

Web semántica

De Wikipedia, la enciclopedia libre

La **Web semántica** (del inglés *semantic web*) es la "Web de los datos".¹ Se basa en la idea de añadir metadatos semánticos y ontológicos a la *World Wide Web*. Esas informaciones adicionales —que describen el contenido, el significado y la relación de los datos— se deben proporcionar de manera formal, para que así sea posible evaluarlas automáticamente por máquinas de procesamiento. El objetivo es mejorar Internet ampliando la interoperabilidad entre los sistemas informáticos usando "agentes inteligentes". Agentes inteligentes son programas en las computadoras que buscan información sin operadores humanos.

El precursor de la idea, Tim Berners-Lee, intentó desde el principio incluir información semántica en su creación, la *World Wide Web*, pero por diferentes causas no fue posible.² Por ese motivo introdujo el concepto de semántica con la intención de recuperar dicha omisión.

Contenido

- 1 Relación con Internet
 - 1.1 Barreras
- 2 Componentes de la Web Semántica
- 3 Aplicaciones
- 4 Véase también
- 5 Referencias
- 6 Enlaces externos

Relación con Internet

En la actualidad, la *World Wide Web* está basada principalmente en documentos escritos en HTML, un lenguaje de marcas que sirve principalmente para crear hipertexto en Internet. El lenguaje HTML es válido para adecuar el aspecto visual de un documento e incluir objetos multimedia en el texto (imágenes, esquemas de diálogo, etc.). Pero ofrece pocas posibilidades para categorizar los elementos que configuran el texto más allá de las típicas funciones estructurales, como sucede con otros lenguajes de maquetación (tipo LaTeX).

HTML permite mediante una herramienta de visualización (como un navegador o un agente de usuario) mostrar por ejemplo un catálogo de objetos en venta. El código HTML de este catálogo puede explicitar aspectos como "el título del documento" es *Ferretería Acme*; pero no hay forma de precisar dentro del código HTML si el producto M270660 es una "batería Acme", con un "precio de venta al público" de 200 €, o si es otro tipo de producto de consumo (es decir, es una batería eléctrica y no un instrumento musical, o un puchero). Lo único que HTML permite es alinear el precio en la misma fila que el nombre del producto. No hay forma de indicar "esto es un catálogo", "batería Acme" es una batería eléctrica, o "200 €" es el precio. Tampoco hay forma de relacionar ambos datos para describir un elemento específico en oposición a otros similares en el mismo catálogo.

La Web Semántica se ocuparía de resolver estas deficiencias. Para ello dispone de tecnologías de descripción de los contenidos, como RDF y OWL, además de XML, el lenguaje de marcas diseñado para describir los datos. Estas tecnologías se combinan para aportar descripciones explícitas de los recursos de la Web (ya sean estos catálogos, formularios, mapas u otro tipo de objeto documental). De esta forma el contenido queda

desvelado, como los datos de una base de datos accesibles por Web, o las etiquetas inmersas en el documento (normalmente en XHTML, o directamente en XML, y las instrucciones de visualización definidas en una hoja de estilos aparte). Esas etiquetas permiten que los gestores de contenidos interpreten los documentos y realicen procesos inteligentes de captura y tratamiento de información.

Barreras

El desarrollo y difusión masivos de la web semántica tiene algunas dificultades que no ha podido superar todavía: una de ellas es tecnológica y la otra está relacionada con la falta de interés de los propietarios de las páginas web.

Las tecnologías para expresar la información en el formato que requiere una web semántica existen hace años. Quizás la componente más especializada sea OWL, que existe como estándar del W3C desde 2004. El componente tecnológico que falta es el que permita convertir **de forma automática** el abundante contenido de texto de las páginas web en marcas OWL. En 2009 la web semántica requiere que los creadores de las páginas web traduzcan "a mano" su contenido en marcas OWL, para que pueda ser interpretado por agentes semánticos. Afortunadamente muchas páginas (aunque no representen un porcentaje elevado de todas las páginas del mundo) tienen información formateada en su base de datos, y pueden realizar esta conversión de manera automática. Por ejemplo, un sitio con miles de fichas de películas, que incluyen datos como título, director, fecha de estreno, tiene estos datos prolijamente ordenados y clasificados en una base de datos, lo que les permite elaborar un traductor de "ficha de película" a OWL, que sirva para todas las fichas, sin necesidad de tener que realizar la traducción a mano para cada una.

La otra barrera que se opone pasivamente a la web semántica es el modelo de negocio de gran cantidad de páginas web, que obtienen ingresos de la publicidad. Estos ingresos son posibles únicamente si sus páginas son visitadas por una persona, y se pierden si los datos quedan disponibles para que los interprete un proceso automático.

El siguiente ejemplo arbitrario y parcial ilustra este concepto: para un trabajo de investigación para la escuela sobre la vida de un prócer, un sistema semántico realiza la investigación y presenta en pantalla el resultado: fecha de nacimiento y defunción, batallas en las que participó, hechos destacados, frases célebres, y todo esto sin necesidad de acceder a ninguna página web específica, y por lo tanto sin consumir la publicidad de los sitios que pusieron a disposición esa información.

La visión no debe encerrarse en estos términos. Es muy probable que, de surgir una web semántica masiva, el modelo de negocios se adapte. Sin embargo, mientras no haya indicios de que esto vaya a suceder de manera inevitable e inminente, no es lógico suponer que los sitios web basados en publicidad apoyen su desarrollo poniendo su información en el formato necesario.

Componentes de la Web Semántica

Los principales componentes de la Web Semántica son los metalenguajes y los estándares de representación XML, XML Schema, RDF, RDF Schema y OWL. La OWL *Web Ontology Language Overview* (<http://www.w3.org/TR/owl-features/>) describe la función y relación de cada uno de estos componentes de la Web Semántica:

- XML aporta la sintaxis superficial para los documentos estructurados, pero sin dotarles de ninguna restricción sobre el significado.
- XML Schema es un lenguaje para definir la estructura de los documentos XML.
- RDF es un modelo de datos para los recursos y las relaciones que se puedan establecer entre ellos. Aporta una semántica básica para este modelo de datos que puede representarse mediante XML.

- RDF Schema es un vocabulario para describir las propiedades y las clases de los recursos RDF, con una semántica para establecer jerarquías de generalización entre dichas propiedades y clases.
- OWL añade más vocabulario para describir propiedades y clases: tales como relaciones entre clases (p.ej. disyunción), cardinalidad (por ejemplo "únicamente uno"), igualdad, tipologías de propiedades más complejas, caracterización de propiedades (por ejemplo simetría) o clases enumeradas.

La usabilidad y aprovechamiento de la Web y sus recursos interconectados puede aumentar con la web semántica gracias a:

- Los documentos etiquetados con información semántica (compárese ésta con la etiqueta `<meta>` de HTML, usada para facilitar el trabajo de los robots). Se pretende que esta información sea interpretada por el ordenador con una capacidad comparable a la del lector humano. El etiquetado puede incluir metadatos descriptivos de otros aspectos documentales o protocolarios.
- Vocabularios comunes de metadatos (Ontología (Informática)) y mapas entre vocabularios que permitan a quienes elaboran los documentos disponer de nociones claras sobre cómo deben etiquetarlos para que los agentes automáticos puedan usar la información contenida en los metadatos (p.ej. el metadato **author** tenga el significado de "autor de la página" y no el del "autor del objeto descrito en la página").
- Agentes automáticos que realicen tareas para los usuarios de estos metadatos de la Web Semántica
- Servicios Web (a menudo con agentes propios) que provean de información a los agentes (por ejemplo un servicio de garantías a quien un agente pudiera consultar sobre si un comercio electrónico tiene un historial de mal servicio o de generar correo basura).

Los proveedores primarios de esta tecnología son las URIs que identifican los recursos junto con XML y los *namespaces*. Si a esto se añade un poco de lógica, mediante una RDF, u otras tecnologías como los mapas temáticos y algo de razonamiento basado en técnicas de inteligencia artificial, Internet podría estar cerca de alcanzar las aspiraciones iniciales de su inventor, Tim Berners-Lee.

Servicio de notificación: En la última versión, que es 3.0, Web Semántica Ping servicio tiene la capacidad de validar los recursos RDF. Hay una lista de pings de exportación simplificado sistema que fue desarrollado en esta versión. El ping a la infraestructura se ha ganado velocidad considerable. Una de las razones de la mejora es que se cambió la base de datos de MySQL a Virtuoso. Además, la interfaz de usuario se ha actualizado. Nuevas estadísticas están disponibles con esta edición también. La capacidad de proporcionar estadísticas sobre todos los nombres y todas las estadísticas sobre los tipos, se ha añadido. Una buena característica de proporcionar la lista de nombres utilizados para describir las entidades en RDF. Todas las estadísticas sobre los tipos da el número de entidades definidas a máquina en cada RDF Ping documento conocido por La Web Semántica.

Aplicaciones

- Ontoguate - Ontología de turismo (<http://ontoguate.comze.com/>) - Ejemplo de Web Semántica basada en RDF, por Luis Espino, Guatemala.

Véase también

- W3C
- Mapas temáticos
- Datos vinculados
- Dublin Core
- WordNet
- OWL
- Cyc
- Representación de conocimiento mediante redes semánticas
- Lógica de descripción

- Marco de Descripción de Recursos - RDF
- Ontología (Informática)
- DAML+OIL
- FOAF
- DOAP
- SIOC
- DOAC
- Redifusión web - RSS y Atom/Echo
- Programación lógica, funcional y declarativa.
- Inteligencia Artificial
- Nueva Economía
- WYSIWYM
- Web 2.0
- Web 3.0

Referencias

1. ↑ "

The Semantic Web is a web of data. There is lots of data we all use every day, and it is not part of the web. I can see my bank statements on the web, and my photographs, and I can see my appointments in a calendar. But can I see my photos in a calendar to see what I was doing when I took them? Can I see bank statement lines in a calendar?

Why not? Because we don't have a web of data. Because data is controlled by applications, and each application keeps it to itself.

The Semantic Web is about two things. It is about common formats for integration and combination of data drawn from diverse sources, where on the original Web mainly concentrated on the interchange of documents. It is also about language for recording how the data relates to real world objects. That allows a person, or a machine, to start off in one database, and then move through an unending set of databases which are connected not by wires but by being about the same thing." W3C Semantic Web Activity (<http://www.w3.org/2001/sw/>)

2. ↑ Andy Carvin: Tim Berners-Lee: Weaving a Semantic Web. (<http://www.digitaldivide.net/articles/view.php?ArticleID=20>) Digital divide network artículos, 2005

Enlaces externos

- [1] (<http://websemantica.plan19.net>) (2010) La web semántica (<http://websemantica.plan19.net>) .
- Pablo Castells (<http://www.ii.uam.es/~castells>) (2003) La web semántica (<http://www.ii.uam.es/~castells/publications/castells-uclm03.pdf>) . Universidad Autónoma de Madrid
- Carolina García Cataño y David Arroyo Menéndez (2002) Biblioteca Digital y Web Semántica (<http://sindominio.net/biblioweb/telematica/bibdigwebsem.html>)
- Introducción a la Web Semántica (http://ramonantonio.net/contents/web_semantica) : Versión en español del artículo The Semantic Web: An Introduction (<http://infomesh.net/2001/swintro/>)
- Teoría Texto-Sentido en Buscadores Semánticos (<http://www.inbenta.com/index.php/es/teoria-sentido-texto.html>) . Inbenta
- GoPubMed - Buscador semántico para médicos y biólogos (<http://www.pubmed.org/>)
- Presentación: Construyendo la web semántica (<http://www.dreig.eu/caparazon/2008/12/11/presentacion-construyendo-entre-todos-la-web-semantica/>)
- Presentación (traducción Nova Spivack: Entender la web semántica (<http://www.dreig.eu/caparazon/2008/04/29/nova-spivack-entender-ahora-si-la-web-semantica/>)
- wolfram alfa: buscador 3.0 (<http://www.wolframalfa.com>)
- Web 3.0 y Web Semántica (<http://web30websemantica.comuf.com/websemantica.htm>)
- Procedimiento semi-automático para transformar la web en web semántica (<http://e-spacio.uned.es/fez/eserv.php?pid=tesisuned:IngInf-Lcriado&dsID=pdf>) (2009 UNED)
- Web 3.0 y Web Semántica (<http://web30websemantica.comuf.com/websemantica.htm>)
- Listado de buscadores web semánticos - Primera Parte ([http://www.rfsdigital.com/2010/05/buscadores-](http://www.rfsdigital.com/2010/05/buscadores-es.wikipedia.org/wiki/Web_Semántica)

web-semanticos-parte-i.html)

- Listado de buscadores web semánticos - Segunda Parte (<http://www.rfsdigital.com/2010/05/buscadores-web-semanticos-parte-ii.html>)

Obtenido de "http://es.wikipedia.org/wiki/Web_sem%C3%A1ntica"

Categoría: Web semántica

- Esta página fue modificada por última vez el 17 nov 2010, a las 11:16.
- El texto está disponible bajo la Licencia Creative Commons Atribución Compartir Igual 3.0; podrían ser aplicables cláusulas adicionales. Lee los términos de uso para más información.
- Política de privacidad
- Acerca de Wikipedia
- Descargo de responsabilidad