

---

## **DESCRIPCIÓN SEMÁNTICA DE FUENTES DE INFORMACIÓN APLICADA A LA GESTIÓN DE SITIOS WEB**

### **Resumen / Abstract**

Las representaciones en que se basa la web actual, básicamente describen la forma de mostrar la información, pero no su significado. Esto implica que las búsquedas realizadas fundamenten sus procesamientos en palabras claves sin ninguna relación semántica con los contenidos de las páginas. La web semántica, como una nueva visión, propone añadir un componente descriptivo a los recursos disponibles, que extienda las representaciones vigentes. En el presente trabajo, se expone un estudio sobre gestión de sitios, web semántica y sus tecnologías, como base para una propuesta de descripción de un sitio web, como un proceso más natural que tenga en cuenta la semántica del contenido gestionado.

*The representations in which today's Web is based basically describe the way the information is shown, but not its meaning. This brings about that the searches carried out make their processes based in key words without any semantic relationship with the pages' contents. The semantic Web, as a new vision, suggests the addition of the descriptive component to the resources available, in order to extend the current representations. The present work deals with a study on web site positioning, semantic Web and its technologies, as basis for suggesting a description of a Web site, as a more natural process that takes into account the semantics of the managed content.*

### **Palabras clave / Key words**

Posicionamiento, buscadores, recuperación de información, web semántica, ontologías

*Search engine optimization, search engines, information retrieval, semantic web, ontologies*

---

**Alicia Enríquez Perdomo**, Ingeniera Informática, Instructora, Centro de Estudios de Ingeniería de Sistemas (CEIS), Instituto Superior Politécnico José Antonio Echeverría, Cujae Ciudad de La Habana, Cuba  
e-mail: aenriquez@ceis.cujae.edu.cu

**Félix Oscar Fernández Peña**, Ingeniero Informático, Máster en Telemática, Asistente, CEIS, Instituto Superior Politécnico José Antonio Echeverría, Ciudad de La Habana, Cuba  
e-mail: felix@ceis.cujae.edu.cu

## **INTRODUCCIÓN**

Desde el surgimiento de la web, Internet ha cobrado un auge mucho mayor. La web ha sido indiscutiblemente la aplicación que ha universalizado la red de redes, la que le confiere un especial atractivo entre los más diversos usuarios. Numerosas entidades empresariales, educativas y de disímiles ámbitos, se han insertado en este mundo a través de su presencia en la web, ya sea con sitios estáticos o dinámicos, pero que les reservan un espacio en esta telaraña de información. Sin embargo, no basta con tener un sitio web excelente desde el punto de vista de la programación, el diseño y el contenido, si no está visible para su público objetivo.

Una de las posibilidades que brinda la web y que resulta especialmente atractiva para los usuarios, es sin duda la recuperación de información.<sup>1</sup> El número de documentos estáticos existentes en Internet en el año 2003 ascendía a la cifra de tres billones, utilizados además por un total de 200 millones de usuarios a nivel internacional.<sup>2</sup> No obstante, este gran volumen de documentación tiene implicaciones desfavorables cuando se trata de encontrar y acceder a una información específica.

La mayor parte del tráfico que se pueda generar hacia un sitio se logra a partir de los principales buscadores establecidos en Internet. Alrededor del 90 % de los internautas los utiliza habitualmente como método para buscar información,<sup>3</sup> sin embargo, es una realidad objetiva que un elevado

porcentaje de los sitios web existentes en Internet no son visibles a través de ellos, y que utilizándolos no siempre se logran los resultados esperados a partir de una determinada búsqueda.

El modo en que estas se realizan, hace que se obtengan resultados imprecisos, y que muchas veces no se relacionen con la necesidad que la originó. Los motores actuales basan su procesamiento en la busca de palabras que se correspondan con la consulta que hizo el usuario. Esto puede traer consigo la recuperación de información con esas palabras pero utilizadas en un contexto completamente diferente del que se está indagando, o la omisión de información importante, que aunque no incluya los mismo términos, sí tiene otros que conceptualmente significan lo mismo.<sup>4</sup>

## GESTIÓN DE SITIOS

Aún con estos inconvenientes, siguen siendo los buscadores la puerta de entrada a Internet cuando se necesita encontrar información. "Alrededor del 90 % del tráfico generado por búsquedas proviene de 8 o 10 de los principales motores".<sup>3</sup> Ubicarse por tanto entre ellos, es de vital importancia para ser visitado. Este proceso que se realiza para indexar un sitio web en los motores de búsqueda fundamentales, se denomina gestión, y su realización de manera eficaz implica el establecimiento del sitio gestionado entre las primeras 30 posiciones de estos buscadores esenciales.<sup>3</sup> No basta por tanto publicar un sitio en Internet para que se genere tráfico hacia él, es necesaria además una gestión eficiente que garantice un alto número de visitantes.

El proceso de gestión de un sitio web comienza desde que surge la necesidad de que sea visto, o sea, desde el mismo momento en que se concibe su diseño e implementación, y se mantiene vigente mientras el sitio está publicado en Internet.

La manera en que se realizan las búsquedas actualmente (correspondencia textual entre términos buscados y la información a mostrar), provoca que si se desea aparecer entre los primeros resultados de una determinada consulta, debe prestarse singular atención a las palabras que se incluyen en las páginas a indexar. Las palabras o frases claves de una página web, son aquellas que describen mejor su contenido y las que con mayor probabilidad utilizarán los usuarios para realizar sus búsquedas sobre algún material similar a dicha página.<sup>3</sup> Ante estas condiciones, la gestión debe estar encaminada a una buena selección de palabras y frases claves que describan correctamente al sitio, así como su ubicación en lugares vitales dentro de la página que son analizados por los buscadores.

La selección de los términos y la ubicación estratégica en el contenido documental del sitio web se enfrenta además a la dificultad que genera el frecuente cambio de los parámetros que miden el posicionamiento, por parte de los robots y arañas que se encargan de indexar las bases de datos de los buscadores. Esto hace necesario el monitoreo constante del ciclo de vida del sitio web en su interacción con la red, para chequear qué posición ocupa en los resultados, y evitar que con los cambios de métricas sea relegado a posiciones desventajosas.

Sin embargo, aún cuando se realice una gestión efectiva, existen las limitaciones ya mencionadas en los mecanismos de recuperación de información, que están condicionadas por las características de la web actual. Estas limitaciones expresivas hacen que los elementos de semántica que manejan los buscadores se reduzcan a palabras claves sin ninguna relación entre ellas, que además no permiten tener un dominio sobre el significado que expresan.<sup>5</sup>

El problema principal que origina esta situación es la incapacidad de las representaciones en que se basa la web actual para expresar conceptos y sus relaciones entre ellos. O sea, la carencia de una descripción semántica de los contenidos presentes en la web. No basta obtener un conjunto de términos descriptivos de un sitio, si no se cuenta de forma explícita con la relación que los une, de manera que añadiendo un componente semántico en la inferencia y descripción de estos contenidos, se favorezca la gestión que se haga a posteriori.

## HERRAMIENTAS PARA

### EL POSICIONAMIENTO DE SITIOS WEB

En la tarea de posicionamiento de un sitio web en los buscadores existentes, deben seguirse una serie de pasos con el objetivo de organizar y hacer más eficiente la labor del experto. Existe un conjunto de herramientas, alguna de ellas gratuitas, que pueden significar un apoyo importante, logrando agilizar el trabajo para que este no resulte largo y tedioso.

Una primera etapa de la gestión de un sitio es la definición de las palabras claves a emplear para la descripción de su contenido.<sup>3</sup> De esta manera, quien realiza el proceso de gestión del sitio, debe definir estos términos para luego apoyado en herramientas existentes, nutrirse de otros que guarden relación con los definidos en un inicio. En la actualidad existen softwares especializados en sugerir términos a partir de otros que el usuario proporciona. Una herramienta que brinda esta funcionalidad es el Wordtracker, ([www.wordtracker.com](http://www.wordtracker.com)). Otra con este mismo fin es Goto.com, aunque con menos funcionalidades que la anterior.

Para etapas posteriores de gestión y monitoreo, hay disponibles otros recursos de software que también pueden resultar provechosos. Entre ellos se destaca el WebPosition Gold (<http://www.webposition.com>), el cual permite hacer el análisis general de un sitio, y que integra a su vez varias herramientas muy útiles con este objetivo. Una de las posibilidades de este software es el estudio de las páginas, a partir del cual brinda consejos sobre cómo optimizarlas para buscadores específicos, teniendo en cuenta las métricas que estos valoran.

Otros software existentes proveen también a los especialistas en posicionamiento, de importantes posibilidades de análisis y acción. Resultan interesantes las variantes de construir y(o) generar etiquetas a adicionar en una página (AccuTagger1.01), determinar el límite de repetición de palabras claves, así como su relevancia (Meta Medic 3.2), y a partir del análisis de una página, sugerir las posibles palabras claves en función del texto (SpiderView Density Analyzer 1.1), entre otras.<sup>3</sup>

En ninguno de los casos anteriormente citados, el experto cuenta con una lista de los términos utilizados en la gestión del

sitio. Las funcionalidades que ofrecen estas herramientas, aún cuando son realmente atractivas y apoyan el proceso de gestión, no permiten tener una visión general y centralizada de las palabras claves empleadas y la relación que guardan unas y otras, así como su distribución entre las diferentes páginas del sitio. De esta manera, si bien contribuyen con su utilidad a agilizar la labor de posicionamiento de un sitio, no permiten administrar los contenidos que lo describen y por tanto hacer un análisis semántico de la información que contiene.

## WEB SEMÁNTICA

Ante los elementos que han caracterizado a la web desde su surgimiento, donde destacan la carencia de una semántica explícitamente definida, la inexistente catalogación de sus contenidos, así como el crecimiento desorganizado de los recursos disponibles, existe una nueva visión que propone la organización y clasificación de los recursos a partir de una estructura semántica explícita.<sup>5</sup> A esta nueva visión se le ha llamado web semántica.<sup>6</sup>

"La web semántica es una extensión de la actual web, cuyo objetivo es que no solo los humanos, sino también las máquinas, sean capaces de comprender el contenido de los documentos"<sup>7</sup> Esta nueva perspectiva de la web aporta un enfoque basado en los contenidos, con descripciones explícitas de significados, que debe permitir a las máquinas procesar y "entender" la información existente. Mucho más que su función actual de mostrarla de la manera especificada por los humanos.<sup>6</sup>

## CAPAS DE LA WEB SEMÁNTICA

Los principios de la web semántica son implementados a través de varias capas de tecnologías y estándares. Estas capas se muestran en la figura 1. La capa de Unicode y URI (Uniform Resource Identifier) posibilita que se utilicen conjuntos de caracteres internacionales y que se puedan identificar inequívocamente los recursos de la web.

Los URI son la base que sostiene a la web, y cobran particular importancia en el contexto de la web semántica, donde todos los recursos tendrían el suyo propio, desde un microondas, hasta una página personal.<sup>7</sup>

La capa de XML (eXtensible Markup Language) con los espacios de nombres (NS, Name Space) y las definiciones de XML Schemas "es la base sintáctica sobre la que se sustentan el resto de las capas."<sup>7</sup> Los elementos anteriores aseguran que se puedan integrar las definiciones de la web semántica con los otros estándares basados en XML. Este lenguaje provee una sintaxis para la definición de la estructura y semántica de los datos, (aunque no aporta nada sobre el significado de dichas estructuras),<sup>2,6</sup> y permite además definir otros lenguajes de etiquetado validándolos mediante DTDs (Document Type Definitions) o XML Schemas.<sup>7</sup> Con estos últimos, se pueden acordar de antemano las estructuras que se van a utilizar, así como manejar tipos de datos primitivos y derivados.<sup>5</sup>

El significado del que carecen las estructuras XML puede ser expresado en la capa superior, a través de RDF (Resource Description Framework), lenguaje de propósito general para representar información en la web<sup>8</sup> y que es creado por medio de sintaxis XML. Con él, es posible hacer declaraciones de objetos con URIs, mediante enunciados en forma de tripletas sujeto-predicado-objeto, donde sujeto y predicado son URIs y objeto puede ser un URI o un valor literal. Estas tripletas pueden ser escritas utilizando etiquetas XML.<sup>7,9,10</sup> Por su parte, los RDF Schema, proporcionan un mecanismo para definir clases, jerarquía de herencia entre clases, objetos y propiedades, así como el dominio y restricciones de estas últimas. Esta es la capa que permite clasificar los recursos y enlaces según el tipo que sean. La siguiente, **vocabularios de ontologías**, da soporte a la evolución de estos vocabularios, en la misma medida en que se pueden definir relaciones entre diferentes conceptos.

La capa de **firma digital** proporciona una de las características principales de la web semántica. Según la referencia 7, las firmas digitales se pueden definir como "bloques de datos cifrados que las computadoras y agentes podrán usar para verificar que la información adjunta ha sido proporcionada por una fuente fiable", lo cual permitiría detectar alteraciones en documentos

Y para finalizar, se encuentran las capas superiores de **lógica**, **pruebas** y **confianza**, aún en investigación. La capa **lógica** posibilita la declaración de reglas, mientras que la de **pruebas** las ejecuta y evalúa junto con la de **confianza**, para determinar si se puede confiar o no en las pruebas proporcionadas.<sup>11</sup>

## ONTOLOGÍAS

Con el objetivo de que la web semántica pueda hacerse realidad, se necesita que el conocimiento esté representado de forma que sea entendible por las computadoras, esté consensuado, y sea reutilizable. Un medio para representar este conocimiento lo proveen las ontologías, uno de los conceptos más importantes vinculados a la web

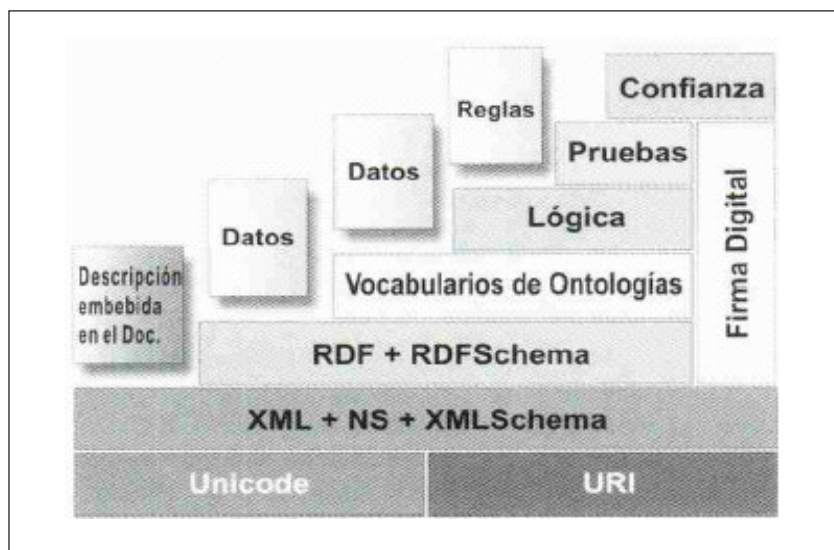


Fig. 2 Capas de la web semántica.<sup>7</sup>

semántica. Este es un término utilizado desde mucho antes en el contexto de la filosofía, y que se ha extendido al ámbito computacional. Se ha convertido en un tópico de investigación importante dentro de la inteligencia artificial desde la década de los noventa.<sup>2</sup>

Existen muchas definiciones de ontologías, una de las más importantes y representativas es la que aporta Gruber,<sup>12</sup> según la cual, una ontología es una especificación formal y explícita de una conceptualización compartida. Esto implica la definición formal de relaciones entre conceptos, a partir de taxonomías y un conjunto de reglas de inferencia.<sup>2</sup>

### Componentes

Para representar el conocimiento de algún dominio específico, las ontologías deben contar con los siguientes componentes básicos:<sup>13</sup>

- **Conceptos:** Ideas básicas que se pretenden formalizar. Pueden ser clases, métodos, planes, estrategias, procesos de razonamiento, etcétera.

- **Relaciones:** Representan la interacción y correspondencia entre los conceptos de un dominio. Por ejemplo: *subclase-de*, *parte-de*, *conectado-a*, etcétera.

- **Funciones:** Tipo de relación donde un elemento es identificado a través del cálculo de una función que involucra varios elementos de la ontología. Por ejemplo: *categorizar-clase*, *asignar-fecha*, etcétera.

- **Instancias:** Representaciones de objetos determinados de un concepto.

- **Axiomas:** Teoremas que se declaran a partir de relaciones que deben cumplir los elementos de la ontología. Por ejemplo: "Si A y B son de la clase C, entonces A no es subclase de B", "Para todo A que cumpla la condición C1, A es B", etcétera.

### Lenguajes de implementación y ambientes de desarrollo

Para trabajar con ontologías en una aplicación, es necesario evaluar sus necesidades reales y a partir de ahí determinar cuán compleja resultará la ontología a diseñar, pues una muy sofisticada puede implicar un costo excesivo para una aplicación que no requiera su uso. De cualquier forma, cuando se decide utilizarlas, hay dos elementos fundamentales sobre los cuales decidir: el lenguaje de definición y el ambiente de desarrollo.<sup>14</sup>

**Lenguaje de definición:** Aún cuando se esté trabajando con una ontología simple debe tenerse en cuenta, pero más aún si se está considerando el uso de una más compleja. Debe ser considerado especialmente el poder expresivo de sus representaciones, y que sea epistemológicamente adecuado. O sea, capaz de expresar los conceptos del dominio.

Algunos de estos lenguajes para su uso en la web, son: SHOE (Simple HTML Ontology Extensions), OIL (Ontology Inference Layer), Daml+OIL, y OWL (Web Ontology Language) derivado de Daml+OIL.<sup>7</sup>

Los lenguajes más comunes para representar las ontologías son RDF y OWL, aunque cada uno de ellos ofrece diferentes primitivas de modelación, y de esta manera, diferentes niveles de complejidad.<sup>15</sup>

- **Ambiente de desarrollo:** Debe analizarse cómo las ontologías utilizadas van a modificarse y recibir mantenimiento. Si esto es responsabilidad de un experto en la materia vinculada al dominio, en vez de un experto informático, es más importante aún una adecuada selección.

Existe un cierto número de herramientas disponibles para estos fines, entre ellas ontolingua, chimaera, OilEd, y Prótegé.<sup>14,16</sup>

## APLICACIÓN DE LAS ONTOLOGÍAS EN EL CONTEXTO DE LA GESTIÓN SEMÁNTICA DE SITIOS WEB

Como se explicó con anterioridad, el proceso de búsqueda de la manera en que se realiza en la actualidad, funciona a partir de la correspondencia entre los términos que el usuario introduce en el buscador, y las páginas que posteriormente aparecen entre los resultados devueltos. Esta correspondencia es puramente sintáctica, o sea, los buscadores analizan las páginas para detectar la aparición en ellas de textos contenedores de estas palabras.

La precisión de estos métodos de búsqueda puede ser por tanto bajo, pues están fundamentados en palabras en vez de conceptos subyacentes. Una posible solución a este problema es la recuperación de información basada en ontologías. Esta variante está basada en la organización jerárquica de los contenidos a través de categorías de información. La utilización de ontologías para lograr esto, permite obtener diversos beneficios:<sup>17</sup>

- Las ontologías pueden ser utilizadas para describir el dominio de conocimiento y la terminología de una aplicación en más detalle. Por ejemplo, se pueden definir relaciones entre categorías de contenidos.

- Pueden ser empleadas para crear anotaciones semánticas más precisas en términos de un dominio de conocimiento.

- Con la ayuda de las ontologías el usuario puede expresar las consultas con más precisión y sin ambigüedades, lo cual conduce a mejores índices de precisión y retentiva.

De esta manera, el dominio del conocimiento, las anotaciones que se hagan y la recuperación de información estarían basadas en estructuras semánticas ricas en ontologías, en vez de simples clasificaciones de palabras claves.

### Propuesta de aplicación en la gestión de contenidos de un sitio web

Sobre la base de los beneficios antes expuestos, y teniendo en cuenta la necesidad actual de dotar a los recursos existentes en la web de una descripción semántica, se propone una herramienta que permita administrar los contenidos de un sitio web a través del uso de ontologías, para la descripción de su información y recursos disponibles.

Se plantea la organización de los recursos en forma de una ontología, de manera que el usuario experto (alguien con conocimientos acerca de las ideas contenidas en el sitio), pueda ver los recursos existentes y la relación presente entre ellos. La visualización en forma de árbol de directorios, donde se muestren no solo las páginas sino además los enlaces físicos entre ellas, y su relación con otros recursos (por ejemplo, las imágenes

incluidas), constituiría un significativo apoyo a la labor de administración.

Dentro de la propuesta está incluida además, la posibilidad de que el experto pueda asociar a cada recurso una descripción, así como establecer enlaces semánticos entre páginas de contenidos relacionados, para chequear los enlaces físicos que deberían establecerse entre estos documentos. Así podría detectarse la no correspondencia entre unos y otros, y por tanto, la existencia de enlaces rotos, o no definidos por programación.

Muy importante resulta también la posibilidad de definir las palabras claves por página, de manera que cada una pueda almacenarse asociada al documento que describe. Esto es particularmente útil desde el punto de vista del posicionamiento web, pues permite tener la relación completa de conceptos que definen la temática del sitio, de manera que en todo momento pueda conocerse los términos actuales, eliminar e incorporar nuevos en función de la actualización del contenido.

Con esta aplicación, estaría disponible de manera centralizada la definición de palabras claves que se tiene de un sitio, así como la organización y relación entre sus recursos.

#### Evaluación de la propuesta

Con el objetivo de enfrentar el trabajo de posicionamiento de un sitio en los buscadores, es necesario el dominio de la información que se maneja en sus páginas. Incluso es importante conocer la relación que existe entre unas y otras, no solo desde el punto de vista de la navegación, sino también desde el punto de vista semántico. O sea, cómo están relacionados sus contenidos, para así estar en mejores condiciones de gestionarlos.

Para la realización de esta labor, resulta de utilidad tener el control sobre la información que aparece en las páginas, y poder definir a partir de los conocimientos acerca de ellas, cuáles son los términos o palabras claves que las describen. Más allá de esta ventaja, también es importante conocer el empleo que se hace de los recursos disponibles, de forma tal que puedan ser aprovechados eficazmente, a partir de saber cuáles se utilizan y cuáles se encuentran aislados dentro del sitio, por no tener relación con ningún otro.

Con esta propuesta se logra describir semánticamente el contenido de un sitio web, y a partir de ahí obtener palabras y frases claves más descriptivas que impacten significativamente en la gestión del sitio contra los motores de búsqueda. Si bien esto no garantiza por sí solo el posicionamiento del sitio en los primeros lugares de un buscador, es un punto de partida para lograr que los términos a emplear se obtengan a través de un proceso más natural que tenga en cuenta la semántica del contenido gestionado más que el proceso sintáctico que implementan las herramientas de gestión utilizadas en la actualidad.

## CONCLUSIONES

La web semántica constituye un salto cualitativo con respecto a la web actual. Su expansión y estandarización crecerán en la

misma medida en que aumenten las aplicaciones que resulten de ella. La semántica explícita que puede dotarle a las estructuras vigentes y los procesamientos que pueden hacerse a partir de ella, son por sí mismos un gran paso de avance para suplir las limitaciones actuales a las que se enfrenta la web.

La gestión de sitios es un proceso que tiene significativa trascendencia en la actualidad, debido al gran número de sitios que hoy se encuentran disponibles en la web. La definición adecuada de términos y palabras claves que describan sus contenidos, es sumamente importante en el logro de un buen posicionamiento.

Utilizar las ontologías como elemento medular de la propuesta realizada, dota a la gestión que se haga de los contenidos del sitio, de un componente semántico. Es posible lograr esto a través del uso de las tecnologías actuales vinculadas a la web semántica, que puede aportar una mayor organización y control sobre la información existente en el sitio. □

## REFERENCIAS

1. FININ, T. AND J. MAYFIELD: *Information Rretrieval on the Semantic Web*. [http://ebiquity.umbc.edu/v2.1/\\_file\\_directory\\_/papers/121.pdf](http://ebiquity.umbc.edu/v2.1/_file_directory_/papers/121.pdf). Consultada: 15/12/2004.
2. FENSEL, D.: *Spinning the Semantic Web*, 2003. <http://mitpress.mit.edu/books/chapters/0262062321intro1.pdf>. Consultada: 10/01/2005.
3. MARTÍNEZ PRIETO, J. *et al.*: *Gestión de sitios web*, ISBN: 959-261-035-5, 2002.
4. SILVA MUÑOZ, L.: *Representación de ontologías en la web semántica*. [http://www.inf.ufrgs.br/~clesio/cmp151/cmp15120021/artigo\\_lydia.pdf](http://www.inf.ufrgs.br/~clesio/cmp151/cmp15120021/artigo_lydia.pdf). Consultada: 19/10/2004.
5. CASTELLS, P.: *La web semántica*, 2004. <http://www.ii.uam.es/~castells/docencia/semanticweb/castells-uclm03.pdf>. 2004. Consultada: 10/01/2005.
6. BERNERS-LEE, T. *et al.*: "The semantic web. Scientific American", *Scientific American Inc.* Vol. 284, No. 5, 2001. <http://search.epnet.com/login.aspx?direct=true&authtype=cookie,ip,url,uid&db=afh&an=4328935>. 2001.
7. PEIS, E. *et al.*: "Análisis de la web semántica: estado actual y requisitos futuros", *El profesional de la información*, Vol. 12, 2003.
8. BECKETT, D.: *RDF/XML Syntax Specification*, 2004. <http://www.w3.org/tr/rdf-syntax-grammar/>. Consultada: 18/05/2005.
9. KLYNE, G., AND J. J. CARROLL: *Resource Description Framework (RDF): Concepts and Abstract Syntax*, W3C Recommendation, February 10, 2004.
10. CUENCA GRAU, B.: *A possible Simplification of the Semantic Web Architecture*, 13th International Conference on World Wide Web, WWW 2004, New York, NY, USA. 2004.
11. KOIVUNEN, M. R. AND E. MILLER: *W3C Semantic Web Activity*, <http://www.w3.org/2001/12/semweb-fin/w3csw>. 2001.
12. GRUBER, T. R.: *A Translation Approach to Portable Ontology Specifications*. <http://www.tomgruber.org/writing/ontolingua-kaj-1993.pdf>. 1993. Consultada: 12/12/2004.

13. **LOZANO TELLO, A.:** *Ontologías en la web semántica*, 2001. <http://www.informandote.com/jornadasingweb/articulos/jiw02.pdf>. 2001. Consultada: 19/10/2004
14. **NOY, N. F. AND D. I. McGUINNESS:** *Ontology Development 101: A Guide to Creating your First Ontology*, <http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html>. Consultada: 12/12/2004.
15. **EHRING, M.:** *Ontology Mapping -an Integrated Approach*, First European Semantic Web Symposium, ESWS 2004, Heraklion, Crete, Greece, Springer-Verlag, 2004.
16. **BROEKSTRA, J., et al.:** Adding formal semantics to the web2000. <http://www.ontoknowledge.org/oil/extending-rdfs.pdf>. Consultado: 12/09/2005.
17. **HYVÖNEN, E. et al.:** *Application of Ontology Techniques to View-Based Semantic Search and Browsing*. First, European Semantic Web Symposium, ESWS 2004, Heraklion, Crete, Greece, Springer-Verlag, 2004.

## GLOSARIO

**Araña:** Programa que busca páginas web de forma automática y las ofrece a los motores de búsqueda. (Se denomina "araña" porque rastrea la web.) Dado que la mayoría de las páginas web contienen vínculos con otras páginas, una araña puede iniciarse prácticamente en cualquier lugar. Cuando esta reconoce un

vínculo con otra página, abandona la página donde se encuentra y va a buscar la vinculada. Los grandes motores de búsqueda disponen de varias arañas funcionando simultáneamente. Las arañas también se denominan rastreadores.

**Ontología:** Colección de expresiones escritas en un lenguaje como RDF que define las relaciones entre conceptos y especifica reglas lógicas para inferir datos a partir de ellos.

**RDF:** Resource Description Framework. Lenguaje para definir información en la web. Provee de una tecnología para expresar el significado de términos y conceptos de manera que una computadora pueda procesarlos.

**Robot:** Programa que se ejecuta automáticamente sin necesidad de intervención humana. Un robot suele poseer algún grado de inteligencia artificial, de modo que es capaz de adaptarse a distintas situaciones. Dos tipos de robot comunes son los agentes y las arañas. Los robots también se denominan *bot*.

**Taxonomía:** Una taxonomía es una jerarquía semántica en la cual entidades de información son relacionadas ya sea por subclasificaciones o subclases.

**URI:** Universal Resource Identifier. Define o especifica una entidad, aunque no necesariamente nombrando su localización en la web. El tipo más conocido de URI son las URLs.

**XML:** eXtensible Markup Language. Lenguaje de marcas como el HTML, que permite definir etiquetas propias para un usuario.

***Disponemos de un departamento informatizado,  
dotado con tecnologías que nos permiten realizar  
todo el proceso de edición de revistas científicas así  
como de otros materiales.***

***Visítenos!!!***

