

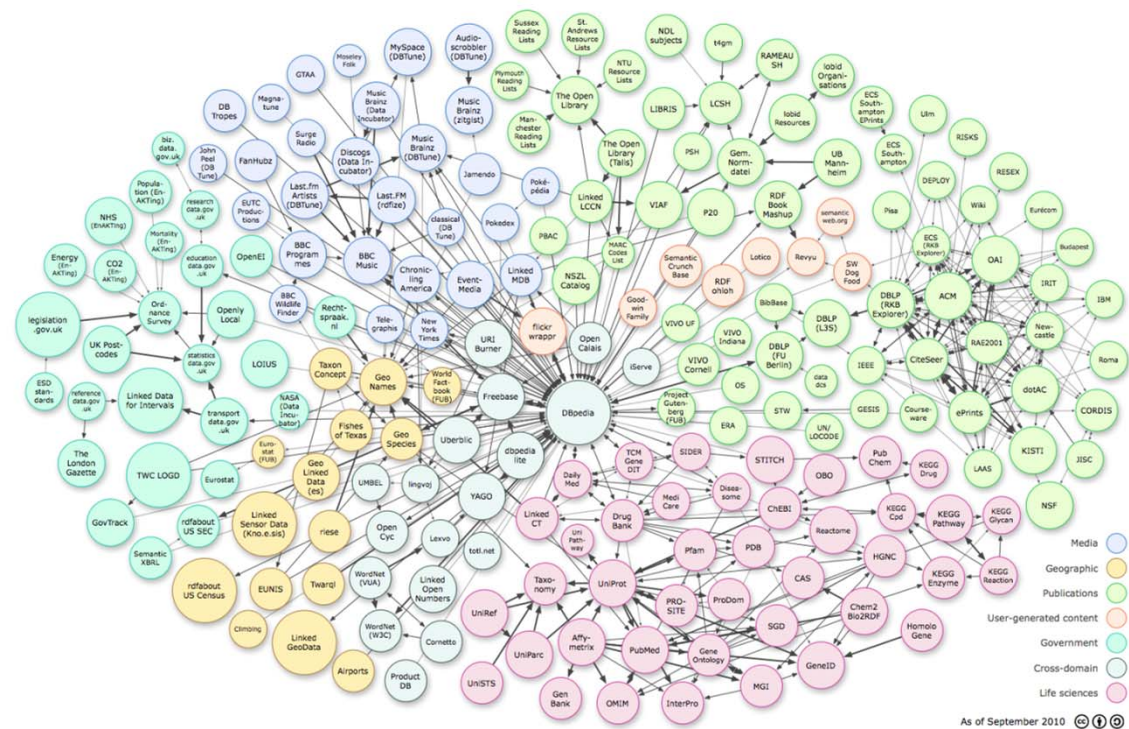
Introducción a la Web Semántica

M. Mercedes Martínez

Dep. Informática (U. Valladolid, España)

Linked Data

- Método para publicar datos estructurados, de modo que estén **relacionados** y sean **más útiles**.



Linked Data

- Principios (Berners-Lee, 2006):
 1. Use [URIs](#) to identify things.
 2. Use [HTTP](#) URIs so that these things can be referred to and looked up ("[dereferenced](#)") by people and [user agents](#).
 3. Provide useful information about the thing when its URI is dereferenced, using standard formats such as [RDF/XML](#).
 4. Include links to other, related URIs in the exposed data to improve discovery of other related information on the Web.

Linked Data

- **Principios (según los redefinió Berners-Lee en 2009):**
 1. Todas las entidades conceptuales deben tener nombres (URI) cuyo inicio sea HTTP
 2. Los datos que se reciben cuando se resuelve una URI deben ser datos en un formato estándar, como RDF/XML
 - Aunque sea el más utilizado, RDF/XML no es obligatorio
 3. Deben contener relaciones con otras cosas, las cuales a su vez estarán referenciadas con una URI, como se indica en 1

URI *dereferenciable*

- Use HTTP URIs so that these things can be referred to and looked up ("dereferenced") by people and user agents:
un cliente http puede obtener, usando la URI que identifica un recurso, una descripción del recurso (RDF)
 - El servidor web debe soportar el tipo MIME application/rdf+xml
 - El servidor web debe soportar el redireccionamiento 303 (reenviar al cliente al sitio donde se encuentra el documento que satisface los requisitos de su petición: rdf+xml, html+xml, ...)
 - Aplicable a las URI 303

Linked Data

- Se apoya en dos tecnologías web básicas:
 - Uniform Resource Identifiers (URIs)
 - HyperText Transfer Protocol (HTTP)
- Se complementa con una tecnología crítica para la *Web of Data*: RDF
- Los enlaces RDF se expresan como tripletas RDF
 - El sujeto es una URI en el espacio de nombres de un conjunto de datos, y el objeto es otra URI, de otro conjunto de datos

Comparación con la Web tradicional

- En la Web, cualquiera puede publicar, cuando quiera.
 - También es cierto en la Web de Datos, pero aquí se publica RDF en lugar de HTML.
- Para acceder a la Web se usan navegadores web.
 - En la Web de Datos se usan navegadores específicos.
- En la Web las cosas están enlazadas.
 - Son enlaces 'sin tipo'.
 - En la Web de Datos están enlazadas por enlaces 'con tipo'.

Comparación con la Web tradicional

- En la Web se pueden obtener datos estructurados, que consumen aplicaciones web.
 - En la Web las aplicaciones acceden a través de APIs.
 - En la Web de Datos las aplicaciones aprovechan la estructura y semántica de los datos.

Pasos que debemos seguir para publicar RDF

1. Asignar URIs para identificar todos los recursos de los que vayamos a decir algo
2. Escoger los vocabularios para nuestros datos
 1. Mejor si reutilizamos
3. Generar las sentencias RDF que describen nuestros recursos
4. Crear enlaces RDF a otros conjuntos de datos RDF
5. Servir las tripletas RDF en la Web

Recursos de información y otros recursos

- **Recurso de información:**
 - Se identifica con una URI
 - Se devuelve la correspondiente representación cuando se solicita el acceso a su URI
- **Otros**
 - También deben identificarse con URI
 - **No debe utilizarse la misma URI para el recurso y para algo que lo describe** (ej., una persona y su página web)
 - Debe devolverse una descripción RDF como retorno

Tipos de URI

■ 303

- Permiten *negociar el contenido* entre el servidor web y el cliente que accede
 - El servidor puede redirigir al cliente hacia otra URI (*303 See Other*)
 - Tipos de contenido:
 - text/plain
 - **application/rdf+xml**
 - **text/rdf+n3**

<http://www.postgrado.uva/masterinf/WebSem>

■ Hash

- Direccionan *fragmentos* dentro de un recurso

<http://www.postgrado.uva/masterinf.rdf#WebSem>

Cuándo usar cada tipo de URI

- **303**
 - Los recursos forman parte de conjuntos de datos muy grandes
- **Hash**
 - Los conjuntos de datos RDF son pequeños. Ej.: vocabularios.
 - El contenido RDF se embebe dentro de páginas HTML. Ej: RDFa.

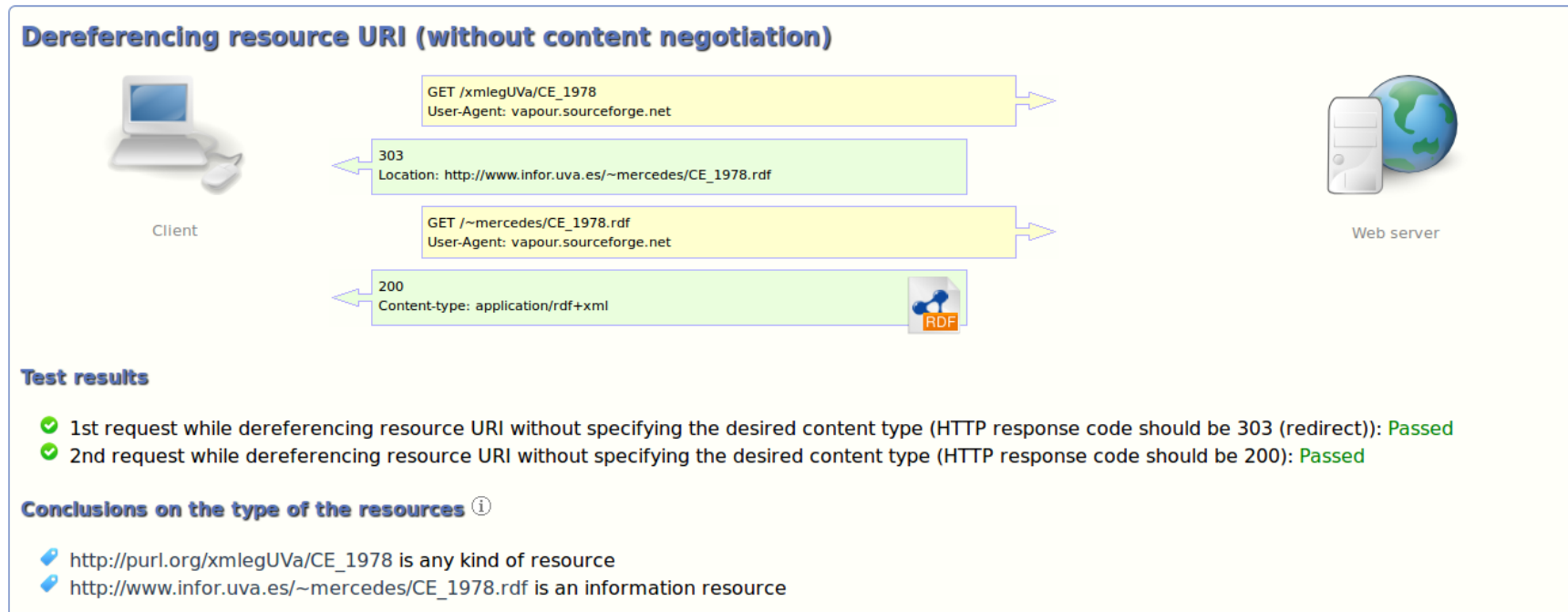
Diálogo entre un cliente y un servidor: Hash URI (I)



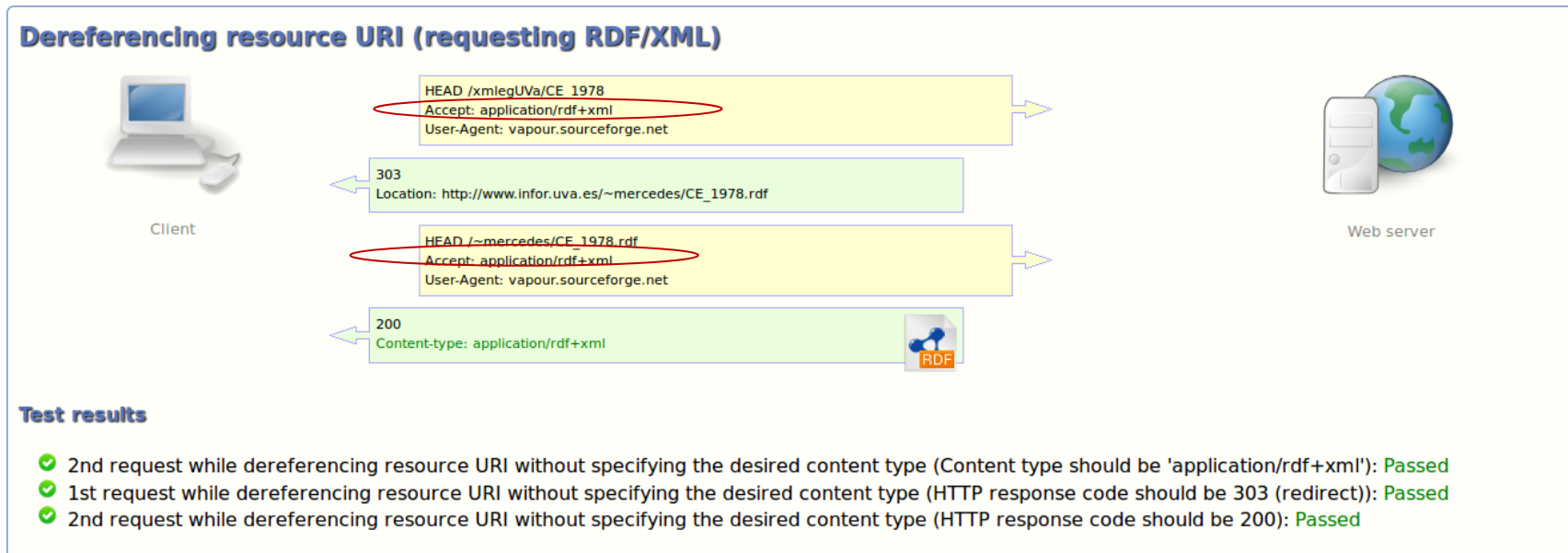
Diálogo entre un cliente y un servidor: Hash URI (II)



Diálogo entre un cliente y un servidor: 303 URI (I)



Diálogo entre un cliente y un servidor: 303 URI (II)



Pautas habituales para las URI

- El recurso

<http://dbpedia.org/resource/Spain>

- La página web que lo describe

<http://dbpedia.org/page/Spain>

- Los datos RDF que lo describen

<http://dbpedia.org/data/Spain>

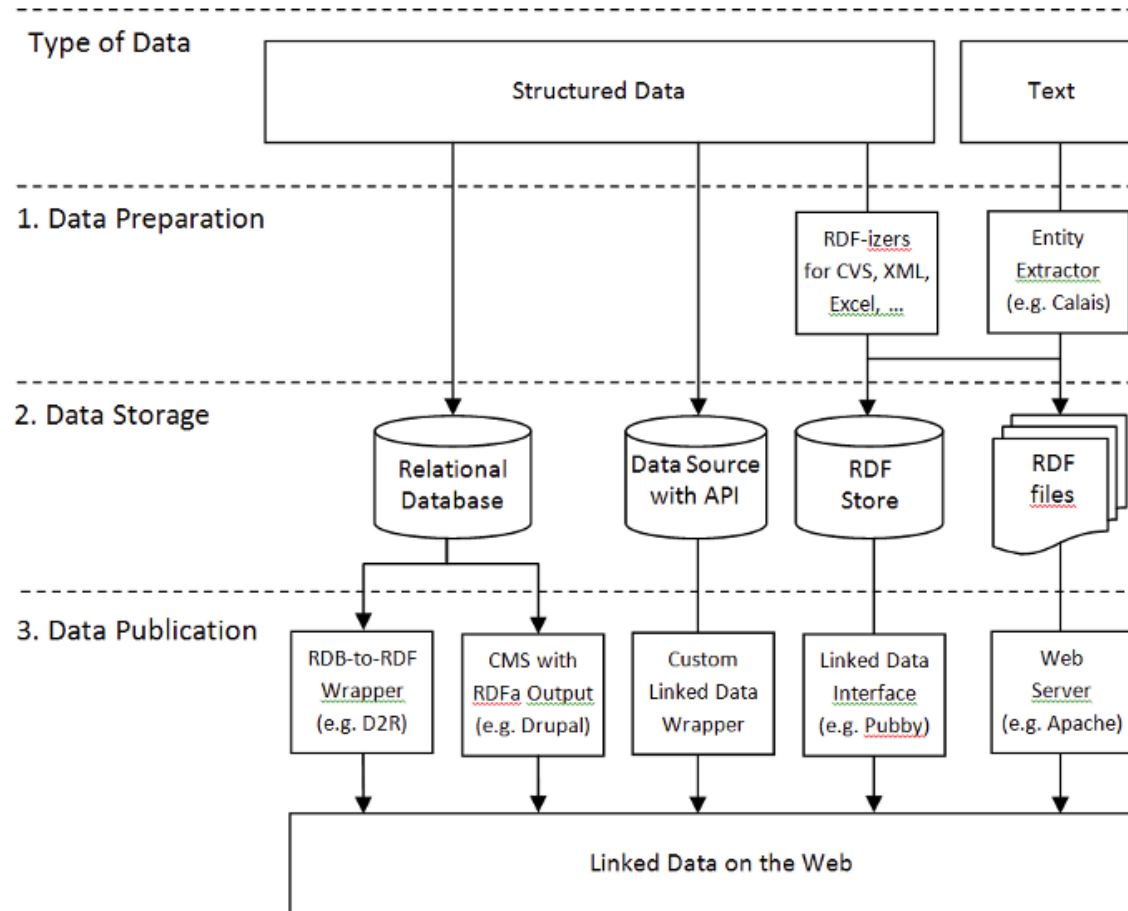
Creando enlaces

- **Modos más habituales:**
 - Expresando relaciones:
 - `foaf:interest` (interés de una persona)
 - `rdfs:seeAlso` (recursos relacionados por cualquier razón)
 - Expresando identidades:
 - `owl:sameAs` (dos URI se refieren al mismo recurso)
 - Expresando la clasificación en base a vocabularios (re)utilizados:
 - `rdf:type` (un recurso es de un cierto tipo)

Metadatos

- **Describen los grafos RDF**
 - *void* (*Vocabulary of Interlinked Datasets*)
 - Estándar para describir conjuntos de datos en el contexto de Linked Data
 - Son datos RDF
- **¿Qué incluyen?**
 - Información sobre el origen (fuente) de los datos
 - *Dublin Core*: dc:creator, dc:publisher, dc:date, ...
 - Licencias asociadas
 - *Creative Commons*, ...

Patrones de publicación



[Heath&Bizer, 2011]

Patrones de publicación (II)

1. Servir los datos como ficheros RDF/XML estáticos
2. Servir los datos como RDF embebido en ficheros HTML
 - RDFa
3. Servir RDF y HTML con scripts a medida, en el servidor
4. Servir datos RDF obtenidos de bases de datos relacionales
 - *A Direct Mapping of Relational Data to RDF. W3C Recommendation 27 September 2012.*
5. Servir datos RDF almacenados en almacenes RDF (Triple Stores)
6. Crear un *wrapper* que accede a las aplicaciones a través de sus API y genera RDF
 - Amazon, Twitter, Freebase API, ...

Referencias

- [Heath & Bizer, 2011] “*Linked Data: Evolving the Web into a Global Data Space* (1st edition)”. Tom Heath and Christian Bizer. Morgan & Claypool. 2011. Accesible en <http://linkeddatabook.com/book>.
- [Yu, 2011] “*A Developers Guide to the Semantic Web*”. Liyang Yu. Springer Publishing Company. 2011.
- linkeddata.org