

Apuntes de la web Semántica

Sesión 1 y 2

La Web Actual

- Web Actual
 - o Web 1.0 Web de enlaces a recursos (páginas web, archivos)
 - o Web 2.0 implementa servicios (redes sociales, blogs, wikis, imágenes, videos, etc.)
 - o Web Sintáctica y no Semántica para las máquinas.
- Problemas de la Web actual
 - o Heterogeneidad, Masiva, cambio rápido, hecha para los humanos.
- ¿Cómo podemos empoderar a una aplicación para que consuma y relacione datos en la Web?
 - o Anotación Semántica
 - o Lenguaje estándar entre máquinas que incorpore el significado de los recursos que gestiona.
 - o Mecanismos y técnicas formales de razonamiento de la información

Ejemplo: Hacer el Ejemplo de mi información en la web actual con HTML.

Introducción a la Web Semántica

- Diferentes tipos de datos y recursos, pero sin significado para las máquinas
- Una máquina necesita un lenguaje de especificación formal de los recursos Web y sus relaciones. Las recomendaciones de la W3C con respecto al uso de RDF y OWL van enfocadas a construir la Web Semántica.

Ejemplo: Hacer Ejemplo de mi Información con semántica propia XMLS.

Web Semántica

- ¿Cómo construir la Web Semántica?
 - o Utilizar adecuadamente las URI y URL. (identificación única de los recursos web)
 - o Crear y utilizar un lenguaje que describa y relacione los recursos de la Web
 - o Debe ser procesable por las máquinas y entendibles a los humanos
 - o Debe ser estándar y que use tecnologías completamente interoperables.
- ¿Cuándo se implemente la WS, que características tendremos en la Web?
 - o Es capaz de organizar la información y estructurarla lógicamente.
 - o Permite integrar recursos web con diferentes formatos.
 - o Permite la interoperabilidad entre dispositivos.
 - o Es capaz de contestar con exactitud lo que el usuario busca.
- Piedra angular de la WS: Metadatos

- Se pueden encontrar en las propiedades de los recursos web (Word, pdf, imágenes, páginas Web, metadatos sociales, micro-formatos, etc.).
- Estructura de un metadato: **Recurso** (sujeto) → **Propiedad** (predicado) → **Valor** (objeto). En dónde cada uno puede ser un URI.
- Se pueden implementar utilizando:
 - Microformatos – Etiquetas de metadatos (**meta**)
 - Mediante especificaciones XHTML llamados (**schema**)
 - Mediante perfiles (**profiles**).
 - Mediante la definición personalizada de metadatos (**archivos RDF**)
- ¿Qué Tecnologías necesita la Web Semántica?
 - XML, RDF , SPARQL, OWL y RDFa

Ejemplo: Hacer una página con metadatos “meta” con las diferentes formas

Sesión 3

¿Qué es el XML?

- XML: eXtensible Markup Language, es un metalenguaje extensible de etiquetas. Permite definir la gramática de un lenguaje particular. Permite la interoperabilidad sintáctica entre las máquinas.

¿Qué es RDF?

- RDF: Resource Description Framework, proporciona información descriptiva sobre los recursos que se encuentran en la Web. es una base para procesar metadatos. Permite la interoperabilidad semántica entre las máquinas. (permite recuperación, catalogación, compartir, calificación, propiedad intelectual y seguridad de los recursos web).

¿Qué es SPARQL?

- SPARQL: Simple Protocol and RDF Query Language, es un lenguaje de consulta sobre RDF, permitiendo recuperar recursos web que cumplan con los filtros correspondientes, aplicando lógica de primer orden para relacionar los recursos.

¿Qué son las ontologías?

- Una ontología es un conjunto de conceptos interrelacionados entre sí, explícitos, compartidos

Taller No 1 Manejo de metadatos de los Recursos Web

Objetivo: El presente taller tiene el fin de que los estudiantes hagan un uso correcto de los metadatos existentes en la Web y de la Anotación de recursos Web

Tiempo: 1 hora

Recursos: Visual Studio, ExifTool, metadata++

Descripción: El taller tiene dos partes, la primera es un ejercicio de búsqueda y recopilación de metadatos de los recursos web existente y la segunda parte es la anotación de recursos web utilizando diferentes técnicas existentes.

Parte I

1. Descargue la herramienta de línea de comando llamada ExifTool (<http://owl.phy.queensu.ca/~phil/exiftool/>), la cual le permitirá editar los metadatos de los archivos.
2. Seleccione dos copias de imágenes de su biblioteca personal, en lo posible que hayan sido posteadas en las redes sociales y a las que se las haya colocado comentarios de las personas en las fotos. También seleccione un pdf y documento de Word.
3. Por medio de línea de comando verificar los metadatos de los archivos anteriores así: exiftool "nombrearchivo" <enter>. Si algunas imágenes no muestran metadatos es porque están encriptadas.
4. Para sistemas operativos Windows se puede descargar también una aplicación que utiliza un GUI llamado metadata++, que utiliza ExifTool como herramienta de análisis en segundo plano (http://www.logipole.com/metadata++_en.htm). Seguir los pasos recomendados para su instalación y después realice una exploración de la herramienta para ver los metadatos, modificarlos y agregar metadatos a los archivos.
5. Crear un proyecto de sitio Web en Visual Studio.
6. Buscar en <https://www.nuget.org> cualquier paquete con Exif. Se desplegará una lista de los posibles paquetes para lectura de metadatos. Busque NexifTool, lea como instalarlo a través de la línea de comando de NUGET e instálelo en el proyecto.
7. Una vez instalado el paquete proceda a crear una página aspnet que permita seleccionar una imagen y analícelo con el paquete tal como se indica en las referencias y documentación del paquete.
 - a. Cree una carpeta de imágenes en el proyecto y copie las imágenes a ser analizadas.
 - b. Coloque un control button en la página aspnet. Cambie las propiedades para que se lea "Leer metadatos"
 - c. A continuación, coloque un control de imagen y cargue una de las imágenes por defecto al control con la propiedad imageUrl.
 - d. Coloque una textbox y la propiedad textMode colóquela Multiline
 - e. Agregue la referencia a la dll importada "using NexifTool;"
 - f. En el evento click del botón coloque el código fuente para obtener los metadatos de la imagen y colocarlos en el textbox.

Parte II

1. Cree una página HTML con un resumen de lo visto hasta ahora en la Tutoría de Web Semántica y coloque Microformatos básicos con etiquetas meta (description, keywords, autor, abstract).
2. Utilizando el mismo HTML usando perfiles, en específico usando el perfil de dublin core para adicionar etiquetas de (date, contributor, rights).
3. Finalmente crear un RDF que utilice dublin core también para colocar metadatos sobre el autor, el título y la descripción del texto del HTML.