

Unos apuntes sobre temas avanzados

Yannis Dimitriadis, Pablo de la Fuente

Palencia, 22/02/2016, 10:00-14:00

Unos temas “avanzados” a tratar

- Aplicaciones y arquitecturas de BBDD
- El protocolo OAI-PMH
- Open and Linked Data

Aplicaciones de BBDD

- Interactuando **directamente** con el SGBD (MySQL) mediante **interfaz Web** (phpMyAdmin) o **ssh** (terminal remoto y línea de comandos)
- Escribiendo **sentencias SQL** o ejecutando **programas SQL** (scripts) con sentencias SQL
- Usando una **aplicación “compilada”** en la cual se han “embebido” sentencias SQL dentro de un programa escrito en un lenguaje genérico (Python, C, C++)

Aplicaciones de BBDD

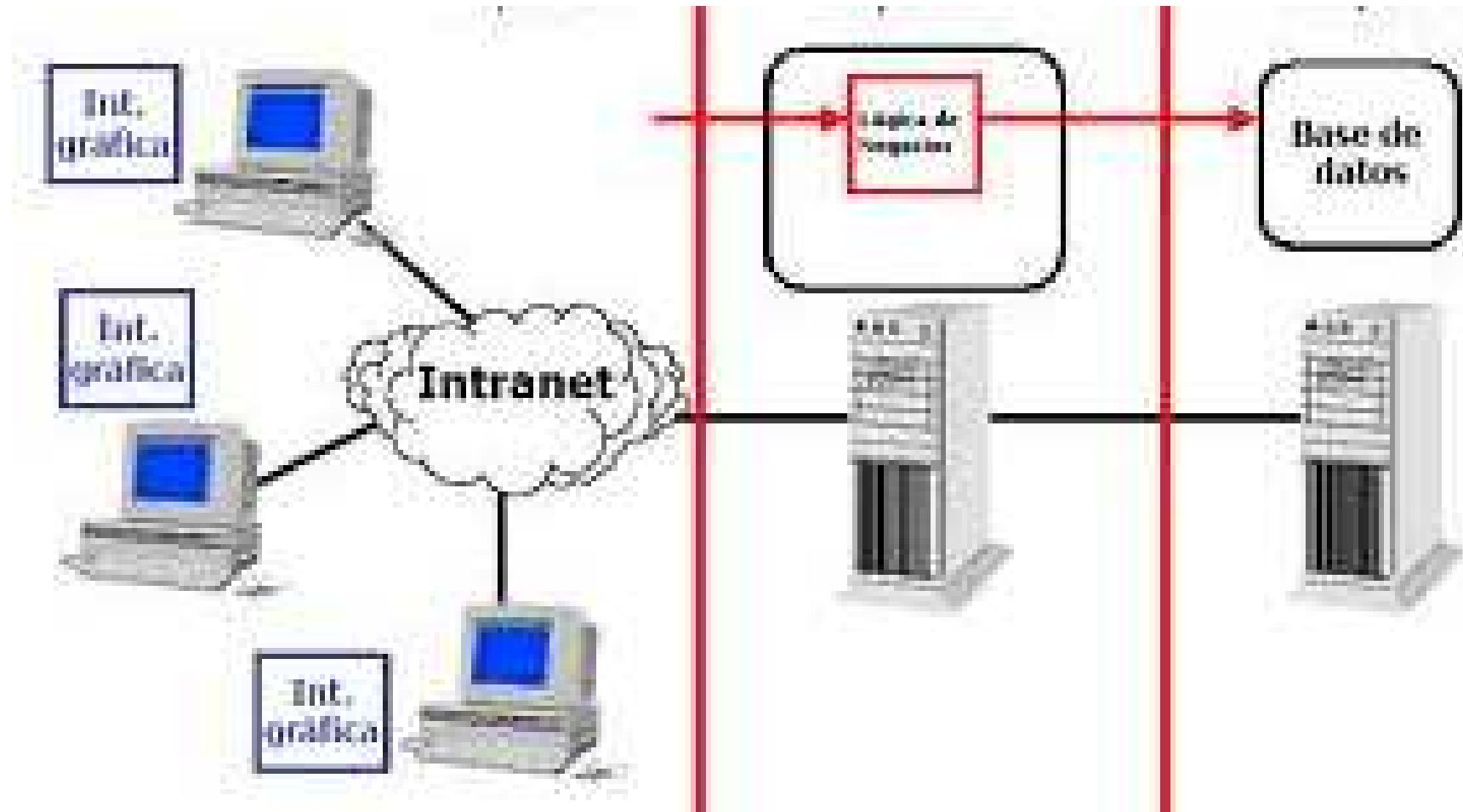
- Interactuando **directamente** con el SGBD (MySQL) mediante **interfaz Web** (phpMyAdmin) o **ssh** (terminal remoto y línea de comandos)
- Escribiendo **sentencias SQL** o ejecutando **programas SQL** (scripts) con sentencias SQL
- Usando una **aplicación “compilada”** en la cual se han “embebido” sentencias SQL dentro de un programa escrito en un lenguaje genérico (Python, C, C++)

Arquitecturas de BBDD

- Como usuarios podemos interactuar usando sistema local (PC propio) o mediante conexión con sistema remoto
- Típicamente se usan arquitecturas cliente/servidor de 3 capas
 - Cliente de usuario vs. servidores Web (Apache) y de BBDD (phpMyAdmin)
 - Capa de cliente (sistema local) vs. capa de interfaz Web (Front-end), lógica de negocio y de gestión de datos (Back-end)

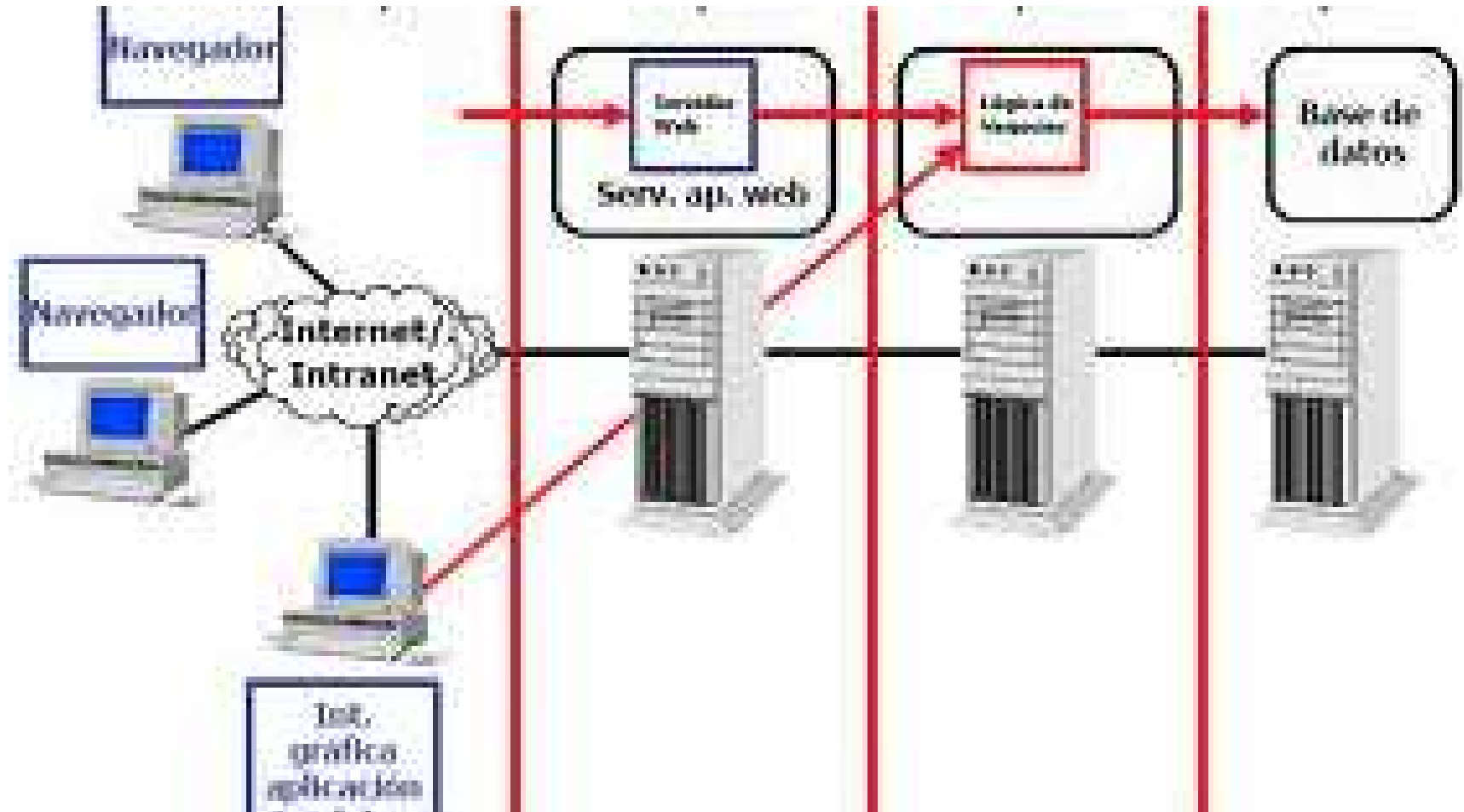
Arquitecturas de 3 capas

<https://karlamendozaaplicacionesdistribuidas.wordpress.com/2013/03/09/1-1-3-aplicaciones-de-23-y-n-capas/>



Arquitecturas de 4 capas

<https://karlamendozaaplicacionesdistribuidas.wordpress.com/2013/03/09/1-1-3-aplicaciones-de-23-y-n-capas/>



Protocolo OAI-PMH

La disponibilidad de numerosas recursos digitales que pueden estar en distintos formatos plantea el reto de la interoperabilidad. La idea básica es que alguien pueda acceder no sólo a la información sobre un objeto digital sino al contenido de dicho objeto, independientemente del sistema que lo contenga.

Una de las propuestas de estándares de interoperabilidad más extendida actualmente es la desarrollada por OAI (*Open Archives Initiative*) para facilitar la diseminación eficiente de contenidos en Internet.

Open se refiere al punto de vista de la arquitectura del sistema, definiendo interfaces que faciliten la disponibilidad de los contenidos de varios proveedores sin implicar la gratuidad ni el acceso ilimitado.

La OAI no define o prescribe ningún esquema para la gestión de derechos. Los temas relacionados con restricciones en el acceso y gestión de la propiedad intelectual son responsabilidad de los proveedores de datos.

Qué es un *Archivo Abierto*?. Cualquier sistema, basado en el WWW, que se puede acceder a través de la interfaz del Protocolo para Recolección de Meta-datos (PMH)., que cumple con las pautas dictadas por la OAI. No tiene implicaciones respecto a:

- Almacenamiento físico de los datos.
- Formatos internos de los datos y sus metadatos.
- Control de acceso al servidor

Recolección frente a federación

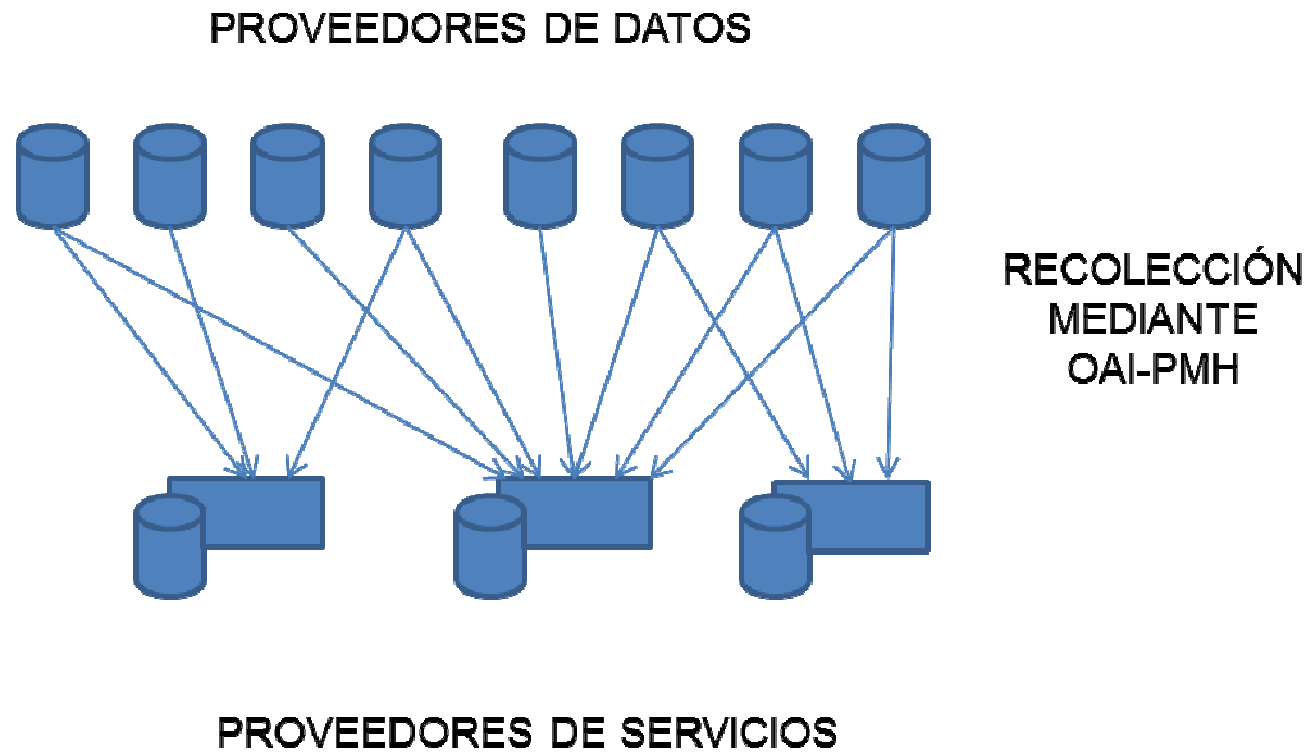
Meta-datos frente a los datos

Los datos describen objetos digitales o representaciones digitales de objetos. Los meta-datos contienen información acerca del objeto (e.ej. título, autor, etc.). OAI se centra en los meta-datos, con el acuerdo implícito que los meta-datos normalmente contienen enlaces útiles a la fuente de los objetos digitales.

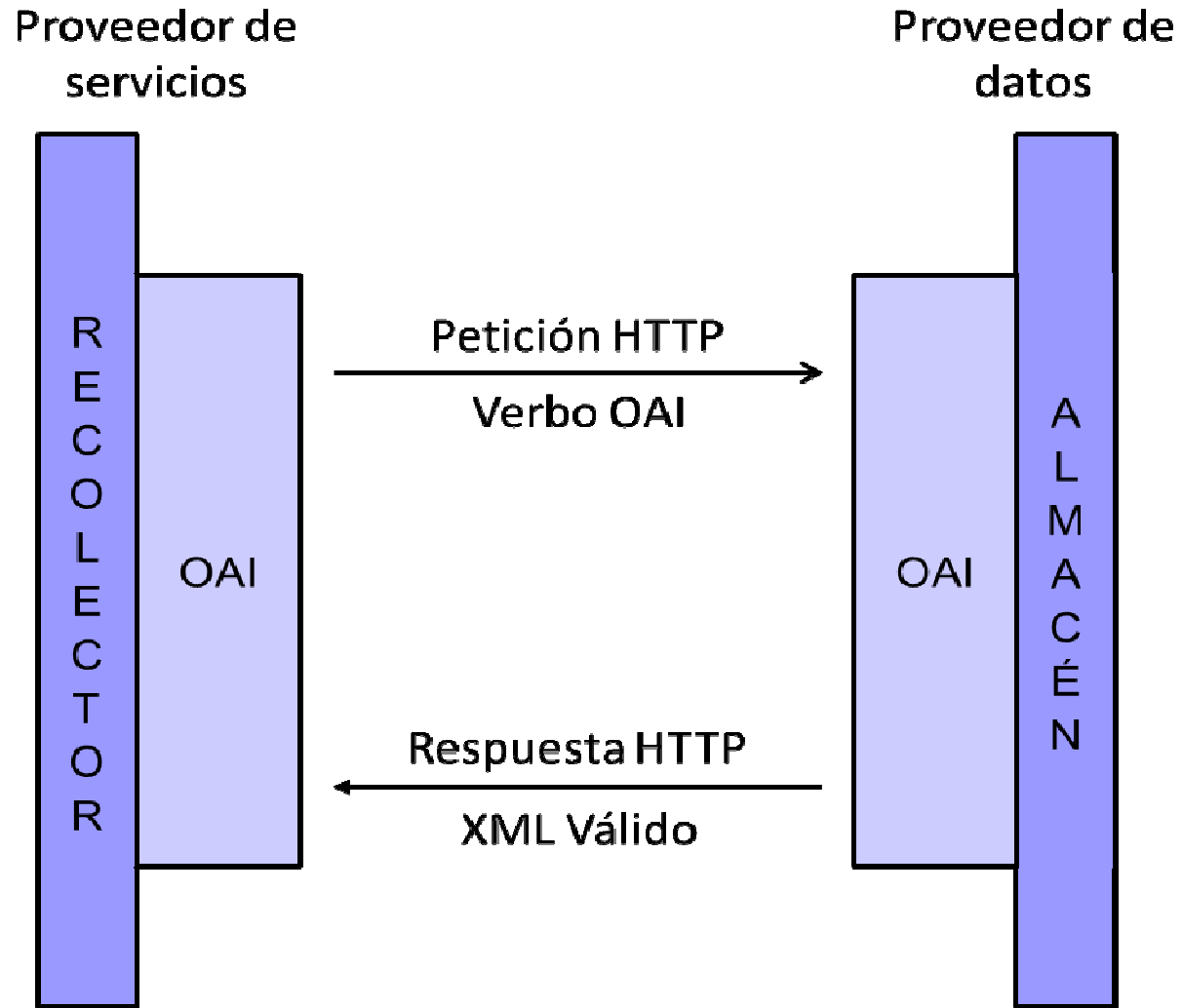
Proveedores de datos y servicios

Los proveedores de datos son entidades que poseen datos/meta-datos y que están dispuestos a compartirlos (interiormente o exteriormente) vía protocolos bien definidos de OAI (p.ej. servidores de bases de datos).

Protocolo OAI-PMH



Protocolo OAI-PMH



Peticiones y respuestas en OAI

Ejemplo de protocolo OAI-PMH

<http://www.europeana.eu/portal/>

Linked Open Data

En informática, los **datos enlazados** o **datos vinculados** (a menudo capitalizados como Linked Data, en inglés) describe un método de publicación de datos estructurados para que puedan ser interconectados y más útiles. Se basa en tecnologías Web estándar, tales como HTTP, RDF y los URI, pero en vez de utilizarlos para servir páginas web para los lectores humanos, las extiende para compartir información de una manera que puede ser leída automáticamente por ordenadores. Esto permite que sean conectados y consultados datos de diferentes fuentes.

Tim Berners-Lee, director del Consorcio de la World Wide Web, acuñó el término en una nota de diseño que trataba de cuestiones relativas al proyecto de Web Semántica.

La evolución de la iniciativa LOD (datos abiertos enlazados) puede verse en:

<http://lod-cloud.net/>

Las ideas fundamentales para publicar datos abiertos están en:

<http://5stardata.info/es/>

RDF. Introducción

Resource Description Framework. RDF (<https://www.w3.org/TR/2004/REC-rdf-primer-20040210/>)

Lenguaje natural

Gramática

RDF

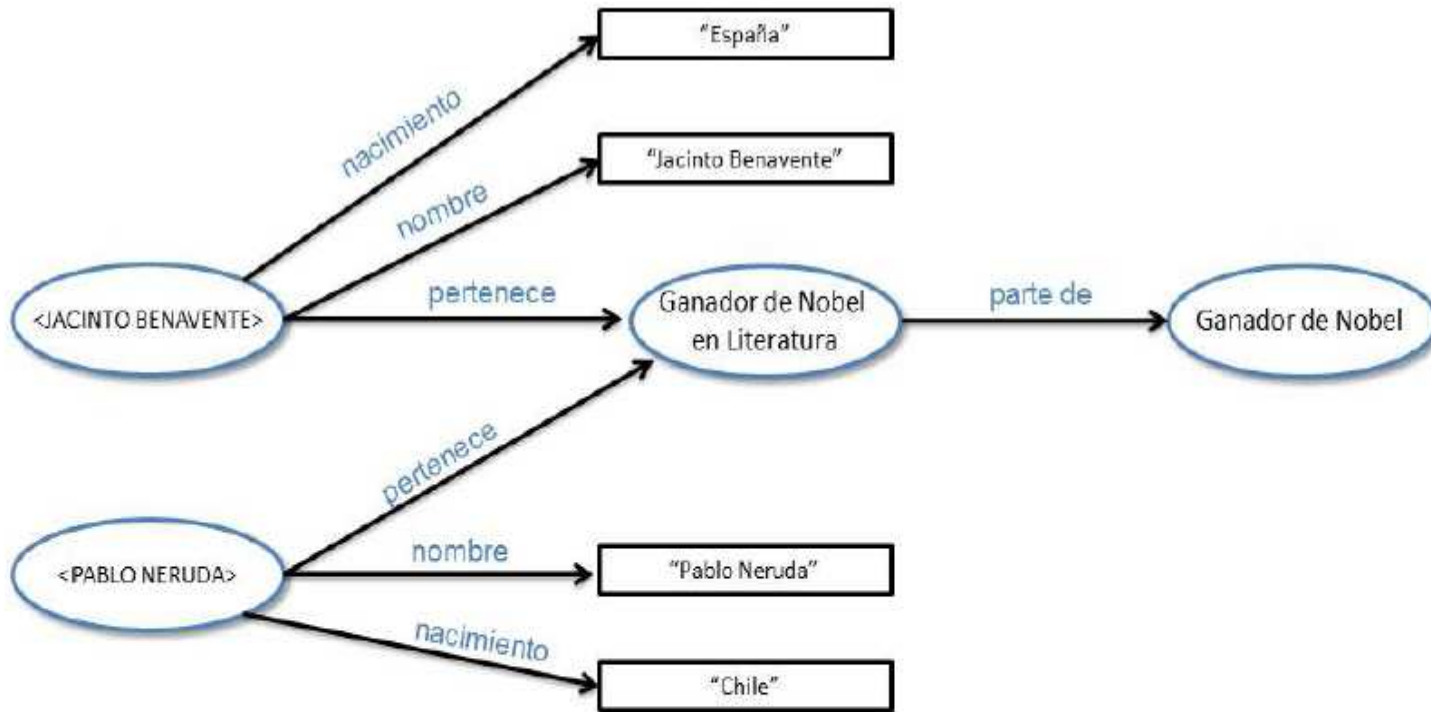
Jacinto Benavente nació en España
Pablo Neruda nació en Chile

Sujeto Predicado

Sujeto Propiedad Objeto

Nota en RDF se suele utilizar Sujeto Predicado Objeto (SPO)

RDF. Introducción



¿Algún problema?

- ¿"Nacimiento" es una fecha o un lugar?
- "Nacimiento" o "ha nacido", "birth", "birth place", ...
- ¿Cómo identifico unívocamente un recurso?
- ...

RDF. Introducción

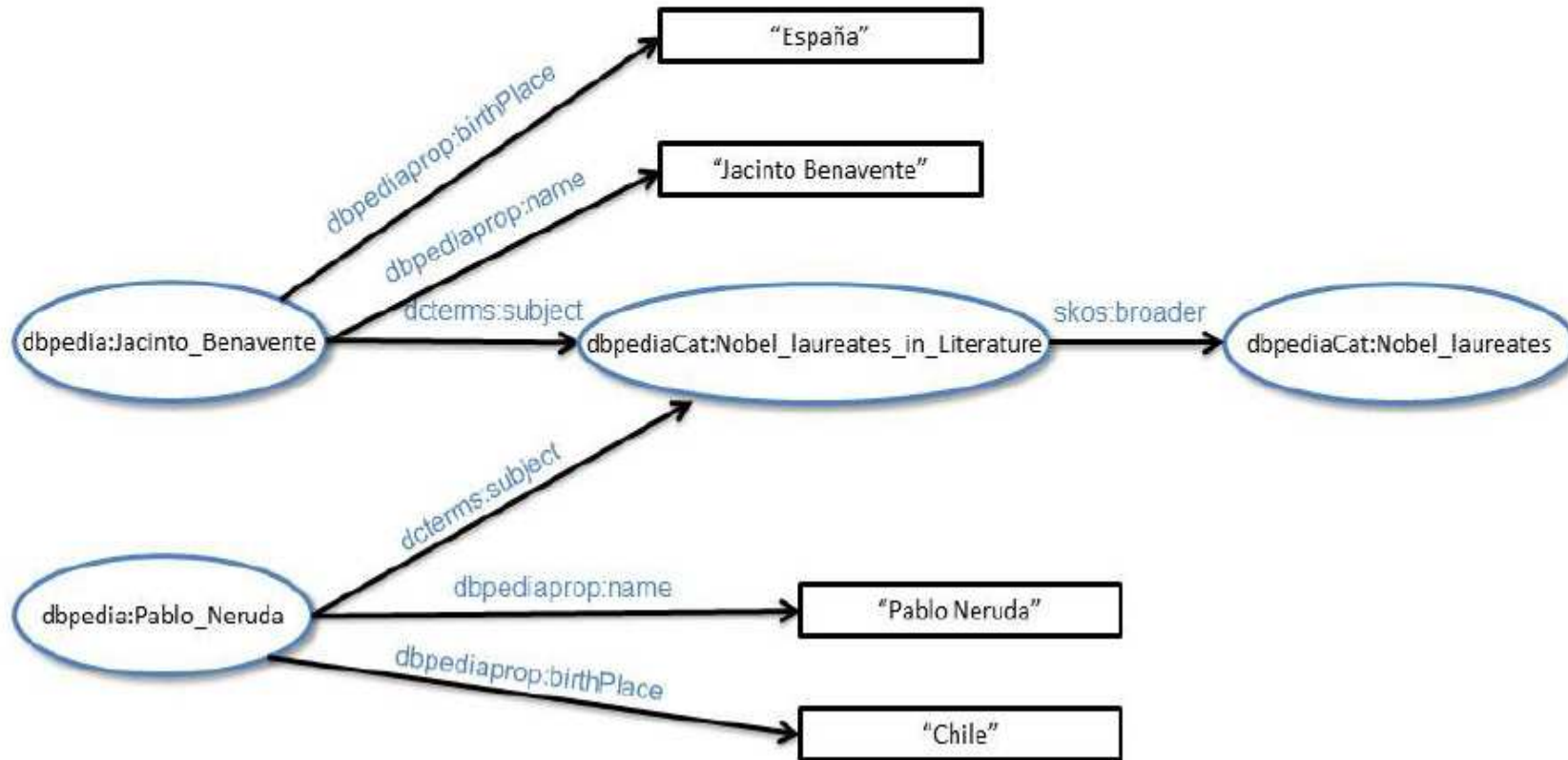
$$\frac{\text{Modelo E/R}}{\text{SQL}} = \frac{\text{Modelo RDF}}{\text{Lenguaje RDF}}$$



- **Recursos y Propiedades: URIs**
 - `<JACINTO BENAVENTE >` → `http://dbpedia.org/resource/Jacinto_Benavente`
 - `nombre` → `http://dbpedia.org/property/name`
- **Propiedades bien caracterizadas. Definición de Vocabularios RDF (RDFS)**
 - **dublin core:** libros, bibliografía
 - Ej: `http://purl.org/dc/elements/1.1/title`
 - **foaf:** personas, entidades y sus relaciones
 - **skos:** categorías, divisiones
 - ...

RDF. Introducción

```
@PREFIX dbpedia: <http://dbpedia.org/resource/>  
@PREFIX dbpedia:prop: <http://dbpedia.org/property/>  
@PREFIX dbpedia:category: <http://dbpedia.org/resource/Category:>  
@PREFIX skos: <http://www.w3.org/2004/02/skos/core#>  
@PREFIX dcterm: <http://purl.org/dc/terms/>
```



Nota: estas últimas transparencias están sacadas de un curso impartido por Javier D. Fernández García y Miguel Ángel Martínez Prieto en el Seminario Aprendiendo a nadar en el diluvio de datos en marzo de 2012.