JIIDE 2014

V Jornadas Ibéricas de
Infraestruturas de Dados Espaciais

Lisboa | 5-7 novembro 2014

GOVERNO DE PORTUGAL

d.gTerritório
Direção-Geral do Território

INSTITUTO GEOGRAFICO NACIONAL

GOBIERNO DE ESPAÑA MINISTERIO DE FOMENTO

COMPETE

QUADRO DE REFERÊNCIA ESTRATÉGICO NACIONAL

UNIÃO EUROPEIA
Fundo Europeu de Desenvolvimento Regional

IAAA
Grupo de Sistemas de Información Avanzados
Universidad Zaragoza

GeoLab

Planet Data

SEVENTH FRAMEWORK PROGRAMME

Upgrading maps with Linked Data

Lopez-Pellicer, Francisco J
Lacasta, Javier
Rentería, Walter

Barrera, Jesús
Lopez-de-Larrinzar, Juan
Agudo, Jose M

IAAA, Universidad de Zaragoza

GeoSpatiumLab

IAAA
Grupo de Sistemas de Información Avanzados
Universidad Zaragoza

The Linked Map project

- **Our vision**
  - Explore new applications of **Linked Data** in **Geographic Information**

- **Our goals**
  - **Semantic upgrade** of OGC WMS (namely LMS)
  - **Semantic integration** of authoritative GI datasets and VGI **with provenance**
  - **Crowdsourcing platform for QA** of semantic integration

The Linked Map project

- **The project was funded by PlanetData**

- **PlanetData (2010-2014)**
  - http://www.planet-data.eu/
  - FP7 Network of Excelence, State-of-art of large-scale public data management
  - **PlanetData Call 2 (2014):** Call for short term projects (1 year)

- **Partners**
  - **IAAA Lab** (Universidad Zaragoza, Spain) http://iaaa.unizar.es/ Research Lab, OGC, INSPIRE, ISO, SDI initiatives
  - **GeoSpatiumLab** (Zaragoza, Spain) http://www.geoslab.com/ SME, focused on GI
  - **CNIG** (Spain) http://www.cnig.es/ National Geographic Institute, data provider

The Linked Map project

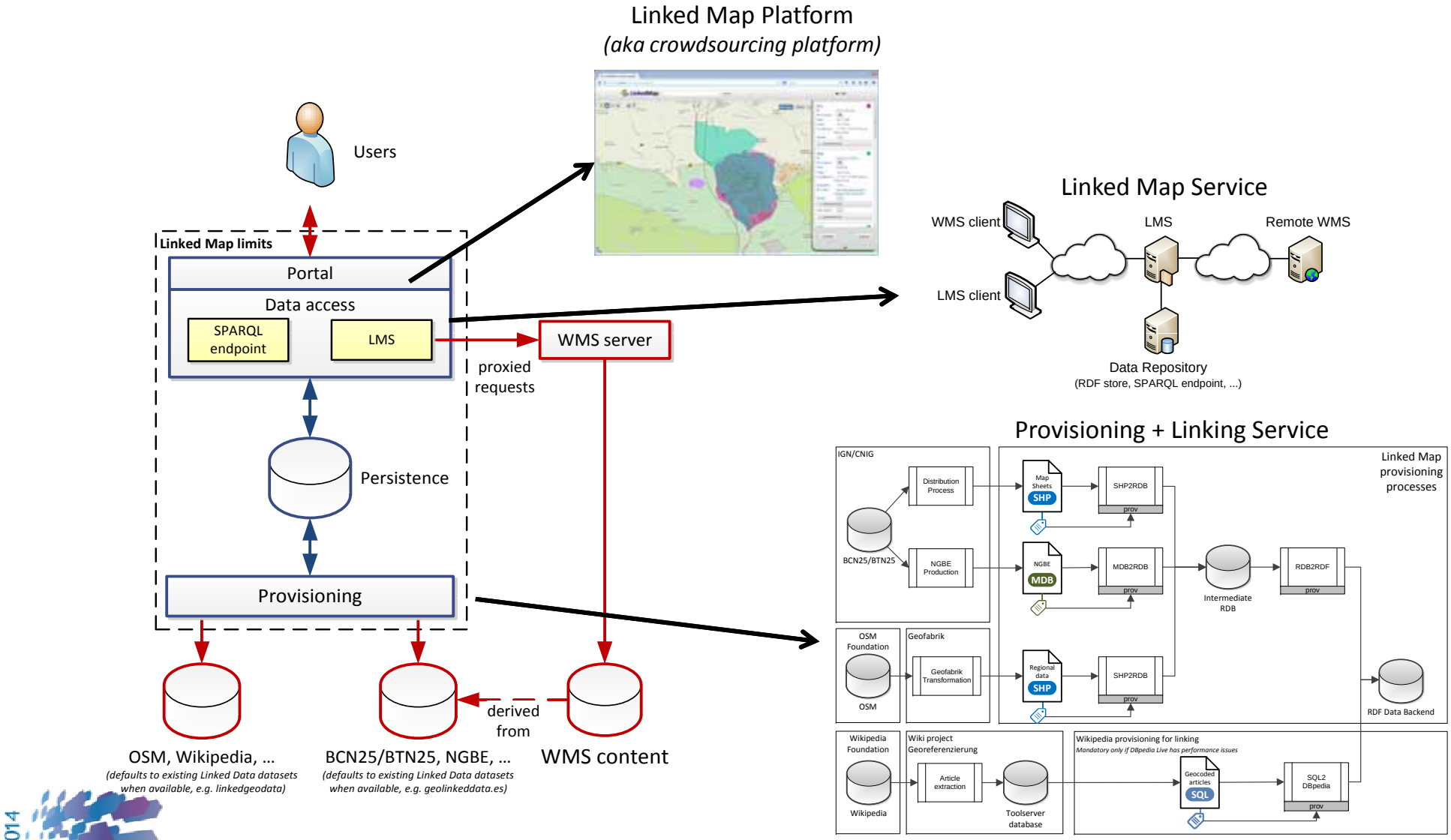
- **The motivating challenge**
  - **VGI and GI integration**
    - "Is VGI data believable for updating official maps?"
    - "Can crowdsourcing be useful for assessing this challenge?"
  - Key point: problem **relevant for GI producers**

- **Our contributions**
  - "Simple" semantic integration VGI & GI data → **Use of W3C PROV data**
  - Semantic upgrade of WMS → **WMS behaving as a LD server**
  - Crowdsourcing platform for integration QA → **Challenges in the use of LD with maps**
  - Evaluation experiments → **Nichesourcing evaluation**

The Linked Map project

- **Additional scenario details**
  - **Data**
    - GI: BCN/BTN (national map of Spain, provided by CNIG)
    - VGI: OpenStreetMap, Wikipedia
  - **Scope**
    - Spatial: Mainland Spain & Balearic Islands
    - Subject: OSM feature types (roads, buildings, ...)

Approach and implementation: architecture



IAAA
Grupo de Sistemas de Información Avanzados
Universidad Zaragoza

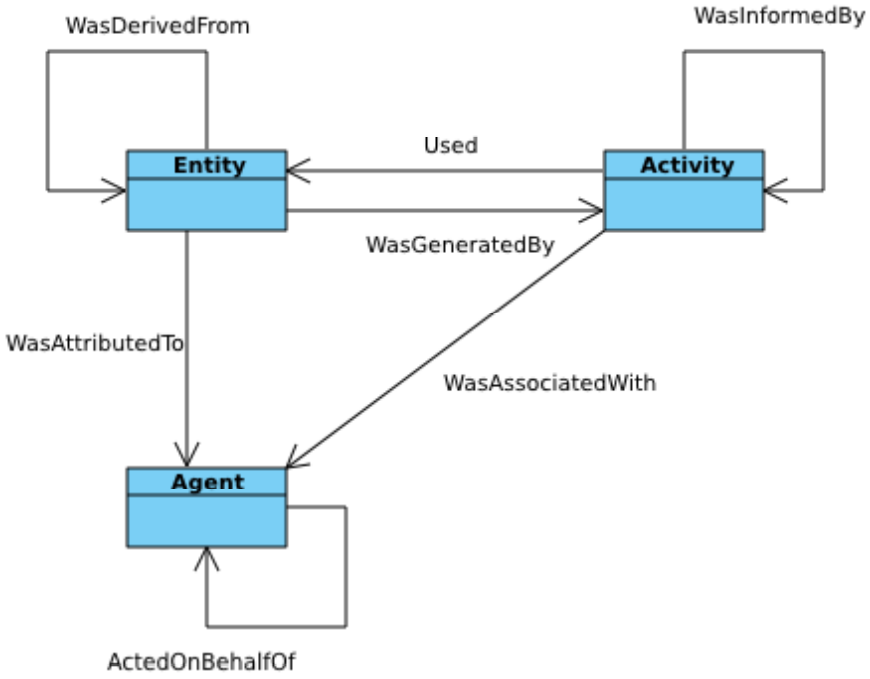
Approach and implementation: semantic integration

- **Tasks**
  - Transform datsets into LD → **GeoKettle** (ETL) + **MorphRDB** (tool)+ **GeoSPARL** (vocabulary)
  - Simple integration (name, type, location) → **Silk**
  - Provenance → Transformation + Integration → **W3C PROV** extended
  - Storage and GeoSPARQL access → **STRABON**

- **Results**
  - BCN/BTN 25: 13M triples, OSM: 35M triples, Enrichment: 3M triples
  - **Provenance tracking at feature level encoded in PROV-XML in RDBMS & RDF**
    - For each feature: File source, row source, transformation script, generation date ....

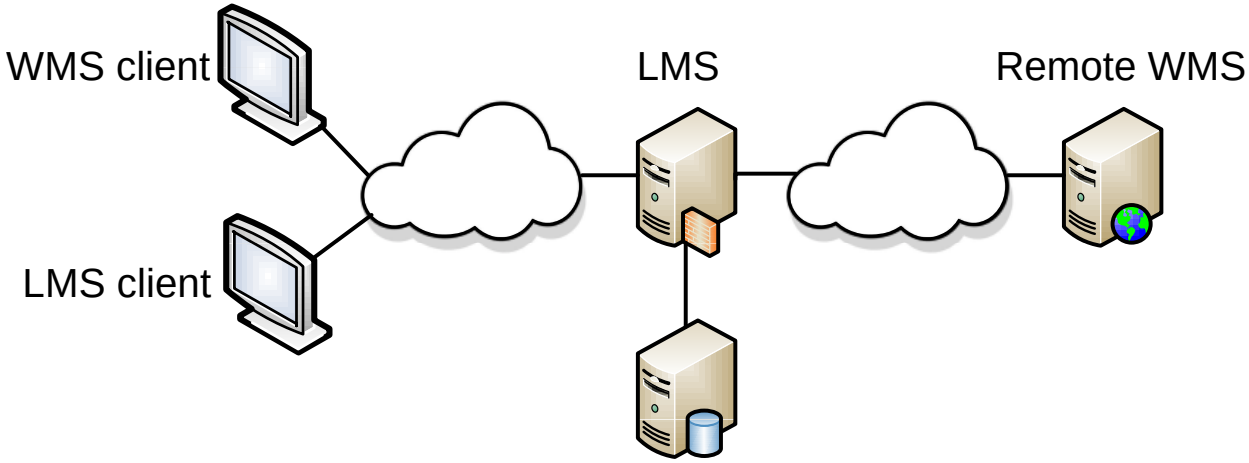
Approach and implementation: semantic integration

- **Track provenance at feature level with PROV**
  - Family of W3C Recommendations and Notes
    - http://www.w3.org/TR/prov-overview/
  - Goal: enable publication and interchange of provenance on the Web
    - Emerging standard (2013)
    - Available in RDF and XML
  - Compatible with ISO 19115 lineage model
    - All ISO 19115 lineage models can be represented
    - ISO metadata record = PROV bundle + explicit primarty topic
    - "scope" (in the sense of view/selector) is not part of PROV but can be added



Approach and implementation: semantic upgrade

- **Transparent enablement WMS (Linked Map Service - LMS)**
  - WMS 1.3.0 **reverse proxy**
  - Semantic upgrade of WMS requests by **content negotiation**
    - e.g. GetMap: return RDF resources spatially related to the map
  - **Web links headers** (RFC 5899) link both WMS and semantic responses



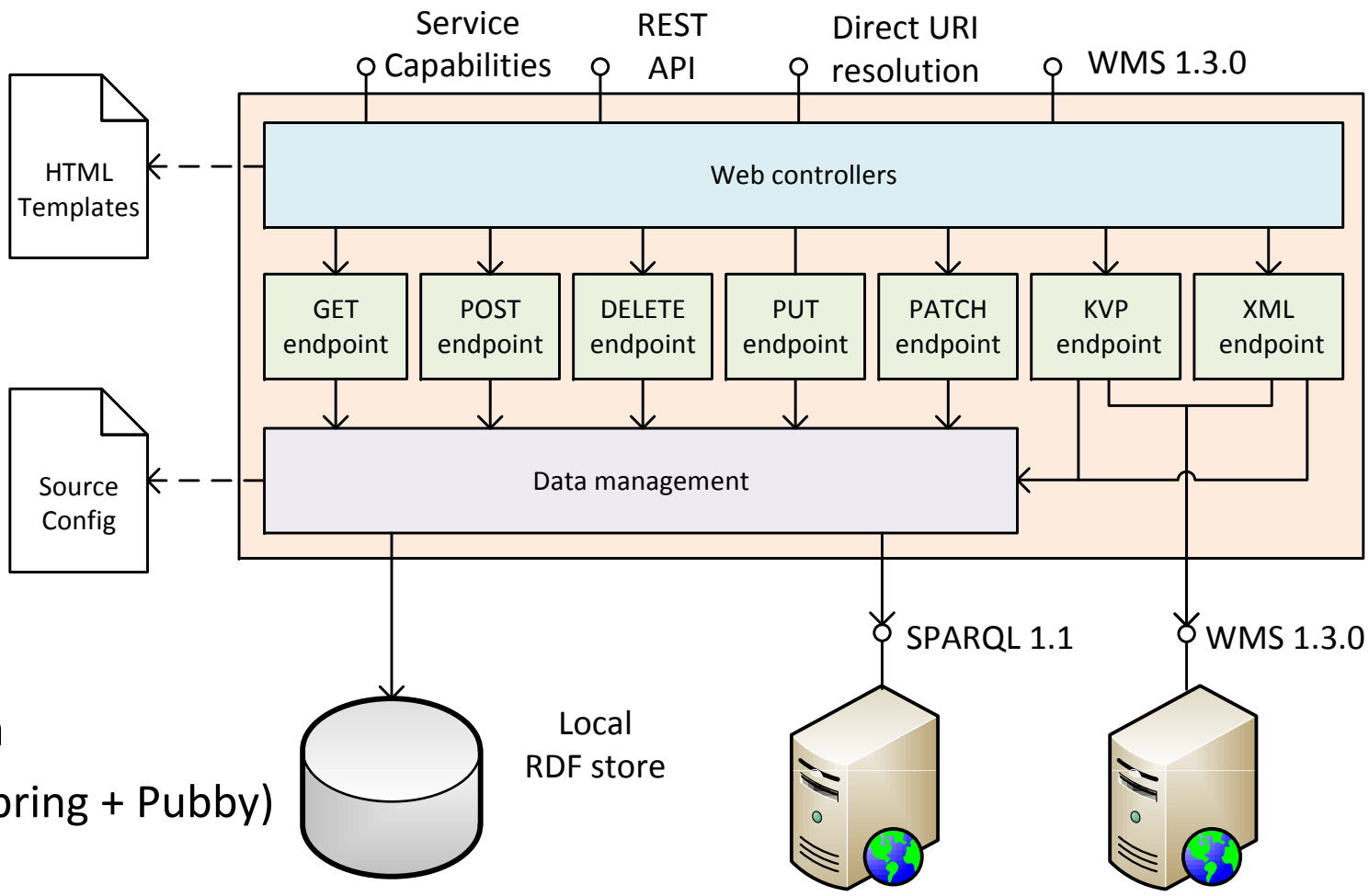
Approach and implementation: semantic upgrade

- **LMS is from the point of view of a WMS client**
  - A **WMS server**
  - WMS response headers contain **Web links** to alternate representations of capabilities, map images, etc;
    - A PNG map contains links to a JPEG map (WMS request) or RDF (machine processable representation)

- **LMS is from the point of view of a generic REST client**
  - A **Linked Data endpoint that contains as resources map tiles and data**
    - Each possible GetMap, GetFeatureInfo or GetCapabilities request denotes a resource
    - Resource **representations contain links to equivalent KVP WMS requests**
  - Response headers contain **Web links** to alternate representations, including KVP WMS requests

Approach and implementation: semantic upgrade

- **Architecture**



- **Implementation**
  - Java based (Spring + Pubby)

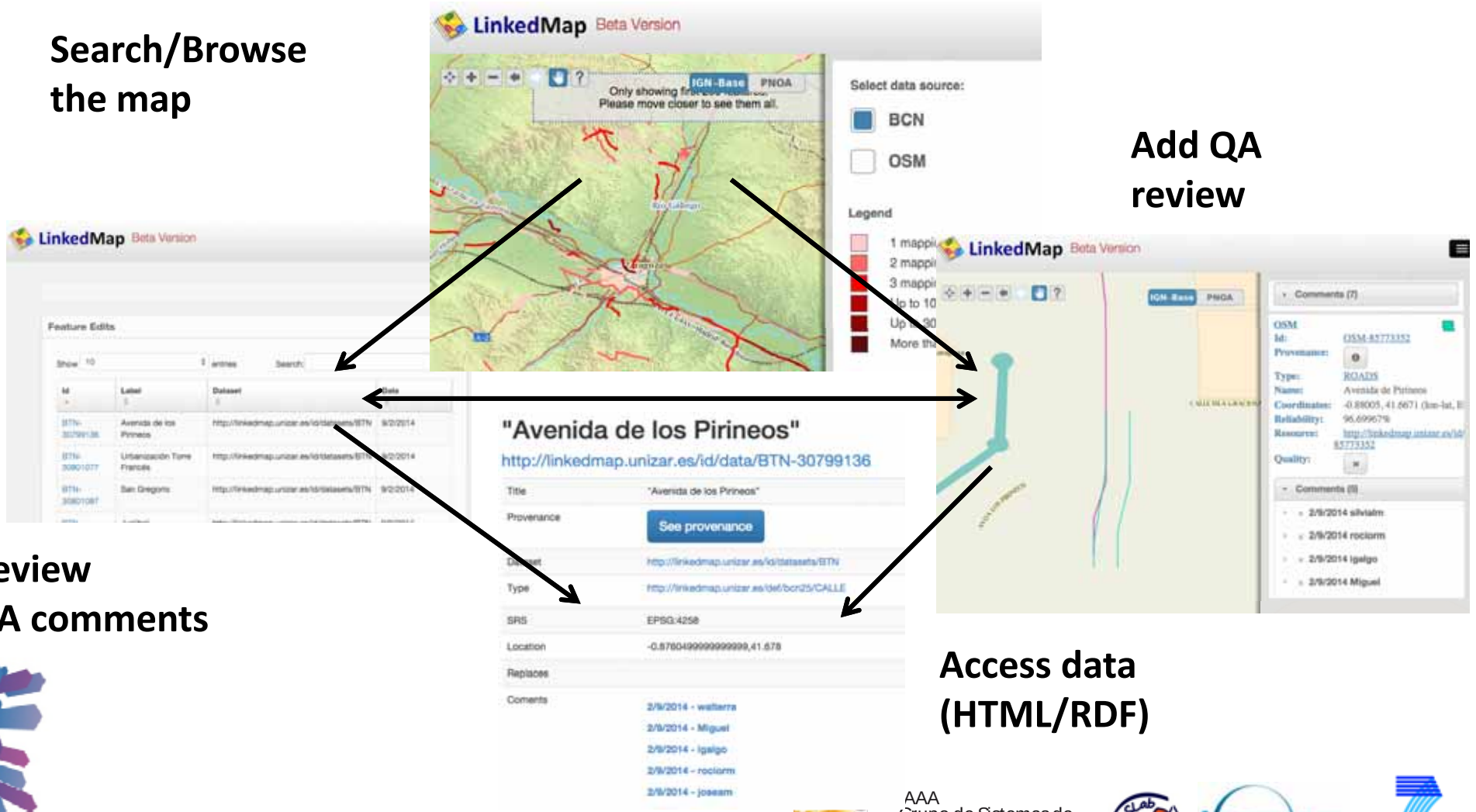
Approach and implementation: crodwourcing platform

- **Frontend**
  - Map client
  - Layer of linked features
  - User can add, review QA of links & features, browse data
- **Backend**
  - LMS + GeoSPARQL endpoint
- **Evaluation experiments**
  - **Open**
    - Can the platform (and the task) engage a big community? → Failled to create a critical mass
  - **Focused**
    - **Nichesourcing**: Can be used by **a small engaged but professional community** for QA of data and links? → Yes, but be careful with authority bias

The platform is available at ...

- http://linkedmap.unizar.es/crowdsourcing-platform/

Search/Browse the map



Add QA review

Review QA comments

Access data (HTML/RDF)

Lessons learned

- **W3C PROV**
  - It is possible to track ETL geo workflows with PROV
  - ISO 19115 compatible, can be stored in RDF as XML literal (as Geo)
- **Semantic upgrade of geoservices**
  - Services can be upgraded transparently using (carefully) existing IETF/W3C standards and best practices
  - Unexpected opportunities: RFC 5899 can increase discoverability of OGC web services in search engines
- **Use of LD with maps**
  - Flexible and rich data model
  - Experiments do not take full advantage of semantics (e.g. Inference)

Concluding remarks

- Years ago the **OGC** standards **were ahead of its time**
- Can we afirm this today? Yes, no? Are OGC standadization programs slow?
- I think that we **can upgrade** them without breaking by using wisely W3C/IETF standards
- Do you agree? What do you **think**?

If you whish know more:

http://linkedmap.unizar.es/

Francisco J Lopez-Pellicer

fjlopez@unizar.es

IAAA, Universidad de Zaragoza