

JIIDE 2014

V Jornadas Ibéricas de Infraestruturas de Dados Espaciais

Lisboa | 5-7 novembro 2014



IIAA
Grupo de Sistemas de
Información Avanzados
Universidad Zaragoza



JIIDE 2014

Una experiencia de interoperabilidad ISO-OGC-W3C-IETF

El proyecto Linked Map

Francisco J. Lopez-Pellicer
IAAA, Universidad de Zaragoza



IAAA
Grupo de Sistemas de
Información Avanzados
Universidad Zaragoza



El proyecto Linked Map en 2 minutos



LinkedMap

<http://linkedmap.unizar.es/>



El proyecto Linked Map en 2 minutos

- **PlanetData (2010-2014)**
 - <http://www.planet-data.eu/>
 - FP7 Network of Excellence
 - Estado del arte de la gestión de grandes volúmenes de datos del sector público
 - Tópicos: representación, asoramiento de su calidad, procedencia, privacidad, confianza, licencias ...
- **PlanetData Call 2 (2014)**
 - Llamada para proyectos de corta duración
 - Tópicos: privacidad, confianza, dinamicidad, ...
 - Linked Map fue financiado por PlanetData Call 2



El proyecto Linked Map en 2 minutos

- **Visión**
 - Resolver problemas de interoperabilidad con la ayuda de Linked Data
 - Foco en IG + VGI + evaluación de la calidad de la integración
- **Objetivos**
 - Convertir datasets geográficos oficiales y VGI a **RDF** usando **R2RML**, **GeoSPARQL** y **WKT**
 - Documentar su procedencia a nivel de objeto geográfico con **PROV**. El modelo se mapea a **RDF** con **PROV-O** y se serializa con **PROV-XML**
 - Desarrollar un **WMS** que ofrezca simultáneamente **Linked Data**



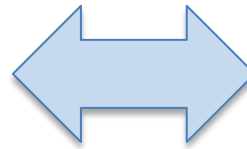
El proyecto Linked Map en 2 minutos

- **IAAA (Universidad de Zaragoza)**
 - <http://iaaa.unizar.es/>
 - Grupo de investigación, OGC, INSPIRE, ISO, IDE
- **GeoSpatiumLab (Zaragoza)**
 - <http://www.geoslab.com/>
 - PYME, enfocada en GI
- **CNIG**
 - <http://www.cnig.es/>
 - IGN, proveedor de datos oficiales



¿Por qué es una experiencia de interoperabilidad?

OGC[®]
Making location count.
www.opengeospatial.org



W3C[®]



I E T F[®]



IAAA
Grupo de Sistemas de
Información Avanzados
Universidad Zaragoza

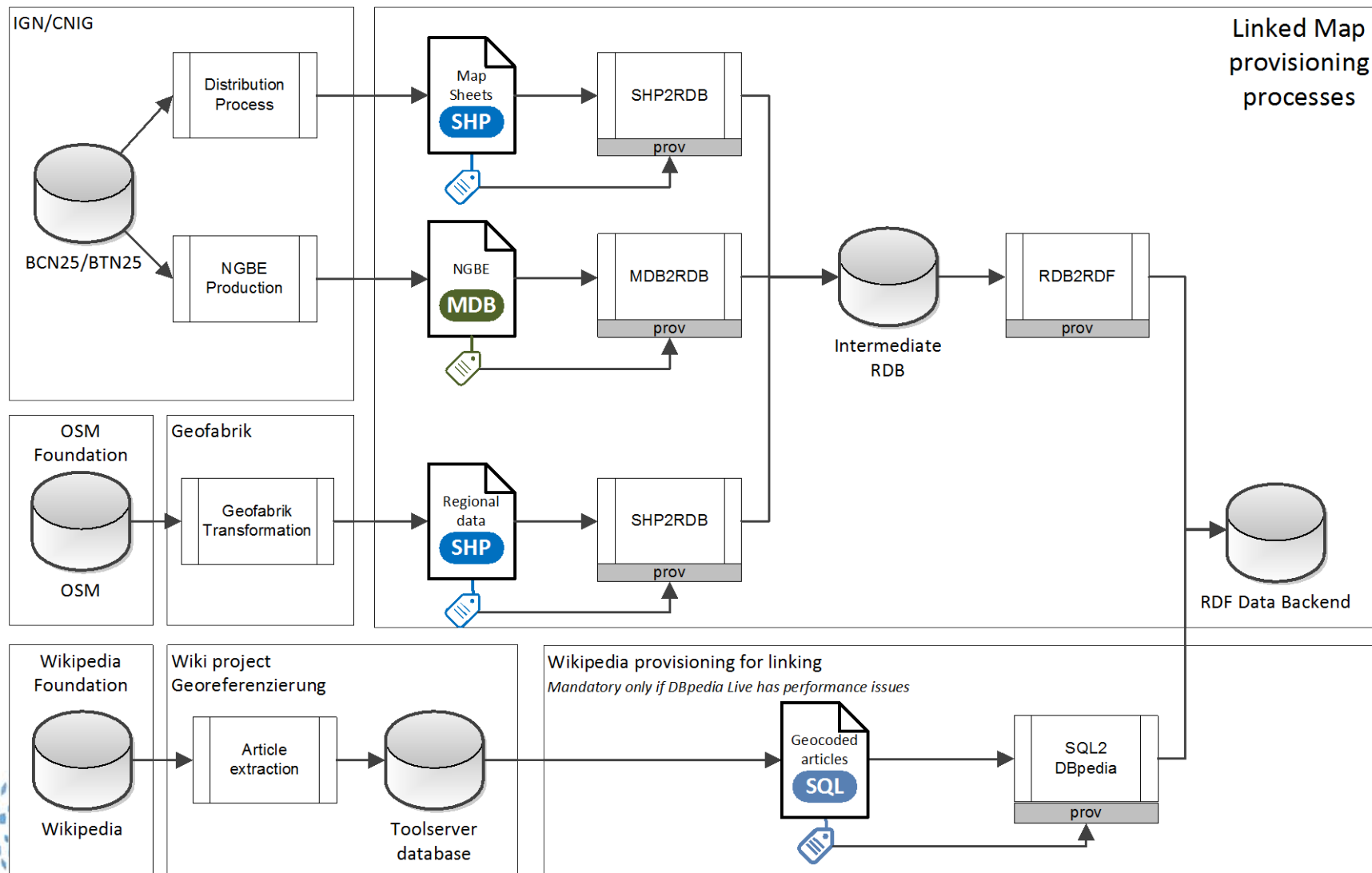


La relevancia de la estandarización

1. Convertir datasets geográficos oficiales y VGI a **W3C RDF** usando **W3C R2RML**, **OGC GeoSPARQL** y **OGC WKT (ISO 19125-1, ISO 13249)**
2. Documentar su linaje a nivel de objeto geográfico con **W3C PROV-DM** si es compatible con **ISO 19115**, **ISO 19115-1** e **ISO 19115-2**. El modelo se mapea a **W3C RDF** con **W3C PROV-O** y se serializa con **W3C PROV-XML**
3. Desarrollar un **OGC WMS (ISO 19128)** que ofrezca simultáneamente **Linked Data** de acuerdo con diversos **W3C Notes** y **W3C Recomendations** y aplicando **IETF RFC 5899 Web Linking**



1. Conversión de datasets geográficos



1. Conversión de datasets geográficos

- **BD original → Formato exportación → BD espacial → RDF**
 - Restringido a España península e Islas Baleares
 - BCN/BTN 25 (13M triples) y OpenStreetMap (35M triples)
 - GeoKettle + MorphRDB + Strabon
 - Linaje de grano fino: **PROV-DM** serializado como **PROV-XML** (SQL y RDF)
 - Geometría: **SF** serializado como **WKT** (SQL y RDF)
 - **GeoSPARQL**
- **RDF → RDF enriquecido**
 - Silk
 - Simple: nombre, geometría, tipos
 - Enlaces y calidad del enlace almacenado como RDF
 - 3M triples



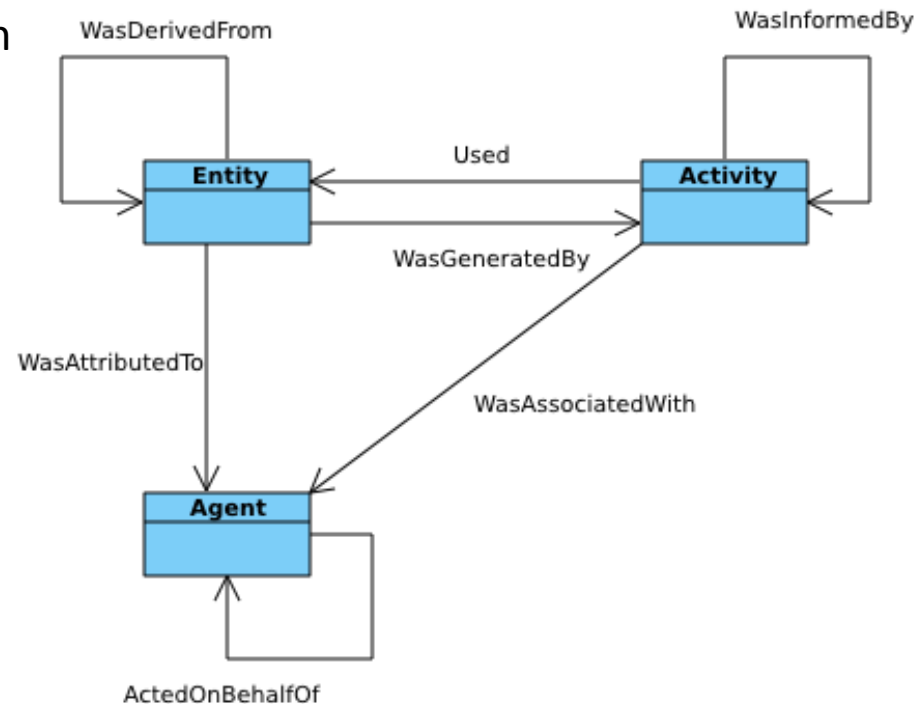
2. Gestión del linaje con PROV

- **W3C PROV**

- Familia de recomendaciones y notas de W3C
- <http://www.w3.org/TR/prov-overview/>
- Objetivo: publicación e intercambio de información de procedencia en la Web
- Serializaciones en RDF, XML y notación
- Reciente (2013)

- **PROV DM**

- <http://www.w3.org/TR/prov-dm/>
- Tipos: Entity, Activity, Agent
- Relaciones: Generation, Usage, ...
- Fácil de extender
- Bundles: procedencia de procedencias
- Colecciones



2. Gestión del linaje con PROV

- ¿Son los modelos de linaje de PROV e ISO 19115 compatibles?
 - ISO → PROV extendido; PROV → ISO se pide información

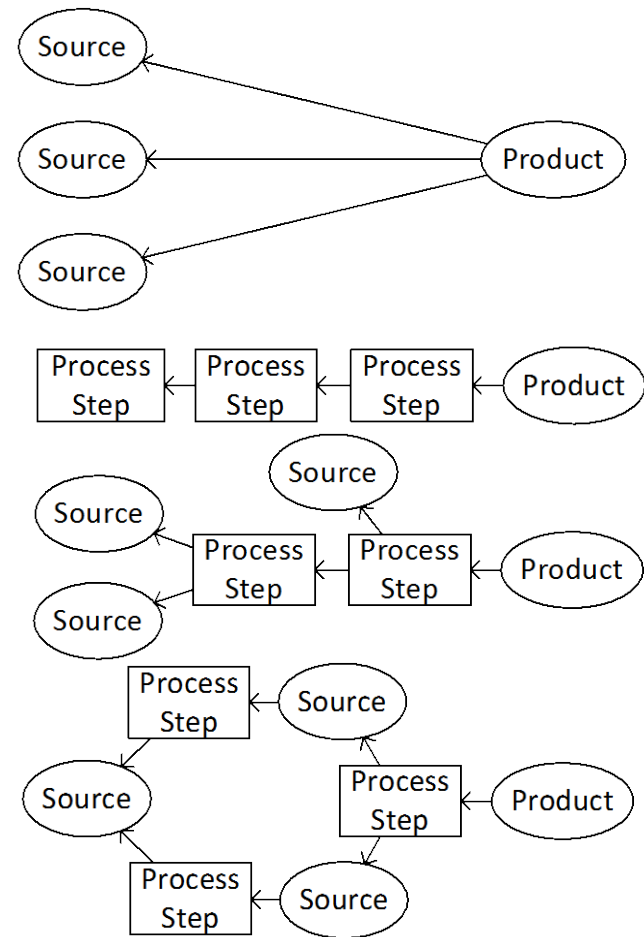
- **Conceptos**

ISO 19115	PROV-DM
Source - as input - as output	Entity - Usage - Generation
Process Step - date time - processor - processing, algorithm, software	Activity - start time, end time - Association with Agent - Association with Agent (optional) with Plan

- **PROV extendido**
 - Registro de metadato = PROV bundle + un tópico primario
 - “Scope” en el sentido ISO 19115 (vista/selección)

2. Gestión del linaje con PROV

- **Ejemplos de linajes ISO 19115 representables**
 - 19115: linaje como una lista de fuentes → “was derived from”
 - 19115: linaje como una lista de procesos → “was generated by, was informed by”
 - 19115: linaje como una lista de procesos con menciones a las fuentes → “was generated by, was informed by, used”
 - 19115-2: flujo de procesos con fuentes entrada/salida → “was generated by”, “used”



2. Gestión del linaje con PROV

- **Ejemplo en notación humana**

```
bundle ex:transformation1
```

```
entity(ex:dataset1, [prov:type="dataset"])
```

```
entity(ex:spain, [prov:type="dataset",
  dct:spatial=ex:Spain])
```

```
li:scope(ex:dataset2, ex:dataset1, ex:spain)
```

```
entity(ex:errodata1, [prov:type="dataset"])
```

```
activity(ex:process1, [prov:type="data transformation"])
```

```
wasGeneratedBy(ex:dataset3, ex:process1)
```

```
wasGeneratedBy(ex:erroldata1, ex:process1)
```

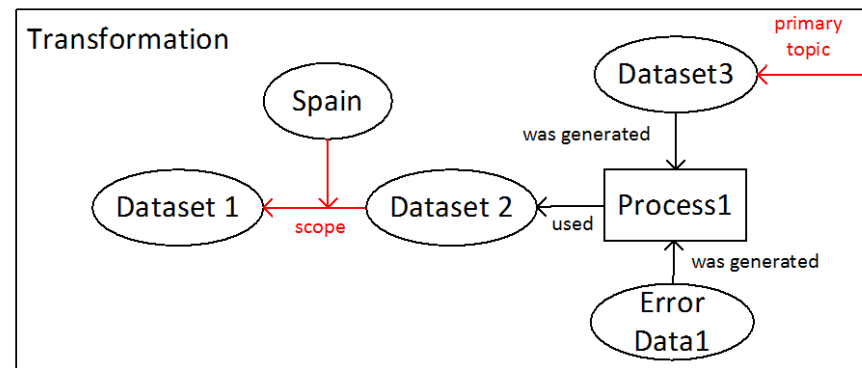
```
used(ex:process1, ex:dataset2)
```

```
li:primaryTopic(ex:transformation1, ex:dataset3)
```

```
endBundle
```

- **Descripción tan fina como se requiera**

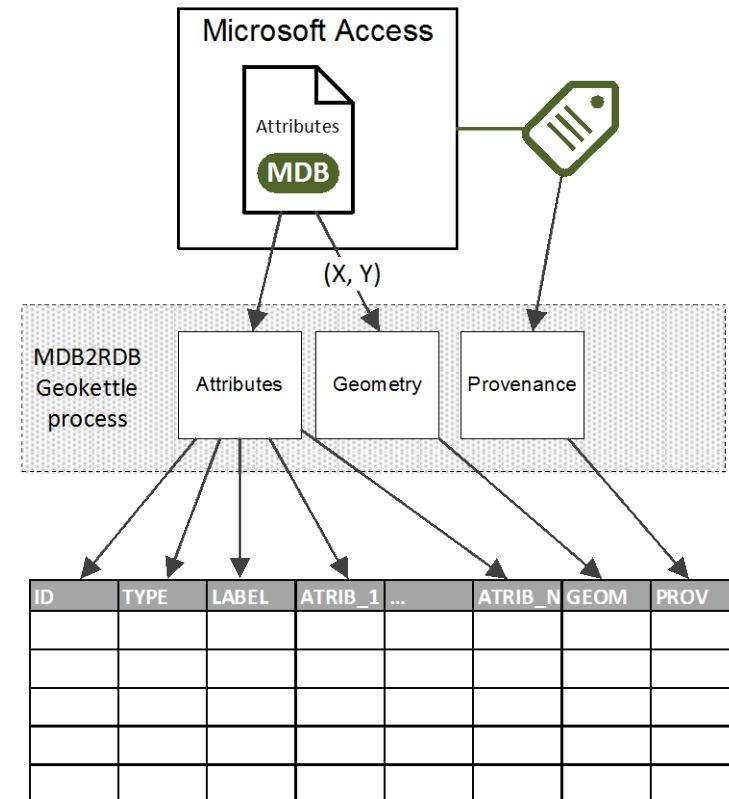
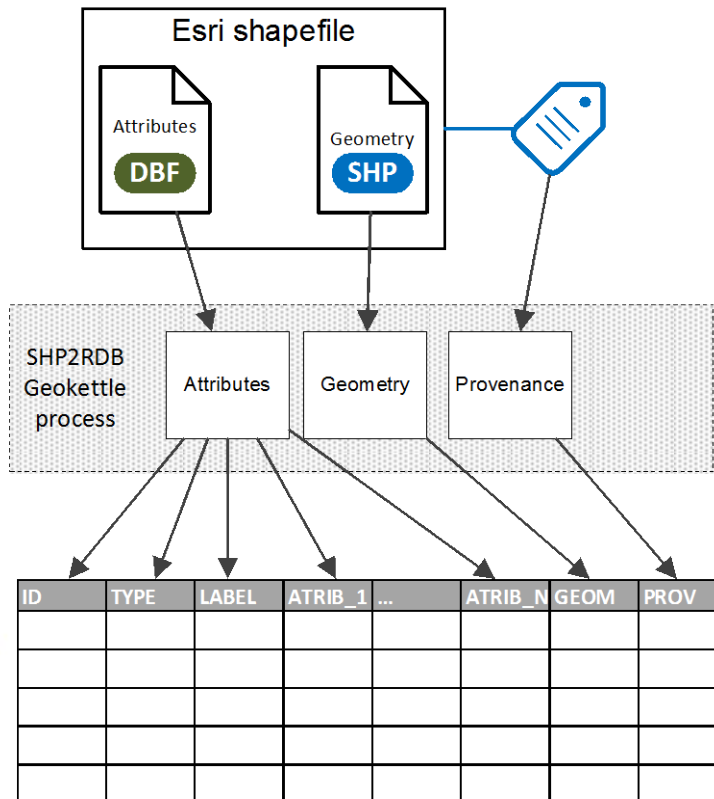
- En nuestro proyecto a nivel de objeto espacial



2. Gestión del linaje con PROV

- **Uso práctico**

- BD original → Formato exportación → BD espacial
- W3C PROV-XML templates (<http://www.w3.org/TR/prov-xml/>)



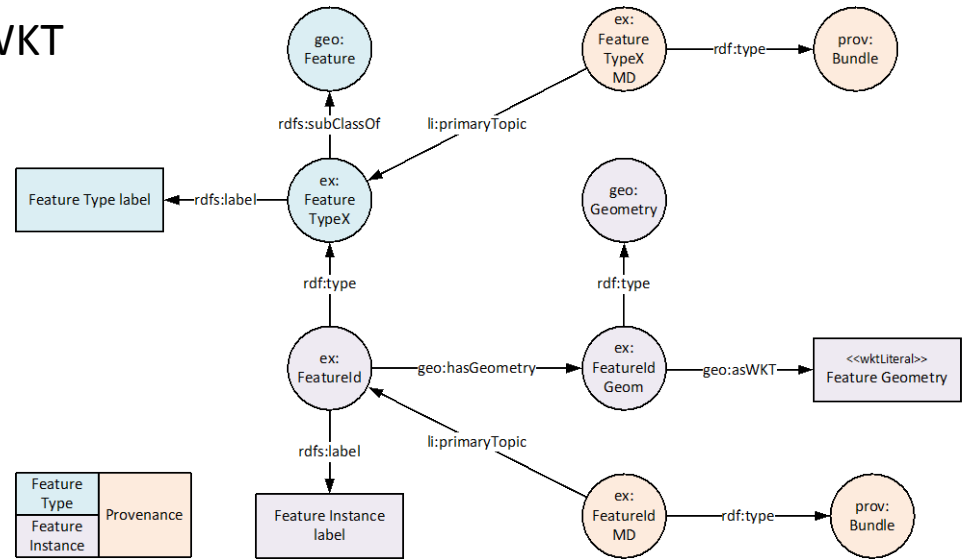
2. Gestión del linaje con PROV

- **Uso práctico**

- BD espacial → RDF espacial
- W3C PROV-O bundles (<http://www.w3.org/TR/prov-o/>)
- Aproximación similar al uso de WKT

```
dt:BCN-14361923 a geo:Feature ;
  rdfs:label "San Vitero" ;
  dct:isPartOf ds:BCN .
geo:hasGeometry [
  a geo:Geometry ;
  geo:asWKT "...^^geo:wktLiteral> ;
] .
```

```
pr:14361923 a prov:Bundle ;
  li:primaryTopic pr:14361923 ; // a Linked Map extension
  prov:value "<prov:document (...)>(…)</prov:document>^^rdf:XMLLiteral .
```



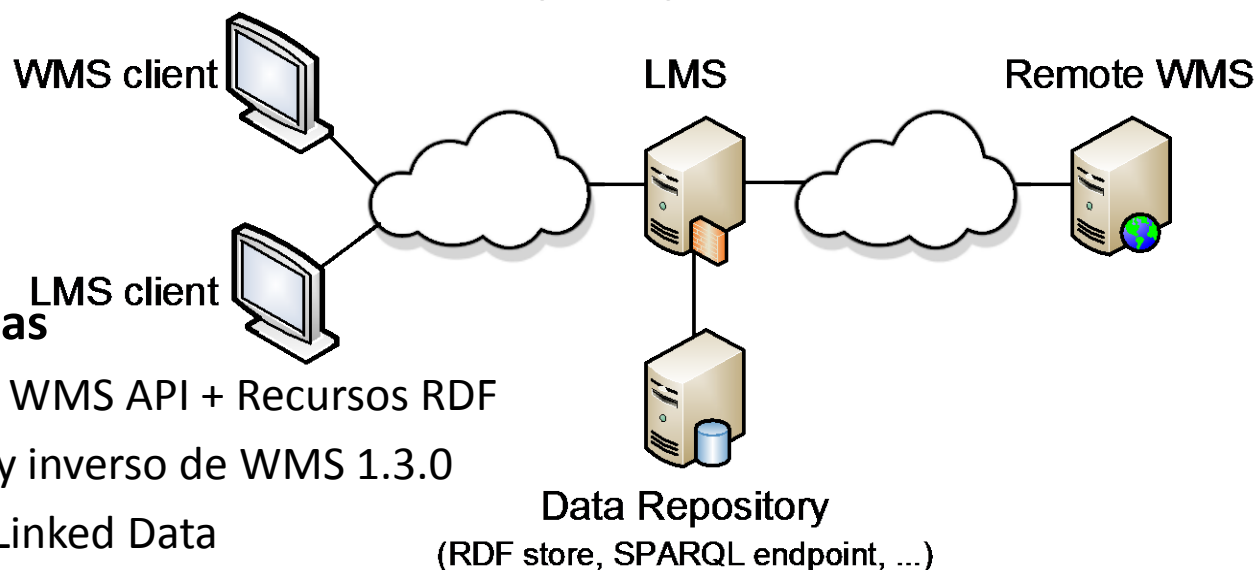
Feature Type	Provenance
Feature Instance	

3. Desarrollo de un WMS con Linked Data (LMS)

- **Visión**

- **Características planeadas**

- Dos espacios de URIs: WMS API + Recursos RDF
 - WMS API → Proxy inverso de WMS 1.3.0
 - Recursos RDF → Linked Data
- WMS API es un Recurso RDF por negociación de contenidos
 - P.E. KVP GetMap + “Accept: text/turtle” → devuelve colección de recursos relacionados espacialmente con el mapa en RDF
- Descubrimiento de recursos via RDF 5899 Web Links (HTTP headers)
 - Bidireccional: WMS API ↔ Recursos RDF
- Read-write Linked Data server
 - POST/DELETE/PUT/PATCH restringido a algunos recursos

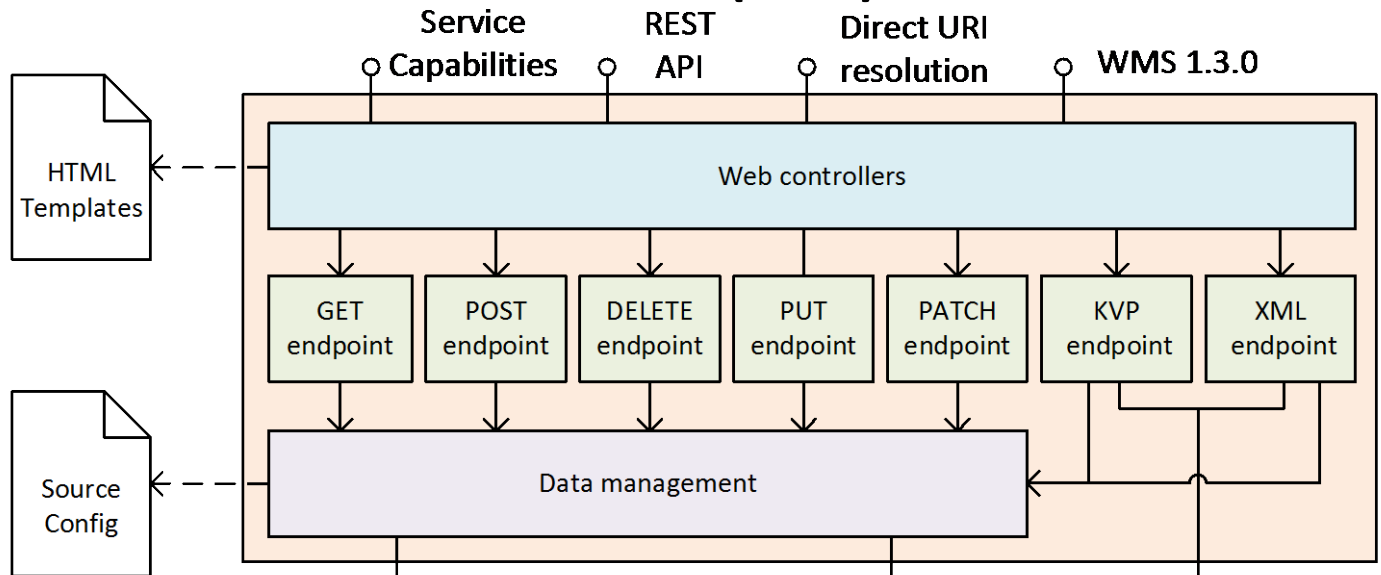


3. Desarrollo de un WMS con Linked Data (LMS)

- **LMS es desde el punto de vista de un cliente WMS es un servidor WMS estándar pero ...**
 - Las cabeceras HTTP de las respuestas contienen Web links a representaciones alternativas: e.g. PNG → enlace a JPEG, enlace a RDF/XML y enlace a Turtle
- **LMS es desde el punto de vista de un cliente REST es un servidor Linked Data pero ...**
 - Las cabeceras HTTP de las respuestas contienen Web links a representaciones alternativas: e.g. Turtle → enlace a RDF/XML, enlace a JPEG y enlace a PNG
- **Un cliente podría utilizar las dos visiones solo interpretando las cabeceras HTTP de las respuestas**

3. Desarrollo de un WMS con Linked Data (LMS)

- Arquitectura**



- Implementación**

- Java
- Spring boot (core)
- Thymeleaf (vistas)
- Pubby (SPARQL)



3. Desarrollo de un WMS con Linked Data (LMS)

- **Estado actual**
 - Proxy inverso WMS 1.3.0
 - WMS API como recurso RDF
 - Web Links
 - Linked Data server: modo lectura más soporte limitado POST
 - Prototipo parte de la plataforma Linked Map
 - <http://linkedmap.unizar.es/crowdsourcing-platform>
- **Versión independiente del proyecto en desarrollo**



Prototipo y caso de aplicación

- Plataforma para la evaluación de la calidad del enriquecimiento mediante crowdsourcing <http://linkedmap.unizar.es/crowdsourcing-platform/>

Buscar en el mapa

Añadir una evaluación de calidad

Revisar los comentarios de otros

Acceder a los datos (HTML/RDF)

Id	Label	Dataset	Date
BTN-30799136	Avenida de los Pirineos	http://linkedmap.unizar.es/id/datasets/BTN	9/2/2014
BTN-30801077	Urbanización Torre Francés	http://linkedmap.unizar.es/id/datasets/BTN	9/2/2014
BTN-30801087	San Gregorio	http://linkedmap.unizar.es/id/datasets/BTN	9/2/2014

"Avenida de los Pirineos"
<http://linkedmap.unizar.es/id/data/BTN-30799136>

Comments (7)

- OSM
- Id: OSM-85773352
- Provenance: [icon]
- Type: ROADS
- Name: Avenida de Pirineos
- Coordinates: -0.88005, 41.6671 (lon-lat, E)
- Reliability: 96.69967%
- Resource: <http://linkedmap.unizar.es/id/85773352>
- Quality: [icon]

Comments (5)

- 2/9/2014 silvialm
- 2/9/2014 rociorm
- 2/9/2014 lgalgo
- 2/9/2014 Miguel



Ideas para una discusión posterior

- ¿Es **W3C R2RML** un estándar adecuado para describir la transformación de datos espaciales?
- ¿Deben los **IETF Web Links** utilizarse en las respuestas de servidores **OGC** para enlazar recursos relacionados?
- Dado que los buscadores utilizan los **IETF Web Links** ¿puede su uso facilitar el descubrimiento de los servicios **OGC** en los buscadores?
- ¿Es viable incorporar **W3C PROV** a los flujos de trabajo actuales?
- ¿Hay escenarios donde un servidor **Linked Data** que pueda negociar contenidos servidos por servicios **OGC** tiene sentido? ¿Y en el caso contrario? ¿Bastaría solo con añadir **IETF Web Links** y olvidarnos de la negociación?



Sumario

- **Nuestra página de proyecto**
 - <http://linkedmap.unizar.es>
- **Nuestro proyecto de demostración**
 - <http://linkedmap.unizar.es/crowdsourcing-platform/>
- **Ideas clave**
 - Transformación simple de un Mapa nacional / VGI → RDF
 - Anotación del linaje a nivel de objeto espacial con PROV
 - Proxy transparente semántico con Web Links para WMS
- **Contacto**
 - Dr. Francisco J. Lopez-Pellicer <mailto:fjlopez@unizar.es>
 - IAAA (Universidad de Zaragoza) <http://iaaa.unizar.es/>

