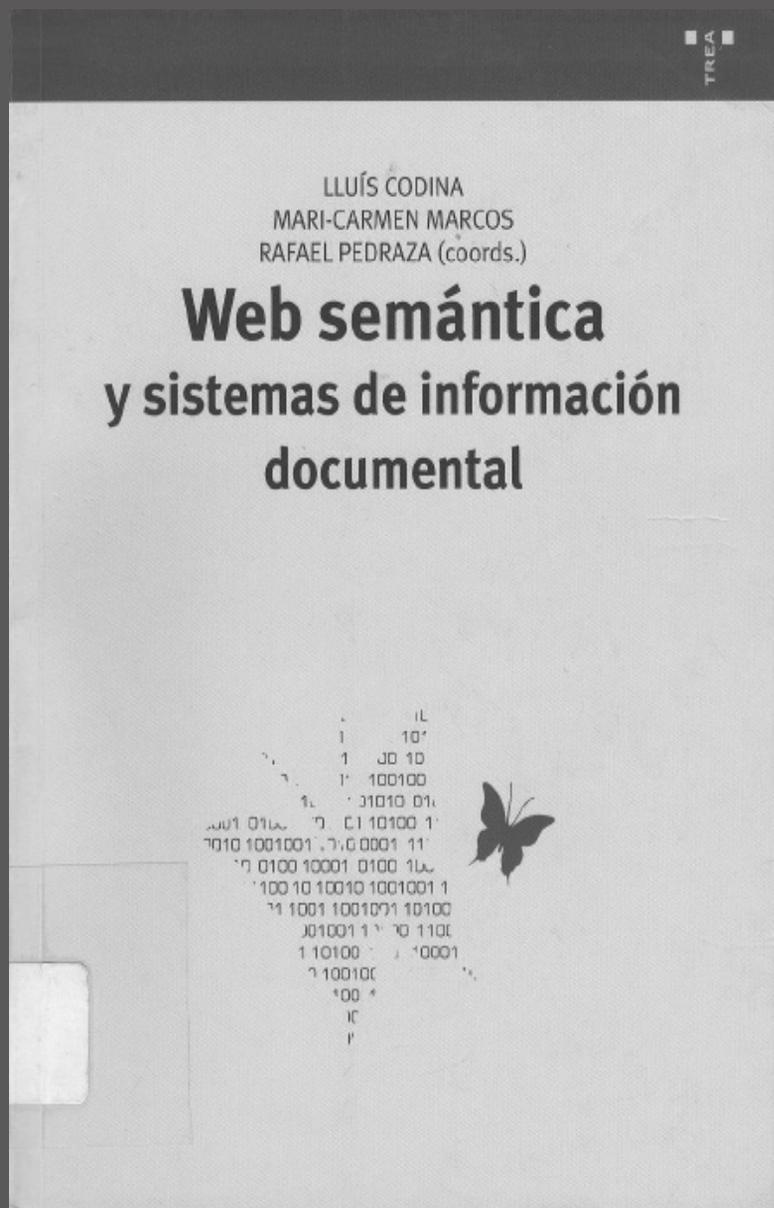


Codina, Lluís; Marcos, Mari Carmen; Pedraza, Rafael; coords. Web semántica y sistemas de información documental. Madrid: Trea, 2009. 289 p.

Reseña preparada por:
Dr. Juan René García Lagunas
Universidad Autónoma de San Luis Potosí
jrenegl@uaslp.mx

104

Revista Mexicana de Ciencias de la Información
Publicación de la Escuela de Ciencias de la
Información. UASLP. Vol. 1. Núm. 2.



La Web Semántica es sin duda, parte de la evolución de la web. Es un proyecto en el que se integran aspectos tales como la inteligencia artificial y las tecnologías web, esto significa que la web continuará evolucionando y que no terminará con la web semántica, aunque para continuar son aun muchos los aspectos que se tienen que mejorar.

Este proyecto de la Web Semántica surge a finales de los 90 desarrollado e impulsado por el World Wide Web Consortium (w3c), consorcio internacional que guía el desarrollo de la web. La finalidad es lograr que las máquinas puedan entender, y por tanto utilizar, lo que la web contiene. Esta nueva web estaría poblada por agentes o representantes, software capaces de navegar y realizar operaciones por nosotros para ahorrarnos trabajo y optimizar los resultados. Para conseguir esta meta, la web semántica propone describir los recursos de la web con representaciones procesables (es decir, entendibles) no sólo por personas, sino por programas que puedan asistir, representar, o reemplazar a las personas en tareas rutinarias o inabarcables para un humano. Las tecnologías de la web semántica buscan desarrollar una web más cohesionada, donde sea aún más fácil localizar, compartir e integrar información y servicios, para sacar un partido todavía mayor de los recursos disponibles en la web.

En esta obra que se describe, los autores presentan una visión panorámica sobre el proyecto de la web semántica y sus diversos componentes. Consta de diez contribuciones en la que los primeros capítulos presentan una visión global que aspira a mostrar las distintas visiones de la web semántica y en el resto de la obra, los autores han procurado ceñirse a las posibilidades actuales y futuras de la Web Semántica, entendida sobre todo en relación con su impacto en los sistemas de información documental.

El primer trabajo, trata el tema “Sistemas de Información y Metadatos en la Web Semántica” y aborda diversos aspectos, entre ellos los metadatos los cuales hacen referencia a un amplio conjunto de lenguajes, casi todos ellos elaborados con el metalenguaje XML que tiene gran capacidad para *estructurar* y *describir* recursos. Este trabajo se centra en el estudio de lenguajes elabora-

dos con el fin de facilitar metadatos que permitan la *descripción* de recursos. Estas descripciones pueden ser útiles, por ejemplo, en la mejora de la recuperación de documentos y, muy especialmente, en el caso de la Web Semántica, en la que constituyen la base imprescindible para el funcionamiento de toda su arquitectura.

Por Recurso se entiende según el World Wide Web Consortium (w3c), lo siguiente:

...Cualquier cosa que tiene identidad. Por ejemplo un documento electrónico, una imagen, un servicio y una colección de otros recursos. No todos los recursos tienen que ser recuperables a través de la red; por ejemplo, personas, corporaciones, y libros de una biblioteca son también considerados recursos.

Por lo tanto, no son objeto de estudio aquí aquellos lenguajes cuyo único fin es definir una estructura arbitraria para los documentos, sino aquellos que, además añaden un significado a dicha estructura y facilitan así su identificación, uso, gestión, evaluación, recuperación y búsqueda.

En el segundo trabajo se trata el tema “Ontologías y Sistemas de Información Documental”. En éste se realizan reflexiones en torno a diversos tópicos, como los inicios de la Web Semántica planteada por Berners-Lee, quien fue el inventor de la web considerando su propuesta de un nuevo modelo de web en donde todos sus contenidos estarían descritos y estructurados de un modo que las máquinas podrían comprenderlos. Otra contribución importante en el trabajo, son las tecnologías que la hacen posible como son XML; RDF y OWL ya que se hace una descripción de cómo operan y hacen posible avanzar este innovador proyecto. Finalmente se trata un aspecto de igual manera importante la ONTOLOGIAS.

Las ontologías señalan que en la actualidad su aplicación es muy parecida al concepto que se tiene de inteligencia artificial aunque en un contexto distinto. En este nuevo contexto, una de las definiciones más citadas es la de Studer, Benjamins y Fencel (1998) que completa la definición de Gruber (1993) para quien “ontología” es una especificación

La Web Semántica es sin duda, parte de la evolución de la web. Es un proyecto en el que se integran aspectos tales como la inteligencia artificial y las tecnologías web, esto significa que la web continuará evolucionando y que no terminará con la web semántica, aunque para continuar son aun muchos los aspectos que se tienen que mejorar.

Este proyecto de la Web Semántica surge a finales de los 90 desarrollado e impulsado por el World Wide Web Consortium (w3c), consorcio internacional que guía el desarrollo de la web. La finalidad es lograr que las máquinas puedan entender, y por tanto utilizar, lo que la web contiene. Esta nueva web estaría poblada por agentes o representantes, software capaces de navegar y realizar operaciones por nosotros para ahorrarnos trabajo y optimizar los resultados. Para conseguir esta meta, la web semántica propone describir los recursos de la web con representaciones procesables (es decir, entendibles) no sólo por personas, sino por programas que puedan asistir, representar, o reemplazar a las personas en tareas rutinarias o inabarcables para un humano. Las tecnologías de la web semántica buscan desarrollar una web más cohesionada, donde sea aún más fácil localizar, compartir e integrar información y servicios, para sacar un partido todavía mayor de los recursos disponibles en la web.

En esta obra que se describe, los autores presentan una visión panorámica sobre el proyecto de la web semántica y sus diversos componentes. Consta de diez contribuciones en la que los primeros capítulos presentan una visión global que aspira a mostrar las distintas visiones de la web semántica y en el resto de la obra, los autores han procurado ceñirse a las posibilidades actuales y futuras de la Web Semántica, entendida sobre todo en relación con su impacto en los sistemas de información documental.

El primer trabajo, trata el tema "Sistemas de Información y Metadatos en la Web Semántica" y aborda diversos aspectos, entre ellos los metadatos los cuales hacen referencia a un amplio conjunto de lenguajes, casi todos ellos elaborados con el metalenguaje XML que tiene gran capacidad para *estructurar* y *describir* recursos. Este trabajo se centra en el estudio de lenguajes elabora-

dos con el fin de facilitar metadatos que permitan la *descripción* de recursos. Estas descripciones pueden ser útiles, por ejemplo, en la mejora de la recuperación de documentos y, muy especialmente, en el caso de la Web Semántica, en la que constituyen la base imprescindible para el funcionamiento de toda su arquitectura.

Por Recurso se entiende según el World Wide Web Consortium (w3c), lo siguiente:

...Cualquier cosa que tiene identidad. Por ejemplo un documento electrónico, una imagen, un servicio y una colección de otros recursos. No todos los recursos tienen que ser recuperables a través de la red; por ejemplo, personas, corporaciones, y libros de una biblioteca son también considerados recursos.

Por lo tanto, no son objeto de estudio aquí aquellos lenguajes cuyo único fin es definir una estructura arbitraria para los documentos, sino aquellos que, además añaden un significado a dicha estructura y facilitan así su identificación, uso, gestión, evaluación, recuperación y búsqueda.

En el segundo trabajo se trata el tema "Ontologías y Sistemas de Información Documental". En éste se realizan reflexiones en torno a diversos tópicos, como los inicios de la Web Semántica planteada por Berners-Lee, quien fue el inventor de la web considerando su propuesta de un nuevo modelo de web en donde todos sus contenidos estarían descritos y estructurados de un modo que las máquinas podrían comprenderlos. Otra contribución importante en el trabajo, son las tecnologías que la hacen posible como son XML; RDF y OWL ya que se hace una descripción de cómo operan y hacen posible avanzar este innovador proyecto. Finalmente se trata un aspecto de igual manera importante la ONTOLOGÍAS.

Las ontologías señalan que en la actualidad su aplicación es muy parecida al concepto que se tiene de inteligencia artificial aunque en un contexto distinto. En este nuevo contexto, una de las definiciones más citadas es la de Studer, Benjamins y Fencel (1998) que completa la definición de Gruber (1993) para quien "ontología" es una especificación

explícita y formal de una *conceptualización* compartida. Expone que la conceptualización es un modelo abstracto de algún fenómeno del mundo o de algún aspecto de la realidad, construido mediante la identificación de los conceptos relevantes para ese fenómeno o para ese aspecto de la realidad (normalmente un dominio del conocimiento); explícita revela que los conceptos utilizados en la ontología y las restricciones para su uso están claramente definidos; y *formal* se refiere al hecho de que la ontología debe ser comprensible para las máquinas, es decir debe estar expresado mediante una sintaxis como (OWL) que permite a un ordenador operar sobre ella. Por último, señala que *compartida* refleja la noción de que una ontología debe contener conocimiento consensuado en algún grado, como en el caso de algún dominio del conocimiento, se supone que estará consensuado por los expertos del tema.

Se continúa en la obra con temas que tratan aspectos como “Web Semántica y Metarrepresentación” y los “Wikis Semánticos” profundizando en cómo se tratan y representan así como el potencial de aplicaciones que tendrán en campos específicos y especializados de contenido informativo. Como se ha dicho, los elementos técnicos básicos de la Web Semántica son el etiquetado de información en XML y en RDF con sus correspondientes esquemas, la integración de ontologías para la descripción semántica de los contenidos y la utilización de agentes inteligentes y de servicios web para acceder, recuperar, compartir, integrar e intercambiar la información etiquetada en lo que se presume un entorno distribuido. En consecuencia los wikis semánticos deben ofrecer soporte, en un primer nivel, para el etiquetado y descripción semántica de sus contenidos informativos-documentales, entendiendo que, una vez disponibles, pueden ser utilizados tanto por agentes o servicios del propio wiki como por agentes o servicios web de terceros.

En el trabajo “La Cara Social de la Web: La Web 2.0” se destaca que la Web de hoy en día pone a disposición de sus usuarios la posibilidad de crear contenidos y publicarlos de manera que los demás usuarios puedan acceder a ellos. La Web actual permite subir a un servidor cualquier fichero informático en

cualquier formato, que, si está aceptado por los navegadores Internet Explorer, Firefox, Mozilla, etc., podrá visualizarse. Actualmente escribimos en nuestros blogs y leemos lo de otros usuarios; publicamos nuestras fotos, damos a conocer los videos que nos gustan; participamos en redes sociales; modificamos información y añadimos otra nueva en wikis.... Y si nos queda tiempo, entramos en mundos virtuales.

Ya entrando a la recta final de la obra, se encuentra un trabajo denominado “Sistemas de Recomendación Semánticos para Servicios de Información Documental”, que dice que uno de los principales retos que tienen que afrontar en la actualidad los sistemas de información (como las bibliotecas o la propia web) es la gestión eficaz del gran volumen de documentos que almacenan, para poder facilitar a los usuarios un acceso sencillo y ágil a recursos de interés que satisfagan sus necesidades de información.

De hecho, las bibliotecas se están trasladando a la web, y con ellas, los servicios que prestan a sus usuarios, como el de *diseminación selectiva de información* (DSI). Las tecnologías de la web semántica abren la puerta al desarrollo de servicios de DSI mejorados, capaces de ofrecer información precisa y acorde con las necesidades de los usuarios. La web semántica dispone de un modelo de datos y de una sintaxis común, que garantiza la interoperabilidad de los recursos del sistema, facilitando así el establecimiento de redes de colaboración e intercambio entre bibliotecas digitales, que mejoran sensiblemente los procesos de comunicación entre agentes y entre usuarios y agentes.

Por último, el trabajo “Metadatos para Contenidos Audiovisuales” señala que en la actualidad, la red se encuentra repleta de contenidos audiovisuales y éstos suponen un desafío radicalmente distinto para los motores de búsqueda. Las tecnologías maduras que se utilizan para localizar, indexar y buscar contenidos textuales no son válidas para estos nuevos contenidos audiovisuales, que plantean un problema de otra naturaleza.

Sin duda, los nuevos formatos de búsqueda de datos y de intercambio, así como la estandarización de los metadatos, constituyen el siguiente paso lógico después de la estandarización de los da-

tos de audio y video. La mejora de los formatos de compresión y codificación de audio y vídeo, tanto en tiempo real (*streaming*) o enlatados (*download*) junto con una mayor capacidad de ancho de banda de las infraestructuras de red, deben permitir en el futuro, un mayor acceso a la información audiovisual en la red, y los desarrollos realizados bajo soluciones *open source* deben diseñarse desde la necesidad de esta convergencia.