



LAS ONTOLOGÍAS. NUEVOS RETOS.

Lic. Anisleiby Fernández Hernández; Lic. Sergio Carbonell de la Fe;

MSc. Yudeisy Pérez González; Lic. Tatiana Villalón Aguilera.

*Universidad de las Ciencias Informáticas
Carretera de San Antonio de los Baños Km 21/2. Torrens Boyeros. Ciudad Habana. Cuba
ani@uci.cu*

RESUMEN

Se realiza una aproximación a los cambios que se han venido acrecentando en la web actual donde la recuperación de la información en este contexto se hace eminente, y donde para ello los sistemas de representación y organización de la información juegan un papel fundamental, así como sistemas más complejos capaces de representar el conocimiento.

Se valora la posición y el rol de las Ciencias de la Información y el papel que debe seguir el profesional de la información ante estos nuevos retos que se le imponen. Se presentan breves cuestiones teóricas sobre el fenómeno de las Ontologías, vistas como sistemas de representación del conocimiento. Se abordan determinados elementos de forma breve sobre los tesauros, las taxonomías y su relación con las Ontologías, como sistemas de representación y organización de la información y del conocimiento.

Por último se muestra a través de la visualización de redes y gráficos algunos datos vinculados a la producción científica en el dominio de las Ontologías, dentro de las Ciencias de la Información y las Ciencias de la Computación. Descritas en el Journal Citation Report en el período de 1998-2007, para un total de 240 registros obtenidos.

ABSTRACT

It carried out an approach to the changes that have been increasing in the current web in which information retrieval in this context is eminent, and where to do so the systems of representation and organization of information play a key role, as well as systems More complex capable of representing knowledge.

Ranked the position and role of Information Sciences and the role they must follow the professional information to these new challenges imposed on it. Presents brief theoretical questions about the phenomenon of Ontologies viewed as a system of knowledge representation. Certain elements are addressed in a brief on thesauri, taxonomies and their relationship with Ontologies, as systems of representation and organization of information and knowledge.



Finally shown through the display of some graphics and data networks linked to the scientific production in the domain of Ontologies, within the Information Science and Computer Science. Described in the Journal Citation Report for the period 1998-2007, a total of 240 records obtained.

PALABRAS CLAVES

Ontologías, taxonomías, tesauros, recuperación de la información, sistemas de representación del conocimiento.



INTRODUCCIÓN

El final del siglo XX fue testigo del inicio de una nueva era: la era de la información. A partir de entonces hemos sido testigos de los cambios ocurridos en el entorno web, sobre todo con la difusión de la información al alcance de todos, donde asistimos cada día al crecimiento del volumen de información presentada en la Web.

Existen enormes dificultades para la recuperación de información pertinente contenida en este ámbito, el cual es un fiel reflejo del crecimiento exponencial de la información científica y técnica, y de su carácter acumulativo.

El crecimiento rápido y continuo del volumen de información hace que cada vez se torne más difícil de encontrar, organizar, acceder o de ofrecer información requerida para los usuarios.

Este hecho condujo a la web del usuario, el problema del exceso de información, o lo que es lo mismo, el fenómeno denominado “infoxicación”, acompañado de desorientación y la consiguiente dificultad de acceso a la información.

Estamos inmersos, indudablemente con la llegada de internet, en una avalancha de cambios en todos los órdenes: Ha cambiado la forma de los soportar información, esta se encuentra en el ciberespacio accesible y libre para todo tipo de usuarios, las organizaciones ya no son dueñas absolutas de estos recursos, sino que estos son susceptibles de ser compartidos. Ha cambiado el formato de la información, aparece el hipertexto con la presencia simultánea de elementos media.

Se ha producido un cambio en la forma de lectura, se rompe su linealidad para darle paso al hiperenlace. El usuario se ha hecho protagonista de todos los procesos informativos, se convierte en consumidor y creador de información.

En muchas ocasiones información altamente relevante se encuentra dispersa en diferentes recursos de información. Los usuarios por ende tienen que depender de una cantidad sustancial de tiempo en poder localizar esa información que se encuentra dispersa, y no sólo eso sino que muchas veces pasan mucho tiempo leyendo para descubrir que estos recursos verdaderamente encajen o estén relacionados con la estructura del área general del problema. Por tanto la información recuperada muchas veces no satisface las perspectivas del usuario. Nos encontramos ante un volumen informativo que supera con creces sus posibilidades de captación y entendimiento.



El profesional de la información. Nuevos Retos.

Con la llegada y desarrollo de esta “Era” o también llamada “Sociedad de la Información” muchos de los antiguos paradigmas en el tratamiento de la información se han modificado, abriéndose y adaptándose a los nuevos cambios. Evidentemente la organización y representación de la información se ha visto sometida a diferentes transformaciones para asumir el contexto de las NTICS y la web. Para las Ciencias de la Información y sus profesionales (los profesionales de la Información) se ha convertido en un reto ir a la par de los profundos cambios tecnológicos y desarrollar un pensamiento encauzado hacia la utilización de estos en el ejercicio de la profesión.

Lógicamente, para muchos, aún los esquemas mentales surgidos en la evolución de la disciplina no han sufrido alteración, algunas veces sólo se ha adaptado la misma manera de hacer y pensar pero ahora aplicado a la nueva tecnología. Precisamente el rompimiento de esquemas mentales es lo que garantizará el cambio y reconocimiento de nuestra especialidad en estos tiempos.

Al hablar de cambios tecnológicos, de Era de la Información, estamos sin dudas en el umbral de la inteligencia artificial. Logrando sistemas automatizados cada vez más inteligentes. Sistemas que entiendan las necesidades más específicas en el hombre en este sentido.

La creación de sistemas de información inteligentes lleva implícito el cambio en las formas de representar la información. Nuevas tendencias aparecen en este espacio. Se trata de representar la mente y el funcionamiento del cerebro humano. Se encaminan los esfuerzos hacia lograr la representación de asociaciones, interrelaciones entre conceptos, analogías e inferencias, que den respuestas a las necesidades informativas más particulares. Reflexionar acerca de la creación y desarrollo de nuevos sistemas de representación de información más avanzados capaces de representar el pensamiento humano, capaces de representar el conocimiento y su papel en el cambio de paradigmas en las Ciencias de la Información.

En este entorno, pues lógicamente, ha cambiado la forma de organizar y representar información para que esta pueda ser recuperada. La tendencia está en lograr sistemas de representación que respondan a estos cambios, que hagan verdaderamente útiles los recursos informativos, que hagan frente incluso, a uno de los grandes problemas de la web, mencionado anteriormente: “la infoxicación”. La tendencia se orienta a la creación de sistemas que organicen y representen la información, pero que sean capaces de añadir valor, que sean capaces de generar nueva información, de ofrecer soluciones a problemas puntuales. El reto de las Ciencias de la Información en estos tiempos es la creación de sistemas de representación y organización de la información capaces de crear por si mismos nuevo conocimiento para la toma de decisiones.

Es esta precisamente la mirada en el siglo XXI, una mirada que debe ser asumida por las Ciencias de la Información.



Las ontologías constituyen lenguajes y a su vez sistemas de representación de la información y del conocimiento, que describen diferentes recursos de información, para su posterior recuperación, lo que también se ha venido usando para mejorar la recuperación de información en la web, en términos de efectividad, rapidez y facilidad de acceso a la información, a lo que comúnmente se denomina “la web semántica”.

Los profesionales a su vez cada día más entienden que es urgente la construcción y aplicación de sistemas de organización y representación del conocimiento y esto se puede apreciar a través del incremento de los trabajos que aparecen en las revistas científicas y especializadas.

DESARROLLO

Las ontologías. Breves cuestiones teóricas.

Aunque aparentemente nuevo, producto a sus nuevas formas de uso, el término ontologías proviene del campo de la filosofía y la epistemología por tanto, una ontología se considera como una explicación sistemática de la Existencia y está relacionada con el estudio del ente, la comprensión del ser y la existencia humana.

Muchos autores han contextualizado a las ontologías, en el marco de estos tiempos, como un lenguaje para la representación de conceptos de la vida cotidiana, donde se encuentran establecidas relaciones e inferencias. Por tanto las Ontologías estudian los tipos de objetos que pueblan o que describen una realidad, dígase dominio, así como sus propiedades y relaciones entre estos objetos (conceptos).

Toda ontología representa cierta visión del mundo con respecto a un dominio. Se pueden definir tantas ontologías como sea posible, de hecho un mismo dominio puede ser interpretado de diferentes maneras, lo cual depende de la visión y de los tipos de conceptos que se manejen a la hora de representarlo.

Aunque existen muy diversas definiciones del concepto de ontología, una de las más ampliamente aceptadas es la siguiente: *“Una ontología es la especificación de una conceptualización”*. (Gruber, 1993).

Sobre esta línea de pensamiento está el concepto de Breuker *“Una ontología es una representación explícita de una conceptualización cognitiva, es decir, la descripción de los componentes de conocimiento relevantes en el ámbito de la modelización”*. (Breuker, 1999).

Otros autores la enfocan hacia *“un instrumento que define los términos básicos y relaciones a partir del vocabulario de un área así como las reglas de combinación de estos términos y relaciones para definir extensiones a un vocabulario”*. (Neches, 1991).



Desde este mismo enfoque: como herramienta informativa “ *Una ontología es una base de datos que describe los conceptos generales o sobre un dominio, algunas de sus propiedades y cómo los conceptos se relacionan unos con otros*”. (Weingand, 1997).

Otros autores también se han referido a las ontologías como “*una especificación formal de una conceptualización compartida*”. (Borst, 1997).

"Una ontología es una base de datos que describe los conceptos generales o sobre un dominio, algunas de sus propiedades y cómo los conceptos se relacionan unos con otros". (Weingand, 1997).

Y por último lo cual no quiere decir que existan más conceptualizaciones dadas por otros autores, sólo que no es nuestro objetivo realizar un amplio espectro conceptual sobre el tema, sino los rasgos más distintivos, *una ontología necesariamente incluirá un vocabulario de términos y una especificación de su significado (definiciones e interrelaciones entre conceptos) que impone estructura al dominio y restringe las posibles interpretaciones*. (Uschold, 1996).

Lógicamente, las primeras aproximaciones al mundo de las ontologías también han sufrido cambios con el desarrollo acelerado de la NTICS y el perfeccionamiento de los sistemas.

Vale destacar que aún las ontologías suelen confundir a muchos, ya que comienzan a despuntar como otra forma de representación de la información no manejada ni conocida hasta el momento.

Actualmente las organizaciones y entidades que desarrollan ontologías emplean mayormente el lenguaje OWL, lenguaje compatible con la arquitectura de WWW y que genera aplicaciones para la Web Semántica.

Según OWL “*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 un dominio y también el conocimiento que extiende los dominios. En este sentido, hacen el conocimiento reutilizable.*” (W3C, 2004).

Partiendo de esta posición, el concepto ontología se hace más acabado, y emerge entonces la idea de la representación del conocimiento.

Las ontologías presentan, como todo sistema de representación, diferentes componentes.



Aunque algunos lo tratan con diversos matices, mayormente se han tomado como elementos de las ontologías, los siguientes dados por (Gruber, 1995):

A) Conceptos: son las ideas básicas que se intentan formalizar. Los conceptos pueden ser clases de objetos, métodos, planes, estrategias, procesos de razonamiento, etc. Las clases en una ontología se suelen organizar en taxonomías a las que se les pueden aplicar los mecanismos de herencia.

B) Relaciones: representan la interacción y enlace entre los conceptos de un dominio. Suelen formar la taxonomía del dominio. Por ejemplo: subclase-de, parte-de, parte-exhaustiva-de, conectado-a, etc.

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

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

E) Reglas de restricción o axiomas: son teoremas que se declaran sobre relaciones que deben cumplir los elementos de la ontología. Por ejemplo: "Si A y B son de la clase C, entonces A no es subclase de B", "Para todo A que cumpla la condición B1, A es C", etc.

Los axiomas, junto con la herencia de conceptos, permiten inferir conocimiento que no esté indicado explícitamente en la taxonomía de conceptos.

Una vez que se ha abordado el problema terminológico es pertinente aproximarse a la cuestión tipológica, deudora en gran medida de la concepción que se tenga sobre las mismas. Se habla entonces de clasificación de ontologías.

Existen diversos tipos atendiendo a diferentes aspectos. Básicamente este tipo de sistema se ha clasificado atendiendo a:

Según el área o magnitud de conocimiento.

(Steve, 1998) distinguen tres tipos fundamentales de ontologías

Ontologías generales: son las ontologías de nivel más alto ya que describen conceptos generales (espacio, tiempo, materia, objeto, etc.)

Ontologías de dominio: describen el vocabulario de un dominio concreto del conocimiento.



Ontologías específicas: son ontologías especializadas que describen los conceptos para un campo limitado del conocimiento o una aplicación concreta.

Según al tipo de agente al que se destinen

Ontologías lingüísticas: se vinculan a aspectos lingüísticos, esto es, a aspectos gramáticos, semánticos y sintácticos destinados a su utilización por los seres humanos.

Ontologías no lingüísticas: destinadas a ser utilizadas por robots y agentes inteligentes.

Ontologías mixtas: combinan las características de las anteriores.

Según el grado de abstracción y razonamiento lógico

Ontologías descriptivas: incluyen descripciones, taxonomías de conceptos, relaciones entre los conceptos y propiedades, pero no permiten inferencias lógicas.

Ontologías lógicas: permiten inferencias lógicas mediante la utilización de una serie de componentes como la inclusión de axiomas, etc.

Como es evidente, alrededor de las ontologías ha surgido toda una plataforma teórica que facilita su posterior aplicación, desde los conceptos menos elaborados hasta tipos de lenguajes utilizados en las ontologías, patrones comunes para su funcionamiento, elementos de obligada utilización en su diseño, así como aplicaciones e importancia en la web, especialmente en la web semántica.

Dentro de los lenguajes conocidos hasta el momento aparecen: SHOE, OIL, DAML, KIE, OWL. Este último está siendo muy empleado por las organizaciones, ya que es un lenguaje de etiquetado semántico para publicar y compartir ontologías en la Web y principalmente se adapta a las necesidades específicas de cada usuario.

Entre las tantas de las funciones de las ontologías en la especialidad se destacan su empleo en la ingeniería del conocimiento, en la representación del conocimiento, en la creación de sistemas de información cooperativos, gestión del conocimiento. Tienen gran aplicación, además, dentro del mundo empresarial, en el comercio electrónico, cada vez más extendido en la red. Dentro de la informática, pues facilita el razonamiento automático, al poseer lenguajes entendibles por las máquinas.

Además de los tipos y los conceptos también están presentes en la literatura algunas metodologías y herramientas para el desarrollo de ontologías. Entre las metodologías más usadas podemos destacar las siguientes:



- **Diligence:** Metodología de desarrollo de ontologías basada en la colaboración de múltiples participantes, de manera que la creación de una ontología se concibe como un proceso social, distribuido y muy poco controlado.

- **Competency Questions:** La metodología de las “Preguntas Relevantes” o “Preguntas de Verificación” es una de las más sencillas de aplicar. Consiste en determinar el dominio y el alcance de la ontología mediante la lista de preguntas que el sistema debería ser capaz de contestar. Las respuestas a estas preguntas sugieren lo que podrían ser las instancias de la ontología, a partir de las cuales se deducirían (generalizando) las clases de la misma. (Grüninger, 1995).

- **Methontology:** Desarrollado en la Universidad Politécnica de Madrid. Propone un ciclo de vida de construcción de la ontología basado en prototipos evolutivos, porque esto permite agregar, cambiar y remover términos en cada nueva versión (prototipo). Para cada prototipo, el proceso consta de los siguientes pasos esenciales: Especificación, Conceptualización, Formalización, Implementación y Mantenimiento.

- **On-To-Knowledge:** Desarrollado por la Universidad de Karlsruhe, hace hincapié en las aplicaciones futuras de la ontología a la hora de diseñarla. Los pasos esenciales de que consta son: Estudio de viabilidad, inicio, refinamiento, evaluación y mantenimiento.

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

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

En cuanto a las diversas herramientas informáticas empleadas actualmente en el desarrollo de ontologías, destacaríamos las siguientes: Protegé KAON, WebOnto, OntoEdit, Ontolingua y el OntoSaurus.



Las ontologías tratan de imitar lo que a cada instante se produce en el cerebro humano: esta asociación de ideas, analogías de conceptos, de relaciones para arribar a conclusiones que indican la comprensión de información hasta obtener conocimiento, conocimiento que a su vez es aplicado para solucionar situaciones y tomar decisiones, que es lo que nos convierte en seres inteligentes. Es precisamente aquí donde surge la inteligencia artificial, en la capacidad de los sistemas de actuar como agentes inteligentes. Sin lugar a dudas, las ontologías constituyen la base para la evolución de máquinas pensantes.

A las principales funciones de las ontologías vistas con anterioridad, se suman otras aplicaciones para la web. Se señalan mayormente:

- ◆ Servir de herramienta para la adquisición de información.
- ◆ Servir de herramientas de referencia en la construcción de sistemas de bases de conocimiento que aporten consistencia, fiabilidad y falta de ambigüedad a la hora de recuperar información.
- ◆ Normalizar los atributos de los metadatos aplicables a los documentos.
- ◆ Crear una red de relaciones que aporte especificación y fiabilidad.
- ◆ Permitir compartir conocimiento.
- ◆ Posibilitar el trabajo cooperativo al funcionar como soporte común de conocimiento entre organizaciones, comunidades científicas, etc.
- ◆ Permitir la integración de diferentes perspectivas de usuarios.
- ◆ Permitir el tratamiento ponderado del conocimiento para recuperar información de forma automatizada.
- ◆ Permitir la construcción automatizada de mapas conceptuales y mapas temáticos.
- ◆ Permitir la reutilización del conocimiento existente en nuevos sistemas.
- ◆ Permitir la interoperatividad entre sistemas distintos.
- ◆ Establecer modelos normativos que permitan la creación de la semántica de un sistema y un modelo para poder extenderlo y transformarlo entre diferentes contextos.
- ◆ Servir de base para la construcción de lenguajes de representación del conocimiento.

Una de las mayores aplicaciones de las ontologías está dada en que hacen posible el funcionamiento de la Web semántica, ya que una de las condiciones para que esta funcione es que la información se represente a través de ellas. La Web semántica supondrá un salto cualitativo en el ambiente del ciberespacio, dará solución a los problemas de sobresaturación y desorden de información. Ofrecerá un espacio donde la información tenga un significado bien definido, de manera que pueda ser interpretada tanto por humanos como por máquinas. La Web semántica, será capaz de lograr que los softwares procesen su contenido, combinarlo y realizar deducciones lógicas para resolver problemas cotidianos automáticamente.



Las ontologías, constituyen la herramienta fundamental para el desarrollo de esta nueva web. En este sentido, es evidente, además, sus características de capturar el conocimiento de un modo genérico y formal, de manera tal que este puede ser compartido y reutilizado, estas características, de permitir la reutilidad del conocimiento, hace que las ontologías se proyecten hacia un sistema de organización de marcada trascendencia.

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

- ◆ Pueden existir ontologías múltiples: si el propósito de una ontología es hacer explícito algún punto de vista, en algunos casos, necesitamos combinar dos o más ontologías. Cada ontología introduce conceptualizaciones específicas.
- ◆ Se pueden identificar distintos niveles de abstracción estableciendo una topología de ontologías: se puede caracterizar una red de ontologías usando multiplicidad y abstracción. Al no poder realizar una descripción completa del mundo, se puede pensar una estrategia de construcción gradual que vaya de abajo hacia arriba.
- ◆ Multiplicidad de la representación: un concepto puede ser representado de muchas formas, por lo que pueden coexistir múltiples representaciones del mismo concepto.
- ◆ Mapeo de ontologías: se pueden establecer las relaciones entre los elementos de una o más ontologías para establecer generalizaciones, especializaciones, conexiones, etc.

El trabajo en la creación de ontologías está tomando auge. Muchas organizaciones se encuentran involucradas. Han salido a la luz numerosos proyectos para el desarrollo de ontologías y todo lo relacionado a su funcionamiento y aplicación. Entre estos:

Annotea: <http://www.w3.org/2001/Annotea/>

OntoWeb: <http://ontoweb.org/>

SchemaWeb: <http://www.schemaweb.info/>

W3C Semantic Web: <http://www.w3.org/2001/sw>.

Es necesario que las organizaciones se involucren en un marco de trabajo común para el desarrollo de ontologías, que tengan en cuenta elementos de suma importancia: adquisición del conocimiento, edición, fusión, integración, navegación, diseño, evaluación, traducción a diferentes lenguajes y formatos, e intercambio de contenido con otros instrumentos.

Es fundamental reflexionar en este sentido, la tendencia, que se hace cada vez más creciente a gestionar el conocimiento, ha traído aparejado el desarrollo de nuevos sistemas de representación, encauzados a favorecer la toma de decisiones, a ser capaces



de ofrecer al usuario una solución a sus necesidades más particulares; sistemas proactivos, capaces de ofrecer información con alto valor agregado, y de generar en si mismos una nueva información.

Partiendo de este punto, valdría la pena cuestionarse la función de los sistemas de representación de información. Sin lugar a dudas, los tradicionales sistemas de representación, los cuáles solo se orientaban a ofrecer respuestas concretas a una búsqueda terminológica, se han quedado atrás; sus sucesores son capaces de comprender el intelecto y proyectarse hacia la solución de problemas, los nuevos sistemas de representación son capaces de generar en si mismos un nuevo conocimiento.

Sistemas de representación, organización de la información y del conocimiento. Tesoros y Ontologías.

Desde los comienzos de esta especialidad, (la Ciencia de la Información), cuando aún no se conocía su denominación, se hizo necesaria la representación y organización de la información para su localización y recuperación. Desde el papiro hasta la biblioteca virtual se han creado instrumentos para clasificar, organizar y representar el contenido de los documentos. No ha sido tarea sencilla para el mundo bibliotecológico desarrollar sistemas de representación y organización de la información, tomando en consideración la complejidad de la lengua y el establecimiento de recursos lingüísticos adaptables a diferentes contextos.

Asimismo, con el desarrollo las TICS, comienzan a surgir nuevas formas de representación y organización de la información. Se ha de partir de que con la presencia de la Web, aparece una nueva visión de cómo asumir el paradigma informacional y con este el ciclo de vida de la información. Se produce una ruptura en la linealidad de este ciclo, los procesos van a darse de forma iterativa y no con una diferenciación marcada en cada una de las etapas, sino que estas se suceden todo el tiempo. La tendencia de los sistemas es ser cada vez más abiertos, interactivos, dinámicos, de ofrecer el mundo hipermedial ante los ojos de los navegantes en todas sus dimensiones.

En esta tarea de establecer un lenguaje controlado sobre la base del natural, donde se evidencia una gran variedad de modelos, de autores, de contextos, han surgido diferentes sistemas para la clasificación: Dewey, CDU, Bliss, Clasificación de Ranganathan, entre otros por todos conocidos.

Explicitar toda una serie de transformaciones lingüísticas y representarlas para poder lograr la búsqueda y recuperación de información de manera eficiente, se ha convertido en una de las áreas de mayor enfoque dentro de la especialidad. Una de las herramientas de control terminológico más utilizadas lo constituyen los tesauros, estructurados bajo un control léxico-semántico más riguroso, y una presencia de relaciones cada vez



mayor. Los tesauros también han venido perfeccionándose y naturalmente, al insertarse en el ambiente digital han alcanzado otra dimensión siendo adaptables al hipertexto y a la hipermedia. De igual forma, bajo el ambiente digital han venido desarrollándose los metadatos, como lenguaje que posibilita la descripción de los recursos de la red.

A tal efecto, un tesoro debe tener una buena estructura jerárquica, una adecuada base léxica y ser permeable a estándares web.

En el análisis conceptual sobre ontologías resulta importante destacar que estas han sido objeto de un sinnúmero de controversias, entre estas el cuestionamiento de que si ontologías y tesauros constituyen básicamente el mismo sistema. Entorno al problema existen varias posturas, aunque no se adopta ninguna como verdad absoluta, es meritorio estudiar todos los criterios.

Muchos autores han abordado la problemática de las similitudes entre ontologías y tesauros y la posibilidad de que estas herramientas se fundan en una. Algunos criterios lo reflejan de esta forma:

“Las ontologías son instrumentos claramente conectados con los agentes expertos en su objetivo de filtrar información y que, en muchos casos, pueden derivar en tesauros.” (García, 2001).

“Las ontologías son instrumentos adecuados para las definiciones de un vocabulario de representación y coinciden con los tesauros y con las clasificaciones en su capacidad para representar el contenido de un documento a través de la abstracción y de las relaciones entre conceptos” (Qin, 2002).

Según estas posturas en algún punto los tesauros y las ontologías se igualan, no existe una marcada distinción entre ellos que pueda colocar a alguno en un nivel superior. Básicamente presentan las mismas funciones y características.

En contraposición, otros autores establecen marcadas diferencias entre ambos instrumentos, tomando como referencia a Antonio García Jiménez se aprecian en las ontologías elementos muy particulares que hasta cierto punto reflejan superioridad. En este sentido:

- ◆ Las ontologías presentan un nivel más alto de concepción y de descripción del vocabulario.
- ◆ Presentan un desarrollo semántico más profundo y un mayor cuidado en su descripción.
- ◆ Hacen uso de la lógica en la descripción de situaciones.
- ◆ Permiten la reusabilidad y la posibilidad del trabajo en sistemas heterogéneos

Otros autores de igual manera han analizado puntos de convergencia y divergencia entre estos, centrando su atención en aspectos que enriquecen a las ontologías como



lenguajes de representación de la información. En esta vertiente aparecen (Ding y Foo, 2002) los cuales se apoyan en los siguientes elementos:

1. Una ontología puede estar elaborada de acuerdo con diferentes requerimientos y, al mismo tiempo, puede funcionar como un esquema de base de datos, como una auténtica base de conocimiento, para definir varias tareas o aplicaciones
2. Una ontología potencia la comunicación entre humanos y ordenadores mientras que un vocabulario convencional sólo permite la comunicación entre seres humanos.
3. Una ontología promueve la normalización y reutilización de la representación de la información mediante la identificación del conocimiento común y compartido
4. Las ontologías añaden valor a los tesauros tradicionales a través de una semántica más profunda, así como desde un prisma conceptual, relacional e informático. De hecho, una mayor profundidad semántica pueden implicar niveles más profundos de jerarquía, unas enriquecidas relaciones entre clases y conceptos, así como la capacidad de formular reglas de inferencia.

Lógicamente, tanto los tesauros como las ontologías parten de la representación conceptual del mundo, se utilizan para estructurar conceptualmente áreas del conocimiento a través de un vocabulario. Si embargo en el caso de las ontologías, estos conceptos se hacen más complejos, introducen mayor nivel de profundización semántica, la descripción lógica y semántica puede ser interpretada por las computadoras y los usuarios, permiten la interoperabilidad entre sistemas distintos. Pero el valor y particularidad de las ontologías está precisamente, en su capacidad para crear relaciones, establecer inferencias, en tratar de llegar a funcionar como la mente humana.

Taxonomías y Ontologías

Crear una página Web fácil de usar no es tarea fácil. Una experiencia interactiva positiva se deriva, en muchos casos de la correcta organización de los contenidos de una Web, y de taxonomías que así lo permitan y faciliten. (Fernández, 2007).

Si queremos diseñar una página que el usuario final sepa usar, genere experiencia interactiva positiva y a su vez resulte un proyecto rentable, no debemos perder el papel fundamental que juegan las taxonomías en este sentido. Viéndolas como estructuras predeterminadas que se usan para dividir un área temática (en este caso el área temática se refiere al contenido del sitio) Y estas áreas temáticas en otras áreas progresivamente más pequeñas y así sucesivamente organizar y estructurar en clases a partir de propiedades de conceptos, por similitud de contenidos. Se requiere entonces de un



análisis conceptual que diferencie estos conceptos (clases) y a su vez los agrupe según sus características. (Fernández, 2007).

Como podemos apreciar las taxonomías constituyen otro de los lenguajes utilizados en la Web. Una taxonomía es una jerarquía semántica, son excelentes mecanismos para clasificar entidades de información. Las taxonomías incorporan relaciones de equivalencia y jerarquía, aunque mayormente hacen énfasis en la organización de sus componentes, logrando mejoras en la navegación y el desarrollo de sistemas de búsqueda de información.

Asociado a las taxonomías y a estas recientes formas de organización y representación de la información y del conocimiento ha surgido una nueva generación de sistemas, sistemas que están centrados en la organización del conocimiento principalmente, que son capaces de analizar de pensar y ofrecer soluciones. Aparecen, entonces, las ontologías, que en este sentido este fenómeno se ha hecho popular en gran parte, porque tienen como objetivo promover un entendimiento común y compartido en un campo, que pueden comunicarse entre las personas y los sistemas de aplicación, además como una herramienta para la gestión del conocimiento que sin duda alguna constituye la base para la llamada inteligencia artificial.

Las taxonomías representan la manera en que se organizan clases y subclases dentro de una ontología.

Por otra parte a partir de la base de un tesoro se puede construir una ontología, de ahí que los tesauros, las taxonomías y las ontologías estén estrechamente relacionados.

Como colofón a todos estos elementos que se han venido presentando podemos agregar que con la introducción de las ontologías como herramientas para representar el conocimiento, las formas de hacer y pensar, obligatoriamente, deben cambiar. Deseamos sistemas acabados, capaces de entender la mente humana, de brindar la respuesta exacta a nuestras necesidades y problemas, pero los esquemas mentales siguen siendo los mismos y mayormente se trabaja en adaptar los antiguos sistemas a las nuevas tecnologías.

La tendencia cada vez hacia lo ergonómico, hacia diseños que sean adaptables a las características más particulares del ser humano, debe ser aplicada con prontitud al mundo de las Ciencias de la Información. El empeño de trabajar en la satisfacción de las necesidades del usuario, que sea este la piedra angular de las organizaciones y servicios de información; de tener posiciones proactivas y altamente eficientes, ha llegado a su más alta expresión con el surgimiento de sistemas cada vez más avanzados.



Visibilidad de la producción científica en el dominio de las ontologías.

Para la visualización de los datos que se muestran a continuación se realizó un estudio para analizar el comportamiento de las Ontologías como dominio científico dentro de las Ciencias de la Información y las Ciencias de la Computación.

Se realizó una búsqueda bibliográfica en las bases de datos del WOS, usando el Science Citation Index (SCI) y Social Science Citation Index (SSCI) del ISI, para un total de 240 registros obtenidos.

Los términos usados para hacer la estrategia de búsqueda y recuperación de los registros fueron (ontologies or ontology) en el campo Topic y el análisis se realizó limitando la búsqueda en el período que comprende desde 1998-2007 en el campo Date of Publication.

Los softwares usados fueron el sitkis 1.6 para el procesamiento de datos bibliométricos y el ucinet.

El sitkis es un sistema que permite crear una base de datos relacional sobre JAVA y Access, a partir del cual se normalizaron los datos, se obtuvo conteos de frecuencia preliminares y se diseñaron y exportaron matrices para UCINET.

Las tablas y gráficos básicos para indicadores bibliométricos simples de actividad se generaron en Excel 2007. Se emplearon además indicadores biométricos relacionales (indicadores de coocurrencia) para analizar en lo fundamental el comportamiento de la colaboración (co-autoría tanto de autores individuales como institucionales), la estructura de la base intelectual del dominio a través de la co-citación de autores y revistas. Las redes y visualizaciones de información se generaron y analizaron utilizando UCINET Y NETDRAW así como Pajek.

En este caso sólo nos adentraremos a analizar los autores más productivos, la co-citación de documentos, la productividad por años y las instituciones más productivas, así como las Universidades, Empresas y Centros de Investigación.

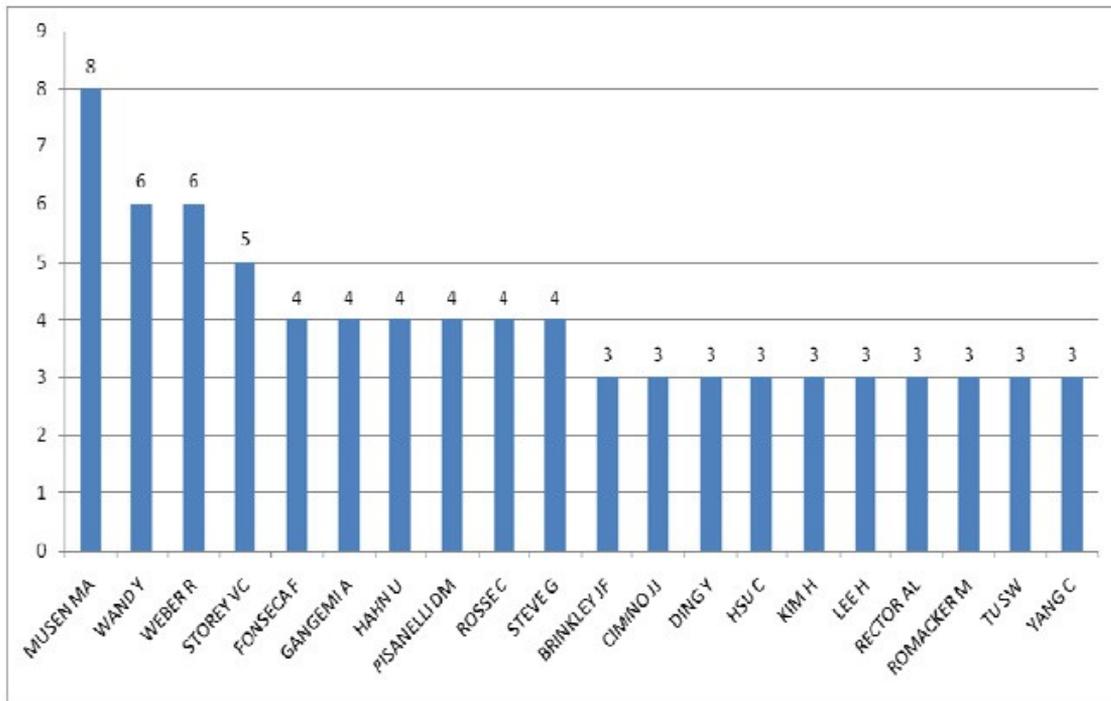


Figura1.- Productividad entre autores.

En el gráfico de la (figura 1) se puede observar la productividad por autores. En una escala del 0-9 se puede observar que el autor con una mayor producción resultó ser Musen Ma con 8 trabajos bajo su autoría, seguido por Wand Y, con 6; Weber R con 5, Fonseca F, Gangemia A, entre otros autores con 4 y Brinkley JF, Cimino JJ entre otros con 3 respectivamente.

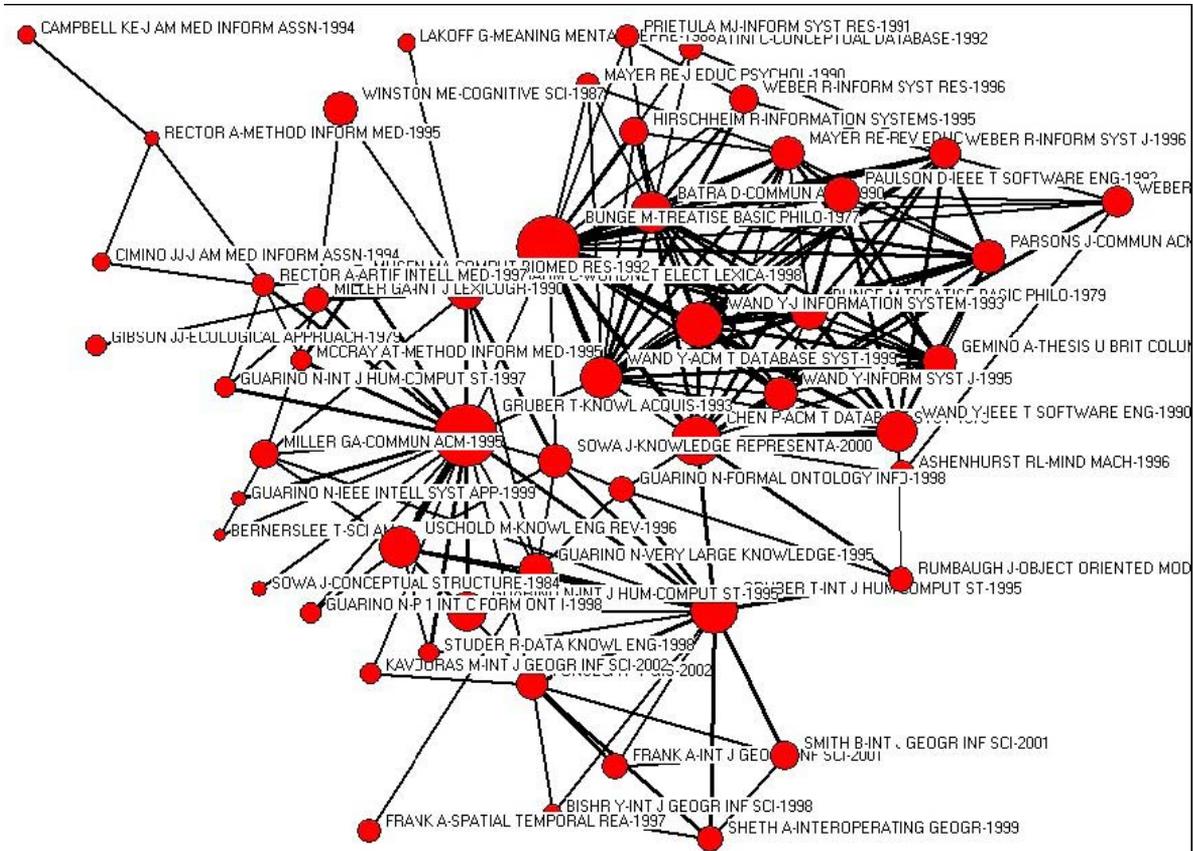


Figura2.- Co-citación de documentos citados.

En la red anterior se muestra la co-citación de documentos citados. La mayor centralidad la ocupan el documento de Bunge que lleva por título “Treatise on Basic Philosophy: Ontology II”. Reidel publicado en 1979 y el de Gruber que lleva por título “A translation approach to portable ontologies”. Knowledge acquisition publicado en el año 1993, que constituye además uno de los clásicos en esta materia, por ser el documento donde se emite uno de los conceptos más usados y mejor entendido en el tema de las Ontologías, después del documento de Bunge que se publica en 1979 como se explica anteriormente. Se observa que en el documento de Bunge todas las relaciones que establece son relaciones fuertes, lo cual podemos observar a través de la intensidad de sus enlaces con otros documentos, y que por tanto y demás es uno de los documentos más citados en este campo por su nivel de importancia, al igual que el de Gruber, donde se ven más relaciones con otros documentos en esta red. Es importante destacar en este sentido que ambos documentos son co-citados entre sí.

Los documentos que llevan por autores tales como Gibson JJ, Frank A, Campbell KE y Lakoff G mantienen sólo una conexión, lo cual significa que son los documentos que menor número de cocitación presentan. Los nodos de mayor tamaño observados identifican a los documentos productores que se encuentran entre los más citados, o sea

mayor cantidad de citas recibidas y la intensidad de los enlaces entre ellos representa el nivel de co-citación.

Como podemos observar los documentos que llevan como autores a Gruber T y Bunge M, juegan un papel central dentro del clúster atendiendo a la posición, al tamaño del nodo y al número de enlaces que mantienen con estos. La intensidad de los enlaces demuestra que son los documentos que mayor número de veces aparecen co-citados entre sí, por lo que existe una gran cercanía entre estos trabajos que desarrollan, en este caso los tratados sobre el tema de las Ontologías. El número de enlaces, demostrado con su intensidad demuestra cuales son los documentos que mayor cocitación mantienen con estos. Los documentos que llevan por autores a Chen P y Wand Y, también mantienen una posición bastante visible dentro de los clúster.

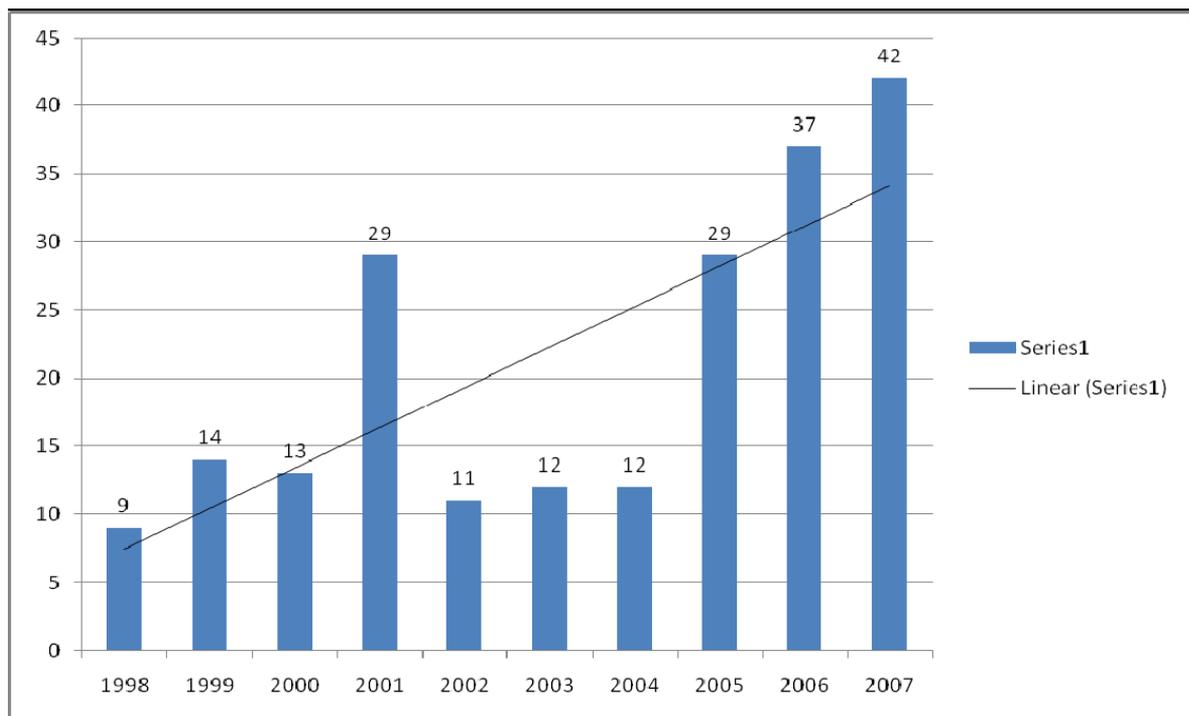


Figura3.- Productividad por años.

En la figura anterior se muestra cuales han sido los años más productivos en cuanto a la producción científica en Ontologías. Como se puede apreciar el año donde mayor producción se alcanzó fue el 2007 para un 42, siguiéndole el año 2006 con un 37. El año 2001 y 2005 mantuvieron un mismo índice de productividad con un 29. Por lo que se puede observar que ha habido un incremento notable en el año 2007 con respecto a los años interiores.

También podemos observar que los años 1998 y 2002 respectivamente fueron los años que menor índice de productividad alcanzaron para un 9 y un 11. Se muestra un

aumento exponencial en el año 2001 en comparación con los años anteriores, 1998, 1999 y 2000, de un 29. Al igual que los años posteriores 2002, 2003 y 2004 pero que en este caso, fueron en descenso, con respecto al año 2001. Después de este año la producción disminuyó y no fue hasta el 2005 con un 29, que no aumentaron los niveles de productividad, hasta llegar al 2007 con un 42.

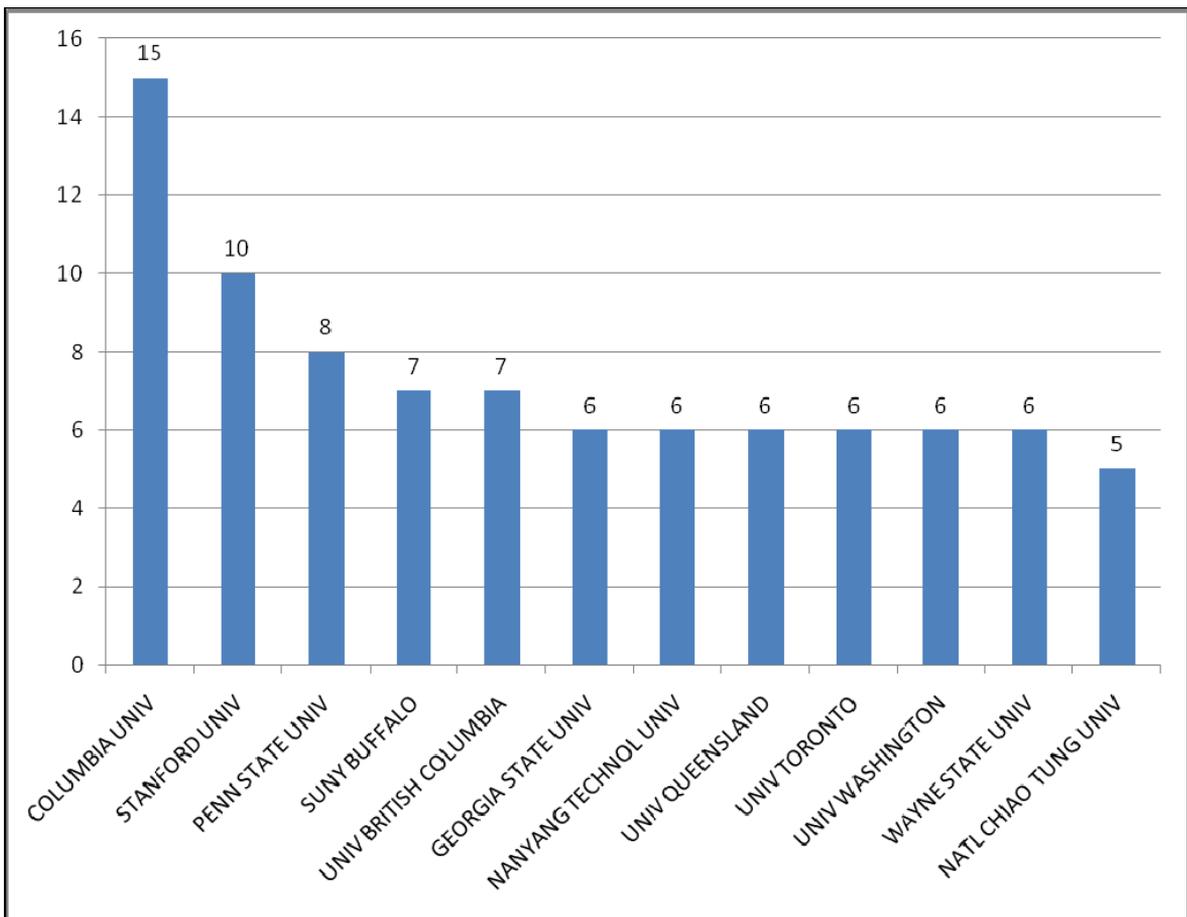


Figura4.- Productividad por Instituciones.

Podremos observar las Instituciones que mayor grado de productividad alcanzaron en la (figura 4). La Universidad de Columbia con un 15 fue la Universidad que mayor grado de productividad alcanzó. Le continúa la Universidad de Stanford con un 10. La Universidad de Penn State con un 8 y las demás Universidades Suny Buffalo, la British Columbia, la Georgia State con un 7 y 6. Por último la Universidad de Natl Chiao Tung con el menor grado de productividad para un 5. Se puede agregar que las Universidades que mayor grado de productividad alcanzan son Universidades de Estados Unidos, por lo que los índices de productividad en este país son los más elevados en comparación a otros países.

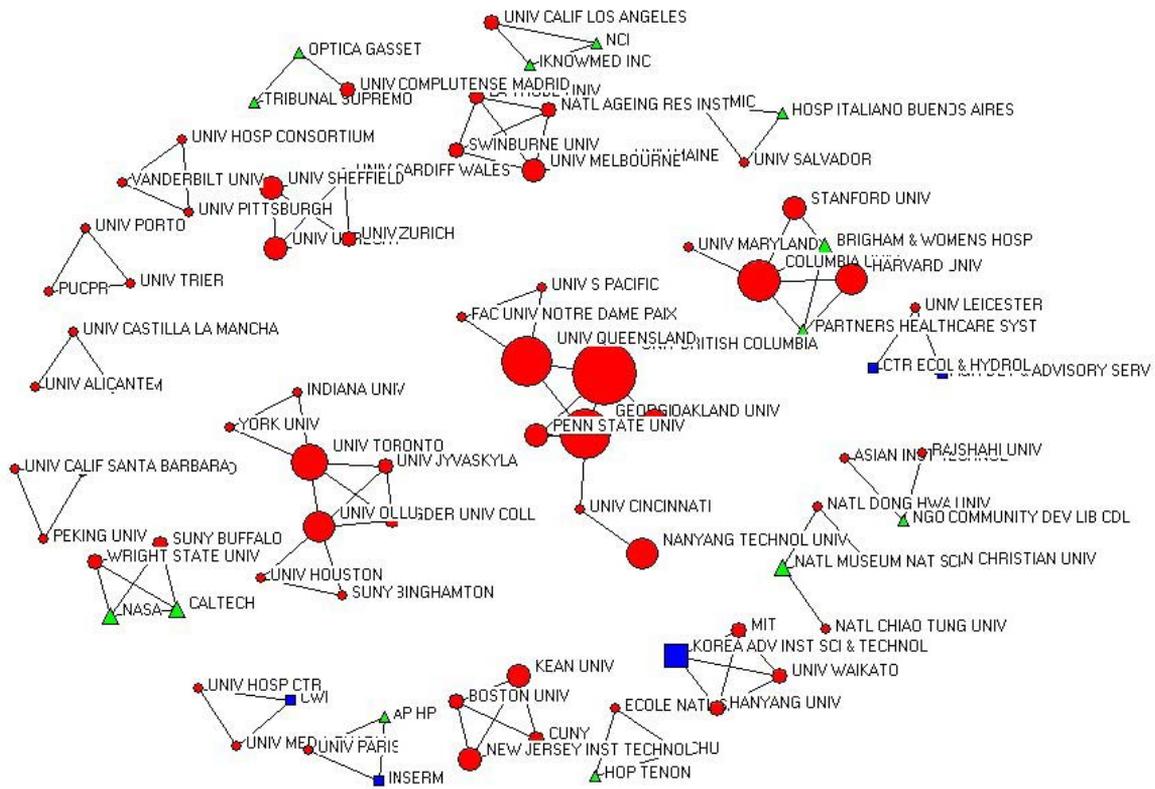


Figura 5.- Colaboración entre Instituciones.

La figura anterior muestra cual ha sido la colaboración entre las instituciones. Los nodos rojos representan las Universidades, los cuadrados azules los Centros de Investigación y los triángulos verdes las Empresas.

Se puede observar que la colaboración se establece entre Universidades, Centros de Investigación y Empresas, donde las Universidades alcanzan el mayor número de colaboración por ser la que mayor presencia tienen.

Los nodos rojos más grandes son los que mayor visibilidad obtienen, observándose que en ese sentido las Universidades tales como Queensland, Georgia y Toronto son las que mayor número de colaboración alcanzan. También podemos apreciar que las Universidades mencionadas anteriormente junto con otras relacionadas y representadas en sus clúster no presentan ninguna colaboración con el resto de las Instituciones, dígame, Centros de Investigación y Empresas, ya que lo hacen sólo entre ellas.



El Centro de Investigación de Korea es el que mayor grado de colaboración muestra, en comparación con los demás centros y su colaboración la establece solamente con Universidades.

La Universidad de Harvard y la de Columbia también muestran una alta presencia, representado por su grado nodal. Se puede observar que estas Universidades a diferencia de los demás clúster mencionados anteriormente representados por otras Universidades si establecen colaboración con otras Instituciones y preferentemente con Empresas.

CONCLUSIONES

Las ontologías se usan para favorecer la comunicación entre personas, organizaciones y aplicaciones, lograr la interoperabilidad entre sistemas informáticos, razonar automáticamente y para la ingeniería de software. Proporcionan una comprensión común de un dominio determinado describiendo esta realidad, de modo que se eliminan confusiones conceptuales y terminológicas, a la vez que quedan definidas para futuras investigaciones.

Las ontologías serán imprescindibles en la Web semántica y esto se ha venido demostrando con el día a día, y un ejemplo de ello además de otros muchos a los cuales pudiéramos referirnos, es el número de publicaciones que existen sobre la temática, donde cada vez se enfatiza aún más en el fenómeno de las Ontologías, alcanzando un mayor auge e importancia; tanto para la web como para Instituciones, dígase de cualquier tipo, Empresas, Universidades, Centros de Investigación etc.

La problemática de la indización, recuperación y divulgación de la información depositada en Internet u otros repositorios se debe abordar desde la perspectiva de la construcción de Ontologías.

Estas podrán permitir una gestión rápida, eficaz, pertinaz y permanentemente actualizada de las necesidades a cubrir.

Mediante ellas, será mucho más fácil recuperar información relacionada temáticamente, aun cuando no existan enlaces directos entre las páginas web.

En el campo de la informática, las ontologías sirven para traducir los términos usados por una aplicación a otra ya que las aplicaciones pueden estar escritas en distintos lenguajes de programación.

El hecho de que puedan ser reutilizables da una cierta medida de lo útil que pueden resultar las Ontologías. Con ello se preservaría el conocimiento percedero de los expertos en cualquier campo de aplicación.



La descripción de los conceptos y sus relaciones pueden ser empleadas de forma más general que los actuales tesauros, de manera que se pueden establecer un mayor número de relaciones distintas, y tener definidas cada una de ellas.

Autores tales como Musen MA, Gruber, Lee S, Bunge M, Wand Y, Uschold M, Warino, entre otros, además de ser muy productivos, mantienen un alto grado de colaboración con otros autores de la rama.

Los centros que hoy en día se adentran más en el mundo de las ontologías son las Universidades, además de ser las que mayor número de colaboración presentan entre sí y con otras Instituciones, Centros de Investigación y Empresas.

Por la presencia de la productividad por años vista en el gráfico correspondiente a la misma se puede enfatizar que la producción en el mundo de las Ontologías ha cobrado un mayor auge en estos últimos años, lo cual puede ir incrementándose notablemente en los años posteriores. En este trabajo sólo analizamos la producción científica hasta el 2007, pero por las experiencias vividas en el día a día nos hemos percatado que cada día son más las publicaciones científicas de esta temática en diferentes revistas de alto grado de visibilidad en el mundo.

En la Disciplina de Ciencias de la Información se necesita a pesar de la importancia que se le viene dando al tema profundizar en estos aspectos y ampliar aún más el espectro en este sentido. Los profesionales de la Información como bien lo expresa la palabra son los encargados por regla con el tratamiento de la información y con todo lo que de ella deviene, por tanto se necesitan profesionales bien preparados, con mente abierta al cambio y con ansias de investigar en estos campos que para nada son nuevos para ellos. Es un reto que se impone.

Comúnmente nos encontramos en muchas organizaciones diferentes aplicaciones informáticas jugando papeles diferentes y tenemos un poco agobiados a los usuarios de las mismas porque el uso excesivo de las TIC lejos de facilitarnos la tarea diaria, lo que hace es complicarnos la misma.

REFERENCIAS BIBLIOGRÁFICAS

ABEL, M. (2001). "Estudo da perícia em petrografia sedimentar e sua importância para a engenharia de conhecimento". In: Programa de PG em Ciência da Computação. UFRGS: Porto Alegre. p. 239.

BEALE, S.; S. NIRENBURG & K. MAHEH (1995). "Semantic Analysis in the Mikrokosmos Machine Translation Project", en Proceedings of the Second Symposium



on Natural Language Processing (SNLP-95), Agosto 2-4. Kaser Sart University, Bangkok, Tailandia.

BENJAMINS, V.; D. FENSEL & A. GOMEZ (1999). "Knowledge Management Through Ontologies".[en línea]. Consulta 07/07/08]. Disponible en Web <<http://www.aifb.unikarlsruhe.de/WBS/broker/inhalt-wp>>.

BORST, W. (1997): "Construction of Engineering Ontologies". University of Twente, Enschede, NL- Center for Telematica and Information Technology.

BUNGE, M. (1979). Treatise on Basic Philosophy: Ontology II. Reidel

BREUKER. A. (1999). "Indexing problem solving methods for reuse". In D. Fensel and R. Studer, Editors, Knowledge Acquisition, Modelling and Management, pg 315-322. Springer, 1999.

DAVENPORT, T.; L. PRUSAK. (1998): "Working Knowledge. Harvard Business School Press. Boston".

DING, Y; S. FOO. (2002). "Ontology research and development". Part 1-a review of ontology generation. Journal of Information Science 2002, vol.28, no.2, p 123-136

FERNANDEZ, A. (2007). "Organización de los contenidos en los sitios Web: las taxonomías". [en línea]. [Consulta 14 /07/08]. Disponible en Web:<http://bvs.sld.cu/revistas/aci/vol15_05_07/aci12507.htm>.

GARCIA, A. (2001). "Principios de lenguajes epistemográficos. La representación del conocimiento sobre patrimonio histórico Andaluz". Sevilla: Instituto Andaluz de Patrimonio histórico,

GUARINO, N. (1998). "Formal Ontologies and Information Systems". En: N. Guarino (ed.): 3-15.

GRUBER, T. (1995). "Towards principles for the design of ontologies used for knowledge sharing". [en línea]. International Journal of Human-Computer Studies, 1995, 43(4-5): 907-928. [Consulta 06/07/08]. Disponible en Web: <http://tomgruber.org/writing/onto-design.htm>.

GRUBER, T. (1993). "A translation approach to portable ontologies". Knowledge acquisition, 5(2):199-220.

GRÜNINGER, M.; M. FOX. (1998). "Methodology for the design and evaluation of ontologies". [en línea]. In: WORKSHOP ON BASIC ONTOLOGICAL ISSUES IN

KNOWLEDGE SHARING. Montreal, 1995. patrimonio histórico Andaluz. Sevilla: Instituto Andaluz de Patrimonio histórico.

GRÜNINGER, M.; M. FOX. (1995). "Methodology for the Design and Evaluation of Ontologies". [en línea]. En: Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal. NOY, Natalya F.; MCGUINNESS, Deborah L. Ontology Development 101: A Guide to Creating Your First Ontology. [Consulta: 06/07/08]. Disponible en Web: <<http://wwwksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html>>.

JOHN F. "Hierarchies of Categories. Ontology". [en línea]. [Consulta: 19/03/07]. Disponible en Web: <<http://www.jfsowa.com/ontology/index.htm>>

NECHES, R. (1991). "Enabling Technology for Knowledge Sharing". AI Magazine. 12(3), 1991, 36-56.

NEWELL, A. (1982). The knowledge level, in Artificial Intelligence. Amsterdam. p. 87-127. QIN, J.; S. PALING. (2002). "Converting a controlled vocabulary into an ontology: The case of Gem". [en línea]. Information research, 2000-01, vol. 6, no. 2. [Consulta: 19/05/07]. Disponible en Web: <[http:// informationr. net/ir/6-2/paper94.html](http://informationr.net/ir/6-2/paper94.html)>.

STEVE, G.; A. GANGEMI, D. PISANELLI (1998). "Integrating Medical Terminologies with ONIONS Methodology". [en línea]. [Consulta: 19/03/07]. Disponible en Web: <<http://saussure.irmkant.rm.cnr.it>>.

USCHOLD, M.; M. GRUNINGER. (1996). "Ontologies: principles, methods and applications" [en línea] Pennsylvania: School of Information Sciences and Technology (IST). Pennsylvania State University. [Consulta: 19/03/07]. Disponible en Web: <<http://citeseer.nj.nec.com/uschold96ontology.html>>.

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

W3C (2004). "W3C - World Wide Web Consortium". [en línea]. [Consulta: 19/03/07]. Disponible en Web : <<http://www.w3.org>>.