

Web Semántica Geo Espacial

M. Andrea Rodríguez-Tastets

Universidad de Concepción, Chile
Centro de Investigación de la Web, Univ. de Chile
andrea@udec.cl

Tópicos Emergentes de la Web
Santiago - 2007

Outline

Geo-espacial + Web Semántica

Desafíos

Contenido en Datos no Estructurados
Integración de información

Explotando la Geo Información

Geo-referenciación
Consultas con Operadores Espaciales

Integración de Datos Basada en Ontologías

Representación de Información geográfica
Ontologías
Integración de Ontologías

Resumen

Visión de la Web por Tim Berners-Lee (2001)

“**En la oficina del doctor**, Lucy se comunica con su agente de Web Semántica a través de su browser en su dispositivo de bolsillo. El agente rápidamente recupera información acerca del tratamiento indicado a la madre desde el agente del doctor, encuentra varias listas de proveedores y revisa si ellos están dentro del plan del seguro médico de la madre, se encuentran **dentro de un radio de 20 millas** de **su hogar** y con un buen historial de servicios.”

Web Semántica versus web semántica

	Web Semántica	web semántica
Filosofía	Construir un formato común para representar el significado. Uso de ontologías para entender el contenido de la Web	Codificar el contenido de la Web con tags especiales
Lenguajes	RDF, RDFS, OWL	Microformatos basado en XHTML
Semántica	Definida por una ontología	Modelo semántico no formal

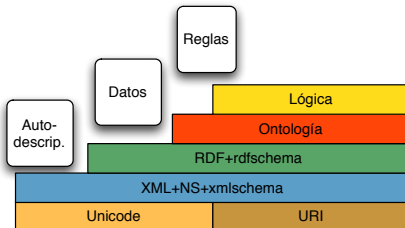
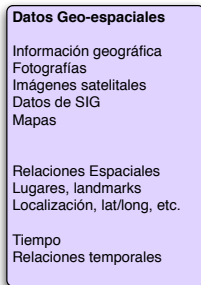
Información geográfica se refiere a datos con una localización en la superficie terrestre.

Ejemplo de información geográfica

- ▶ Datos censales
- ▶ División político administrativa
- ▶ Datos de relieve de la superficie
- ▶ Redes de transportes
- ▶ Información textual con referencial geográficas
- ▶ Fotografías aéreas
- ▶ Imágenes satelitales



Geo-espacial + Web Semántica



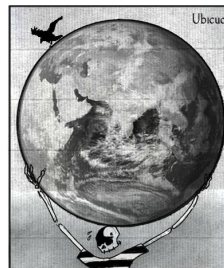
Contenido en Datos
no Estructurados
Integración de
información

Geo-referenciación
Consultas con
Operadores Espaciales

Representación de
Información
geográfica
Ontologías
Integración de
Ontologías

'Localización' o referencia geográfica es común en la Web

- ▶ ¿ Dónde se ubica el Teatro Municipal?
- ▶ ¿ Dónde trabajas?
- ▶ ¿ Dónde se encuentra el retén de policía más cercano desde donde yo estoy?
- ▶ ¿ Dónde se ubica Torres del Paine?
- ▶ ¿ Con quién limita Chile en el Norte?
- ▶ ¿ Cómo está el clima en Concepción hoy?
- ▶



Contenido en Datos
no Estructurados
Integración de
información

Geo-referenciación
Consultas con
Operadores Espaciales

Representación de
Información
geográfica
Ontologías
Integración de
Ontologías

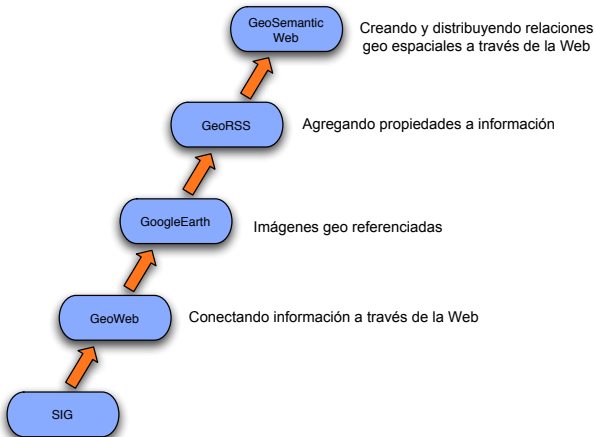
Componente	Descripción
Geo-espacial	Mapas y visualización de mapas Representación geométricas Relaciones espaciales Sistemas de referencias y coordenadas
Web	Datos distribuidos (localmente creado/globalmente accesado) Servicios compartidos (acoplados estrechamente o no a los datos geo espaciales) Inter operabilidad entre vendedores, tecnologías y arquitecturas
Semántica	Descubrir información o conocimiento implícito Inter operabilidad entre comunidades y dominios Razonamiento automático e inferencia
Geo Semántica	Reconocimiento de información geográfica Reconocimiento de asociaciones espaciales Razonamiento espacial

Razonamiento Espacial

- ▶ Inferencia de **Información Desconocida**
SI Chile **está al oeste** de Argentina y
Isla de Pascua **es parte** de Chile **ENTONCES**
Isla de Pascua **está al oeste** de Argentina

- ▶ Evaluación de **Consistencia**
SI Chile **está al oeste** de Argentina y
Isla de Pascua **es parte** de Chile **ENTONCES**
Isla de Pascua **debe estar al oeste** de Argentina

Evolución



Contenido en Datos no Estructurados
Integración de información

Geo-referenciación
Consultas con Operadores Espaciales

Representación de Información geográfica
Ontologías
Integración de Ontologías

Datos no Estructurados

Cómo extraer conocimiento geo espacial de documentos Web
(noticias, blogs, imágenes)

Hoy, probablemente, muchos chilenos pueden acordarse dónde se encontraban durante los terremotos ocurridos en las ciudades de Concepción y Valdivia en 1960 y Santiago en 1985, e incluso los de más edad recordarán el de Chillán en 1939. En este sentido, los terremotos han sido una constante en toda la historia de Chile. Ubicado en el llamado Cinturón de Fuego del Pacífico, Chile es una de las regiones más sísmicas del planeta. Bajo su territorio convergen la placa de Nazca y la placa continental americana, provocando periódicamente movimientos telúricos de diversa magnitud que en ocasiones provocan gigantescas catástrofes.

Tiempo

Hoy, probablemente, muchos chilenos pueden acordarse dónde se encontraban durante los terremotos ocurridos en las ciudades de Concepción y Valdivia en 1960 y Santiago en 1985, e incluso los de más edad recordarán el de Chillán en 1939. En este sentido, los terremotos han sido una constante en toda la historia de Chile. Ubicado en el llamado Cinturón de Fuego del Pacífico, Chile es una de las regiones más sísmicas del planeta. Bajo su territorio convergen la placa de Nazca y la placa continental americana, provocando periódicamente movimientos telúricos de diversa magnitud que en ocasiones provocan gigantescas catástrofes.

Localizaciones geográfica

Hoy, probablemente, muchos chilenos pueden acordarse dónde se encontraban durante los terremotos ocurridos en las ciudades de **Concepción** y **Valdivia** en 1960 y **Santiago** en 1985, e incluso los de más edad recordarán el de **Chillán** en 1939. En este sentido, los terremotos han sido una constante en toda la historia de **Chile**. Ubicado en el llamado Cinturón de Fuego del Pacífico, **Chile** es una de las regiones más sísmicas del planeta. Bajo su territorio convergen la placa de Nazca y la placa continental americana, provocando periódicamente movimientos telúricos de diversa magnitud que en ocasiones provocan gigantescas catástrofes.

+Entidades geográficas

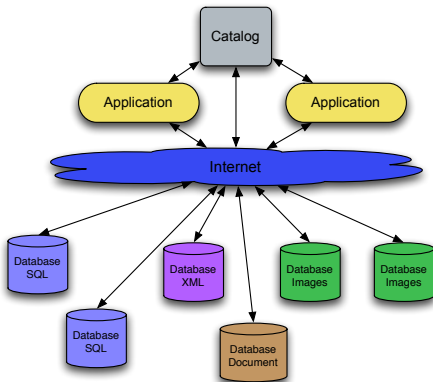
Hoy, probablemente, muchos chilenos pueden acordarse dónde se encontraban durante los terremotos ocurridos en las **ciudades** de **Concepción** y **Valdivia** en 1960 y **Santiago** en 1985, e incluso los de más edad recordarán el de **Chillán** en 1939. En este sentido, los terremotos han sido una constante en toda la historia de Chile. Ubicado en el llamado **Cinturón de Fuego del Pacífico**, **Chile** es una de las regiones más sísmicas del planeta. Bajo su territorio convergen la **placa de Nazca** y la **placa continental americana**, provocando periódicamente movimientos telúricos de diversa magnitud que en ocasiones provocan gigantescas catástrofes.

Información no Estructurada

Cómo extraer conocimiento geo espacial de documentos Web (noticias, blogs, imágenes):

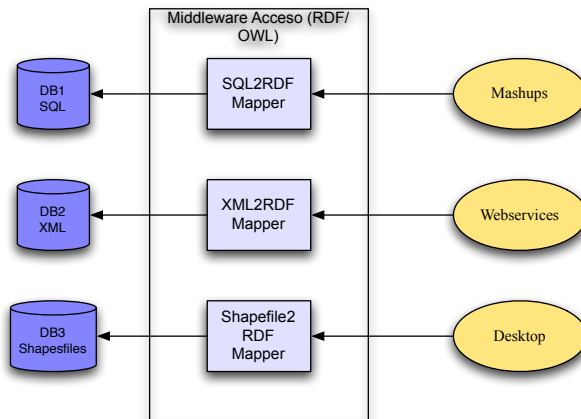
- ▶ Extraer geo-referencias automáticamente del análisis textual
- ▶ Extraer geo-referencias asignadas a través de marcadores especiales (tagging)

Contenido en Datos Estructurados



Cómo permitir que datos puedan ser integrados o que sistemas inter operen (bases de datos, XML, gazetter, etc)

Integración Dinámica



Gereferenciación por Marcadores

Insertando marcadores especiales a través de lenguajes de representación de información geográfica permite asignar una localización espacial de referencia a un documento textual en XHTML.

```
<html xmlns:foaf="http://xmlns.com/foaf/0.1#  
xmlns:geo="http://www.w3.org/2003/01/geo/wgs84_pos#>
```

```
...  
<p class="foaf:Person about="#me">  
<a rel="foaf:homepage" href="http://harry.hchen1.com">Harry Chen</a> was born in Shanghai, China.  
He moved to Hong Kong with his parents when he was ten. During the last year of his high school, he studied  
in the US as an exchange student . He completed undergraduate and graduate studies in Computer Science at  
the <a title="UMBCrel="foaf:schoolHomepage" href="http://www.umbc.edu">University of Maryland,  
Baltimore County</a> . He was awarded a PhD fellowship from HP Labs for his work on intelligent agents  
in mobile and pervasive computing.
```

```
...  
Currently he lives in Columbia, Maryland (  
<div rel="foaf:base_near">  
< span class="geo:Point">  
    <span property="geo:lat">39.207806</span> ,  
    < span property="geo:long">-76.825475</span>  
) with his wife Gigi.  
</span>  
</div>  
</p>
```

From Harry Chen (2007), Geospatial Semantic Web, Inv ited Talk.

Más allá de una búsqueda por Palabras Claves

Uno puede extender actuales facilidades de sistemas de recuperación de información en la Web incorporando operadores espaciales:

- ▶ Recuperar hoteles en Púcon y ciudades adyacentes
- ▶ Recuperar Ski Resorts localizados a 100 Km de Santiago
- ▶ Recuperar aeropuertos y puertos de Chile que estén a una distancia de 100 km uno del otro

Recuperar holetes en Púcon y ciudades adyacentes

Keyword: Hoteles and (Pucón or Villarrica or Caburga...)

Query: SELECT d.url
 FROM document d
 SUCH THAT d MENTIONS "hotels"
 WHERE d.g_reference IN "Pucón" or
 d.g_reference EQUAL "Pucón" or
 d.g_reference ADJACENT "Pucón"

Recuperar aeropuertos y puertos de Chile que estén a una distancia de 100 km uno del otro

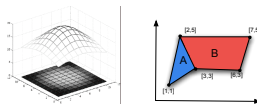
Query:

```
SELECT      d1.url,d2.url
FROM        document d1
           SUCH THAT d1 MENTIONS "aeropuerto",
           document d2
           SUCH THAT d2 MENTIONS "puerto",
WHERE       d1.g_reference IN "Chile" and
           d2.g_reference IN "Chile" and
           Distance(d1.g_reference,d2.g_reference) < 100
```

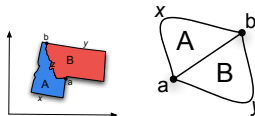
Diversidad en Información geográfica

- ▶ Mapas, Imágenes, Texto

- ▶ Superficie versus Objetos



- ▶ Geometría versus Topología



- ▶ Diversidad en Sistemas de Referencias

Estándares para información geográfica

La diferencia en la descripción de contenido espacial es la necesidad de especificar características geométricas (forma, localización, etc.).

Actualmente existen muchas formas de describir información geográfica:

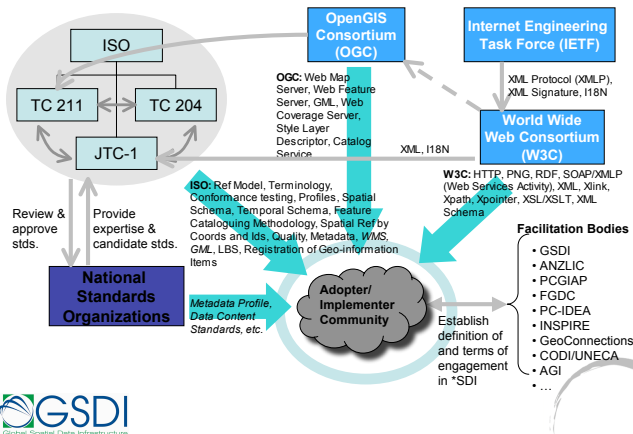
- ▶ GML (Geographical Markup Language, definido por OGC)
- ▶ W3C geo ontología de W3C
- ▶ GeoRSS
- ▶ Geonames ontology
- ▶ Google KML usado por Google Earth y Google Maps

Contenido en Datos no Estructurados
Integración de información

Geo-referenciación
Consultas con Operadores Espaciales

Representación de Información geográfica
Ontologías

Integración de Ontologías



W3C Geo

```

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:geo="http://www.w3.org/2003/01/geo/wgs84_pos#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns="http://xmlns.com/foaf/0.1/" >
  <Person>
    <name>Dan Brickley</name>
    <homepage dc:title="Dan's home page"
      rdf:resource="http://danbri.org/" />
    <based_near geo:lat="51.47026" geo:long="-2.59466" />
    <rdfs:seeAlso rdf:resource="http://danbri.org/foaf.rdf" />
  </Person>
</rdf:RDF>

```

Subject	Predicate	Object
:_a	rdf:type	foaf:Person
:_a	foaf:name	"Dan Brickley"
:_a	foaf:based_near	:_geo
:_geo	geo:long	"-2.5966"
:_geo	geo:lat	"51.47026"

Origen: <http://www.w3c.org/2003/01/geo/>

GeoRSS

GeoRSS es un lenguaje para incluir la descripción de localización en RSS.

Tres tipos de formatos son:

- ▶ GeoRSS Simple
- ▶ W3C Geo
- ▶ GML Geometry

Simple	GML
<code><georss:point> 45,256 - 71,92 </georss:point></code>	<code><georss:where></code>
<code><georss:line> 45,256 - 110,45</code>	<code><gml:Polygon></code>
<code>46,46 - 109,48</code>	<code><gml:exterior></code>
<code>43,84 - 109,86 </georss:line></code>	<code><gml:LinearRing></code>
<code><georss:polygon> 45,256 - 110,45</code>	<code><gml:posList></code>
<code>46,46 - 109,48</code>	<code>45.256 -110.45 46.46 -109.48</code>
<code>43,84 - 109,86</code>	<code>43.84 -109.86 45.256 -110.45</code>
<code>45,256 - 110,45 </georss:polygon></code>	<code></gml:posList></code>
<code><georss:box> 42,943 - 71,032</code>	<code></gml:LinearRing></code>
<code>43,039 - 69,856 </georss:box></code>	<code></gml:exterior></code>
<code><georss:featuretypeptag>city</georss:featuretypeptag></code>	<code></gml:Polygon></code>
<code><georss:relationshipptag>is_contained_within<</code>	<code></georss:where></code>
<code>/georss:relationshipptag></code>	

Elementos de GeoRSS:

- ▶ Point
- ▶ Line
- ▶ Polygon
- ▶ Box
- ▶ Elevation
- ▶ Radius
- ▶ Feature Type
- ▶ Feature Name
- ▶ Relationship Tags

<http://www.georss.org/simple.html>

Geo-espacial +
Web Semántica

Desafíos

Contenido en Datos
no Estructurados
Integración de
información

Explotando la Geo
Información

Geo-referenciación
Consultas con
Operadores Espaciales

Integración de
Datos Basada en
Ontologías

Representación de
Información
geográfica
Ontologías
Integración de
Ontologías

Resumen

Ontología

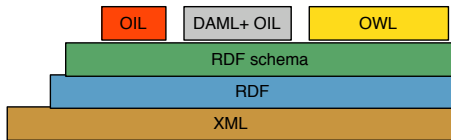
- ▶ Teoría lógica que especifica de forma explícita y acotada una conceptualización, entendiendo por conceptualización una estructura semántica intencional que codifica las reglas implícitas que restringen la estructura de una pieza de la realidad.
- ▶ Es un tipo de base de conocimiento que contiene información independiente de un estado.
- ▶ En su forma más simple es una jerarquía de conceptos creados por un proceso de generalización.
- ▶ Es una forma más compleja es un especificación que contiene axiomas que relacionan conceptos y restringen interpretaciones.

Posibles componentes de especificación

- ▶ Una jerarquía de categorías elementales.
- ▶ Una jerarquía de enlaces semánticos para conectar las categorías elementales.
- ▶ Un conjunto de definiciones de conceptos compuestos en términos de conceptos elementales.
- ▶ Un conjunto de axiomas acerca de los conceptos.
- ▶ Un conjunto de restricciones utilizadas para realizar el enlace semántico de objetos.

Lenguajes de Especificación

Cómo hacer que la gente comparta ontologías y estándares. Los estándares actuales abarcan diferentes niveles de especificación.



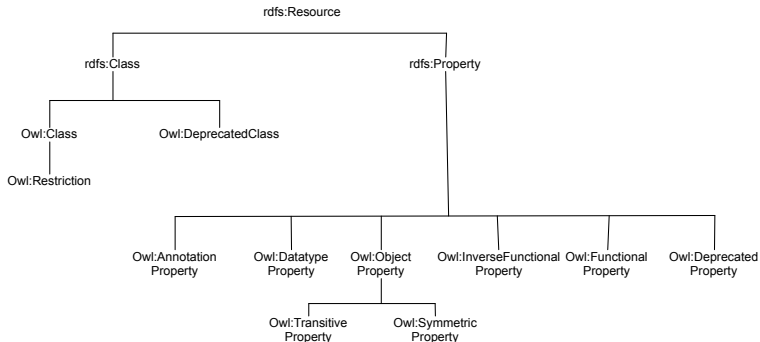
Contenido en Datos
no Estructurados
Integración de
información

Geo-referenciación
Consultas con
Operadores Espaciales

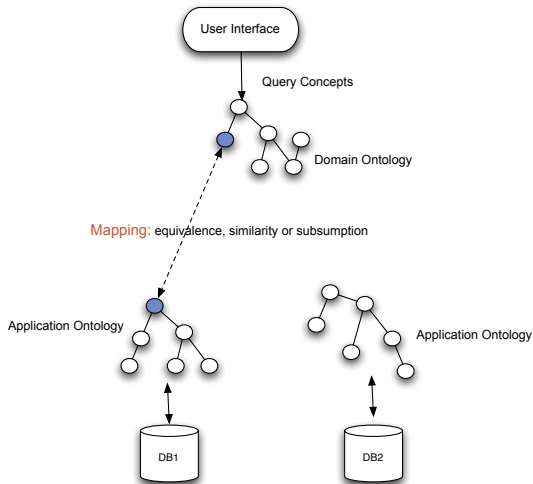
Representación de
Información
geográfica

Integración de
Ontologías

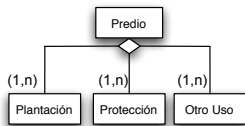
OWL como una extensión de RDF



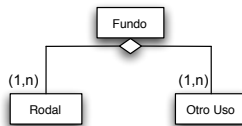
Rol de Ontologías en la Integración



Ejemplo: Incendios versus Manejo Forestal

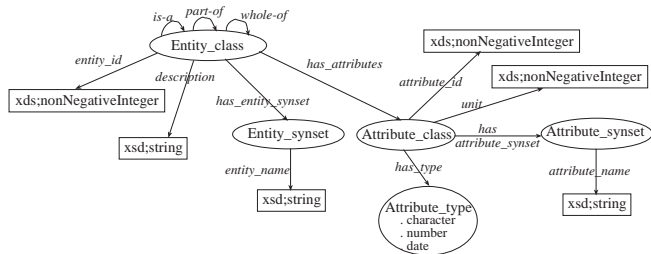


Para Control de Incendio Forestales



Para Manejo Forestal

Grafo RDF: Ontología para un Dominio Geográfico



Geo-espacial +
Web Semántica

Desafíos

Contenido en Datos
no Estructurados
Integración de
información

Explotando la Geo
Información

Geo-referenciación
Consultas con
Operadores Espaciales

Integración de
Datos Basada en
Ontologías

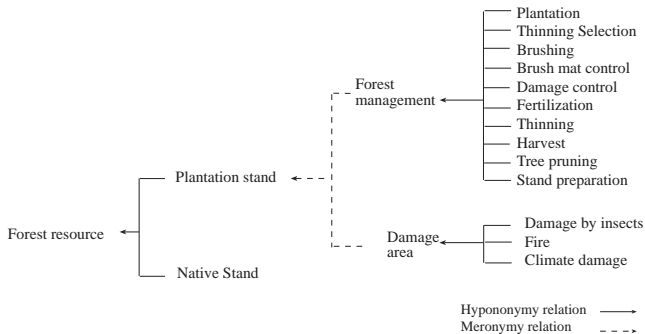
Representación de
Información
geográfica

Ontologías

Integración de
Ontologías

Resumen

Ontología para Dominio Forestal



Integrar Ontologías

A través de comparaciones que determinan métricas de similaridad.

- ▶ Comparación léxica de nombre de conceptos
- ▶ Comparación de atributos entre conceptos usando funciones de conjunto
- ▶ Comparación de vecindad conceptual.
- ▶ Comparación de relaciones espaciales entre conceptos geográficos

Y entonces...

- ▶ Aplicaciones y datos geo-espaciales son parte de la Web y aumenta su productividad
- ▶ Diferentes tipos de información geográfica existe y, por lo tanto, diferentes formas de representación (texto, mapas, imágenes, etc..)
- ▶ Las operaciones sobre datos geográficos están directamente relacionados con los servicios de Web y determinan restricciones en la forma de representar datos
- ▶ Para que haya una mayor productividad de la Web en el contexto de información geográfica, es necesario crear representaciones de información de recursos que están débilmente organizados y enlazar estos recursos a formalizaciones dadas por Web Semántica.
- ▶ El problema de integración conlleva problemas de heterogeneidad semántica y problemas de consistencia.
- ▶ Cómo hacer que un usuario sepa dónde buscar información, lo que implica descubrir servicios de información