

Escuela Politécnica de Ingeniería
Grado de Ingeniería Informática en Tecnologías de la Información

Tecnologías Emergentes

Tecnologías Web



Guía docente

- El entorno Web X.0
- Web Semántica y metadatos

Cronograma y esquema

- 3 horas de teoría
 - Fundamentos
 - Web 1.0, 2.0 y 3.0: Evolución de características
 - Tecnologías: Evolución con web X.0
 - Herramientas: Evolución con web X.0 + Tecnologías
 - Aplicación: Desarrollo de aplicaciones (usos) aplicando tecnologías y herramientas (web 3.0, web semántica)
- 2 horas de prácticas (una sesión)
 - Práctica vinculada a aspectos de web 3.0, web semántica: Web de datos y SPARQL

Contenido

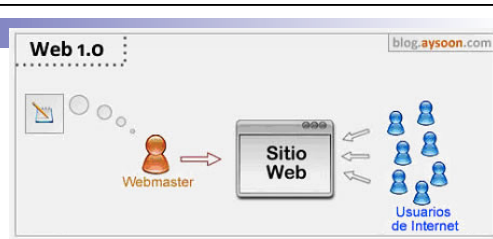
- Web 1.0, Web 1.x, Web 2.0, Web 3.0, ...
 - Características
 - Categoría
 - Evolución
- Web semántica
 - Generalidades
 - Información y consultas
 - Microformat, Microdata, RDF, ...
 - SPARQL

EVOLUCIÓN DE LA WEB

Web 1.0



Web 1.0



■ Características:

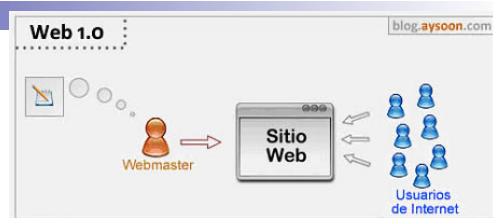
- Web unidireccional, de lectura.
- Contenidos estáticos: una vez publicados, no se actualizaban salvo que el "webmaster" los modifique y vuelva a subir la web de nuevo a internet.

■ Objetivo:

- Carácter principalmente divulgativo: documentos e información principalmente cultural. Desventaja: contenidos que rápidamente quedaban anticuados al ser complejo actualizarlos.

- Esta Web fue creada alrededor del año 1989, aunque se hizo pública más tarde, 1991. Son páginas que sólo muestran contenido y donde los usuarios no pueden comentar, agregar imágenes o videos, sólo pueden leer.

Web 1.x

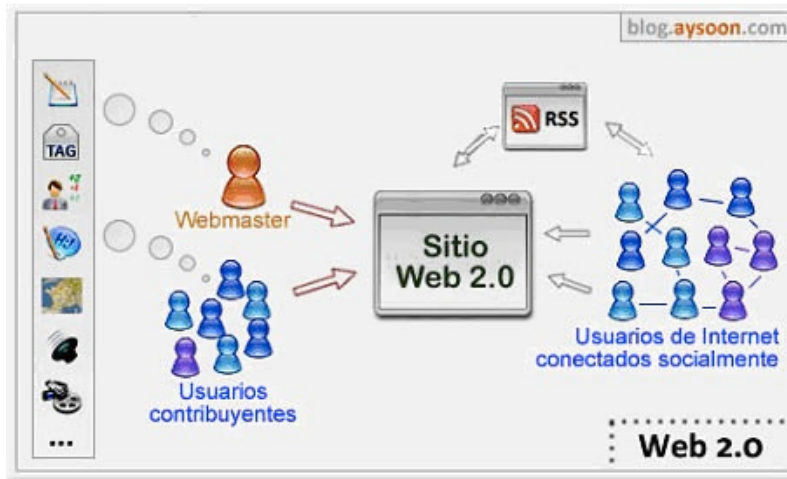


■ Características

- Reconocida en algunos casos como un estado avanzado de la Web 1.0
- Web unidireccional pero DINÁMICA: el administrador tiene herramientas para manipular los contenidos sin que éste tenga conocimientos de programación o Internet.

- Primeros "portales" de Internet, sitios web basados esencialmente en contenidos dinámicos, con acceso a bases de datos, publicados por la empresa que los gestiona.

Web 2.0



Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

9

Web 2.0

■ Características:

- Web de lectura y escritura.
- Permite la colaboración, comunicación e intercambio de información entre usuarios. Cualquier persona es capaz de agregar contenido sin la necesidad de ser el creador de la Web.
- Redes sociales, Wikis, Foros, Blogs
- Sustentada bajo unas conexiones a internet evolucionadas (ADSL), y herramientas para desarrollar web, mejores servidores, etc.

■ Objetivo:

- Compartición del conocimiento, Web colaborativa, con usuarios "no técnicos".

- El término nace a mediados del 2004, pero aparece entre los años 1999 (con sitios como Napster y Blogger) y 2001 (con Wikipedia entre otros). Es la generación de "hoy", la que usamos, donde los usuarios son los creadores de contenidos: cualquier sitio Web donde los usuarios pueden interactuar con la página y con las demás personas creando perfiles, agregando comentarios, imágenes o creando cualquier clase de contenido.

Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

10

Web 2.0. <http://internality.com/web20>

Este diagrama basado en el *Creative Commons de los Participantes*, se permite la copia, distribución y modificación por cualquier medio siempre que se mantenga el reconocimiento a sus autores. El contenido puede estar sujeto a modificaciones posteriores sin aviso previo.

Weblogs
Páginas web que contienen anotaciones ordenadas cronológicamente.

Redes de blogs
Sistemas de creación y alojamiento: Blogger, WordPress, La Cuchara, Windows Live Spaces, etc.

Lectores RSS y servicios relacionados
Código que permite al navegador conectarse mediante suscripciones a los contenidos RSS.

Agregadores
Servicios que agrupan información procedente de varios fuentes: Digg, MyScoop, etc.

Noticias y contenidos votados por usuarios
Servicios que permiten a los usuarios votar y recomendar contenido: Digg, MyScoop, etc.

Recomendaciones de contenidos
Sistemas de recomendación de contenido: Last.fm, Pandora, etc.

Búsquedas 2.0
Búsquedas personalizadas: WebCrawler, etc.

Comunidades Móviles y de Acceso
Comunidades de usuarios móviles: Facebook, etc.

Compartir videos
Plataformas de video: YouTube, etc.

Compartir fotos
Plataformas de fotos: Flickr, etc.

Podcasting
Archivos de audio que se distribuyen mediante suscripción a los usuarios.

Vídeos blogs
Archivos de vídeo que se distribuyen mediante suscripción.

Wikis
Contenido creado por los usuarios que puede ser editado y compartido.

Aplicaciones en línea
Puntos de acceso al contenido de aplicaciones, como correo y calendarios.

Páginas de inicio personalizadas
Páginas de inicio personalizadas con contenido y servicios.

Aplicaciones sobre Mapas
Aplicaciones que utilizan mapas: Google Maps, etc.

Redes sociales personales
Redes sociales para compartir información personal: Orkut, etc.

Redes sociales profesionales/Networking
Redes sociales para profesionales: LinkedIn, etc.

Mircosociedades
Pequeñas comunidades de usuarios: Del.icio.us, etc.

Mircosociedades Sociales y tagging
Redes sociales con etiquetas: Flickr, etc.

Conceptos clave:
 - **«conversaciones»**: Actitud de las empresas y organizaciones a relacionarse declaradamente y de forma transparente con los consumidores gracias a los nuevos canales de comunicación y canales alternativos.
 - **microformatos**: Código que permite añadir significado semántico a un contenido de texto de modo que las ordenaciones puedan entenderse mejor y utilizarse adecuadamente.
 - **reputación / confianza**: Cantidad de actividad en un entorno, su credibilidad influye en lo que le rodea, especialmente en la cantidad de actividad y confianza que es capaz de generar a su alrededor.
 - **contenido generado por el usuario**: La información generada, subida y compartida por los individuos tiene que seguir nuevos servicios basados principalmente en ese tipo de contenido.
 - **participación**: A través de la información, la actividad de los usuarios puede y minimizar los costes y ayuda a los demás y a uno mismo, sean personas o empresas.
 - **filtrado colaborativo**: Cuando muchos usuarios aportan sus gustos sobre ciertos tipos de contenido se pueden crear modelos y predicciones individuales basadas en la opinión de muchos usuarios.
 - **periodismo ciudadano**: Lo que se convirtió en un fenómeno que ha permitido a los ciudadanos, a través de los blogs, vídeos, las cámaras, etc. y compartirlos, crear un nuevo tipo de periodismo más colectivo y directo que el tradicional.
 - **hacerlo divertido**: Muchos servicios online de sus aplicaciones online como el contenido personal que está disponible y se puede «ver» desde cualquier lugar.
 - **la web como plataforma**: Muchos servicios online de sus aplicaciones online como el contenido personal que está disponible y se puede «ver» desde cualquier lugar.
 - **comunicación**: Transferir información en cualquier formato, a los usuarios de cualquier forma simple y fácil, según sea una de las prioridades de los usuarios.
 - **larga cola**: Concepto económico que afirma que la demanda colectiva de lo menos popular, en su conjunto, puede llegar a ser mayor que la de los grandes éxitos, creando el efecto.
 - **computación social**: Utilización de la colección para realizar tareas de comunicación colectiva o complejas, mediante el apoyo de los usuarios, que a veces son intrínsecamente humanas y no mecánicas.
 - **datos e información: lo más importante**: El contenido es el rey porque existen nuevas posibilidades de compartir, desde el mundo físico, desde internet, desde el móvil, etc.
 - **simplicidad**: El contenido es el rey porque existen nuevas posibilidades de compartir, desde el mundo físico, desde internet, desde el móvil, etc.
 - **compartir**: Permite la complejidad de servicios con muchas características y funciones, la simplicidad es necesaria para que sea usable y accesible.
 - **interfases «ricas»**: Formas avanzadas de que un usuario interactúa con una aplicación o página web, ofreciendo funciones y nuevas posibilidades de interacción en tiempo real y de manera dinámica.
 - **tags**: Etiquetas que permiten clasificar y organizar la información y que sirven para clasificarla, generalmente de modo informal.
 - **folksonomías**: Metodología de clasificación de la información que utiliza etiquetas informales de modo descentralizado sobre algunos recursos tales como blogs, páginas, vídeos o textos.
 - **software social**: Herramientas que basan su actividad en las necesidades o flujos de comunicación de los usuarios y que normalmente forman una comunidad de usuarios comunes.
 - **redes sociales**: Redes en las que estructura los nodos individuales son personas que mantienen relaciones, tales como amistad, intereses comunes o fines comerciales.

Web 2.0

Fundación Orange. Web 2.0 (2007)
Curso 2013-2014
Tecnologías Web. Tecnologías Emergentes

Web 3.0

documents on the Web ...

created by webapps from ...

data in traditional databases

JSON

RDF/N3

Triplify

LinkedData

Semantic Web

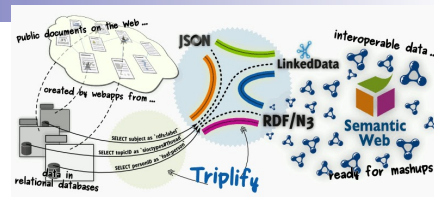
interoperable data

ready for mash

```
SELECT subject as "rdfs:label"
SELECT topicID as "sioc:types#thread"
SELECT personID as "foaf:person"
```

Curso 2013-2014
Tecnologías Web. Tecnologías Emergentes

Web 3.0



■ Características

- web de la nube: aprovecha la nube para prestar servicios al usuario, eliminando la necesidad de disponer de sistemas operativos y de capacidad de almacenamiento.
- web de las aplicaciones
- web multidispositivo: Ordenadores, Tablets, Smartphones, y más... mapas interactivos, automóviles, electrodomésticos, ...

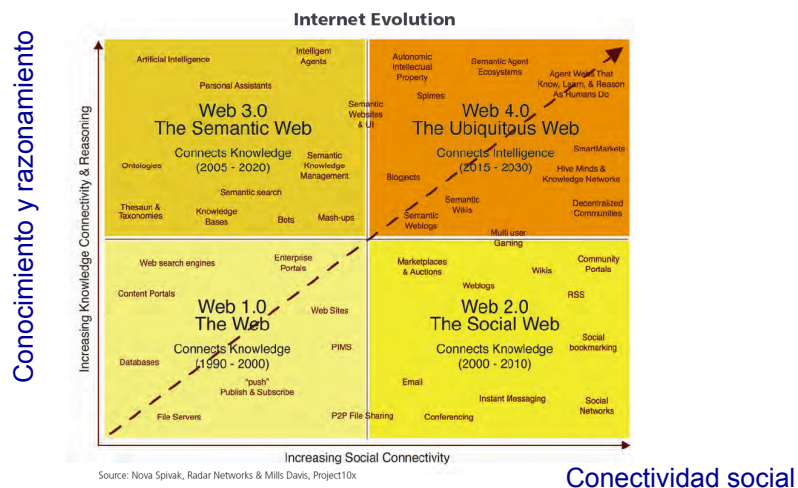
■ Objetivo:

- Combinación del contenido semántico, inteligencia artificial, inteligencia colectiva y gestión del conocimiento.

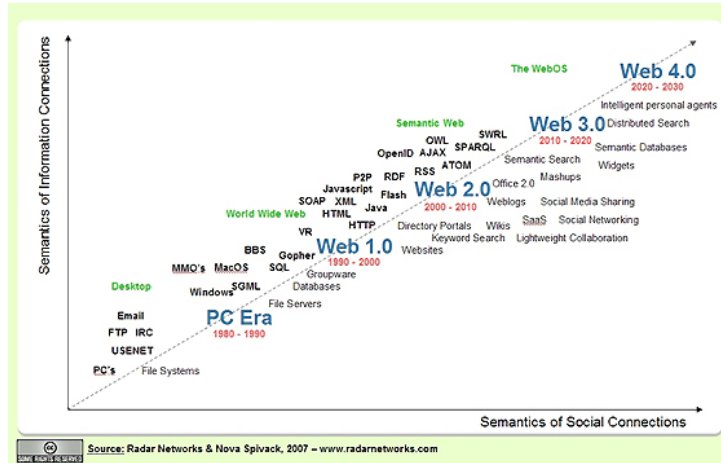
■ WEB SEMÁNTICA

- Interpretación correcta de las palabras del usuario para mostrar información con contenido relevante a la búsqueda. Por ejemplo buscando "trasmordos entre oviedo y Málaga en tren":
 - Ahora Google muestra una serie de resultados donde se pueden encontrar agencias de viajes, buscador de vuelos, noticias e información de Oviedo, Málaga, trenes...
 - Web 3.0 listaría los trasmordos de las combinaciones de trenes desde Oviedo a Málaga.
- Será como crear contenido con solo dar los metadatos: la Web buscará toda la información posible para que al final mostrar resultados relevantes sobre el tema

Clasificación de la web

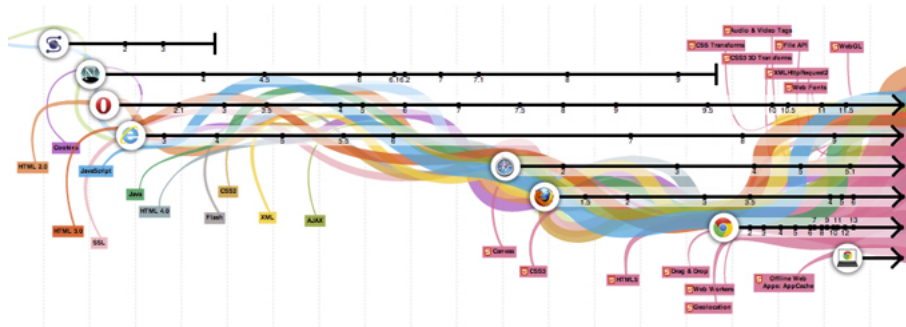


Evolución de la web (I)



Evolución de la web (II)

- <http://evolutionofweb.appspot.com>



¿Web 3.0 = Web Semántica?

- Web 3.0: entorno en el que las aplicaciones y usuarios intercambian datos, los procesan y se infiere nuevo conocimiento
- Web Semántica según W3C:
“The term “Semantic Web” refers to W3C’s vision of the **Web of linked data**. Semantic Web technologies enable people to create data stores on the Web, build vocabularies, and write rules for handling data. Linked data are empowered by technologies such as RDF, SPARQL, OWL, and SKOS”
- Web 3.0 lo conforman las aplicaciones que utilizan la web semántica y donde se aplican sus tecnologías

WEB SEMÁNTICA

Incorporamos contenido semántico a las aplicaciones

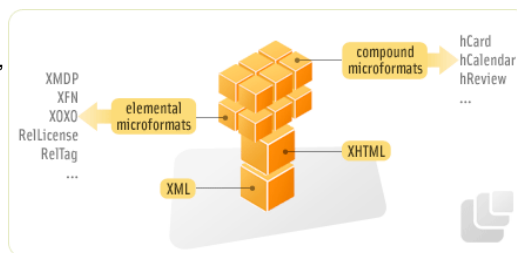
- Posibilidades
 - Microformats
 - Microdata
 - RDF + RDFs + OWL + SPARQL + ...

Contenido Semántico

MICROFORMATOS

Microformatos

- Elementos que permiten agregar contenido semántico a una página web, mediante nombre de clase específicos:
 - Atributo **class**: indica el nombre de la clase para un elemento
 - Atributo **rel**: indica en un enlace la relación entre dos documentos (el actual, es decir el que enlaza, con el documento destino)
- Actualmente permiten añadir contenido de:
 - eventos, información de contacto, direcciones, localización, ...
- Conjuntos de formatos de datos abiertos y simples, desarrollados sobre estándares ya existentes de W3C, ampliamente adoptados, incluyendo HTML5, XHTML y CSS3.
- Ver microformats.org



Microformatos. Ejemplo(I)

```
<div id="hcard-María-José-Suárez-Cabal" class="vcard">
  <span class="fn n">
    <span class="given-name">María</span>
    <span class="additional-name">José</span>
    <span class="family-name">Suárez-Cabal</span>
  </span>
  <div class="org">Universidad de Oviedo</div>
  <a class="email" href="mailto:cabal@uniovi.es">cabal@uniovi.es</a>
  <div class="adr">
    <div class="street-address">Campus de Gijón</div>
    <span class="Locality">Gijón</span>,
    <span class="region">Asturias</span>,
    <span class="postal-code">33240</span>
    <span class="country-name">España</span>
  </div>
  <p style="font-size: smaller;"> This <a href="http://microformats.org/wiki/hcard">hCard</a> created with the <a href="http://microformats.org/code/hcard/creator">hCard creator</a>.</p>
</div>
```

Microformatos. Ejemplo (II)

<http://microformats.org/code/hcard/creator>

hCard Creator

hCard-o-matic

given name	<input type="text" value="María"/>
middle name	<input type="text" value="José"/>
family name	<input type="text" value="Suárez-Cabal"/>
organization	<input type="text" value="Universidad de Oviedo"/>
street	<input type="text" value="Campus de Gijón"/>
city	<input type="text" value="Gijón"/>
state/province	<input type="text" value="Asturias"/>
postal code	<input type="text" value="33240"/>
country name	<input type="text" value="España"/>
phone	<input type="text"/>
email	<input type="text" value="cabal@uniovi.es"/>
url	<input type="text"/>
photo url	<input type="text"/>
AIM screenname	<input type="text"/>
YIM screenname	<input type="text"/>
Jabber screenname	<input type="text"/>
Categories (comma separated)	<input type="text" value="Restaurar"/>
	<input type="button" value="Build it!"/>

Warning - publishing your email address, phone number or instant messenger screenname on the web can open it up to abuse.

```
code
<div id="hcard-Maria-José-Suárez-Cabal" class="vcard">
<span class="fn n">
<span class="given-name">María</span>
<span class="additional-name">José</span>
<span class="family-name">Suárez-Cabal</span>
</span>
<div class="org">Universidad de Oviedo</div>
<div class="email" href="mailto:cabal@uniovi.es">cabal@uniovi.es</div>
<div class="adr">
<div class="street-address">Campus de Gijón</div>
<span class="locality">Gijón</span>

```

preview

María José Suárez-Cabal
Universidad de Oviedo
cabal@uniovi.es
Campus de Gijón
Gijón, Asturias, 33240 España
This hCard created with the [hCard creator](#).

Microformatos. Ejemplo (III)

The screenshot shows a web browser window with the address bar at localhost:8080. The page content includes a 'Hello World' heading, an hCard microformat for María José Suárez-Cabal, and a Google search results snippet for the same hCard. The hCard contains the following information: given name: María, middle name: José, family name: Suárez-Cabal, organization: Universidad de Oviedo, street: Campus de Gijón, city: Gijón, state/province: Asturias, postal code: 33240, country name: España, email: cabal@uniovi.es. The Google search results snippet shows the name, organization, email, and address, along with a map icon.

Microformatos desarrollados

- Microformatos estables. Establecen especificaciones para:
 - hCalendar: eventos
 - hCard: personas y organizaciones
 - rel-license: contenidos de licencias
 - rel-nofollow: advertencia de que el enlace es de una tercera parte con contenido no verificado
 - rel-tag: la página que contiene el enlace está relacionada con la etiqueta indicada
 - XFN, XMDP, XOXO: definen otros vocabularios de relaciones sociales, enlaces, perfiles, y otras características
- Otros microformatos, que aún no son estables y que podrían cambiar o actualizarse:
 - adr: direcciones postales
 - geo: coordenadas geográficas
 - hAtom, hNews: posts en blogs, noticias
 - hListing, hProducts: listas de productos o servicios
 - hMedia: videos, imágenes, audio
 - hRecipe: cocina, recetas
 - hResume: curriculum
 - hReview, hReview-aggregate: críticas de artículos, servicios, rangos, ...
 - rel-author: página del autor
 - rel-home: homepage
 - Rel-payment: mecanismo de pago

Ventajas del uso de microformatos

- Aplicación práctica para problemas concretos.
- Usados por herramientas que permiten extraer información.
- Facilidad de desarrollo.
- Compatibles con estándares de W3C, utilizando los elementos y atributos especificados.
- Útiles para el marcado semántico, algunos los llaman “semántica en minúsculas”.

Contenido Semántico

MICRODATA

Microdata

- Nuevas etiquetas de (x)HTML que permiten incorporar contenido semántico.
- En concreto, se trata de añadir atributos en etiquetas como ``, `<div>`, `<a>`,... para que el buscador entienda, interprete y muestre la información de manera adecuada:

```
<!DOCTYPE html>
<html>
<head><title></title></head>
<body>
  <div itemscope itemtype="http://data-vocabulary.org/Person">
    <span itemprop="name">María Jose Suarez Cabal</span> Podeis encontrarne en:
    <a href="http://grupos.uniovi.es/web/cabalu" itemprop="url">
      grupos.uniovi.es/web/cabalu</a>
    Trabajo de <span itemprop="title">Profesor Contratado Doctor</span>
    en la <span itemprop="affiliation">Universidad de Oviedo</span>.
  </div>
</body>
</html>
```

Microdata. Atributos

- <http://www.w3.org/html/wg/drafts/microdata/master/> (especificación HTML Microdata Nightly)
- **itemscope**: Crea un ítem e indica que el elemento es parte de un grupo de "microdata"
- **itemtype**: URL que proporciona el contexto para el microdata (vocabulario) y sus propiedades
- **itemid**: Identificador único
- **itemprop**: Indica una etiqueta que contiene el valor de la propiedad especificada. Los valores de las propiedades, normalmente son cadenas, pero pueden ser URLs en atributos href y elementos como img y src, que enlazan o incluyen recursos externos
- **itemref**: Proporciona la lista de elementos con propiedades

Microdata. Ejemplo

```
<div itemscope itemtype="http://schema.org/Person">
  <span itemprop="name">Maria Jose Suarez Cabal</span>
  Puede encontrarme en: <a href="http://grupos.uniovi.es/web/cabalu"
  itemprop="url">grupos.uniovi.es/web/cabalu</a>
  Trabajo como <span itemprop="title">Profesor Contratado Doctor</span> en la
  <span itemprop="affiliation" itemscope itemtype="http://schema.org/Organization">
    <span itemprop="name">Universidad de Oviedo</span>,
    <span itemprop="address" itemscope itemtype="http://schema.org/PostalAddress">
      Dirección: <span itemprop="streetAddress">Campus Universitario de Gijón</span>,
        <span itemprop="addressLocality">Gijón</span>,
        <span itemprop="addressRegion">Asturias</span>
        <span itemprop="addressCountry">España</span>.
    </span>
  </span>
  Compañeros: <span itemprop="colleague" itemscope itemtype="http://schema.org/Person">
    <span itemprop="name">Enrique de la Cal</span>
    <span itemprop="name">Yo misma</span>
  </span>
</div>
```

- ✓ Es posible anidar ítems: añadiendo el atributo itemscope a un elemento con itemprop
- ✓ Múltiples propiedades con el mismo nombre y distinto valor

Microdata. Más información

- Microdata habitualmente utilizados (schema.org):
 - Trabajos creativos CreativeWork, Book, Movie, MusicRecording, Recipe, TVSeries...
 - Objetos embebidos: AudioObject, ImageObject, VideoObject
 - Eventos: Event
 - Organizaciones: Organization
 - Personas: Person
 - ...
- Herramientas (<http://schema.rdfs.org/tools.html>)
 - Generación : Schema Creator, The Microdata Generator
 - Validadores: Validator.nu
 - Pruebas: Google's Rich Snippets Testing Tool
<http://www.google.com/snippets/testing/rstest.html>
- Uso de Microformat en Microdata:
 - Se pueden utilizar los vocabularios definidos para microformatos (hCalendar, hCard, hRecipe, hReview, ...) incluyendo la propiedad itemType:
itemtype="http://microformats.org/profile/hCard"
 - Y, luego, utilizando las propiedades definidos en ellas

Microdata. Otro ejemplo

```
<h2>Próxima clase</h2>
<section itemscope itemtype="http://microformats.org/profile/hcalendar#vevent">
  <h3 itemprop="summary">Clase de Teoría</h3>
  <p>
    <strong>Día:</strong>
    <time itemprop="dtstart" datetime="2013-12-03">3 de Diciembre de 2013
  </time>,
  desde <time datetime="2011-12-03T19:00-0800" itemprop="dtstart">7pm</time>
  hasta <time datetime="2013-12-03T20:00-0800" itemprop="dtend">8pm</time>
  </p>
  <p>
    <strong>Donde:</strong> <span itemprop="location">Aula F</span>
  </p>
</section>
<section itemtype="http://microformats.org/profile/hcard" itemscope>
  <strong>Profesora:</strong>
  <h3 itemprop="fn">Maria Jose Suarez</h3>
</section>
```


Contenido Semántico

WEB SEMÁNTICA

A por la web semántica “con mayúsculas”¿?...

Vamos a centrar el tema

- ¿Qué queremos hacer? Supongamos:
 - Aplicación web para una biblioteca donde informar, localizar, mostrar desde diferentes fuentes de información..., sobre libros
 - Selección de diferentes criterios de búsqueda (autor, traductor, editorial, otros libros del mismo autor, ...)

- Desde el punto de vista de obtener la información ¿qué opciones tenemos para llevar a cabo este proyecto?

Opciones:

- Opción 1:
 - Buscamos sitios de interés por la Web y actualizamos nuestra aplicación manualmente con los sitios encontrados
- Opción 2:
 - Buscamos sitios por la Web, pero tenemos herramientas que, a partir de los sitios encontrados, se va actualizando nuestra aplicación
- Opción 3:
 - Buscamos/actualizamos sitios en/desde la Web
- En cualquier caso, ¿cuánto tiempo tardará nuestra aplicación en estar obsoleta?

¿Alguna otra opción?

- ¿ Y si utilizamos información externa y pública, p.e. de wikipedia, Open Linked Data?
- La información, en vez de estar disponible en sitios web, estaría disponible como DATOS:
 - Los datos no serán extraídos con APIs, ni estarán ocultos en sitios web, sino que serán accedidos mediante peticiones HTTP o consultas
- Sería similar a usar la Web como un sistema de gestión de contenidos

Pero... ¿hay datos en la Web?

- La respuesta es:
 - Sí, cada vez hay más datos en la Web
 - Datos gubernamentales, datos relacionados con la salud, conocimiento en general, información de compañías, información de vuelos, restaurantes, ...
 - Y, además, cada vez más aplicaciones acceden a ellos y dependen de la disponibilidad de estos datos

Comparemos...

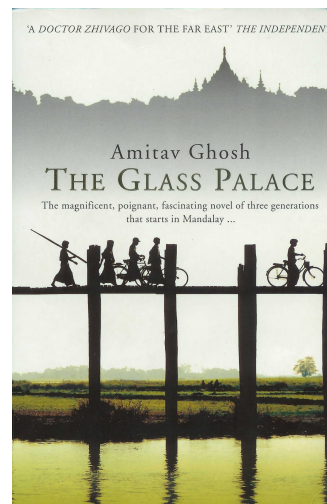
- ¿Qué pasaría si en la Web
 - los documentos estuvieran disponibles para descargarse,
 - pero no hubiera enlaces entre ellos?
- De la misma manera: que los datos estén en la Web, no es suficiente.
- En la Web de Datos, también se necesita una infraestructura de forma que:
 - Los datos estén disponibles Y accesibles mediante las tecnologías web estándar
 - ADEMÁS los datos deben estar enlazados: INTEGRADOS
- Aquí es donde entra la Web Semántica

Si tenemos datos, ¿cómo podemos trabajar con ellos? (I)

1. Mapear los datos disponibles en una representación abstracta
 - Hace que los datos sean independientes de su representación interna
2. Combinar, unir las diferentes representaciones obtenidas y...
3. ¡Ya podemos empezar a consultar información sobre el conjunto!
 - Realmente útil es obtener resultados en consultas de varios datasets que no serían posibles de forma individual

¿Cómo podemos trabajar con los datos? (II)

- Veamos un ejemplo con el que se pretenden ilustrar los principales conceptos de la Web Semántica
- En nuestra aplicación, tendremos formas de acceder a la información



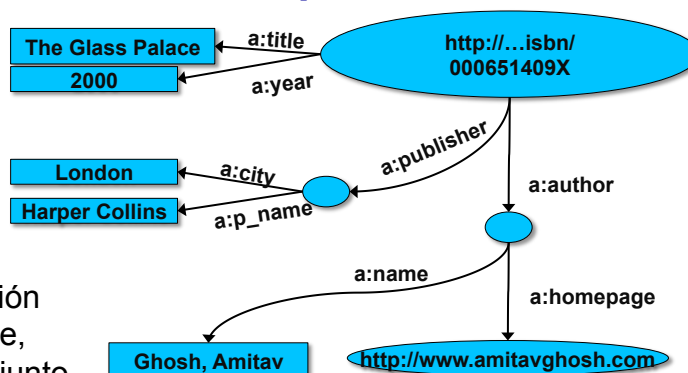
Ejemplo (I). Disponemos de cierta información: dataset “A”

ISBN	Author	Title	Publisher	Year
0006511409X	id_xyz	The Glass Palace	id_qpr	2000

ID	Name	Homepage
id_xyz	Ghosh, Amitav	http://www.amitavghosh.com

ID	Publisher's name	City
id_qpr	Harper Collins	London

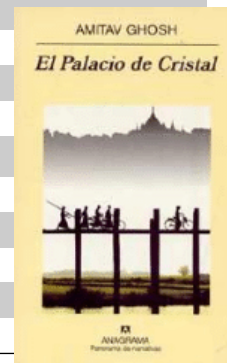
Ejemplo (II). Exportamos los datos (dataset “A”)



Representación independiente, como un conjunto de relaciones

Ejemplo (III). Disponemos de más información (dataset "F")

A	B	C	D
1	ID	Título	Traductor
2	ISBN 9788433969767	El Palacio de Cristal	\$A12\$
3			ISBN 0-00-6511409-X
4			
5			
6	ID	Autor	
7	ISBN 0-00-6511409-X	\$A11\$	
8			
9			
10	Nom		
11	Ghosh, Amitav		
12	Gómez Ibáñez, Benito		

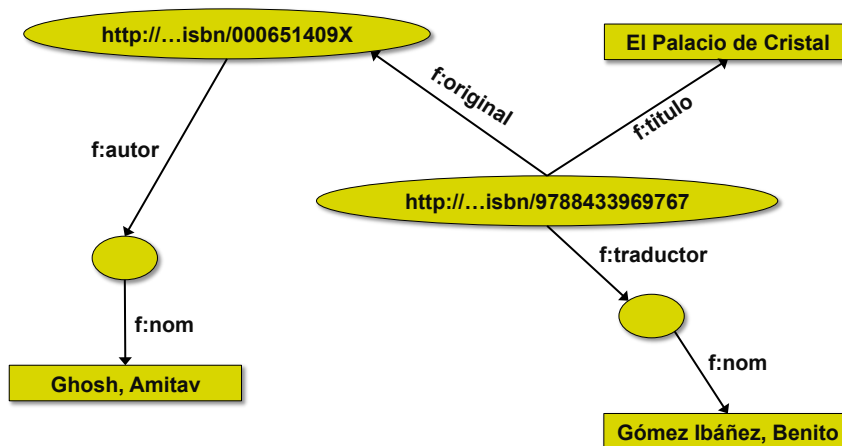


Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

43

Ejemplo (III). Exportamos "F"...

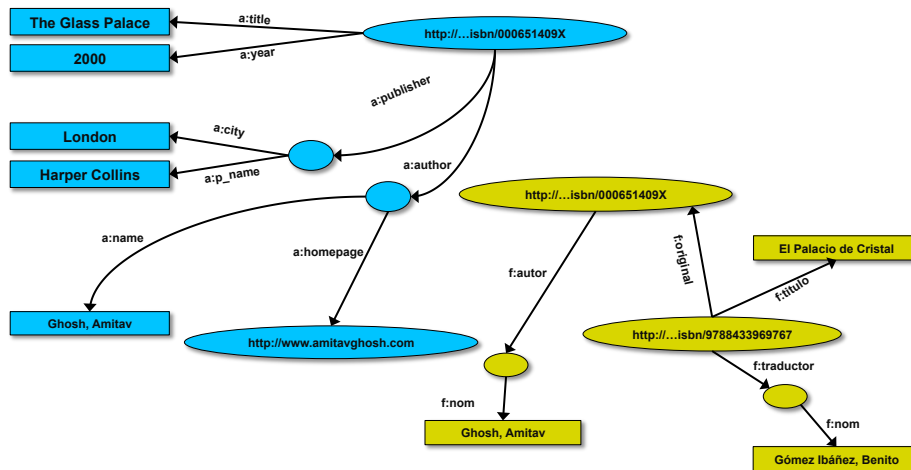


Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

44

Ejemplo (IV). Combinamos los datos

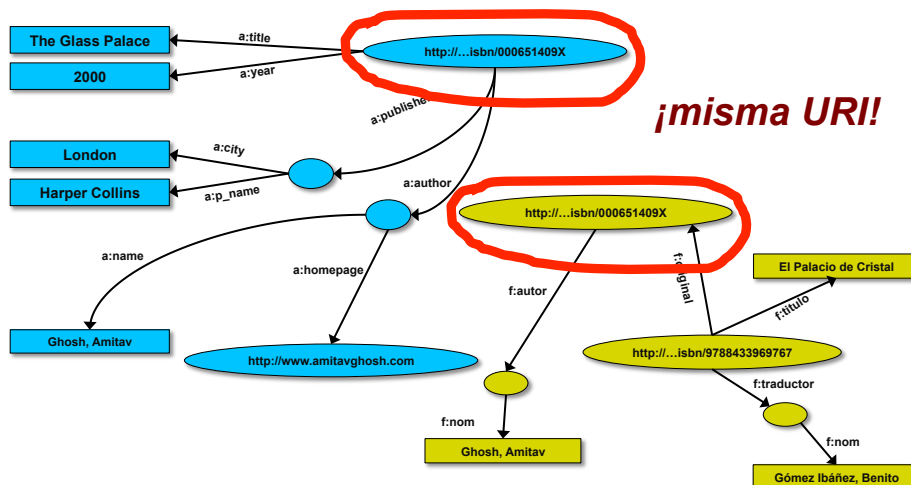


Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

45

Ejemplo (V). Combinamos los datos

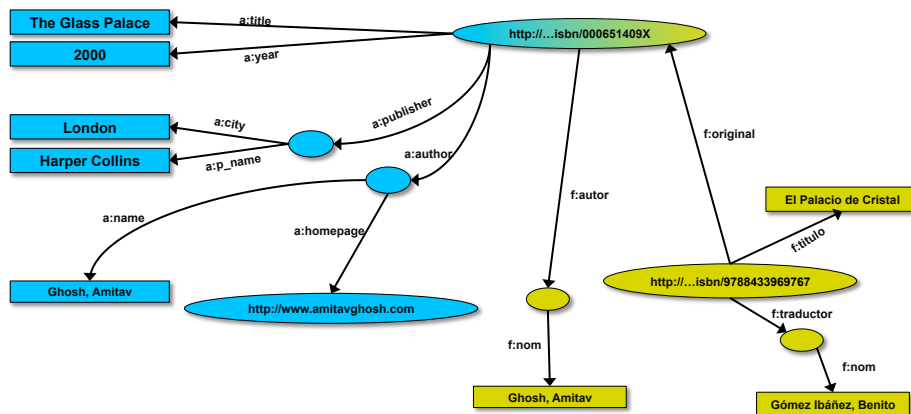


Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

46

Ejemplo (VI). Resultado de la combinación



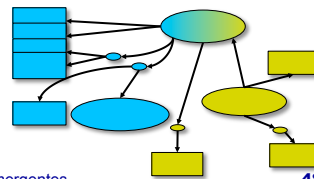
Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

47

Ejemplo (VII). Consultamos sobre los datos combinados...

- Los usuarios del dataset “F” pueden ahora consultar:
 - “dame el título original”
- Esta información no estaba disponible en “F” ...
- ...pero se puede obtener si se considera la combinación de los datasets “F” y “A”



Curso 2013-2014

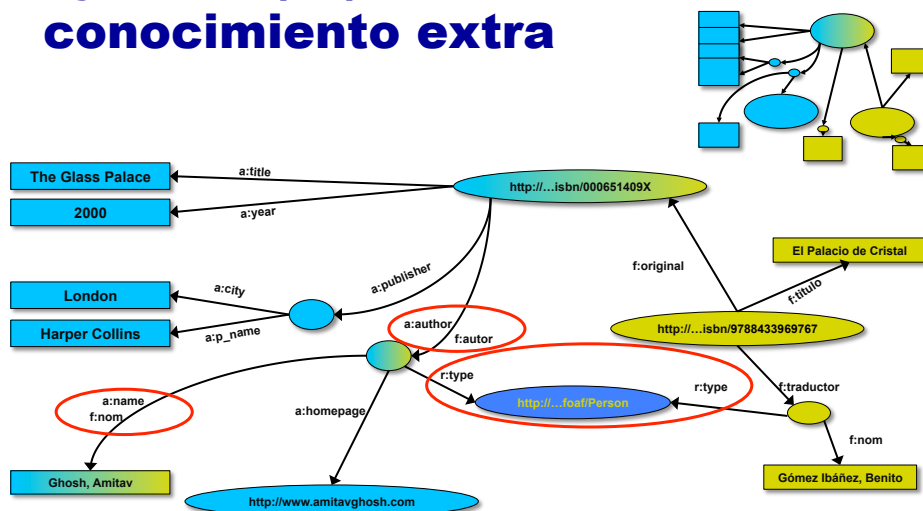
Tecnologías Web. Tecnologías Emergentes

48

Ejemplo (VIII). ¿Podremos generar nueva información?

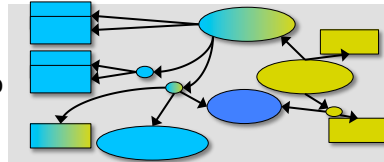
- Parece que a:author y f:autor deberían ser el mismo, aunque no se obtiene directamente de la combinación. Se puede añadir este conocimiento:
 - a:author es el mismo f:autor
- Además, podemos añadir información extra a los datos indicando:
 - a:author y, por tanto, f:autor identifican a una persona.
 - “Person” es un término que podría ya estar definido. Supongamos que se identifica el nombre de una persona y su página web
 - “Person” puede ser usado como una categoría para cierto tipo de recursos, por ejemplo para autor y traductor.

Ejemplo (IX). Añadimos conocimiento extra



Ejemplo (X). Ya podemos hacer consultas más complejas!

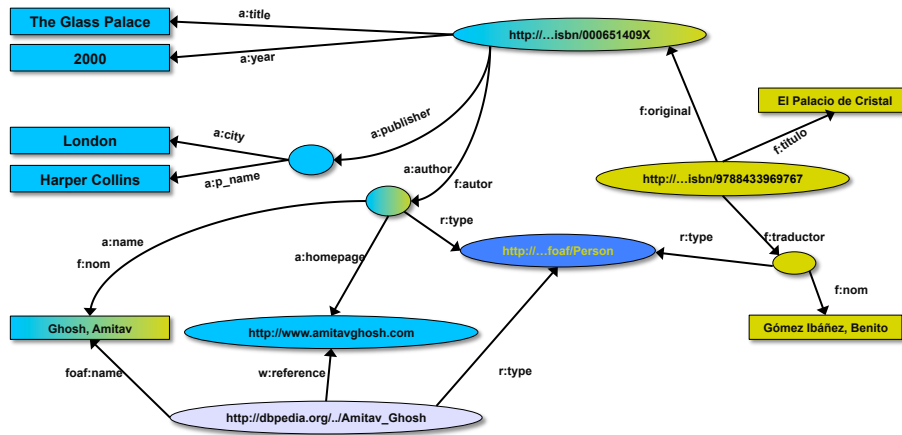
- Un usuario del dataset “F” puede consultar:
 - “cuál es la página web del autor”
- La información no está disponible en “F” (no hay esta información), tampoco en “A” (no tiene la etiqueta “autor”)...
- ...pero puede ser accesible por:
 - ✓ La combinación de los datasets
 - ✓ Identificar a:autor y f:autor como el mismo
 - ✓ Añadir la categoría “Person” y las relaciones de f:autor con ella, por lo que vamos a poder obtener su homepage



Ejemplo (XI). ¿Combinamos con otros datasets?

- Por ejemplo, mediante “Person”, reconocido y utilizado en otros datasets, nuestra información puede ser combinada con otros recursos.
 - Por ejemplo el proyecto “[dbpedia](http://dbpedia.org)” (dbpedia.org) permite extraer como datasets la información disponible en Wikipedia

Ejemplo(XII). Combinamos con los datos de Wikipedia

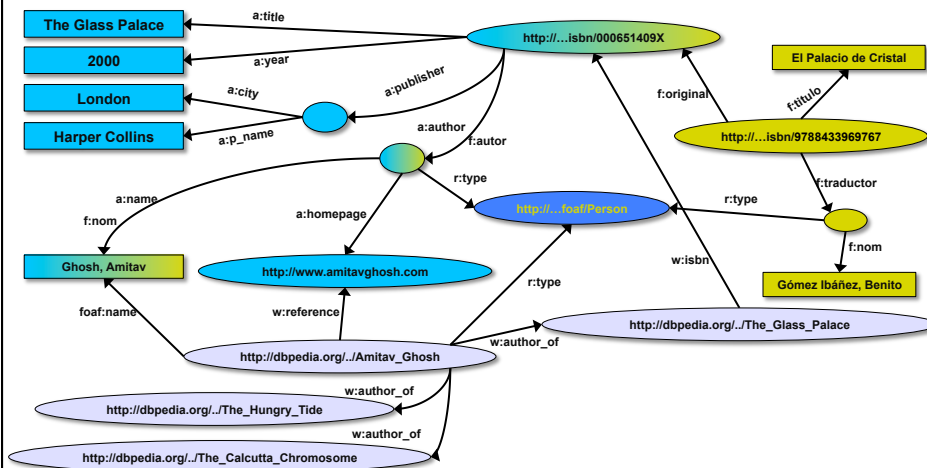


Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

53

Ejemplo(XIII). Y más datos de Wikipedia

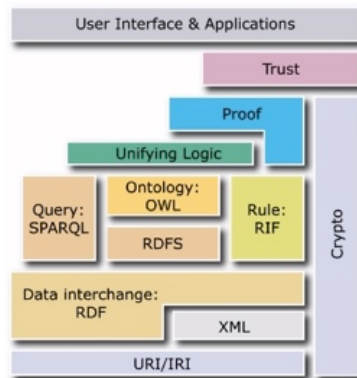
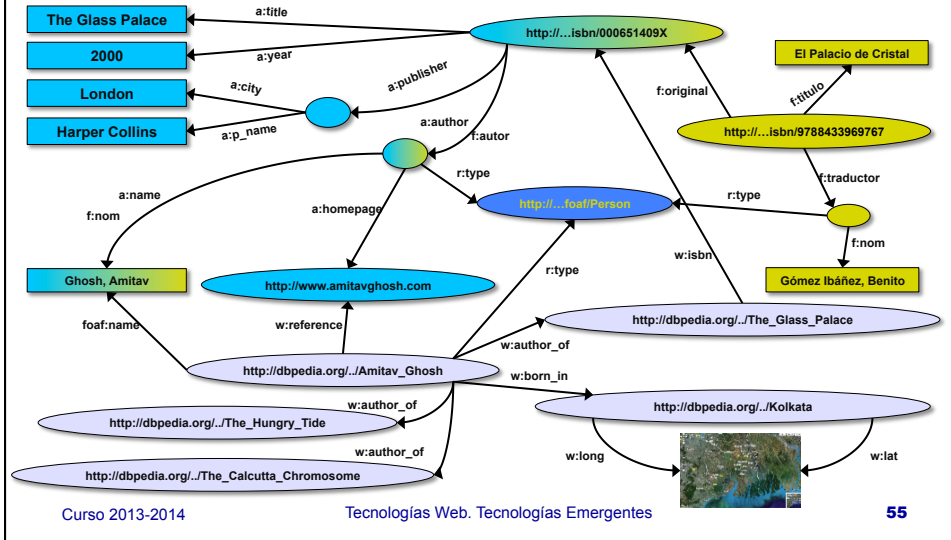


Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

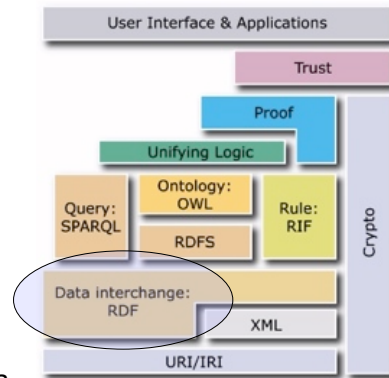
54

Ejemplo(XIV). Y más...



Contenido Semántico. Web Semántica

ARQUITECTURA (W3C)



Contenido Semántico. Web Semántica

RDF

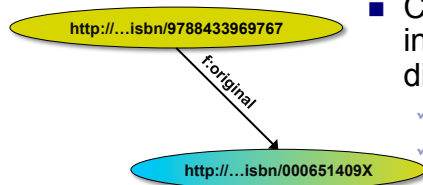
RDF

- Permite describir recursos (identificados por URIs) en la web que forman un grafo etiquetado y dirigido
- Formaliza la información conectando recursos
- *RDF* es un modelo general para estos recursos en formatos legibles por una máquina
- Estándar de W3C (www.w3c.org)

Tripletas RDF (s p o) =sentencias

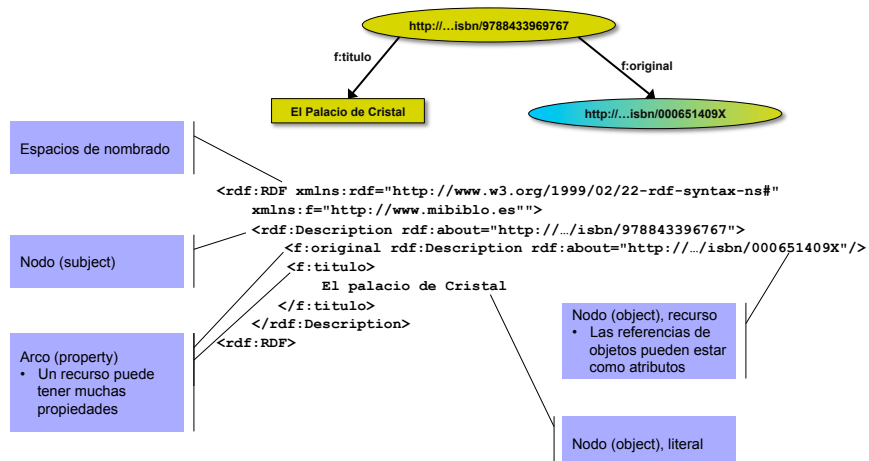
- Una tripleta RDF se define como (s p o) donde:
 - s (subject -- recurso), p (predicate -- propiedad) se representan con URIs
 - o (object -- valor de la propiedad) es URI o literal

(`<http://...isbn...767>`, `<http://.../original>`, `<http://...isbn...409X>`)



- Conjunto de sentencias interconectadas forman un grafo dirigido donde:
 - ✓ s (subject), o (object) son nodos
 - ✓ p (property o predicate) son arcos

RDF/XML



Nodos “Internos”, blank nodes

- Supongamos que tenemos que representar:
 - “El editor es «algo» con nombre y ciudad”
- Si los nodos se identifican con una URI... ¿cuál es la URI de «algo»?



Blank nodes, podemos:

- Crear URI (único) para el recurso (visible en la web)

```
<rdf:Description rdf:about="http://.../isbn/000651409X">
  <a:publisher rdf:resource="urn:uuid:f60ffb40-307d-..."/>
</rdf:Description>
<rdf:Description rdf:about="urn:uuid:f60ffb40-307d-...">
  <a:p_name>HarpersCollins</a:p_name>
  <a:city>London</a:city>
</rdf:Description>
```

- Crear un nodo interno, no visible desde fuera

```
<rdf:Description rdf:about="http://.../isbn/000651409X">
  <a:publisher rdf:nodeID="A234"/>
</rdf:Description>
<rdf:Description rdf:nodeID="A234">
  <a:p_name>HarpersCollins</a:p_name>
  <a:city>London</a:city>
</rdf:Description>
```

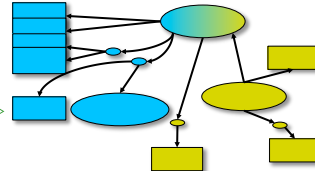
- Dejar que el sistema lo cree

```
<rdf:Description rdf:about="http://.../isbn/000651409X">
  <a:publisher>
    <rdf:Description>
      <a:p_name>HarpersCollins</a:p_name>
      <a:city>London</a:city>
    </rdf:Description>
  </a:publisher>
</rdf:Description>
```

Herramienta.

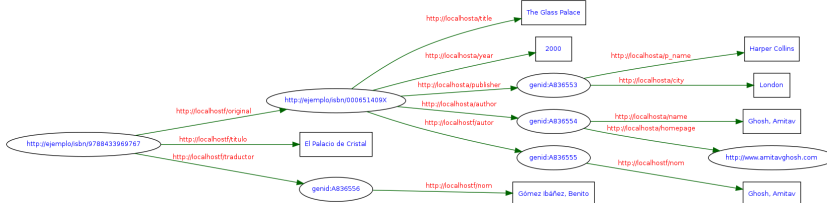
<http://www.w3c.org/RDF/validator>

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:a="http://localhost/a/"
  xmlns:f="http://localhost/f/">
  <rdf:Description rdf:about="http://ejemplo/isbn/000651409X">
    <a:title>The Glass Palace</a:title>
    <a:year>2000</a:year>
    <a:publisher> <rdf:Description>
      <a:p_name>Harper Collins</a:p_name><a:city>London</a:city>
    </rdf:Description> </a:publisher>
    <a:author> <rdf:Description>
      <a:name>Ghosh, Amitav</a:name>
      <a:homepage rdf:resource="http://www.amitavghosh.com"/>
    </rdf:Description> </a:author>
    <f:autor> <rdf:Description> <f:nom>Ghosh, Amitav</f:nom> </rdf:Description> </f:autor>
  </rdf:Description>
  <rdf:Description rdf:about="http://ejemplo/isbn/9788433969767">
    <f:titulo>El Palacio de Cristal</f:titulo>
    <f:traductor> <rdf:Description> <f:nom>Gómez Ibáñez, Benito</f:nom> </rdf:Description>
    </f:traductor>
    <f:original rdf:resource="http://ejemplo/isbn/000651409X"/>
  </rdf:Description>
</rdf:RDF>
```



Validador. Tripletas y grafo

Subject	Predicate	Object
http://ejemplo/isbn/000651409X	http://localhost/a/title	"The Glass Palace"
http://ejemplo/isbn/000651409X	http://localhost/a/year	"2000"
http://ejemplo/isbn/000651409X	http://localhost/a/publisher	genid:A836564
genid:A836564	http://localhost/a/p_name	"Harper Collins"
genid:A836564	http://localhost/a/city	"London"
http://ejemplo/isbn/000651409X	http://localhost/a/author	genid:A836565
genid:A836565	http://localhost/a/name	"Ghosh, Amitav"
genid:A836565	http://localhost/a/homepage	http://www.amitavghosh.com
http://ejemplo/isbn/000651409X	http://localhost/f/autor	genid:A836566
genid:A836566	http://localhost/f/nom	"Ghosh, Amitav"
http://ejemplo/isbn/9788433969767	http://localhost/f/titulo	"El Palacio de Cristal"
http://ejemplo/isbn/9788433969767	http://localhost/f/traductor	genid:A836567
genid:A836567	http://localhost/f/nom	"Gómez Ibáñez, Benito"
http://ejemplo/isbn/9788433969767	http://localhost/f/original	http://ejemplo/isbn/000651409X



RDF en programación

- Utilizando Java+API, como puede ser “Sesame” o “Apache Jena”, es posible:
 - Crea un objeto grafo
 - Parsear un fichero con datos RDF, para almacenarlos en el grafo
 - Añadir o recuperar, mediante métodos de la API:
 - tripletas
 - Para un sujeto, los pares (property,object)
 - Para un objeto, los pares (subject,property)
 - etc.
 - Además, combinar grafos de forma automática

Ejemplo con Jena Apache (I)

<http://jena.apache.org>

```
import com.hp.hpl.jena.rdf.model.Model;
import com.hp.hpl.jena.rdf.model.ModelFactory;
import com.hp.hpl.jena.rdf.model.Property;
import com.hp.hpl.jena.rdf.model.Resource;

public class GrafoRDF {
    Model model; //grafo

    void createModel(){
        String dba="http://localhost/"; String dbf="http://localhost/";

        model = ModelFactory.createDefaultModel(); // grafo vacío

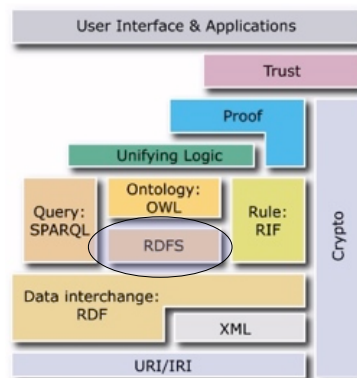
        model.setNsPrefix( "a", dba ); // establecer espacios de nombrado
        model.setNsPrefix( "f", dbf );

        Property atitle = model.createProperty(dba+"title"); // definir propiedades
        Property ayear = model.createProperty(dba+"year");
        Property aauthor = model.createProperty(dba+"author");
        ...
        Property ftitulo = model.createProperty(dbf+"titulo");
        Property fautor = model.createProperty(dbf+"autor");
        Property ftraductor = model.createProperty(dbf+"traductor");
        ...
    }
}
```

Ejemplo con Jena Apache (II)

```
// creación de los recursos añadiendo propiedades en cascada
Resource book = model.createResource("http://ejemplo/isbn/000651409X")
    .addLiteral(atitle, "The Glass Palace")
    .addLiteral(ayear, "2000")
    .addProperty(apublisher,
        model.createResource()
            .addLiteral(ap_name, "Harper Collins")
            .addLiteral(ap_city, "London"))
    .addProperty(aauthor,
        model.createResource()
            .addLiteral(aname, "Ghosh, Amitav")
            .addProperty(ahomepage, "http://www.amitavghosh.com"))
    .addProperty(fautor,
        model.createResource()
            .addLiteral(fnom, "Ghosh, Amitav"));

Resource libro = model.createResource("http://ejemplo/isbn/9788433969767")
    .addLiteral(ftitulo, "El Palacio de Cristal")
    .addProperty(ftraductor,
        model.createResource()
            .addLiteral(fnom, "Gómez Ibáñez, Benito"))
    .addProperty(foriginal, book);
```



Contenido Semántico. Web Semántica

RDF SCHEMA

RDF Schema. RDFs

- Lenguaje que describe un vocabulario, en un dominio de aplicación concreto, utilizado como extensión semántica de RDF:
 - Descripción de grupos de recursos relacionados
 - Relaciones entre estos recursos
- Notación escrita en RDF especificando características, como dominios y rangos de propiedades

RDFs. Clases y propiedades

- CLASE: Agrupación de recursos (las propias clases son recursos)
 - ✓ “todas las cosas” son recursos, instancias de esta clase.
 - ✓ rdfs:Class Clases de recursos, clases RDF
 - ✓ Valores como strings o enteros (literales) son clases
 - ✓ Las propiedades RDF también son clases
- PROPIEDAD: Relación entre “sujeto” y “objeto”, identificadas por URIs
 - ✓ Podemos definir propiedades
 - ✓ Subpropiedades
 - ✓ Subclases
 - ✓ ...

Clases y propiedades

predefinidas <http://www.w3.org/TR/rdf-schema>

✓ Clases

<code>rdfs:Resource</code>	Recurso, cualquier "cosa", instancia de <code>rdfs.Class</code>
<code>rdfs:Literal</code>	Valor: String, entero, ...
<code>rdfs:Class</code>	Clase
<code>rdfs:Property</code>	Propiedad RDF, instancia de <code>rdfs.Class</code>
<code>rdfs:Bag</code> , <code>rdfs:Seq</code> , <code>rdfs:Alt</code>	Contenedores (no ordenados, ordenados, alternativas)
<code>rdfs:Container</code> , <code>rdfs:List</code>	Contenedor, Listas
...	

✓ Propiedades

<code>rdfs:type</code>	Instancia de una clase
<code>rdfs:subClassOf</code>	Subclase de una clase
<code>rdfs:subPropertyOf</code>	Subpropiedad
<code>rdfs:domain</code>	Dominio de una propiedad
<code>rdfs:range</code>	Rango de una propiedad
<code>rdfs:label</code>	Etiqueta para un recursos
<code>rdfs:comment</code>	Descripción de un recurso
...	

Ejemplos de vocabularios

- Dublin Core: autores, fechas, ...
- FOAF: personas
- schema.org: eventos, organizaciones, revisiones, ...
- ...

RDF Schema. Ejemplo

```
<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:foaf="http://xmlns.com/foaf/0.1/">
  <rdfs:Class rdf:about="Book"
    rdfs:label="Book" >
  </rdfs:Class>
  <rdfs:Property rdf:about="Author"
    rdfs:label="Author"
    rdfs:comment="Author of an English book">
  <rdfs:domain rdf:resource="#Book"/>
  <rdfs:SubClassOf rdf:resource="foaf:Person"/>
  </rdfs:Property>
  <rdfs:Property
    rdf:about="Title" rdfs:label="Title"
    rdfs:comment="Title of an English book">
  <rdfs:domain rdf:resource="#Book"/>
  </rdfs:Property>
  <rdfs:Property rdf:about="Title" .../>
  <rdfs:Property rdf:about="Year" .../>
  <rdfs:Property rdf:about="Publisher" .../>
  <rdfs:Property rdf:about="P_name" ...>
  <rdfs:subPropertyOf rdf:resource="#Publisher"/>
  </rdfs:Property>
  <rdfs:Property rdf:about="City" ...>
  <rdfs:subPropertyOf rdf:resource="#Publisher"/>
  </rdfs:Property>
  <rdfs:Class rdf:about="Libro"
    rdfs:label="Libro" >
  </rdfs:Class>
  <rdfs:Property rdf:about="Autor" rdfs:label="Autor"
    rdfs:comment="Autor de un libro inglés".../>
  <rdfs:Property rdf:about="Original" ...>
  <rdfs:domain rdf:resource="#Libro"/>
  <rdfs:range rdf:resource="#Book"/>
  </rdfs:Property>
</rdf:RDF>
```

Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

73

RDF Schema. Ejemplo de uso

```
<rdf:Description rdf:about="http://ejemplo/isbn/000651409X">
  <Title>The Glass Palace</Title>
  <Year>2000</Year>
  <Publisher>
    <rdf:Description>
      <P_name>Harper Collins</P_name>
      <City>London</City>
    </rdf:Description>
  </Publisher>
  <Author>
    <rdf:Description>
      <foaf:name>Ghosh, Amitav</foaf:name>
      <Homepage rdf:resource="http://www.amitavghosh.com"/>
    </rdf:Description>
  </Author>
</rdf:Description>
```

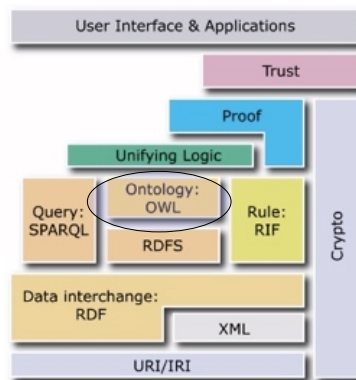
Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

74

RDF Schema. Otro ejemplo

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/">
  <foaf:Person rdf:nodeID="A12">
    <foaf:name>Maria Jose Suarez Cabal</foaf:name>
    <foaf:mbox rdf:resource="mailto:cabal@uniovi.es" />
    <foaf:title>Dra</foaf:title>
    <foaf:givenname>Maria Jose</foaf:givenname>
    <foaf:family_name>Suarez Cabal</foaf:family_name>
    <foaf:homepage rdf:resource="http://directo.uniovi.es/cabalu" />
    <foaf:nick>cabal</foaf:nick>
    <foaf:workplaceHomepage rdf:resource="www.uniovi.es"/>
    <foaf:workInfoHomepage rdf:resource="Universidad de Oviedo"/>
    <foaf:knows>
      <foaf:Person> <foaf:name>Enrique de la Cal</foaf:name> </foaf:Person>
    </foaf:knows>
    <foaf:knows>
      <foaf:Person> <foaf:name>Otra persona</foaf:name> </foaf:Person>
    </foaf:knows>
    <foaf:knows><foaf:Person rdf:nodeID="A12"> </foaf:Person></foaf:knows>
  </foaf:Person>
</rdf:RDF>
```



Contenido Semántico. Web Semántica

ONTOLOGÍAS. OWL

OWL. Lenguaje de ontología web

<http://www.w3.org/TR/owl-ref/>

- *Ontología*: Ciencia de descripción de clases de entidades en el mundo y cómo se relacionan
- OWL: Extiende RDF y RDF Schema para expresar relaciones complejas entre las clases RDFs y más precisión en restricciones de clases y propiedades:
 - Relaciones entre clases
 - Expresar cardinalidad
 - Expresar igualdad
 - Enriquecer el tipado de propiedades
 - Establecer características de propiedades, como simetrías
 - Definir clases enumeradas

Tipos de entidades en OWL

- Instancia: Objetos del dominio
- Propiedad: Enlace de objetos mediante relaciones binarias
- Clase: Conjuntos de objetos con características comunes

Clases (I)

- Pueden ser subclases
- Pueden definirse enumerando el contenido

```
<owl:Class rdf:ID="Currency">  
<owl:oneOf rdf:parseType="Collection">  
  <owl:Thing rdf:ID="£"/>  
  <owl:Thing rdf:ID="€"/>  
  <owl:Thing rdf:ID="$"/>  
</owl:oneOf>  
</owl:Class>
```

- Puede ser unión (owl:unionOf), intersección (owl:intersectOf), complemento (owl:complementOf), de otras clases

```
<owl:Class rdf:ID="Literature">  
<owl:unionOf rdf:parseType="Collection">  
  <owl:Class rdf:about="#Novel"/>  
  <owl:Class rdf:about="#Short_Story"/>  
  <owl:Class rdf:about="#Poetry"/>  
</owl:unionOf>  
</owl:Class>
```

Clases (II). Restricciones de propiedades

- Pueden establecer restricciones en sus **valores**:

owl:allValuesFrom

owl:someValuesFrom

owl:hasValue

owl:hasSelf

```
<owl:Restriction>  
  <owl:onProperty rdf:resource="#hasParent" />  
  <owl:allValuesFrom rdf:resource="#Human" />  
</owl:Restriction>
```

- Pueden establecerse restricciones de **cardinalidad**:

owl:maxCardinality

owl:minCardinality

owl:cardinality

```
<owl:Restriction>  
  <owl:onProperty rdf:resource="#hasParent" />  
  <owl:cardinality  
  rdf:datatype="xsd:nonNegativeInteger">2</owl:cardinality>  
</owl:Restriction>
```


Establecer equivalencias

■ Para clases:

- owl:equivalentClass: clases con los mismos individuos
- owl:disjointWith: sin individuos en común

■ Para propiedades:

- owl:equivalentProperty
- owl:propertyDisjointWith



■ Para instancias:

- owl:sameAs: URIs que se refieren a mismo concepto
- owl:differentFrom: negación de owl:sameAs
- owl:allDifferent: todas son diferentes

Otras propiedades

■ Pueden establecerse las definidas en RDFs (rdfs:subPropertyOf, rdfs:domain, rdfs:range)

■ Otras:

- owl:inverseOf

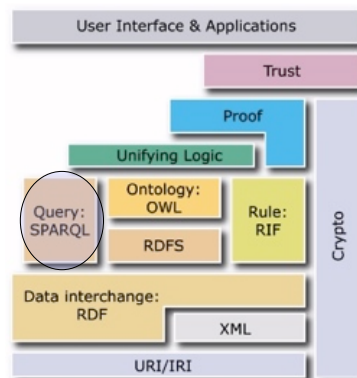
```
<owl:ObjectProperty rdf:ID="hasChild">
  <owl:inverseOf rdf:resource="#hasParent"/>
</owl:ObjectProperty>
```
- owl:FunctionalProperty

```
<owl:ObjectProperty rdf:ID="husband">
  <rdfs:domain rdf:resource="#Woman" />
  <rdfs:range rdf:resource="#Man" />
</owl:ObjectProperty>

<owl:FunctionalProperty rdf:about="#husband" />
```
- owl:InverseFunctionalProperty
- owl:TransitiveProperty
- owl:SymmetricProperty

Ejemplos de ontologías web

- eClassOwl: ontología para productos y servicios.
<http://www.heppnetz.de/projects/eclassowl/>
- Ontología del National Cancer Institute: más de 100.000 clases y utilizado en más de 10 proyectos.
<http://bioportal.bioontology.org/ontologies/NCIT>
- Open Biomedical Ontologies Foundry: colección de ontologías biomédicas. <http://www.obofoundry.org>
- SWRC (Semantic Web for Research Communities): ontología para comunidades de investigación: personas, equipos, publicaciones, relaciones,... <http://ontoware.org/swrc/>
- ...

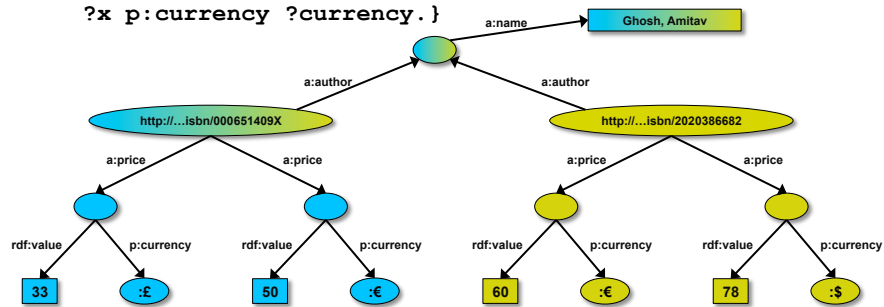


Contenido Semántico. Web Semántica

SPARQL

SPARQL. Lenguaje de consulta dentro de la Web Semántica

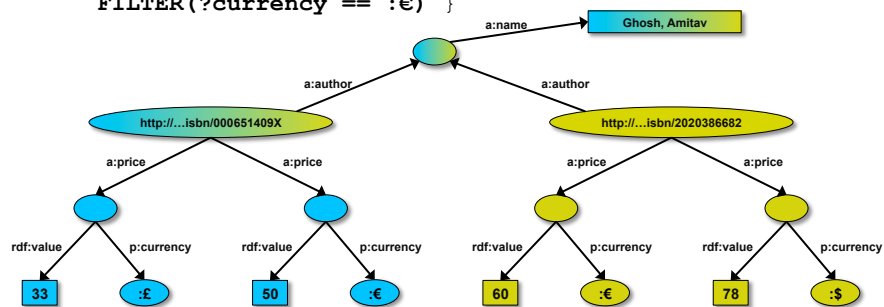
```
SELECT ?isbn ?price ?currency
WHERE { ?isbn a:price ?x.
        ?x rdf:value ?price.
        ?x p:currency ?currency. }
```



Devuelve: [<...409X>,33,:£], [<...409X>,50,:€], [<...6682>,60,:€], [<...6682>,78,:\$]

SPARQL. Filter

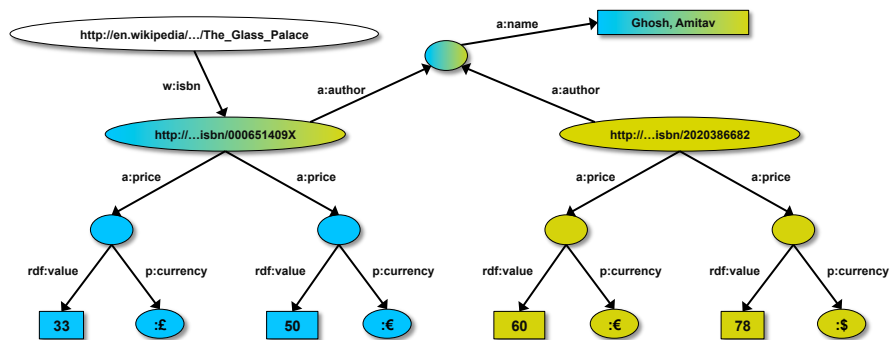
```
SELECT ?isbn ?price ?currency
WHERE { ?isbn a:price ?x.
        ?x rdf:value ?price.
        ?x p:currency ?currency.
        FILTER(?currency == :€) }
```



Devuelve: [<...409X>,50,:€], [<...6682>,60,:€]

SPARQL. "Optional"

```
SELECT ?isbn ?price ?currency ?wiki
WHERE { ?isbn a:price ?x.
        ?x rdf:value ?price.
        ?x p:currency ?currency.
        OPTIONAL ?wiki w:isbn ?isbn. }
```



Devuelve: [[<..09X>,33,:£,<...Palace>], [<..6682>,78,:\$,]]

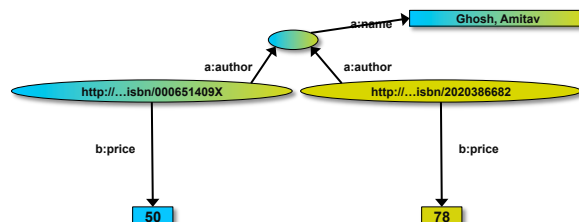
Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

87

SPARQL. Construcción de un nuevo grafo

```
CONSTRUCT { ?isbn b:price ?price.
             ?isbn a:author ?y.
             ?y a:name ?name . }
WHERE { ?isbn a:price ?x. ?x rdf:value ?price.
        ?x p:currency ?currency. ?isbn a:author ?y.
        ?y a:name ?name . FILTER(?currency == :€) }
```



Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

88

Otras características de SPARQL

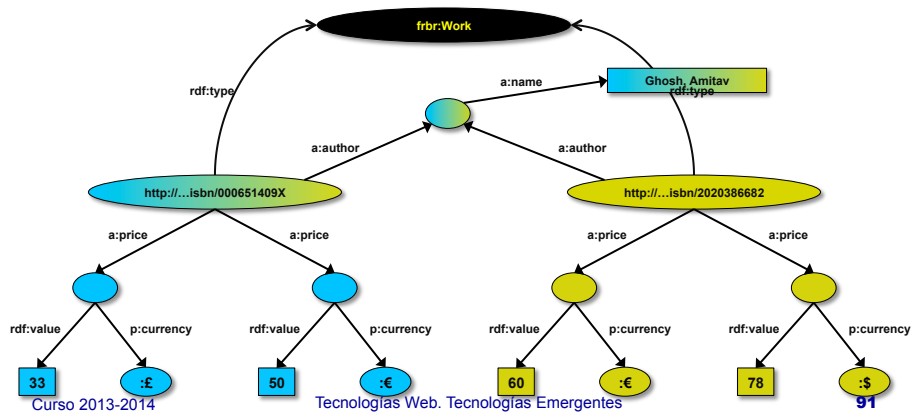
- Limitar el número de resultados devueltos
- Eliminar duplicados
- Ordenar resultados
- Uso de varias fuentes de datos (vía URIs) en la consulta
- Construir un grafo combinando varios resultados
- Uso de patrones para el matching
- Agregación de resultados (min, max, avg, etc.)

SPARQL 1.1 Update

- SPARQL CONSTRUCT devuelve un nuevo grafo, el original permanece sin cambios
- SPARQL 1.1 Update MODIFICA el grafo (dataset) original

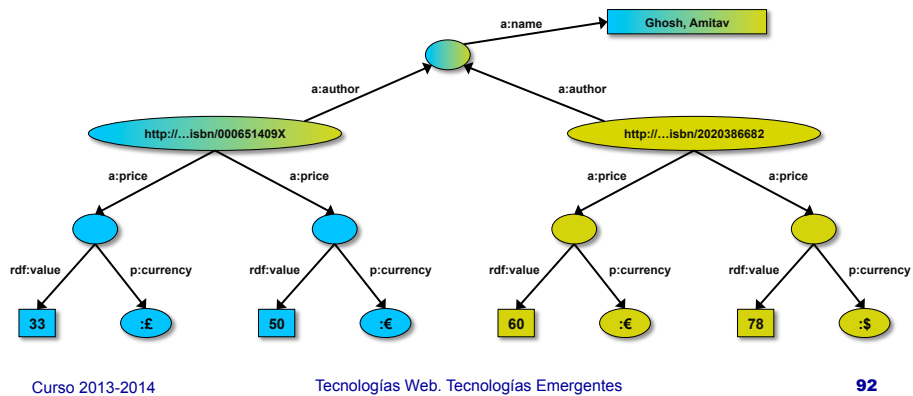
SPARQL 1.1 Update. Insert

```
INSERT {?isbn rdf:type frbr:Work}
WHERE {?isbn a:price ?x.
       ?x rdf:value ?price.
       ?x p:currency ?currency. }
```



SPARQL 1.1 Update. Delete

```
DELETE {?x p:currency ?currency}
WHERE {?isbn a:price ?x.
       ?x rdf:value ?price.
       ?x p:currency ?currency. }
```



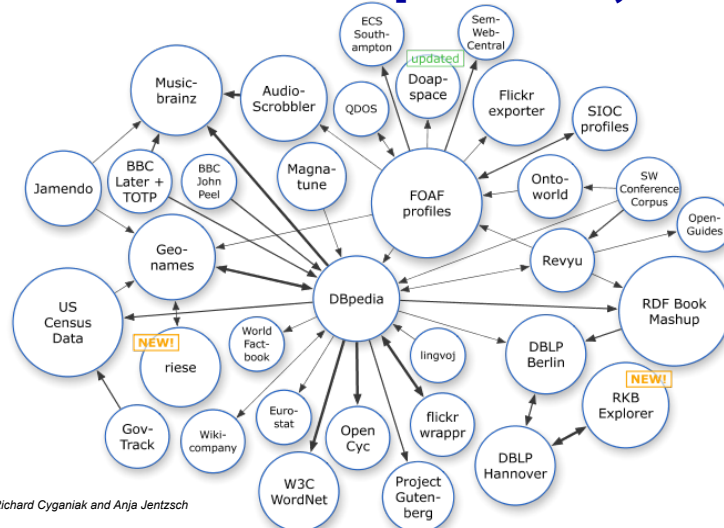
Contenido Semántico. Web Semántica

PROYECTO LINK DATA

Proyecto “Linked Data”

- Uso de la Web para conectar datos relacionados que no estaban previamente enlazados
- Núcleo que permite exponer, compartir y conectar datos, información y conocimiento en la web semántica utilizando URIs y RDF:
 - Referencias estables de recursos
 - Relaciones que pueden ser reutilizadas
- <http://linkeddata.org>

La “nube” Link OpenData, Marzo 2008



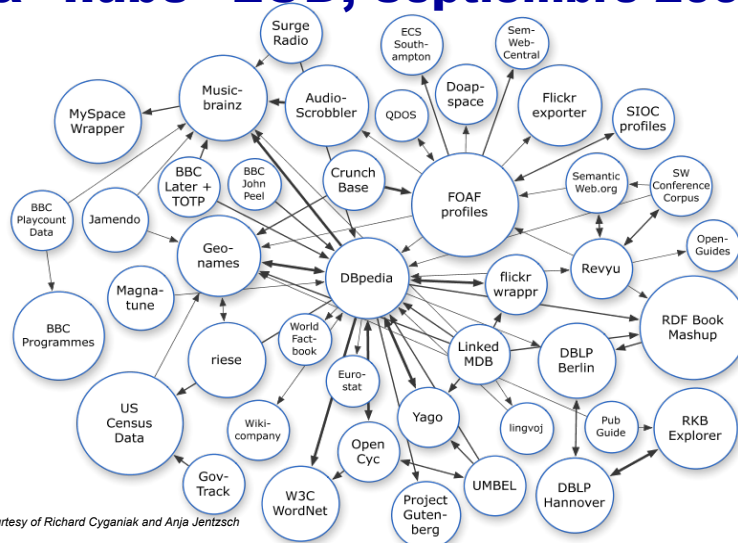
Courtesy of Richard Cyganiak and Anja Jentzsch

Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

95

La “nube” LOD, Septiembre 2008



Courtesy of Richard Cyganiak and Anja Jentzsch

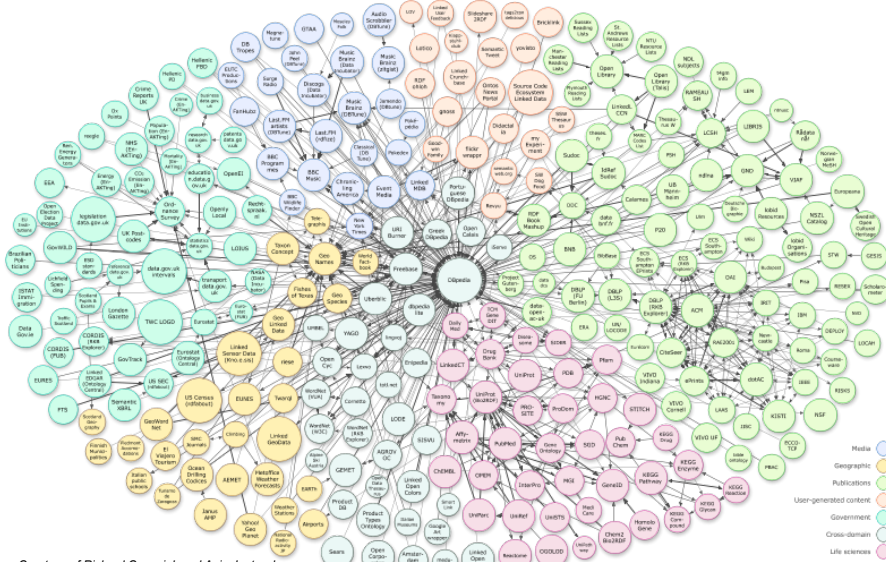
Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

As of September 2008

96

La “nube” LOD, Septiembre 2011



Courtesy of Richard Cyganiak and Anja Jentzsch

Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

As of September 2011

99

Ejemplo. DBpedia

- DBpedia permite:
 - Extraer información estructurada (“infobox”) de Wikipedia
 - Proporciona un “endpoint” para consultas del dataset
 - Permite interconectar el dataset de DBpedia con otros datasets de la Web

- <http://dbpedia.org/About>



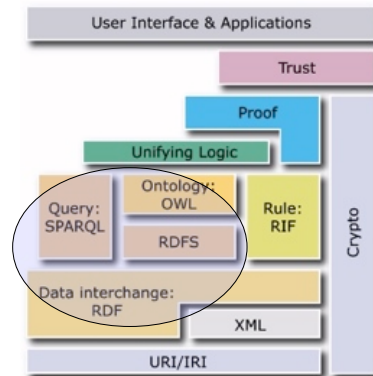
UNIVERSITÄT LEIPZIG



Curso 2013-2014

Tecnologías Web. Tecnologías Emergentes

100



Contenido Semántico. Web Semántica

ALGUNAS CONCLUSIONES

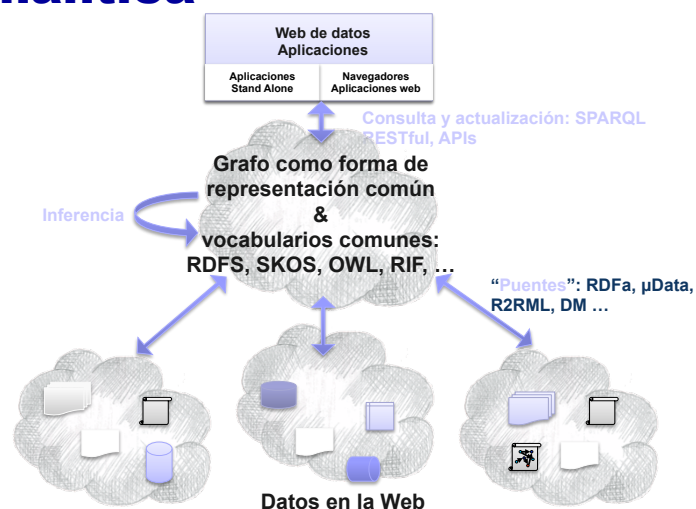
Realmente, tenemos una herramienta muy poderosa...

- Podríamos añadir conocimiento extra a los datasets:
 - Clasificaciones de tipos de datos, información geográfica, etc.
- OJO!: Aquí es donde entran las ontologías, reglas extras, etc. (que pueden ser relativamente simples y pequeñas o muy complejas)
- COMPLEJIDAD

Problemas de la Web Semántica

- Parte de sus funciones están aún en desarrollo
- Desarrollo “libre” de etiquetas, ontologías, ... → diferentes etiquetas para describir lo mismo
- Problema de la identidad: por ejemplo una URL de una empresa ¿representa a una empresa o es una página web sobre ella? Es importante para inferir conocimiento → necesidad de establecer fuentes de confianza
- Necesidad de establecer ontologías para dar significado → requiere trabajo: evolución de los sitios ya existentes adecuándolos a la web semántica y mantenimiento, evolución de las ontologías
- Representación en grafos

La propuesta de la Web Semántica



La Web Semántica, ¿tendrá éxito?

- La abstracción tendrá éxito porque...
 - ... la representación es independiente de la estructura
 - ... un cambio en el esquema (base de datos, xhtml. ...) no afecta al conjunto
 - ... se puede añadir fácilmente nuevos datos, conexiones, etc.
- Además, mediante URIs se puede enlazar cualquier dato con cualquier otro
- ✓ La Web Semántica proporciona las tecnologías que hacen que tal integración sea posible