina y por el hombre al mismo tiempo; pero en la actualidad se está destacando una sobre todas, en razón de su orientación a la estandarización bajo los auspicios de una organización reconocida. En la figura 2 se muestra una evolución de estos lenguajes.

Al principio de la década de los 90, los lenguajes más representativos eran Cycl (Lenat & Guha, 1990), LOOM (McGregor, 1991), Ontolingua (Gruber, 1993), FLogic (Kifer *et al.*, 1995), OCML (Motta, 1999), entre otros.



**Figura 2.** The architecture of levels of T. Berners lee. Fuente: Berners (2000).

La mayoría siguieron una sintaxis basada en LISP y fueron llamados «lenguajes clásicos». Más recientemente, se aprovechó un lenguaje de marcado ampliamente usado en la web (XML) para crear lenguajes para las Ontologías, tales como OML/CKML (Kent, 1998), RDF (Lassila & Swick, 1999), RDF Schema (Brickley & Guha, 1999), SHOE (Luke & Heflin, 2000), OIL (Horroks *et al.*, 2000), DAMIL+OIL (Horroks & van Harmelen, 2001) y OWL. A continuación se detallan algunas de estas:

1. KIF (Knowledge Interchange Format) (Genesereth, 1991). Desarrollado en la Universidad de Stanford, es una lengua centrada en el computador para el intercambio del conocimiento entre programas.

Tiene semántica declarativa (es decir el significado de expresiones en la representación se puede entender sin la ayuda de un intérprete para manipular esas expresiones); es comprensivo lógicamente (es decir, permite la expresión de oraciones arbitrarias en el cálculo de predicado de primer orden); prevé la representación del conocimiento sobre la representación del conocimiento; prevé la representación de la reglas de razonamiento nonmonotonic; y permite la definición de objetos, funciones, y relaciones.

2.DAML+OIL. El lenguaje DAML está siendo desarrollado como una extensión a XML y RDF. El último lanzamiento del lenguaje (DAML+OIL) proporciona un sistema rico de constructores para crear ontologies y marcar información de modo que sea legible y comprensible por la máquina.

3. RDFS. (RDF Schema o Esquema RDF) es una extensión semántica de RDF. Un lenguaje primitivo de ontologías que proporciona los elementos básicos para la descripción de vocabularios. La primera versión fue publicada en abril de 1998 por la W3C, la versión actual de la recomendación fue publicada en febrero de 2004 también por la W3C (Wikipedia, 2008).

4.OWL. OWL es el acrónimo del inglés Ontology Web Language, un lenguaje de marcado para publicar y compartir datos usando ontologías en la Web. OWL tiene como objetivo facilitar un modelo de marcado construido sobre RDF y codificado en XML. Tiene como antecedente DAML+OIL, en los cuales se inspiraron los creadores de OWL para crear el lenguaje. Actualmente, OWL tiene tres variantes: OWL Lite, OWL DL y OWL Full.

OIL, DAML+OIL y OWL están basados en lógica descriptiva (Horrocks, 2005). De todos los lenguajes antes descritos, este último es el más ampliamente usado actualmente, ya que presenta una evolución más completa, esta estandarizado por la W3C logrando mayor expresividad y formalismo, permite usarse en diferentes niveles de experticia y al estar basado en XML, logra una importante portabilidad que le ha permitido estar soportado por la mayoría de las herramientas de creación, edición y visualización de ontologías.

## **ONTOLOGÍAS MÉDICAS**

Primeramente se mostrarán algunas taxonomías de términos médicos, que buscan estandarizar de alguna manera el lenguaje que permita un entendimiento claro y sin ambigüedades. Aunque algunos autores consideran estos como Ontologías, estos carecen de algunos elementos que como ya se detalló, deben estar presentes; especialmente atributos y axiomas. Seguidamente se muestra las ontologías estudiadas en esta investigación y un cuadro comparativo de las mismas.

## **Taxonomías Médicas**

La comunicación no ambigua de conceptos médicos detallados y complejos es crucial hoy en día para sistemas de información médica y entre los mismos especialistas. A continuación presentamos los más representativos (Pérez, 2002):

1. GALEN (Lenat & Guha, 1990) incluye un modelo semánticamente válido de terminología clínica representado en un lenguaje formal, y asociado con el soporte sofisticado para diferentes lenguajes naturales y conversión entre diferentes esquemas de codificación. GALEN se basa en un modelo semántico sólido de terminología clínica llamado GALEN Coding reference (CORE). Este modelo contiene conceptos clínicos elementales (p.ej., fractura, hueso, etc), relaciones que controlan la combinación de conceptos (p.ej., los huesos pueden tener fracturas), y conceptos complejos (p.ej., fractura de la clavícula).

2. El Unified Medical Language System (UMLS) (Pisanelli *et al.*, 1998) es una base de datos diseñada para unificar terminologías biomédicas de fuentes dispares tales como terminologías clínicas, fuentes de drogas, o vocabularios en diversos idiomas. Existen tres fuentes de conocimiento UMLS, a saber: (1) metatesauro, que contiene información semántica sobre conceptos biomédicos, sus nombres y relaciones entre ellos; (2) red semántica, que es una red de categorías generales o tipos semánticos a lo que se asignan todos los conceptos del metatesauro; (3) el lexicón especialista, que contiene información sintáctica sobre términos biomédicos y eventualmente cubrirá la mayoría de términos correspondientes a los nombres de conceptos que aparecen en el metatesauro.

3. MeSH (Medical Subject Heading). es el tesoro de vocabulario controlado de la Librería Nacional de Medicina de los EEUU. Contiene encabezamientos de materias, calificadores (subencabezamientos), definiciones, referencias cruzadas, sinónimos y listas de términos estrechamente relacionados. Consta de más de 33 000 términos ordenados en estructuras jerárquicas llamadas árboles, que se revisan anualmente para asegurar que constituyan un fiel reflejo de la práctica y la terminología médica actual. Presenta una traducción al español que se le conoce como Descriptores en Ciencias de la Salud (DeCS)

4. Kingsbury Center for Cancer Care Glossary. Obtenida del Kingsbury Information Center. Es un glosario básico de términos relacionados con el diagnóstico y el tratamiento del cáncer, dirigido a enfermos y sus familiares.

5. MedicineNet Medical Dictionary, ©Information Network, Inc. Glosario de términos médicos para pacientes y semi-especialistas.

6. Multilingual Glossary of Technical and Popular Medical Terms in nine European Languages. Elaborado por el Heymans Institute of Pharmacology, University of Ghent and Mercator College, Dept. of Applied Linguistics.

Este glosario incluye términos médicos en las nueve lenguas de la UE y, en caso de que exista, añade la denominación correspondiente en lengua general (p. ej. antiséptico – desinfectante).

7. International Classification of Diseases (ICD). Clasificación basada en la versión oficial de enfermedades de la Organización Mundial de la Salud (World Health Organization -WHO). Existen varias versiones, pero la última y más completa es la ICD-10. Está diseñado para clasificar la información relativa a la incidencia de enfermedades e índices de mortalidad y para indizar todo tipo de documentos clínicos para facilitar su clasificación y recuperación. Además de los códigos numéricos para la clasificación de documentos clínicos, el ICD-10 es una completa clasificación de todas las enfermedades conocidas, causas de accidentes y envenenamientos, procedimientos médicos y métodos diagnósticos. El ICD-10 ofrece diferentes estructuraciones de la información (alfabética, tabular, temática) para facilitar la consulta.

Como ya se mencionó, existe un exorbitante número de ontologías en el campo Médico, pero a la hora de buscar una ontología en un campo de la medicina en específico, como por ejemplo en ortopedia, resulta un trabajo un tanto desalentador. Esto es debido a la falta de un repositorio o indexación único a nivel mundial. Existen algunas a nivel europeo (LOA, 2008) (RIHS, 2008) y también existen algunas colecciones de ontologías biomédicas (OWG, 2008) (IO, 2008) (OBO, 2008) (TOBO, 2008) (NCBO, 2008). Incluso algunos investigadores muestran colecciones de ontologías en páginas personales (Golbeck *et al.*, 2003). Resulta interesante el trabajo hecho en (TOBO, 2008), donde se busca crear ontologías biomédicas bajo ciertos lineamientos que permiten una interoperabilidad entre ellas. De ahí que este posea un importante repositorio de ontologías, de aproximadamente 60 de ellas. Incluso estas ontologías son editadas y mostradas bajo un estándar de archivo de texto llamado formato OBO, aunque muchas de ellas tienen sus versiones en OWL.

Todo lo anterior trae como consecuencia la imposibilidad de contabilizar de alguna manera la cantidad total o por área específica de las ontologías médicas existentes, genera redundancia en los esfuerzos, hay una pérdida de conocimiento ya formalizado y experticia en la creación de nuevas ontologías. Es por esta razón que a continuación se presenta un subconjunto de las ontologías médicas, tratando de mostrar las más representativas, tomando en cuenta la cantidad de investigadores inmersos en los desarrollos de las mismas, su visualización en el mundo científico a través de congresos, papers, referencias bibliográficas y demás, y su utilidad en el campo médico (Horrocks, 2005).

### **1.Ontología TAMBIS (Tao).**

Es una ontología para conceptos de Biología Molecular y bioinformática. Fue creada como parte de un completo sistema de consultas que tiene el

mismo nombre. Esta modelada con el lenguaje DAML+OIL. Es usado para:

a. Describir la matadata de fuentes de datos, representando un esquema universal sobre los archivos.

b. Formar consultas expresados en el lenguaje de modelado creado por ellos.

c. Impulsa una interfaz de usuario para la formulación de consulta, dirigida a usuarios que no están preparados para manejar lenguajes de consultas complejos.

d. Mediar entre las distintas fuentes para traducir el modelo mediador al modelo fuente.

Una característica importante de esta ontología es que no contiene instancias, es decir, solo contiene conocimiento acerca de conceptos bioinformáticos y de biología molecular y sus relaciones. Las instancias están representadas en las distintas fuentes de datos externas que utiliza el sistema de consulta. Existe en dos versiones, una pequeña de 250 conceptos y 60 relaciones, y la completa de 1500 conceptos. La dirección electrónica donde se aloja es <http://www.cs.man.ac.uk/~stevensr/tambis/index.html>.

## **2. Ontología RNA(RO).**

Aunque es una ontología aún no creada, se acota por su relevancia. Su finalidad es crear una ontología para entender el funcionamiento de las secuencias RNA en sistemas biológicos. Una de sus intenciones para crearla es utilizar tres ontologías existentes: La ontología Gene (GO), la ontología Sequence (SO) y la ontología estructural usada por la PDB (Protein Data Bank). GO provee un vocabulario controlado para describir Genes y cualidades de productos genéticos en cualquier organismo. Este a su vez está conformado por tres ontologías: una que describe productos genéticos en términos de su proceso biológico asociado, otro para componentes celulares y el último describe funciones moleculares de forma independiente de la especie (<http://www.geneontology.org>). SO busca desarrollar una ontología conveniente para describir secuencias biológicas. Existe una versión ligera de esta ontología llamada SOFA (Sequence Ontology Feature Annotation) (<http://www.sequenceontology.org>). Otro aspecto importante de RO, es que utilizarán un ciclo de construcción bastante sencillo e intuitivo, tal como se aprecia en la figura 3. Esta ontología va a ser escrita en OWL. La página que aloja este macro proyecto es <http://roc.bgsu.edu/>



**Figura 3.** Proceso de desarrollo de ontología de Noy and McGuinness. Fuente: Noy and McGuinness (2001).

### 3. Ontology for Biomedical Investigations (OBI).

Al igual que en caso anterior, esta ontología está en construcción, aunque ya tiene versiones preliminares. Su finalidad principal es desarrollar una ontología integrada para la descripción de investigaciones biológicas y clínicas. Esta modelará el diseño de una investigación, los protocolos e instrumentación usados, el material usado, la data generada y el tipo de análisis hecho. Esta se considera como una ontología OBO (<http://obi.sourceforge.net/index.php>)

### 4. MGED.

Es una ontología que contiene conceptos, definiciones, términos, y recursos para la descripción estandarizada de ejemplos usados en un experimento microarray y es usado como soporte de un sistema de software denominado MAGE v.1. Está dividida en dos partes, un núcleo que es el utilizado para MAGE y una extensión que añadirá nuevas asociaciones y clases no disponibles para el software (<http://mged.sourceforge.net/ontologies/index.php>).

### 5. Ontología de enfermedades (Disease Ontology).

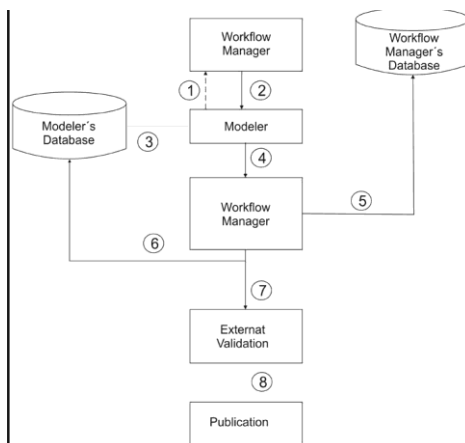
Es un vocabulario médico controlado desarrollado en el campo bioinformático en colaboración con el proyecto NuGene del Centro de Medicina Genética de la Universidad de Northwestern, Chicago. Fue diseñado para facilitar un trazado entre las enfermedades y condiciones asociadas para códigos médicos particulares tal como ICD9CM, SNOMED y otros (<http://diseaseontology.sourceforge.net/>).

## 6. SNOMED (Systematized Nomenclature of Medicine Clinical Terms).

Aunque se pudo haber ubicado como una taxonomía, se decidió colocarlo aquí, ya que parte de él está diseñado siguiendo principios ontológicos y sigue en esa dirección. Esta ontología incluye una terminología básica de cerca de 364.000 conceptos del cuidado médico, con significados únicos y definiciones basados en lógica formal, organizados en jerarquías. A la fecha de enero de 2005, la tabla completa con las descripciones únicas para cada concepto, contenía más de 984.000 descripciones. Aproximadamente 1.45 millones de relaciones semánticas existen para permitir confiabilidad y consistencia de la data recuperada. Está disponible en ediciones de lenguas inglesas, españolas y alemanas. SNOMED International es una división del College of American Pathologists (CAP). Durante varios años, un grupo de médicos y traductores en Buenos Aires, Argentina, han trabajado junto al CAP para desarrollar la edición en español. La calidad del resultado de este trabajo fue validada por un grupo de 14 médicos de 5 países, incluidos Argentina, España, México, Venezuela y Panamá (<http://www.snomed.org/>).

## 7. NCI (National Cancer Institute).

A pesar de que sus comienzos fueron como un metatesauro basado en UMLS, los encargados del mismo decidieron convertirlo en una ontología (Golbeck *et al.*, 2003). Esta contiene terminología producida por el Instituto Nacional de Cáncer. Es uno de los más grandes y entre los tópicos descritos está el de enfermedades, drogas, químicos, diagnósticos, genes, tratamientos, anatomía, organismos y proteínas. En la figura 4 se muestra el proceso que se siguió para construir el tesauro. La ontología tiene aproximadamente 450.000 tripletas en un archivo de más de 33 Mb. Una descripción y versión actual está disponible en <http://www.mindswap.org/2003/CancerOntology>.



**Figura 4.** Versión abreviada del diagrama de Workflow del ciclo de edición y publicación del tesauro NCI. Fuente: Golbeck, 2003.

**2. GALEN.** Esta ontología está presente en dos versiones básicas. La original es producto de la Fundación OpenGalen ([www.opengalen.org](http://www.opengalen.org)) y de ella se generó una versión en OWL (<http://www.co-ode.org/galen/>) manteniendo su estructura original (Rector *et al.*, 1995) (TKBG, 2008). GALEN es el nombre dado a una tecnología que es diseñada para representar la información clínica de una nueva manera, y su intención es «colocar lo clínico en el sitio de trabajo clínico». GALEN produce un sistema de codificación multilingüe computarizado para la medicina, usando un enfoque cualitativo diferentes de los usados anteriormente. Este está dividido en cuatro partes: La ontología de alto nivel, el modelo de referencia común, extensiones detalladas y el modelo de procedimientos quirúrgicos y otros modelos similares que definen conceptos compuestos surgidos de las partes del modelo de referencia común y de sus extensiones. La información anterior se muestra en la tabla 1.

**Tabla 1**  
**Aspectos elementales de las ontologías médicas estudiadas**

Nombre	Lugar, Año.	Lenguaje de Modelado	Metodología Usada	Area de Especialización
TAMBIS	Universidad de Manchester, Reino Unido. 1999	DAML+OIL	?	Información biológica y Bioinformática
RO	Mundial	OWL	Noy McGuiness	and RNA
GO	USA	OWL, RDF	?	Genes
SO	USA	OBO, OWL	?	Secuencias Biológicas
OBI	Mundial	OBO, OWL	Propia	Investigaciones Biológicas y Clínicas
MGED	Mundial	OWL, DAML+OIL, RDFS	?	Experimentos microarray
Disease Ontology	USA	OBO, OWL	?	Enfermedades
SNOMED	USA	Logica Descriptiva - ELH	?	Todo el campo medico
NCI	USA	OWL	Propio	Cáncer
GALEN	España	OWL	Propio	Todo el campo medico

Fuente: Vásquez, Aguilera y Tineo (2010).



## CONCLUSIONES

Es notable la gran cantidad de información que existe sobre ontologías, solamente en el campo médico. Esto puede ser visto como una ventaja, sin embargo, si todos estos modelos de conocimientos no son aprovechados de una manera idónea, de nada sirven los números grandes. Se hace imperiosa la necesidad de lograr un registro mundial de todas estas ontologías, que permita clasificarlas y contabilizarlas para lograr una efectivo reúso de las mismas y evitar redundancia de trabajo al momento de crear nuevas ontologías. Con esto también se lograría un mayor entendimiento en terminología médica y una efectiva utilización de los muchos sistemas que usan estas ontologías para distintos propósitos, pero todos orientados finalmente a buscar el bienestar de algunas iniciativas para indexar las ontologías médicas, estas están aún rezagadas respecto a la velocidad con que se crean nuevas ontologías y no han sido suficientemente publicitadas de manera tal que los ontólogos se ven interesados en indexar sus productos. Otro problema encontrado en esta investigación es el límite difuso que existe para catalogar algo como una ontología o como un tesoro (taxonomía). Existe mucha divergencia en este aspecto, es decir, hace falta una estandarización de cuando catalogar a una colección como ontología o simplemente como una colección de términos que no cumple los requisitos de ontología. Se observa además, que OWL es el lenguaje pionero para modelar las ontologías, ya que la mayoría de ellas, a pesar de tener un lenguaje de modelado distinto a este, también proveen una versión en OWL, asegurando así interoperabilidad y portabilidad de las mismas. Por último, se aprecia una ausencia de información respecto a las metodologías usadas.

## REFERENCIAS

- AI Center, SRI International, Menlo Park (1999) *XOL: An XML-Based Ontology Exchange Language*. Technical Note 559., CA, USA. Karp, P., Chaudhri, V., Thomere, J.
- Aussenac-Gilles, N., Biebow, B., Szulman, S. (2002), *Modelling the travelling domain from a NLP description with Terminae*. Workshop on Evaluation of Ontology Tools, European Knowledge Acquisition Workshop, Sigüenza, Spain.
- Bernaras, A., Laresgoiti, I., y Corera, J. (1996, Agosto), *Building and Reusing Ontologies for Electrical Network Applications*. Proceedings of the European Conference on Artificial Intelligence (ECAI96). Budapest, Hungary.

- Borgo, S. (2004, Julio) *Classifying (Medical) Ontologies*. Ontology Workshop at the Semantic Mining Summer School. Institute for Cognitive Sciences and Technology (ISTC-CNR). Laboratory for Applied Ontology (LOA). Trento-Roma, Italy.
- Bosc, P., Pivert, O. (1995). *SQLf: A Relational Database Language for Fuzzy Querying*. IEEE Transactions on Fuzzy Systems. 3(1), 1-17.
- Brickley, D., Guha, R. (1999) *Resource Description Framework (RDF) Schema Specification*. W3C Proposed Recommendation. [Documento en línea]. Disponible: <http://www.w3.org/TR/PR-rdf-schema>. [consulta: 2008, Febrero 10]
- Departamento de Informática, Universidad de Jaume y Knowledge Bases Group, Universidad de Birmingham. [Datos en línea] *TKBG - Temporal Knowledge Bases Group*. Disponible: <http://krono.act.uji.es/Links/ontologies>. [Consulta: 2008, Abril 04].
- Genesereth, M. (1991). *Knowledge Interchange Format*. En J. Allen et al., (comp). Proceedings of the Second International Conference on the Principles of Knowledge Representation and Reasoning (KR-91). [Libro en línea]. Morgan Kaufman Publishers, San Mateo, CA, USA. 238-249. Disponible: <http://logic.stanford.edu/kif/kif.html>. [consulta: 2008, Febrero 10]
- Golbeck, J., Fragoso, G., Hartel, F., Hendler, J., y Oberthaler J., Parsia, B. (2003). *The National Cancer Institute's Thesaurus and Ontology*. Journal of Web Semantics. 1(1): 75-80.
- Gruber, Thomas. (2004). *Interview to Tom Gruber*. AIS SIGSEMIS Bulletin. 1(3): 4.
- Gruber, Thomas. (1993) *A Translation Approach to Portable Ontology Specifications*. Knowledge Acquisition. 5(2): 199-220.
- Gruninger, M., Fox, M. (1995). The logic of enterprise modeling. En J. Brown & D.O. Sullivan (Comp.). *Reengineering the Enterprise*: (pp. 83-98). London: Chapman & Hall.
- Guarino, N., Welty, C. (2002). *An Overview of OntoClean*, Laboratory for Applied Ontology (ISTC-CNR), Trento, ITALY and IBM Watson Research Center, NY, USA.

- Horrocks, I. (2005). *Ontologies and the Semantic Web*. Report of the Roger Needham Lecture. An account prepared in collaboration with Ian Horrocks by Conrad Taylor of the BCS Electronic Publishing Specialist Group.
- Horrocks, I., y Van Harmelen, F. (2001). *Reference Description of DAML+OIL Ontology Markup Language*. Draft Report.
- Horrocks, I., Fensel, D., Harmelen, F., Decker, S., Erdmann, M., y Klein, M. (2000) *OIL in a Nutshell*. Workshop on applications of ontologies and PSMs. European Conference of Artificial Intelligence (ECAI'00). Berlin, Germany.
- Index of ontology. [Datos en línea]. Disponible: <http://song.cvs.sourceforge.net/song/ontology/>. [consulta: 2008, Abril 04].
- Kent, R. (1998) «*Conceptual Knowledge Markup Language (version 2.0)*». [Documento en línea]. Disponible: <http://wave.eecs.wsu.edu/CKRML/CKML.html>. [consulta: 2008, Febrero 10]
- Kifer, M., Lausen, G., y Wu, J. (1995) *Logical Foundations of Object-Oriented and Frame-Based Languages*. Journal of the ACM. 42(4): 741-843.
- Laboratorio de Ontología Aplicada. [Página Web en Línea]. Disponible: [www.loa-cnr.it](http://www.loa-cnr.it). [consulta: 2008, Abril 03].
- Lassila, O., y Swick, R. (1999). *Resource Description Framework (RDF). Model and Syntax Specification*. W3C Recommendations. [Documento en línea]. Disponible: <http://www.w3.org/TR/PR-rdf-syntax>. [consulta: 2008, Febrero 10]
- Lenat, D., y Guha, R. (1990). *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project*. Editado por Addison-Wesley Publishing Company, Inc., California, USA.
- Lozano, Adolfo. (2002). *Métrica de Idoneidad de Ontologías*. Tesis de Doctorado. Departamento de Informática, Universidad de Extremadura, Extremadura, España.
- Luke, S., y Heflin, J. (2000). *SHOE 1.01. Proposed Specification. SHOE Project*. [Documento en línea]. Disponible: <http://www.cs.umd.edu/projects/plus/SHOE/>. [consulta: 2008, Febrero 10]

- McGregor, R. (1991). *Inside the LOOM classifier*. SIGART bulletin. 2(3): 70-76.
- Middle East Technical University. (2005). *A Roadmap for Interoperability of eHealth Systems*. [Documento en línea]. Disponible: <http://www.srdc.metu.edu.tr/webpage/projects/ride/index.php> [Consulta: 2008, Abril 04]
- Motta, E. (1999) «*Reusable Components for Knowledge Modeling*», Tesis de Docotrado. Knowledge Media Institute, The Open University, UK.
- OBO Download Matrix. [Página Web en Línea]. Disponible: <http://www.berkeleybop.org/ontologies/>. [Consulta: 2008, Abril 04].
- Ontology Working Group. [Página Web en Línea]. Disponible: <http://mged.sourceforge.net/ontologies/OntologyResources.php>. [Consulta: 2008, Abril 04].
- Pérez, Chantal. (2002). Explotación de los córpora textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento. *Estudios de Lingüística Española (ELiEs)*, Universidad de Málaga. [revista en línea], 18. Disponible: <http://elies.rediris.es/elies18/index.html>. [Consulta: 2008, Abril 10]
- Pisanelli, D., Gangemi, A., Steve, G. (1998) *An Ontological Analysis of the UMLS Methatesaurus*. Proceedings of AMIA 98 Conference. Philadelphia, PA, USA.
- Rector, A., Solomon, W., Nowlan, W., Rush, T. (1995). A Terminology Server for Medical Language and Medical Information Systems. *Methods of Information in Medicine*. 34: 147-157.
- Smith, Barry. [Página Web en Línea]. Disponible: <http://ontology.buffalo.edu/smith>. [Consulta: 2008, Abril 03].
- Staab, S., Schnurr, H., Studer, R., Sure, Y.(2001) Knowledge Processes and Ontologies. *IEEE Intelligent Systems*. 16(1): 26-34.
- Stanford University. (1997). *The generic Frame Protocol 2.0*. Technical Report. KSL-97-05. USA. Chaudhri, V., Farquhar, A., Fikes, R., Karp, P. Rice, J.
- Steve, G., Gangemi, A., Pisanelli, D. (1997). Integrating Medical Terminologies with ONIONS Methodology. En Kangassalo H,

Charrel (Comp.) *Information Modelling and Knowledge Bases VIII*. Amsterdam, IOS Press.

Swartout, B., Patil, R. Knight, K., Russ, T. (1997, Marzo). *Toward distributed use of large-scale ontologies*. AAAI-97 Spring Symposium Series on Ontological Engineering. Stanford University, CA, USA.

The National Center for Biomedical Ontology. [Página Web en Línea]. Disponible: <http://www.bioontology.org/index.html>. [Consulta: 2008, Abril 03].

The Open Biomedical Ontologies. [Página Web en Línea]. Disponible: <http://www.obofoundry.org/>. [Consulta: 2008, Abril 03].

Uschold, M., King, M. (1995, Febrero) *Towards a Methodology for Building Ontologies*. Workshop on Basic Ontological Issues in Knowledge Sharing. Montreal, Quebec, Canada.

Valencia, Rafael (2005). *Un Entorno para la Extracción Incremental de Conocimiento desde Texto en Lenguaje Natural*. Tesis de Doctorado. Departamento de la Informática y las Telecomunicaciones. Universidad de Murcia, Murcia, España.

Wikipedia. [Página Web en Línea]. Disponible: [http://es.wikipedia.org/wiki/RDF\\_Schema](http://es.wikipedia.org/wiki/RDF_Schema). [Consulta: 2008, Abril 03].