

METADATOS EN LA WEB

claudio gutierrez

Departamento de Ciencias de la Computación

Universidad de Chile

`cgutierr@dcc.uchile.cl`

Semántica en la Web

- **Comprender** la información
- **Organizar** la información
- **Validar** información
- **Deducir** nueva información

¿Que necesitamos para automatizar los procesos anteriores?

La Web: una analogía

Gigantesca base de datos distribuída

Detalles...

- El administrador cargó los datos antes de tener los esquemas
- No fijó los permisos para crear y modificar

La Web: algunos problemas

- No hay catalogación global
- Débil discriminación sobre la información que existe
- Deficiente integración de información
- Automatización de las funciones anteriores inexistente
- Centros de confianza

Solución: metadatos

- Datos estructurados acerca de datos
- Información descriptiva acerca de un objeto o recurso
- Ejemplo: tarjetas de catálogos de bibliotecas

Metadatos hoy

- Catálogos de propósito general (DCMI, Open Directory Project)
- “Sindicación”, agregación, y rating (PICS, RSS)
- Colecciones personales: Música, Fotos, contactos
- Anotaciones, centros de confianza, acceso automático al conocimiento

Metadatos hoy (cont.)

- **Dublin Core Metadata Initiative**: Vocabulario minimal común para marcar recursos en la web
- **Open Directory Project**: El directorio más grande y comprensivo de la Web; mantenido por editores voluntarios
- **Platform for Internet Content Selection**: Originalmente dirigido a padres y profesores para ayudar a seleccionar contenidos.
- **Platform for Privacy Preferences**: Solución simple y automatizada para controlar el uso de información personal en lo sitios Web visitados.
- **Rich Site Summary**: Descripción de metadatos liviana, multipropósito y extensible para sindicación

Metadatos hoy: problemas

- Metadatos no son extensibles globalmente
- Modelos de metadatos son distintos e incompatibles
- No hay comunicación entre aplicaciones (interoperabilidad)
- Confianza

Metadatos hoy: solución

- Modelo estandarizado de metadatos: **Resource Description Framework (RDF)**
- Vocabulario compartido (Ontologías)



- Lenguaje de metadatos serializado en XML
- Provee interoperabilidad entre aplicaciones (agentes)
- Modelo robusto desde el punto de vista lógico
- Elaborado por W3C y participación de las comunidades de metadatos

Defecto: algo complicado...

Ontología

- Especificación explícita de una conceptualización.
- Entendimiento compartido de algún dominio de interés.
- Especificación formal
 - ... de una conceptualización de un dominio
 - ... de alguna parte de la realidad que es de interés

Objetos: énfasis en dinámica

Ontología: énfasis en la estructura y relaciones

Un ejemplo: Depmark

Se está desarrollando para los Departamentos de Computación de Universidades chilenas:

- Ontología
- Marcado de páginas con metadatos (en RDF)
- Centro de consultas

¿Que puedo hacer yo?

- Difundir la buena nueva...
- Ayudar a estandarizar metadatos
- Desarrollar Ontologías
- Marcar tus datos con RDF
- Desarrollar aplicaciones en comunidades y redes locales

Conclusiones y Referencias

$$\frac{\text{Web Semántica}}{2002} = \frac{\text{Web actual}}{1994}$$

Referencias:

www.dcc.uchile.cl/~cgutierr/websemantica/

www.w3.org/2001/sw/

www.w3.org/sciam.com/special/