

The Web Semantic approach like tool for the e-learning

Ana Vanessa Leguízamo León

Universidad Central de Venezuela.
Apdo. 47002, Los Chaguaramos 1041A, Caracas Venezuela
Universidad de Salamanca
Paseo de Canalejas 169, 37008. Salamanca. España
vleguiza@usal.es

Abstract. This work presents some of the problems of searching the Internet using the current web search engines at any scope, but particularly those with educational contents. The concepts, advantages and disadvantages of the Semantic Web are reviewed in this context, emphasizing the benefits of employing this technology when searching for educational contents. Furthermore, this work presents the necessary changes for these contents to be targeted by semantic searches.

Keywords: Semantic Web, Internet searches, ontologies, e-learning

1. Introducción

En el presente trabajo se muestra una revisión de los problemas que se presentan al realizar búsquedas de diversos materiales y contenidos en Internet, particularmente en el contexto educativo, donde es de gran importancia la validez que estos puedan tener, a fin de garantizar un aprendizaje efectivo en los estudiantes.

Los buscadores actuales realizan las búsquedas textuales, es decir, buscando sobre bases de datos las palabras exactas que se le indican, sin tomar en cuenta el contexto en el que ellas se encuentren. Esto se debe a que los documentos que se encuentran en Internet no comparten un formato común y los ordenadores no pueden procesarlos de manera eficiente.

Como posible solución a este problema, se presenta el concepto de Web Semántica, que fue introducido por Tim Berners-Lee, padre de la WWW, en el año 2001, como forma de encausar la Web hacia donde originalmente fue pensada, documentos con contenido semántico donde los ordenadores funcionen como agentes, que faciliten a los usuarios la búsqueda y obtención de información por significado.

La Web Semántica incluye varios conceptos, entre ellos el de ontologías, como mecanismo para establecer clasificaciones de conceptos y, a partir de allí, incluir metadatos a los documentos que permitan a los ordenadores un mejor procesamiento de la información.

En el área educativa, la Web Semántica añade ventajas orientadas principalmente a la clasificación del conocimiento en ambientes de aprendizaje, donde es posible incorporar relaciones de orden entre los materiales que deben ser consultados por los

2 Ana Vanessa Leguízamo León

estudiantes, preestableciendo así la red de conocimiento que debe ser apropiada sin restar independencia en el proceso.

El siguiente documento está estructurado de la siguiente manera: primero se describen los problemas que se presentan al realizar búsquedas en Internet sobre documentos con los buscadores actuales, en segundo término se definen los conceptos de Web Semántica y ontologías, a continuación se describe brevemente como funciona la Web Semántica, las ventajas y desventajas de esta y los posibles usos en educación en línea. Finalmente se muestran posibles trabajos futuros para incorporar la Web Semántica en el campo de la educación.

2. Problemas en la búsqueda de recursos educativos en la red

Internet se ha convertido en una gran red de recursos de diferentes naturalezas, tanto así que es posible encontrar información de casi cualquier tipo, en cualquier ámbito o contexto. Esta característica tiene un gran potencial, ya que es posible obtener mucha información que está disponible de manera gratuita y que puede ser utilizada inmediatamente. Esta información es colocada en la red por cualquier persona o institución que así lo desee, contribuyendo de esta manera a aumentar la cantidad de recursos disponibles.

No existe un ente regulador de los contenidos que se colocan en Internet, por lo que es difícil determinar la procedencia, fiabilidad y validez de la información allí encontrada. El problema comienza al buscar información en un contexto particular, ya que es necesario contar con aplicaciones que realicen búsquedas sobre esa gran base de datos que constituye Internet. Para ello, existen buscadores que facilitan esta actividad, realizando búsquedas sobre una parte considerable de estos recursos.

Estas búsquedas son realizadas por los buscadores a través de palabras clave que son suministradas por los usuarios. Con estas palabras, el buscador realiza una búsqueda textual sobre los recursos presentes en Internet, es decir, busca las palabras indicadas por los usuarios sobre los materiales disponibles en la red. Por ejemplo, supóngase que se quiere buscar información de la *historia de la informática* en Internet, y para ello se selecciona el buscador Google para realizarla. Se indica al buscador las palabras que corresponden a la búsqueda deseada, y Google devolverá como resultado 1.650.000 páginas web que contienen la frase *historia de la informática*, pero también aquellas que tengan sólo la palabra *historia* y sólo la palabra *informática*. Dentro de estos resultados es posible conseguir páginas web de historia universal, historia del arte, ingeniería informática, cursos de informática, entre otras muchas donde aparezca cualquiera de las palabras clave. Esta búsqueda puede ser restringida, utilizando “ ” para delimitar la frase, en este caso Google sólo retornará páginas web donde se encuentre la frase completa *historia de la informática*. Al realizar esta búsqueda, Google consigue 48.400 resultados donde es posible conseguir páginas web que contengan información de la historia de la informática, pero también aquellas que hagan referencia a la historia de la informática, pero no la desarrollen como tal, o incluso páginas web de cursos donde la historia de la informática sea un tópico a tratar. En cualquiera de los dos casos, la cantidad de información encontrada es inmanejable (1.650.000 y 48.400 resultados), y la única

manera de determinar si la información es útil o no es visitando cada una de las páginas web encontradas.

3. Web Semántica

En el año 2001, Tim Berners-Lee, junto con James Hendler [3] publicaron un artículo en la revista *Nature*, donde expone que la Web fue diseñada como un espacio de información, no sólo entre humanos, sino también dónde pudieran participar los ordenadores con el fin de ayudar a los humanos a conseguir sus objetivos en la red. El mayor obstáculo para que los ordenadores puedan participar ayudando a los usuarios, es que la información que se consigue en Internet está hecha sólo para consumo humano, ya que no se encuentra estructurada, y los datos que allí se contienen no están bien diseñados como para poder ser manipulados por un ordenador. Los autores de este trabajo proponen la incorporación de una nueva tecnología a los documentos que se colocan en la Red, que permita a los usuarios crearlos fácilmente para que estos que puedan ser procesados de manera eficiente por los ordenadores, ser encontrados con mayor rapidez y con mayor porcentaje de acierto. Esta tecnología es lo que se denomina 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 [12]. En la Figura 1 se muestra la evolución de la Web, desde el punto de vista de las tecnologías que se le han incorporado para incrementar su potencialidad.



Fig. 1. Evolución de Web y tecnologías involucradas [2]

4. Ontologías

Las ontologías desde el punto de vista de la filosofía trata la naturaleza de la existencia, los tipos de las cosas que existen. La ontología como disciplina estudia otras teorías. Propone el estudio de todo lo que existe, qué es, cómo es y cómo es posible. La Ontología se ocupa de la definición del ser y de establecer las categorías fundamentales o modos generales de ser de las cosas a partir del estudio de sus propiedades. Los investigadores en Inteligencia Artificial y Web adoptaron el término y establecen que una ontología es un documento que define formalmente relaciones entre conceptos. Nykänen [8], define que una ontología es una especificación formal y explícita de una conceptualización compartida de un dominio particular.

La ontología es la manera más habitual para añadir significado semántico a la Web. Para lograr esto, la ontología debe estar conformada por una taxonomía y un conjunto de reglas de inferencia. La taxonomía define clases de objetos y las relaciones entre ellos. Por ejemplo, podemos definir una clase *Dirección*, que represente un tipo de localidad, y otra clase que represente un *código postal* para cada localidad, como se muestra en a Figura 2 y podríamos definir también la siguiente regla de inferencia: si un código postal está asociado con una localidad y una dirección usa ese código postal, entonces la dirección está asociada a esa localidad. Con esta información, un programa podría inferir que la Universidad de Salamanca se encuentra situada geográficamente en la ciudad de Salamanca, en la comunidad de Castilla y León, en

España, sólo indicando el código postal de ella. Las reglas de inferencia en las ontologías son muy poderosas para expresar relaciones entre objetos.

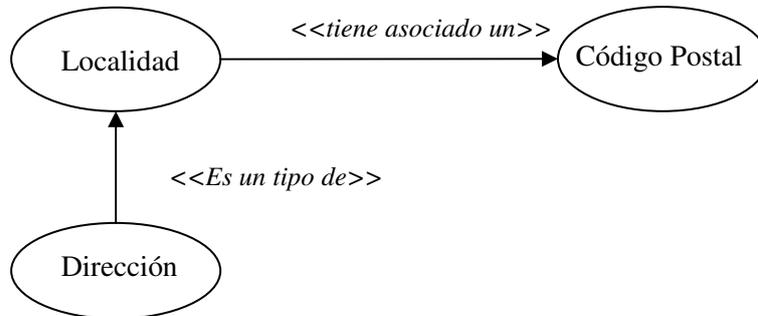


Fig. 2. Ontología para representar una localidad

Las ontologías pueden potenciar el funcionamiento de la Web de muchas maneras, pueden ser utilizadas de manera simple para mejorar la exactitud de las búsquedas en Internet, buscando sólo en aquellas páginas que refieren conceptos precisos de alguna de las palabras solicitadas. Aplicaciones más avanzadas pueden utilizar ontologías para relacionar la información en una página con estructuras de conocimiento y sus reglas de inferencia asociadas [3]; [5].

5. Cómo funciona la Web Semántica

Para incorporar semántica a los documentos colocados en Internet, es necesario definir una ontología y representarla con metadatos e incorporar estos a los documentos, de tal manera que describan, al menos, el contenido, el significado y la relación de los datos allí contenidos. Estos metadatos deben estar bien estructurados y escritos en algún metalenguaje que pueda ser interpretado por un ordenador.

Para la definición de los datos, la Web Semántica utiliza varios elementos, entre ellos XML, XML Schema, el Framework de Descripción de Recursos, RDF (por sus siglas en inglés, *Resource Description Framework*) y el Lenguaje de Ontología Web, OWL (por sus siglas en inglés, *Web Ontology Language*), estos dos últimos estándares 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.

Cada uno de estos elementos tiene una función específica, a saber:

- XML, es utilizado para definir la estructura de los documentos, sin definir significado.
- XML Schema, es el lenguaje utilizado para definir la estructura de los documentos XML

6 Ana Vanessa Leguízamo León

- RDF, que proporciona información descriptiva simple sobre los recursos que se encuentran en la Web. Aporta una semántica básica para el modelo de datos, que puede ser definido en XML.
- El OWL permite desarrollar temas o vocabularios específicos en los que se asocian los recursos definidos. OWL proporciona un lenguaje para definir ontologías estructuradas, que pueden ser utilizadas por 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. Las ontologías incluyen definiciones de conceptos básicos en un campo determinado y la relación entre ellos.

Estos lenguajes son los que permiten a los usuarios incorporar significado semántico a los documentos que incluyen en la Web, permitiendo que éstos sean interpretados por los ordenadores.

6. Ventajas y desventajas de la Web Semántica

La principal ventaja que aporta la Web Semántica es que incorpora contenido semántico a los documentos que se colocan en Internet, permitiendo organizar por conceptos la información que ofrecemos, garantizando búsquedas sobre esta por significado y no por contenido textual.

Entre las desventajas, la primera es que se hace necesario reestructurar los documentos que ya están en Internet, para adaptarlos al formato que propone la Web Semántica, de tal manera que puedan ser interpretados de manera correcta por los ordenadores [4]. Esto incluye los problemas de idiomas, ya cada uno tiene una semántica diferente, lo que lleva a hacer buscadores semánticos por idioma, y más aún, por área de conocimiento [11].

La segunda es en el caso de las ontologías para representar la semántica, ya que en ellas es necesario proveer relaciones de equivalencia entre conceptos, por ejemplo en el caso del código postal, se debe establecer que CP es igual a código postal, y a *zip code*, en el caso del inglés.

7. Posibles usos en educación en línea

La visión de la Web Semántica está basada en dos ideas principales: añadir datos semánticos a los recursos que se encuentran en la Web y la creación de agentes inteligentes capaces de entender y operar con los recursos a un nivel semántico [9].

Los problemas identificados en la primera parte se potencian en el caso de los ambientes de aprendizaje en Internet, dado que para garantizar la correcta apropiación del conocimiento, es necesario que los materiales de los que disponen los aprendices sean válidos, y esto no es siempre posible, considerando la gran cantidad de información que se encuentra en Internet.

Según Koper [7] la importancia de la Web Semántica para la educación depende de cuánto contribuye en la realización de tres objetivos: incrementar la efectividad de la educación, hacer más atractiva la educación y disminuir la carga de trabajo o disminuir los costos.

La Web Semántica puede contribuir principalmente en dos áreas, a saber:

- Ayudando al equipo de trabajo a hacer algunas de sus tareas más flexibles, hacer los aspectos educativos en línea más eficientes y menos aislados, esto incluye las tareas de desarrollo de cursos en línea, ayudar al aprendiz en sus tareas, mejorar las evaluaciones y manejo y administración de cursos.
- Ayudando a las personas que cumplen diferentes roles (aprendices, tutores y proveedores de contenido), a desempeñar sus tareas más efectiva y eficientemente en espacios de aprendizaje grandes, distribuidos, basados en problemas, multi actores, multi recursos, crear y modificar ambientes de aprendizaje centrados en el aprendiz, no lineales y autodirigidos.

En el caso particular de la educación en línea, donde el aprendizaje no es dirigido por el instructor, sino que los estudiantes tienen libertad de acceder al material disponible en un orden no predefinido y componer su aprendizaje de la manera que más se adapte a ellos, puede hacerse necesario incorporar información a los documentos disponibles para la consulta de los estudiantes, que permita indexar y recuperar esta información de manera efectiva.

El uso de metadatos es una respuesta natural para solucionar este problema, y de hecho ya ha sido implementado en librerías en línea desde hace algún tiempo. En *e-learning*, es común que los docentes realicen asociaciones entre sus materiales, o definan las propiedades educativas o pedagógicas de los documentos, o el orden en el que los estudiantes deben acceder a estos materiales. Sin embargo, toda esta información no está reflejada en los documentos cuando se colocan en Internet, por lo que no puede ser manejada automáticamente por los ordenadores. Si se incorporan elementos semánticos a los documentos que colocan los docentes en Internet, reflejando las relaciones entre ellos, una aplicación Web podría actuar como agente, que ayude a los estudiantes a realizar un recorrido efectivo sobre el contenido del curso en línea [1].

Así, las ontologías en ambientes de aprendizaje pueden ser utilizadas para facilitar a los estudiantes el acceso a la información contenida allí, por ejemplo en el caso de aquellos cursos que tienen audiencias heterogéneas, donde cada uno de ellos tiene diferentes niveles de conocimiento en el contexto del curso. Se pueden definir ontologías para clasificar el contenido que deben revisar los estudiantes, dependiendo de su nivel de conocimiento, estableciendo relaciones entre conceptos, como “es parte de”, o “requiere de”. Un ejemplo de estas relaciones puede definirse en función del lenguaje Java, este es parte de los Lenguajes de Programación, y los Lenguajes de Programación para poder ser utilizados requieren conocimientos de algoritmia. Así, un ordenador podría inferir que para utilizar el lenguaje de programación Java, el estudiante debe tener conocimientos de algoritmia, y en caso de no tenerlos, proveérselos. También se pueden definir ontologías para establecer aspectos pedagógicos en los materiales dispuestos, por ejemplo, para clasificarlos como tutoriales, ejemplos, ejercicios, evaluaciones, entre otros. Otro tipo de ontología puede

utilizarse para definir estructuras lógicas entre los materiales, al tener conceptos que son jerárquicos, y la navegación entre ellos se define por relaciones de orden, como anterior, siguiente, inicio o fin de los documentos estudiados [1].

8. Trabajos Futuros

Por lo estudiado a lo largo de este trabajo, y dado que la Web Semántica es un área que apenas comienza a mostrar sus potencialidades, es fácil inferir que la conjunción de esta tecnología con otras áreas disciplinares puede llevar a trabajos interesantes. La educación no escapa de esto, y es por ello que consideramos que la Web Semántica puede contribuir de manera significativa en las siguientes áreas [10], [1]:

- Conducir al aprendiz: describiendo ontologías que permitan enlazar materiales de aprendizaje de diferentes autores. Esto permite diseñar cursos personalizados, adaptables, donde se diseña un camino semántico que garantice que los materiales pueden ser recuperados dentro del contexto del espacio de aprendizaje que se esté estudiando.
- Flexibilizar el acceso: la principal ventaja del *e-learning* es que el conocimiento puede ser accedido en el orden en que lo desee el estudiante, de acuerdo con sus intereses y metas. Sin embargo, se puede definir una semántica apropiada para restringir el acceso a ciertos materiales en los casos en que existan pre-requisitos, manteniendo siempre la posibilidad de acceder al conocimiento de manera no-lineal.
- Integración: La Web Semántica puede proveer una plataforma uniforme para los procesos de negocio en las organizaciones, y se pueden integrar actividades educativas en ese proceso. Esta opción puede ser especialmente valorada en organizaciones donde el personal deba mantenerse constantemente actualizado.

9. Conclusiones

Hasta el momento, Internet se ha convertido en la principal fuente de información en muchos contextos, uno de ellos el educativo, donde es posible encontrar materiales en línea muy recientes, y que están disponibles primero en la red que en las librerías.

Dado que es incalculable la cantidad de información que se encuentra en Internet, era previsible pensar que el problema se dirigiría a la forma eficiente de obtener información útil de esa gran base de datos.

Para facilitar las búsquedas, y orientarlas al significado de la información contenida, surge el concepto de Web Semántica, que tiene como objetivo incorporar elementos que den significado semántico a los documentos contenidos en Internet, y que a la vez, estos elementos puedan ser procesados automáticamente por los ordenadores, permitiendo a éstos participar en el proceso de búsquedas por

significado, trabajo éste que sólo podía ser realizado por el humano que necesitaba una información particular.

La Web Semántica incorpora como ventaja principal, la posibilidad de clasificar los documentos por conceptos, y establecer relaciones entre ellos, pero como desventajas el hecho que los documentos existentes deben ser modificados, y los problemas de idioma, ya que las semánticas son diferentes.

En este trabajo se hizo una revisión de las potencialidades del uso de la Web Semántica al incorporarla en el contexto de la educación en línea, donde podría aportar ventajas significativas a la hora de clasificar el conocimiento a ser adquirido por los estudiantes, permitiendo que el curso que se ofrece se adapte a las características de sus participantes, de acuerdo a su nivel de conocimientos, y lo que necesita para culminar con éxito el curso. Asimismo, la Web Semántica lleva a una nueva manera de concebir la red, donde los contenidos que allí se tienen puedan ser pre-procesados por agentes, que garanticen la significatividad de la información obtenida para el usuario que la solicita.

10. Reconocimientos

Esta investigación es financiada por el Consejo de Desarrollo Científico y Humanístico de la Universidad Central de Venezuela.

11. Referencias

1. ANTONIOU, G. y VAN HARMELEN, F. (2004). *A Semantic Web Primer*. Massachusetts Institute of Technology. USA.
2. BERNERS-LEE, T.; HENDLER, J (2001). *Scientific publishing on the 'semantic web'*. Nature, Abril.
3. BERNERS-LEE, T. HENDLER, J., LASSILA, O. (2001) *The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities*. Scientific American, Mayo, 2001.
4. BOURDA, Y. y DOAN, BICH-LIÊN (2003). *The Semantic Web for Learning Resources*. 3rd IEEE International Conference on Advanced Learning Technologies (ICALT'03)
5. DAVIS, J. STUDER, R. y WARREN, P. (2006). *Semantic Web Technologies, trends and research in ontology-based systems*, Editorial Wiley, Inglaterra.
6. GEROIMENKO, V.; CHAOMEI, Ch (2006). *Visualizing the Semantic Web*. (2a edición). Editorial Springer. Inglaterra.
7. KOPER, R. (2004). *Use of the Semantic Web to Solve Some Basic Problems in Education*. Journal of Interactive Media in Education, 2004 (6).
8. NYKÄNEN, O. (2004). *Metadata for Learning Resources: Technologies and Directions of the Semantic Web – A Brief Review*. IEEE International Conference on Advanced Learning Technologies (ICALT'04).

9. PANTELEYEV, M. PUZANKOV, D. SAZYKIN, P. y SERGEYEV, D. (2002). *Intelligent Educational Environments Based on the Semantic Web Technologies*. IEEE International Conference on Artificial Intelligence Systems (ICAIS'02).
10. SANTOS, J. ANIDO, L. y LLAMAS, M. (2005). *Design of a Semantic Web-based Brokerage Architecture for the E-learning Domain. A Proposal for a Suitable Ontology*. 35th ASEE/IEEE Frontiers in Education Conference IEEE.
11. TJOA, A. ANDJOMSHOAA, A. SHAYEGANFAR, F. y WAGNER, R. (2005). *Semantic Web Challenges and New Requirements* 16th International Workshop on Database and Expert Systems Applications (DEXA'05). IEEE.
12. WORLD WIDE WEB CONSORTIUM W3C (2006). *Guía Breve de Web Semántica*. <http://www.w3c.es/Divulgacion/Guiasbreves/WebSemantica>
Consultado: Marzo 2007.