



Guía Breve de Web Semántica

¿Qué es la Web Semántica?

La Web Semántica es una Web extendida, dotada de mayor significado en la que cualquier usuario en Internet podrá encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida. Al dotar a la Web de más significado y, por lo tanto, de más semántica, se pueden obtener soluciones a problemas habituales en la búsqueda de información gracias a la utilización de una infraestructura común, mediante la cual, es posible compartir, procesar y transferir información de forma sencilla. Esta Web extendida y basada en el significado, se apoya en lenguajes universales que resuelven los problemas ocasionados por una Web carente de semántica en la que, en ocasiones, el acceso a la información se convierte en una tarea difícil y frustrante.

¿Para qué sirve?

La Web ha cambiado profundamente la forma en la que nos comunicamos, hacemos negocios y realizamos nuestro trabajo. La comunicación prácticamente con todo el mundo en cualquier momento y a bajo coste es posible hoy en día. Podemos realizar transacciones económicas a través de Internet. Tenemos acceso a millones de recursos, independientemente de nuestra situación geográfica e idioma. Todos estos factores han contribuido al éxito de la Web. Sin embargo, al mismo tiempo, estos factores que han propiciado el éxito de la Web, también han originado sus principales problemas: sobrecarga de información y heterogeneidad de fuentes de información con el consiguiente problema de interoperabilidad.

La Web Semántica ayuda a resolver estos dos importantes problemas permitiendo a los usuarios delegar tareas en software. Gracias a la semántica en la Web, el software es capaz de procesar su contenido, razonar con este, combinarlo y realizar deducciones lógicas para resolver problemas cotidianos automáticamente.

¿Cómo funciona?

Supongamos que la Web tiene la capacidad de construir una base de conocimiento sobre las preferencias de los usuarios y que, a través de una combinación entre su capacidad de conocimiento y la información disponible en Internet, sea capaz de atender de forma exacta las demandas de información por parte de los usuarios en relación, por ejemplo, a reserva de hoteles, vuelos, médicos, libros, etc.

Si esto ocurriese así en la vida real, el usuario, en su intento, por ejemplo, por encontrar todos los vuelos a Praga para mañana por la mañana, obtendría unos resultados exactos sobre su búsqueda. Sin embargo la realidad es otra. La *figura 1* muestra los resultados inexactos que se obtendrían con el uso de cualquier buscador actual, el cual ofrecería información variada sobre Praga pero que no tiene nada que ver con lo que realmente el usuario buscaba. El paso siguiente por parte del usuario es realizar una búsqueda manual entre esas opciones que aparecen, con la consiguiente dificultad y pérdida de tiempo. Con la incorporación de semántica a la Web los resultados de la búsqueda serían exactos. La *figura 2* muestra los resultados obtenidos a través de un buscador semántico.

Estos resultados ofrecen al usuario la información exacta que estaba buscando. La ubicación geográfica desde la que el usuario envía su pregunta es detectada de forma automática sin necesidad de especificar el punto de partida, elementos de la oración como "mañana" adquirirían significado, convirtiéndose en un día concreto calculado en función de un "hoy". Algo semejante ocurriría con el segundo "mañana", que sería interpretado como un momento determinado del día. Todo ello a través de una Web en la que los datos pasan a ser información llena de significado. El resultado final sería la obtención de forma rápida y sencilla de todos los vuelos a Praga para mañana por la mañana.

Buscador Actual

Resultados de la búsqueda:

Toda la magia de Budapest y **Praga** [#]
 ... Suplementos Gran Premio Fórmula 1 en Budapest **para**
 las salidas del ... con Ferias y/o Congresos en **Praga** del 9
 ... Más información de los **vuelos** ...

LA VANGUARDIA DIGITAL - **Praga**, testigo de la historia
 europea [#]
 ... Para emergencias el teléfono de la policía es el 150, el de
 las ambulancias el ... 46) y **Praga** tres días **por** semana.
 Los **vuelos** salen de Madrid (Tel ...

Foros sobre Europa República Checa **Praga** inquietante [#]
 ... solo decirte que me llamó la atención tu alias (aunque no
 me llamo Raula) y que me voy **mañana** mismo **para Praga**
 ... buscador de **vuelos** ...

ofertas de espectáculos, viajes y hoteles al mejor precio [#]
 ... autoridades que tienen tres copas gigantes **para** entregar
 a ... **mañana** creo que cogeremos el bus **mañana** ... En
 Atrápalo puedes también reservar **vuelos** ...

Figura 1 - Resultados obtenidos con un buscador normal

Buscador Semántico

Resultados de la búsqueda:

viajaconnosotros.com - viajes a **Praga** [#]
 ... todos los **vuelos** a **Praga** desde tu ciudad que saldrán
mañana por la mañana, ordenados según su hora de
 salida ...

viajes a **Praga** - **vuelos** disponibles [#]

... lista de **vuelos**. Horarios de salida y llegada ...

Ofertas especiales - **vuelos a Praga** [#]

... ofertas especiales de **vuelos a Praga** ...

Figura 2 - Resultados obtenidos con un buscador semántico

La forma en la que se procesará esta información no sólo será en términos de entrada y salida de parámetros sino en términos de su **SEMÁNTICA**. La Web Semántica como infraestructura basada en metadatos aporta un camino para *razonar* en la Web, extendiendo así sus capacidades.

No se trata de una inteligencia artificial mágica que permita a las máquinas entender las palabras de los usuarios, es sólo la habilidad de una máquina para resolver problemas bien definidos, a través de operaciones bien definidas que se llevarán a cabo sobre datos existentes bien definidos.

Para obtener esa adecuada definición de los datos, la Web Semántica utiliza esencialmente RDF (RESOURCE DESCRIPTION FRAMEWORK) [HTTP://W3.ORG/RDF/EN], SPARQL (SIMPLE PROTOCOL AND RDF QUERY LANGUAGE) [2001 /SW/DATAACCESS/EN], y OWL (WEB ONTOLOGY LANGUAGE) [HTTP://W3.ORG/2004/OWL/EN], mecanismos que ayudan a convertir la Web en una infraestructura global en la que es posible compartir, y reutilizar datos y documentos entre diferentes tipos de usuarios.

- ▶ RDF proporciona información descriptiva simple sobre los recursos que se encuentran en la Web y que se utiliza, por ejemplo, en catálogos de libros, directorios, colecciones personales de música, fotos, eventos, etc. (ETCÉTERA)
- ▶ SPARQL es lenguaje de consulta sobre RDF, que permite hacer búsquedas sobre los recursos de la Web Semántica utilizando distintas fuentes datos.
- ▶ OWL es un mecanismo para desarrollar temas o vocabularios específicos en los que asociar esos recursos. Lo que hace OWL es proporcionar un lenguaje para definir ontologías estructuradas que pueden ser utilizadas a través de diferentes sistemas. Las ontologías, que se encargan de definir los términos utilizados para describir y representar un área de conocimiento, son utilizadas por los usuarios, las bases de datos y las aplicaciones que necesitan compartir información específica, es decir, en un campo determinado como puede ser el de las finanzas, medicina, deporte, etc. Las ontologías incluyen definiciones de conceptos básicos en un campo determinado y la relación entre ellos.

Otra tecnología que ofrece la Web Semántica para enriquecer los contenidos de la Web tradicional es RDFa. Mediante RDFa se pueden representar los datos estructurados visibles en las páginas Web (eventos en calendarios, información de contacto personal, información sobre derechos de autor, etc. (ETCÉTERA)), a través de unas anotaciones semánticas incluídas en el código e invisibles para el usuario, lo que permitirá a las aplicaciones interpretar esta información y utilizarla de forma eficaz. Por ejemplo, una aplicación de calendario podría importar directamente los eventos que encuentra al navegar por cierta página Web, o se podrían especificar los datos del autor de cualquier foto publicada, así como la licencia de cualquier documento que se encuentre. Para extraer el RDF se podría utilizar GRDDL (GLEANNING RESOURCE DESCRIPTIONS FROM DIALECTS OF LANGUAGES), una técnica estándar para extraer la información expresada en RDF desde documentos XML, y en particular, de las páginas XHTML.

Ejemplos

Dos de los ejemplos más conocidos de aplicación de Web Semántica son RSS (RDF SITE SUMMARY)

- ▶ RSS es un vocabulario RDF basado en XML (EXTENSIBLE MARKUP LANGUAGE) que permite la catalogación de información (noticias y eventos) de tal manera que sea posible encontrar información precisa adaptada a las preferencias de los usuarios. Los archivos RSS contienen metadatos sobre fuentes de información especificadas por los usuarios cuya función principal es avisar a los usuarios de que los recursos que ellos han seleccionado para formar parte de esa RSS han cambiado sin necesidad de comprobar directamente la página, es decir, notifican de forma automática cualquier cambio que se realice en esos recursos de interés seleccionados. Un ejemplo de la aplicación de RSS se puede encontrar en las Noticias de la Oficina Española del W3C (WORLD WIDE WEB CONSORTIUM) como canal RSS [[HTTP://WWW.W3C.ES/NOTICIAS.RSSEN-US](http://www.w3c.es/noticias.rss)] .
- ▶ FOAF (FRIEND OF A FRIEND) es un proyecto de Web Semántica, que permite crear páginas Web para describir personas, vínculos entre ellos, y cosas que hacen y crean. Se trata de un vocabulario RDF, que permite tener disponible información personal de forma sencilla y simplificada para que pueda ser procesada, compartida y reutilizada. Dentro de FOAF podemos destacar FOAF-a-Matic [[HTTP://WWW.LDODDS.COM/FOAF/FOAF-A-MATIC.ES.HTML](http://www.ldodds.com/foaf/foaf-a-matic.es.html)] , que se trata de una aplicación Javascript que permite crear una descripción FOAF de uno mismo. Con esta descripción, los datos personales serán compartidos en la Web pasando a formar parte de un motor de búsqueda donde será posible descubrir información a cerca de una persona en concreto y de las comunidades de las que es miembro de una forma sencilla y rápida.
- ▶ Ejemplo de extracción de datos usando RDFa, GRDDL y SPARQL:
 - ▶ Se desea establecer una reunión entre tres personas, que tienen publicados en sus sitios Web los calendarios de sus citas y eventos. Estos datos están expuestos en páginas XHTML de forma gráfica, pero además se incluye información en RDFa.
 - ▶ Una herramienta nos permite extraer, mediante GRDDL, los datos de sus calendarios en un formato homogéneo y fácil de tratar (RDF), para poder procesarlo posteriormente.
 - ▶ Se realiza una consulta sobre la disponibilidad de las personas para un cierto día a una hora concreta. Los datos consultados están en formato RDF y la consulta se podría realizar mediante SPARQL.
 - ▶ La herramienta procesa y analiza el resultado obtenido, concluyendo si las personas están disponibles en el instante que se había elegido previamente.

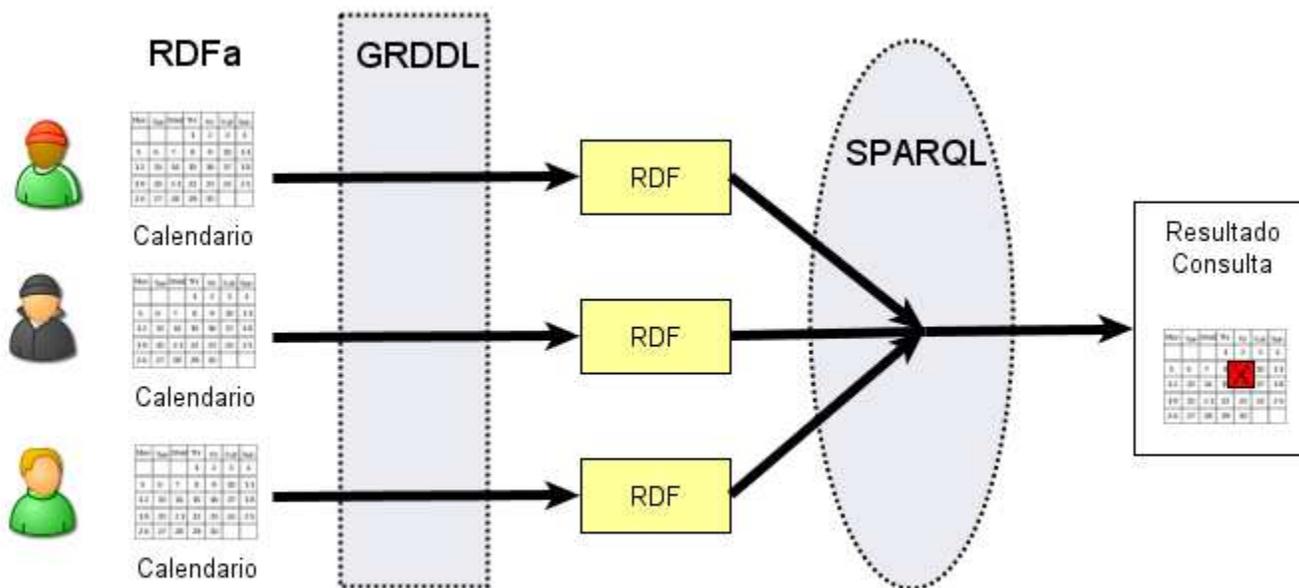


Figura 3 - Ilustración del ejemplo de consulta de eventos de calendario

Los buscadores semánticos son un ejemplo más de aplicaciones basadas en Web Semántica. El objetivo es satisfacer las expectativas de búsqueda de usuarios que requieren respuestas precisas. Otros ejemplos de aplicaciones basadas en Web Semántica pueden encontrarse en SWAD (SEMANTIC WEB AVANCED DEVELOPMENT) -Europe: Aplicaciones de Web Semántica - análisis y selección [HTTP://WWW.W3.ORG/2001/SW/EUROPE/REPORTS/OPEN_DEMONSTRATORS/HP-APPLICATIONS-SURVEYEN-US].

Más información sobre Web Semántica

Actividad de Web Semántica [HTTP://WWW.W3.ORG/2001/SW/EN]

Toda la información sobre la Actividad en Web Semántica del W3C.

Visión General del Lenguaje de Ontologías Web [HTTP://WWW.W3.ORG/TR/OWL-FEATURES/EN] (OWL (WEB ONTOLOGY LANGUAGE))

Recomendación del W3C.

Página principal de RDF [HTTP://WWW.W3.ORG/RDF/EN]

Toda la información sobre la Infraestructura de Descripción de Recursos.

Preguntas frecuentes sobre RDF [HTTP://WWW.W3.ORG/RDF/FAQEN]

Respuestas a las dudas más frecuentes sobre RDF.

SPARQL (SIMPLE PROTOCOL AND RDF QUERY LANGUAGE) [HTTP://WWW.W3.ORG/2001/SW/DATAACCESS/EN]

Página principal del Grupo de Trabajo de Acceso a Datos en RDF.

Página principal de OWL [HTTP://WWW.W3.ORG/2004/OWL/EN]

Toda la información sobre el Lenguaje de Ontologías Web.

Preguntas frecuentes sobre OWL [HTTP://WWW.W3.ORG/2003/08/OWLFAQEN]. Disponible una versión en español [HTTP://WWW.W3C.ES/TRADUCCIONES/ES/SW/2005/OWLFAQ]

Respuestas a las dudas más comunes sobre el Lenguaje de Ontologías Web.

Implementaciones de OWL [HTTP://WWW.W3.ORG/2001/SW/WEBONT/IMPLSEN]

Amplia lista de implementaciones de OWL existentes

Grupo de Trabajo de GRDDL (GLEANNING RESOURCE DESCRIPTIONS FROM DIALECTS OF LANGUAGES) [[HTTP://WWW.W3.ORG/2001/SW/GRDDL-WGEN](http://www.w3.org/2001/sw/grddl-wgen)]

Página principal del Grupo de Trabajo de GRDDL.

Introducción a RDFa [[HTTP://WWW.W3.ORG/TR/XHTML-RDFA-PRIMEREN](http://www.w3.org/TR/xhtml-rdfa-primeren)]

Introducción a RDFa: Introduciendo datos estructurados en Páginas Web.

Wiki de Web Semántica del W3C [[HTTP://ESW.W3.ORG/TOPIC/FRONTPAGEEN](http://esw.w3.org/topic/frontpageen)]

Wiki de Web Semántica, con información práctica sobre tecnologías, herramientas, casos de uso, etc.

SemWebSpain [[HTTP://ESW.W3.ORG/TOPIC/SEMWEBSPAIN](http://esw.w3.org/topic/semwebspain)]

Wiki de Proyectos y Documentos en español o de España sobre la Web Semántica.

SIMILE [[HTTP://SIMILE.MIT.EDU/EN](http://simile.mit.edu/en)]

Proyecto SIMILE: Ejemplos basados en tecnologías de Web Semántica.

SchemaWeb [[HTTP://WWW.SCHEMAWEB.INFO/SCHEMA/BROWSESCHEMA.ASPXEN](http://www.schemaweb.info/schema/browseschema.aspxen)] y **DAML** (DARPA AGENT MARKUP LANGUAGE) [[HTTP://WWW.DAML.ORG/ONTOLOGIES/KEYWORD.HTMLEN](http://www.daml.org/ontologies/keyword.htmlen)]

Repositorios de Ontologías.

Swoogle [[HTTP://SWOOGLE.UMBC.EDU/EN](http://swoogle.umbc.edu/en)]

Buscador temático sobre Web Semántica.

Herramientas de Desarrollo de la Web Semántica [[HTTP://ESW.W3.ORG/TOPIC/SEMANTICWEBTOOLSEN](http://esw.w3.org/topic/semanticwebtoolsen)]

Herramientas de Desarrollo de la Web Semántica (validadores, buscadores, librerías de desarrollo, etc.).