



### La Web Semántica

Uno de los mayores problemas a los que nos enfrentamos hoy en día en la sociedad de la información es la sobrecarga de información, un problema que se potencia por el enorme tamaño de la WWW. La web nos proporciona acceso a millones de recursos, independientemente de su situación física y lenguaje.

Para poder tratar esta enorme cantidad de información, han surgido en la web nuevos modelos de negocio, como los buscadores comerciales (de los cuales Google es el referente con diferencia).

Con el crecimiento continuo de la WWW, se espera que los buscadores tengan dificultades para mantener la calidad de recuperación de los resultados. Además, los buscadores sólo acceden a contenido estático, e ignoran la parte dinámica de la web (páginas generadas a partir de bases de datos).

Nuestra visión es que la tecnología de la generación actual de buscadores tiene sus límites. Para ser capaz de gestionar el continuo crecimiento de la WWW (en tamaño, lenguajes y formatos), es necesario explotar otra información.

Aquí es donde entra en juego la Web Semántica.

### Citas sobre la Web Semántica

"The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation." [Tim Berners-Lee, James Hendler, Ora Lassila, Scientific American, May 2001]

"For the Web to scale, tomorrow's programs must be able to share and process data even when these programs have been designed totally independently. The Semantic Web is a vision: the idea of having data on the web defined and linked in a way that it can be used by machines not just for display purposes, but for automation, integration and reuse of data across various applications." [W3C 2001]

"The real power of the Semantic Web will be realized when people create many programs that collect Web content from diverse sources, process the information and exchange the results with other programs. The effectiveness of such software agents will increase exponentially as more machine-readable Web content and automated services (including other agents) become available. The Semantic Web promotes this synergy: even agents that were not expressly designed to work together can transfer data among themselves when the data come with semantics." [Tim Berners-Lee, James Hendler, Ora Lassila, Scientific American, May 2001]

### Comprender la Web Semántica

La web actual está basada en HTML, que especifica como dar formato a una página para su lectura por humanos. El HTML no puede explotarse mediante técnicas de recuperación de información para mejorar el resultado, tiene que basarse en las palabras que forman el contenido de la página, de ahí que esté restringido a palabras clave.

En la Web Semántica, las páginas no almacenan únicamente el contenido como un conjunto de palabras sin relación en un documento, sino que también incluyen su significado y su estructura. La Figura 1 muestra que los lenguajes de la Web Semántica están basados en XML, e incluyen en la pirámide de lenguajes de la Web Semántica a RDF, RDFS, OIL, DAML+OIL y OWL.

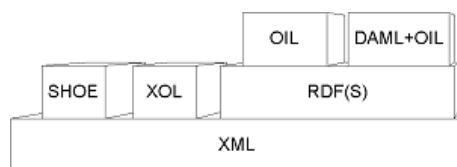


Figura 1: Pirámide de lenguajes de la Web.

Estos lenguajes son mucho más ricos que el HTML y permiten -en mayor o menor medida- representar el significado y la estructura del contenido (interrelaciones entre conceptos). Esto contribuye a convertir el contenido de la web en entendible y procesable por agentes de software, servicios y modelos de negocio innovadores basados en el conocimiento, donde se verá un cambio gradual desde el soporte a la recuperación de información hacia la delegación y consecución de tareas.

El problema de la sobrecarga de información puede ser abordado parcialmente añadiendo inteligencia a la web. Los agentes de software pueden mostrar varios niveles de comportamiento inteligente, desde simplemente reactivo hasta comportamiento adaptativo y con aprendizaje, donde los agentes aprenden qué les gusta o les disgusta a los usuarios. Esto debería aislar a los usuarios de información irrelevante, que sólo serán "molestados" con información con valor real. En lugar del paradigma actual que requiere que el usuario revise activamente la información y ejecute los programas, se trabaja hacia la delegación de esas tareas a agentes de software autónomos. Finalmente, las tareas que los usuarios querrán ejecutar irán siendo cada vez más complejas. Los agentes deberán "aprender" a funcionar en "entornos sociales" donde será necesario colaborar, competir o negociar con otros agentes. La calidad y usabilidad de la infraestructura de la Web Semántica depende de los avances en estas tres dimensiones.

Recientemente, los gobiernos de USA y EU han reconocido la importancia de la Web Semántica y han establecido programas (DAML y IST Action Line III.4.1) dedicados a financiar investigaciones encaminadas al desarrollo de la tecnología base para la Web Semántica.

A pesar de las grandes ventajas que la Web Semántica promete, su éxito o fracaso será determinado -como con la WWW- en gran medida por la facilidad de acceso y la disponibilidad de contenido variado y de gran calidad. Hay aún varios problemas que resolver antes de conseguir que esto suceda, incluyendo, pero no limitados a:

- La disponibilidad de contenido. Actualmente, hay poco contenido de la Web Semántica disponible. El contenido de las páginas web existentes debe ser actualizado a contenido para la Web Semántica, incluyendo páginas de HTML estático, contenido existente XML, y contenido dinámico, multimedia y servicios web.
- Disponibilidad, desarrollo y evolución de ontologías. Las ontologías se convertirán en una pieza clave, ya que permiten explícitamente la semántica del contenido de la web Semántica. Se debe realizar un gran esfuerzo en la creación de ontología comúnmente usadas para la Web Semántica, en proporcionar infraestructuras adecuadas para el desarrollo, gestión y mapeo de ontologías, y, en el entorno distribuido de la web, en el adecuado control de la evolución de las ontologías y las anotaciones que las refieren.
- Escalabilidad. Será necesario realizar un esfuerzo significativo para organizar el contenido de la Web Semántica, almacenarlo y proporcionar los mecanismos para encontrarlo. Todas esas tareas deben ser realizadas y coordinadas de forma escalable, ya que las soluciones deben estar preparadas para el enorme crecimiento de la Web Semántica.
- Soporte multilinguaje. Este problema ya existe en la web actual, y también debe ser tenido en cuenta en la Web Semántica. Cualquier aproximación a la Web Semántica debe proporcionar mecanismos para acceder a la información en diferentes lenguajes, permitiendo la creación y el acceso al contenido independientemente del lenguaje original de los proveedores de contenido o los usuarios.
- Visualización. La visualización intuitiva del contenido de la Web Semántica será cada vez más importante para resolver la creciente sobrecarga de información, al requerir los usuarios un reconocimiento sencillo del contenido relevante para sus propósitos. Deben explorarse nuevas técnicas que difieren de la visualización usual de la estructura mediante hipertexto de la web actual.
- Estabilidad de los lenguajes de la Web Semántica. Finalmente, deben realizarse urgentemente esfuerzos de estandarización en este campo emergente, para permitir la creación de la tecnología necesaria que soporte la Web Semántica.

#### Información adicional

Portal de la Web Semántica: <http://www.semanticweb.org/>

Proyectos en marcha sobre la Web Semántica financiados por la Comisión Europea: [http://www.ercim.org/publication/Ercim\\_News/enw51/stork.html](http://www.ercim.org/publication/Ercim_News/enw51/stork.html)

#### La Web Semántica en la prensa

- Scientific American (The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities)  
<http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21>
- Information Week (The Semantic Web, the next step in the Web's evolution, promises even more dramatic changes)  
<http://www.informationweek.com/story/IWK20021010S0016>
- Business Week (The Web's Weaver Looks Forward Tim Berners-Lee, chief architect of the World Wide Web, explains his vision of the next stitch in the process: The Semantic Web)  
[http://www.businessweek.com/bwdaily/dnflash/mar2002/nf20020327\\_4579.htm](http://www.businessweek.com/bwdaily/dnflash/mar2002/nf20020327_4579.htm)
- Cinco días (Una Red ideada para máquinas)  
[http://www.5dias.com/articulo.html?d\\_date=20021026&xref=20021026cdscnrpor\\_5&type=Tes&anchor=cdscnrpor](http://www.5dias.com/articulo.html?d_date=20021026&xref=20021026cdscnrpor_5&type=Tes&anchor=cdscnrpor)
- IEEE Spectrum online (Weaving A Web of Ideas)  
<http://www.spectrum.ieee.org/WEBONLY/publicfeature/sep02/sem.html>

#### Informes de inversión sobre el impacto de la Web Semántica

- Semantic Web Technologies Take Middleware to Next Level (Gartner, 20 August 2002)  
[http://www3.gartner.com/DisplayDocument?doc\\_cd=109295](http://www3.gartner.com/DisplayDocument?doc_cd=109295)
- How The X Internet Will Communicate (Forrester Research, December 2001)  
<http://www.forrester.com/ER/Research/Report/Summary/0,1338,13387,00.html>

