

Introducción

Web Semántica y Sistemas de Información Documental

(Trea 2009)

Lluís Codina
(UPF)

La web actual está formada en su mayor parte por documentos creados originalmente para que tengan sentido para los seres humanos. Esta afirmación puede parecer una perogrullada porque, precisamente, una definición clásica señala que un documento es una información registrada en un soporte material destinada a un ser humano (al menos en su formato final).

Ahora bien, existe una clase de registros que consisten en conjuntos de datos que no tienen como destinatario a un ser humano, sino a un ordenador (más exactamente a un programa de ordenador).

Al final de la cadena en los dos casos siempre hay un ser humano, claro, pero la cuestión es que hay dos tipos de registros: los que están destinados a seres humanos y que solemos llamarlos *documentos*; y los que están destinados a ordenadores y que solemos llamarlos *datos*.

Hasta ahora, en la web el tipo de registro dominante es el documento en el sentido anterior. Ese documento adopta típicamente la forma de la página web con la que, a día de hoy, una considerable parte del mundo ya está familiarizado. Pero no únicamente, porque en la web actual, podemos navegar entre una variedad amplísima de formatos documentales.

La cuestión es que la Web es capaz de aportar información a personas de todo el planeta para casi cualquier clase de actividad útil y constructiva que estén llevando a cabo (el lado oscuro consiste en que los interesados en actividades destructivas por desgracia también suelen usar la web, pero esto es otro problema).

El proyecto de la web semántica parte de una idea genial: las páginas web (recuerde, cosas destinadas a seres humanos) podrían ser *a la vez* datos (recuerde, cosas destinadas a ordenadores).

¿Qué pasaría si lo anterior fuera cierto?, es decir, ¿qué sucedería si las páginas web pudieran ser procesadas como datos de entrada de programas de ordenador, y no *solamente* como algo destinado a ser mostrado en pantalla para que lo vea un ser humano?

Podrían pasar cosas espectaculares. La actual web 2.0 no es, ciertamente, la web semántica, pero casi todo el mundo reconoce que ha aportado cambios espectaculares y podría ser un buen ejemplo de los cambios que podríamos esperar. Mencionemos dos ejemplos: el monumento al conocimiento humano que supone la Wikipedia y el monumento a la libertad de expresión y creatividad que supone la blogosefera. Desde

este punto de vista, la web semántica podría ser la web 3.0, pero con su propia nueva serie de cambios espectaculares.

Lo que necesitamos para pasar de la web como documentos (la web actual) a la web como datos (la web semántica) es que, en primer lugar, las páginas contengan, además de contenidos para los seres humanos (información), contenidos para los ordenadores (metainformación).

Ahora bien, la web tiene alcance universal, así que necesitamos sistemas de metadatos (o metainformaciones) susceptibles de uso universal e interpretables por ordenadores.

En segundo lugar, necesitamos una especie de súper clasificaciones o súper tesauros formales (ontologías) que ayuden a los ordenadores a realizar inferencias válidas sobre los datos.

A partir de aquí, y en este caso, como estamos haciendo un ejercicio de prospectiva, resulta que la predicción de la clase de cambios que podría aportar la web semántica varía según algunos colectivos. Simplificando mucho (puede que muchísimo) hay al menos dos de ellos: (1) los relacionados con la inteligencia artificial y (2) los relacionados con el procesamiento de la información. El primer colectivo está directamente vinculado con el proyecto original de la web semántica, nacido en el seno del *World Wide Web Consortium*. El segundo colectivo tiene una procedencia diversa: en general, grupos interesados en las brillantes ideas fundacionales pero escépticos sobre las posibilidades de la inteligencia artificial.

El primer grupo (el que hemos identificado con la inteligencia artificial) cree que en el futuro, ordenadores conectados a la web, mediante complejos sistemas de metadatos, ontologías y lógica formal, podrán llevar a cabo “razonamientos” al servicio de los usuarios. Los límites de estos razonamientos no estarían claros. Eventualmente, agentes de software podría llevar a término tareas con éxito, en representación de los usuarios, para las cuales actualmente se requiere inteligencia, o sea tareas que actualmente solamente las personas pueden ejecutar con éxito.

El segundo grupo, el que hemos identificado con el procesamiento de la información, se declara agnóstico, sino abiertamente escéptico, sobre las posibilidades de la inteligencia artificial, pero reconoce y valora los grandes beneficios que el conjunto de normas, lenguajes y nuevos enfoques que el proyecto de la web semántica puede aportar a la nueva web, la que ya está teniendo lugar y la que se está fraguando para los próximos cinco o diez años, llámese web semántica o web 3.0

En esta obra, los autores han procurado ceñirse, en general, a las posibilidades actuales y futuras de la web semántica entendida sobre todo como en este segundo sentido, y sobre todo en su impacto en los sistemas de información documental. No obstante, en los primeros capítulos se ha optado por presentar una visión global que aspira a presentar las distintas visiones de la web semántica.

En este sentido, como *background* general de la obra hemos procurado tener presente la clase de necesidades, tareas y objetivos que suelen tener los profesionales, los estudiosos y los investigadores de los sistemas de información documentales.

Si los autores y los editores de este libro compartimos algunas ideas comunes, estas consisten en que, en primer lugar, e independientemente de cuál de las dos *almas* de la web semántica se haga realidad al final (¿pero habrá un final?), el proyecto del World Wide Web Consortium, debe ser, no solamente conocido por el colectivo de la biblioteconomía-documentación, sino que este colectivo debe formar parte de sus protagonistas.

En segundo lugar, que el conjunto de las nuevas ideas, modelos conceptuales, normas, recomendaciones, lenguajes, etc. de la web semántica está llamado a tener una influencia duradera en muchos aspectos de la biblioteconomía-documentación, lo cual nos lleva de nuevo a la primera idea, a saber, al protagonismo que nuestro colectivo puede tener, y de hecho ya tiene, en el proyecto de la web semántica.

Por **Lluís Codina** (<http://www.lluiscodina.com>)

Introducción al libro: L. Codina, MC Marcos y R. Pedraza (coord.) *Web Semántica y Sistemas de Información Documental*. Gijón: Trea, 2009 (978-84-9704-460-8)