

Impacto de la información semántica en la Usabilidad Web

Pablo M. Vera, Daniel A. Giulianelli, Rocío A. Rodríguez

Departamento de Ingeniería e Investigaciones Tecnológicas

Universidad Nacional de La Matanza

Florencio Varela 1903, San Justo, Buenos Aires, Argentina

4480-8900 interno 8630

pvera@unlam.edu.ar, dgiulian@unlam.edu.ar, rrodri@unlam.edu.ar

Resumen

El presente trabajo muestra una iniciativa para promover y aprovechar el uso de la información semántica en los sitios web desde el punto de vista de la usabilidad del usuario. Permitiéndoles realizar acciones de forma más natural de acuerdo al contenido de la página que esté visitando. Además se aborda el tema de la complejidad que existe actualmente para incorporar la información semántica en los sitios y se presentan alternativas a los complejos formatos como OWL y RDF.

Palabras clave: Usabilidad, Semántica, Microformatos, OWL, RDF

Contexto

El proyecto se desarrolla en la Universidad Nacional de La Matanza (UNLaM), como parte de un equipo de investigación de Ingeniería de Software. Este proyecto es financiado por CyTMA (Ciencia y Tecnología del Partido de La Matanza).

Introducción

Actualmente la inclusión de la información semántica en las páginas web no es una práctica muy difundida. Una de las razones es la falta de

divulgación de sus beneficios que muchas veces sólo se relacionan con mejores resultados de búsqueda mediante la utilización de buscadores semánticos. Otra de las razones, y quizás la principal, es el desconocimiento de los programadores y diseñadores de cómo incorporar dicha información al momento de crear un sitio.

Formatos semánticos existentes

Aquellos que comienzan a indagar en el tema primeramente se encuentran con complejos formatos que requieren la creación de ontologías como ser OWL [W3C04] y RDF [W3C04A].

“OWL (Ontology Web Language) 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”. [W3C10]. Tal como puede verse en esta explicación de la W3C [W3C11A], OWL requiere para su funcionamiento la definición de una ontología específica para el dominio que se está tratando. Esto brinda grandes posibilidades para dotar de significado cualquier dominio que se desee pero su utilización para elementos genéricos puede resultar contradictoria debido a la diversidad de ontologías existentes.

“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”. [W3C10]. Básicamente RDF describe elementos de forma un poco más simple que OWL ya que no dispone de una ontología pero sí cada recurso descripto debe contener una definición en una URL única.

Uno de los objetivos del presente trabajo es encontrar un formato semántico que sea fácil de incorporar a todos los sitios web y que permita describir información útil para mejorar la usabilidad. En un principio se buscará agregar información semántica para:

- Contactos
- Ubicaciones geográficas
- Eventos
- Pagos
- Identificar secciones de un sitio web como por ejemplo su página principal.

Con estos objetivos tanto OWL como RDF resultan ser demasiados complejos para representar esta información y por lo tanto se buscan formatos alternativos como RDFa [W3C08], Microformats [MIC11] y HTML Microdata [W3C11].

Derivado de RDF nace 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.), a través de unas anotaciones semánticas incluidas en el código e invisibles para el usuario” [W3C10].

Los microformatos constituyen un formato basado en XML que permite agregar información semántica de forma sencilla mediante estándares claramente definidos. Entre otras cosas define estándares para:

- Personas y organizaciones
- Eventos
- Opiniones, Rankings y Revisiones
- Categorías, palabras claves y tags
- Listas
- Licencias

Además tiene una serie de definiciones que se encuentran en etapa de revisión pero que son de mucha utilidad como por ejemplo:

- Ubicación geográfica
- Pagos
- Identificación de la página principal del sitio

Microdata es una nueva especificación basada en HTML 5 que permite incorporar información semántica directamente dentro de los tags HTML. Se basa en pares formados por un nombre y un valor. Existen ya estándares de nombres para los ítems entre los que se incluyen entre otros:

- Eventos
- Organizaciones
- Personas
- Productos

- Revisiones
- Breadcrumbs (camino hasta llegar a una página específica a partir de la página principal del sitio)

Forma para mejorar la usabilidad

Los tres formatos presentados permiten incorporar la información semántica necesaria dentro de cualquier sitio web de una forma relativamente sencilla. Pero la pregunta que surge es ¿qué se puede hacer para tomar esa información y hacerle la vida más fácil a los usuarios?

A partir de este interrogante surge la necesidad de contar con herramientas que busquen e interpreten la información semántica en forma automática al navegar por un sitio web sin que el usuario explícitamente deba realizar alguna acción. Lo que se busca es que si la información está disponible, ésta sea aprovechada para facilitar su explotación al usuario indicándole que actividades puede realizar con la misma.

Los navegadores aún no tienen integradas formas de interpretar esta información por lo que se debe recurrir a plugins específicos. Actualmente existen pocos plugins que pueden reconocer información semántica en formato Microformats o RDFa. Microdata es un formato nuevo y por lo tanto es más comprensible que las herramientas estén en etapas tempranas de desarrollo pero los otros formatos ya llevan más tiempo de existencia y sin embargo los plugins existentes son escasos, muy poco difundidos y no tuvieron grandes avances en cuanto a mejoras en la usabilidad.

A continuación se muestra como ejemplo el plugin Operator [KAP10] para el navegador firefox que permite interpretar tanto

Microformats como RDFa. Este plugin interpreta la información semántica existente en la página que se está visualizando y muestra en una barra del navegador las acciones que el usuario puede realizar con dicha información.

La figura 1 muestra una página donde se ha detectado información de 3 contactos y por lo tanto el plugin permite exportar el contacto para ser incorporado en cualquier software de agenda o agregarlo como un nuevo contacto de Yahoo entre otros.

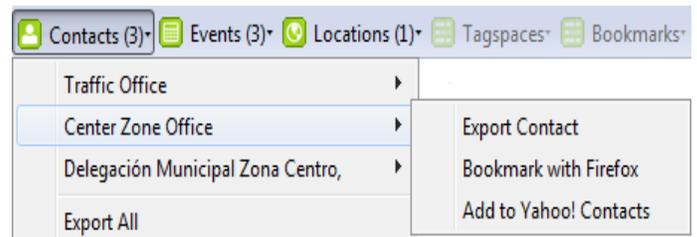


Figura 1: Ejemplo de información de contactos con el plugin Operator

Otro claro ejemplo de la utilidad de este plugin es el presentado en la figura 2 que al detectar una dirección, permite directamente buscarla en un mapa como por ejemplo google maps.

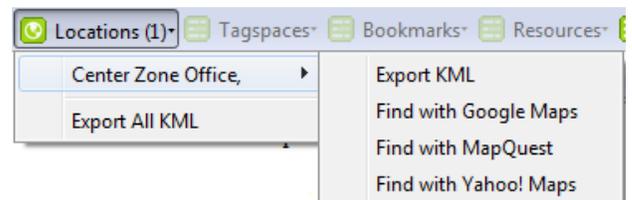


Figura 2: Ejemplo de información de ubicación con el plugin Operator

El plugin Operator facilita las operaciones del usuario haciendo simples cosas que de otra forma requerirían más esfuerzo y conocimiento por ejemplo volviendo al ejemplo de la figura 1 donde se dispone de la información de un contacto que se desea agendar. El no contar con

este plugin implicaría realizar los siguientes pasos:

- Abrir el programa o sitio web donde se encuentra la agenda.
- Crear un nuevo contacto.
- Copiar uno a uno los datos escritos en la página web, debiendo identificar manualmente cada uno.

En cambio con este plugin esta serie de pasos es reemplazada con un único clic sobre la opción correspondiente.

En cuanto al ejemplo presentado en la figura 2 si se quisiera buscar la ubicación geográfica de un lugar la tarea se torna más compleja debido a que, si bien hay sitios como google maps que permiten buscar una ubicación mediante el texto que compone la dirección, es una tarea más engorrosa y propensa a errores. Sin embargo mediante la utilización del plugin se toma la información no visible en la página que incluye la información geográfica de latitud y longitud haciendo que la búsqueda sea exacta en cualquier programa de mapas o dispositivos de posicionamiento.

Si bien este plugin trae las ventajas mencionadas consideramos necesario mejorar el mismo en cuanto a la usabilidad haciendo que sea más fácil para el usuario detectar que la información está presente y mostrar que acciones puede realizar con ella. Actualmente este plugin muestra una barra con números en la parte superior del browser que indica la cantidad de información que encontró en la página, por ejemplo 5 contactos, 3 ubicaciones, 10 eventos. Este esquema de visualización puede hacer que la información disponible sea fácilmente ignorada por no encontrarse en forma directa sobre el contenido visualizado. Además si la página tiene mucha información semántica el plugin la muestra en forma de listado haciendo que se deba buscar en forma manual en el

mismo, por ejemplo, cual es la dirección ó contacto, que el usuario ha leído en el contenido de la página y que le interesa agendar.

Líneas de investigación y desarrollo

Las líneas de investigación del presente trabajo son las siguientes:

- Análisis de los formatos semánticos existentes.
- Análisis de los plugins actuales.
- Desarrollo de nuevos plugins para los principales navegadores que hagan uso de la información semántica mejorando la usabilidad del usuario al automatizar tareas cotidianas.
- Pruebas de Usabilidad.

Resultados Esperados/Obtenidos

A continuación se detallan los objetivos con los que cuenta el presente trabajo de investigación:

- Promover la incorporación de información semántica en los sitios web mediante técnicas sencillas.
- Definir cuales de los formatos se adapta más al objetivo planteado y proponer mejoras al mismo.
- Mejorar la usabilidad al crear nuevos plugins más integrados que muestren directamente junto al contenido de la página las acciones posibles con la información interpretada.

La presentación de este trabajo en WICC va a permitir contribuir a la difusión de la importancia de incorporar información semántica en la web. Poder transmitir la importancia de contar con plugins los cuales permitan mediante la información semántica interactuar con distintos software con los que cuente el usuario. Para finalmente poder

ilustrar la necesidad de tener de forma más usable la información semántica a disposición de los usuarios.

Formación de Recursos Humanos

En esta línea de investigación se encuentran afectados docentes de la universidad, egresados y pasantes.

El tema presentado guarda estrecha relación con la tesis doctoral en curso de uno de los docentes afectados, la cual será presentada ante la Universidad Nacional de La Plata.

Referencias

[W3C04] W3C, “OWL Web Ontology Language Overview”, 2004
<http://www.w3.org/TR/owl-features/>

[W3C04A] W3C, “Resource Description Framework”, 2004
<http://www.w3.org/RDF/>

[W3C10] W3C, “Guía breve de Web Semántica”, 2011
<http://www.w3c.es/Divulgacion/Guiasbreves/WebSemantica>

[W3C11] W3C, “HTML Microdata”, 2011,
<http://www.w3.org/TR/microdata/>

[W3C08] W3C, “RDFa Primer Bridging the Human and Data Webs”, 2008
<http://www.w3.org/TR/xhtml-rdfa-primer/>

[MIC11] Microformats, “About Microformats“, 2011
<http://microformats.org/about>

[W3C11A] W3C España, “Sobre el W3C”
<http://www.w3c.es/Consorcio/>

[KAP10] Kaply M, “Operator Firefox Add-On”
<https://addons.mozilla.org/en-US/firefox/addon/operator/>