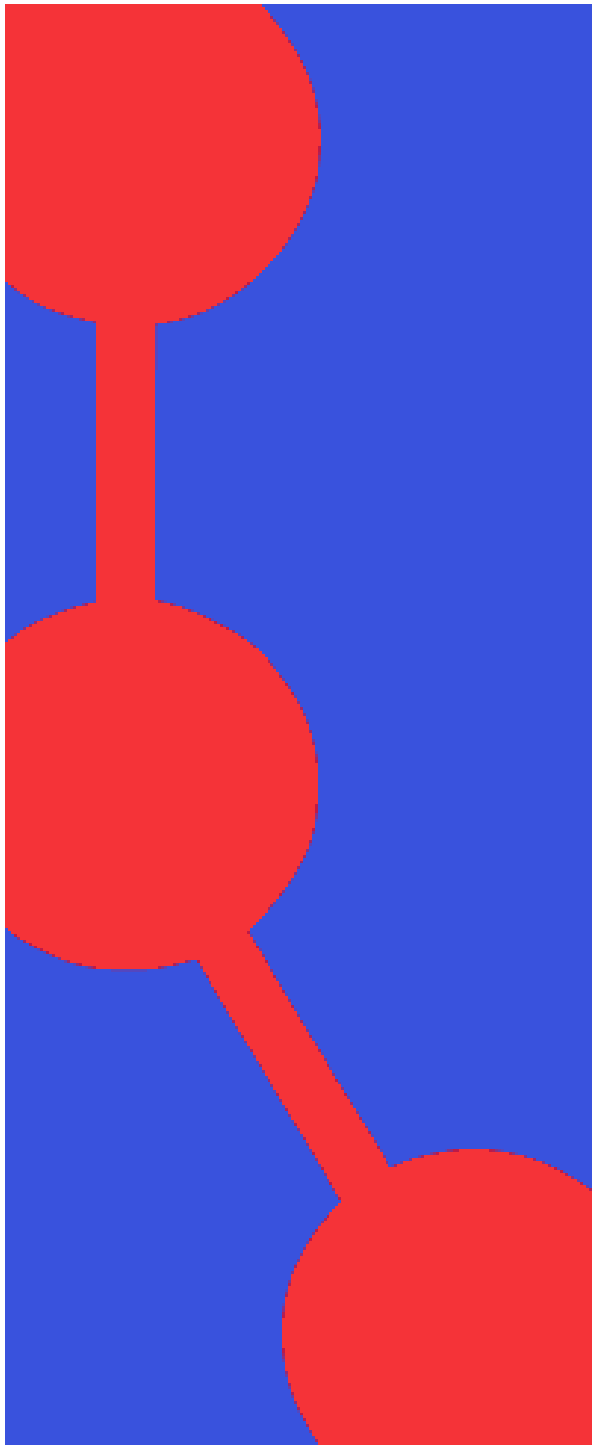


Lift your data

Introduction to Linked Data

Ghislain Atemezing[1] , Boris Villazón-Terrazas[2]

1 EURECOM, France

auguste.atemezing@eurecom.fr

2 Ontology Engineering Group, FI, UPM

bvillazon@fi.upm.es

Slides available at: http://www.slideshare.net/atemezing/

INSPIRE Conference 2012, Istanbul

Acknowledgements: Oscar Corcho, Asunción Gómez-Pérez, Luis Vilches, Raphaël Troncy, Olaf Hartig, Luis Vilches and many others that we may have omitted.

Agenda

- Linked Data
- Geospatial LD Datasets
- 5-star deployment scheme for Linked Open Data

Agenda

- Linked Data
- Geospatial LD Datasets
- 5-star deployment scheme for Linked Open Data

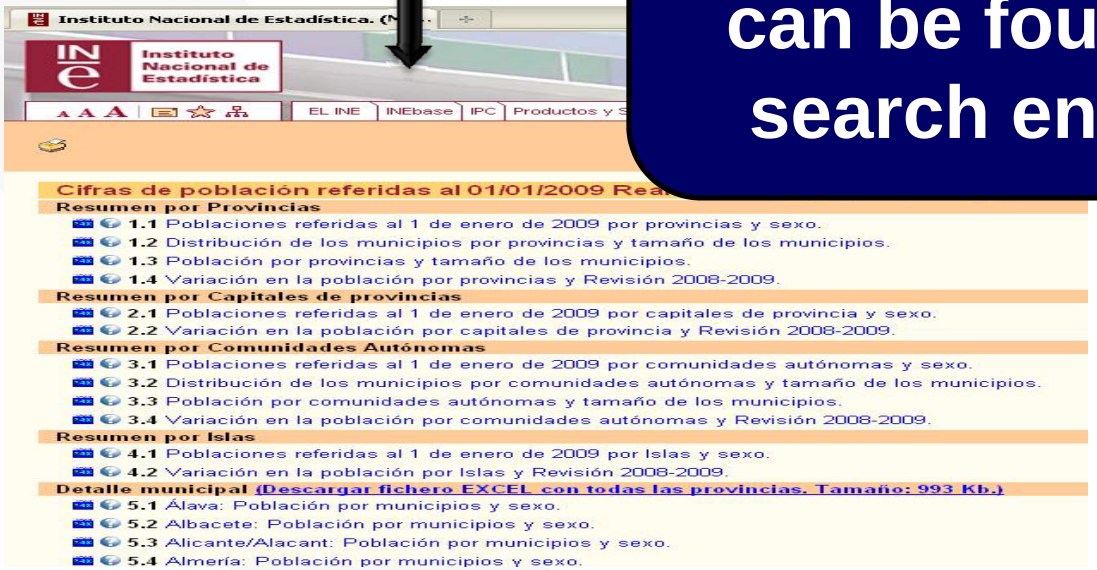
Current Web

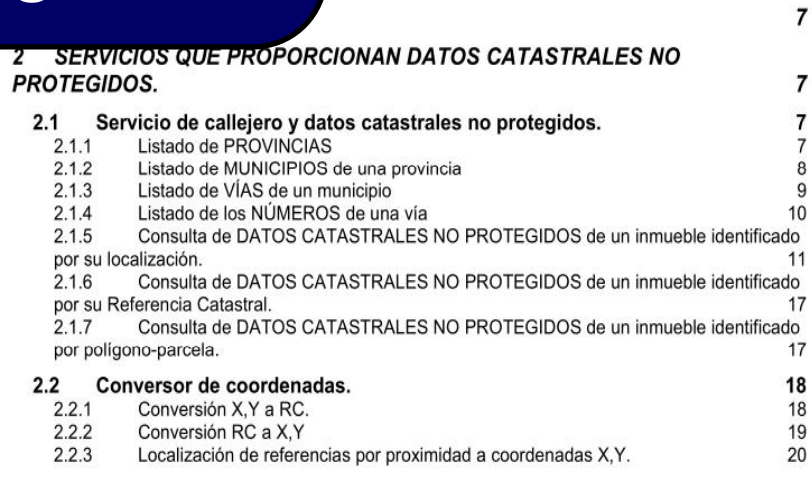
Geospatial Database (Spain)

Statistical Database (Spain)

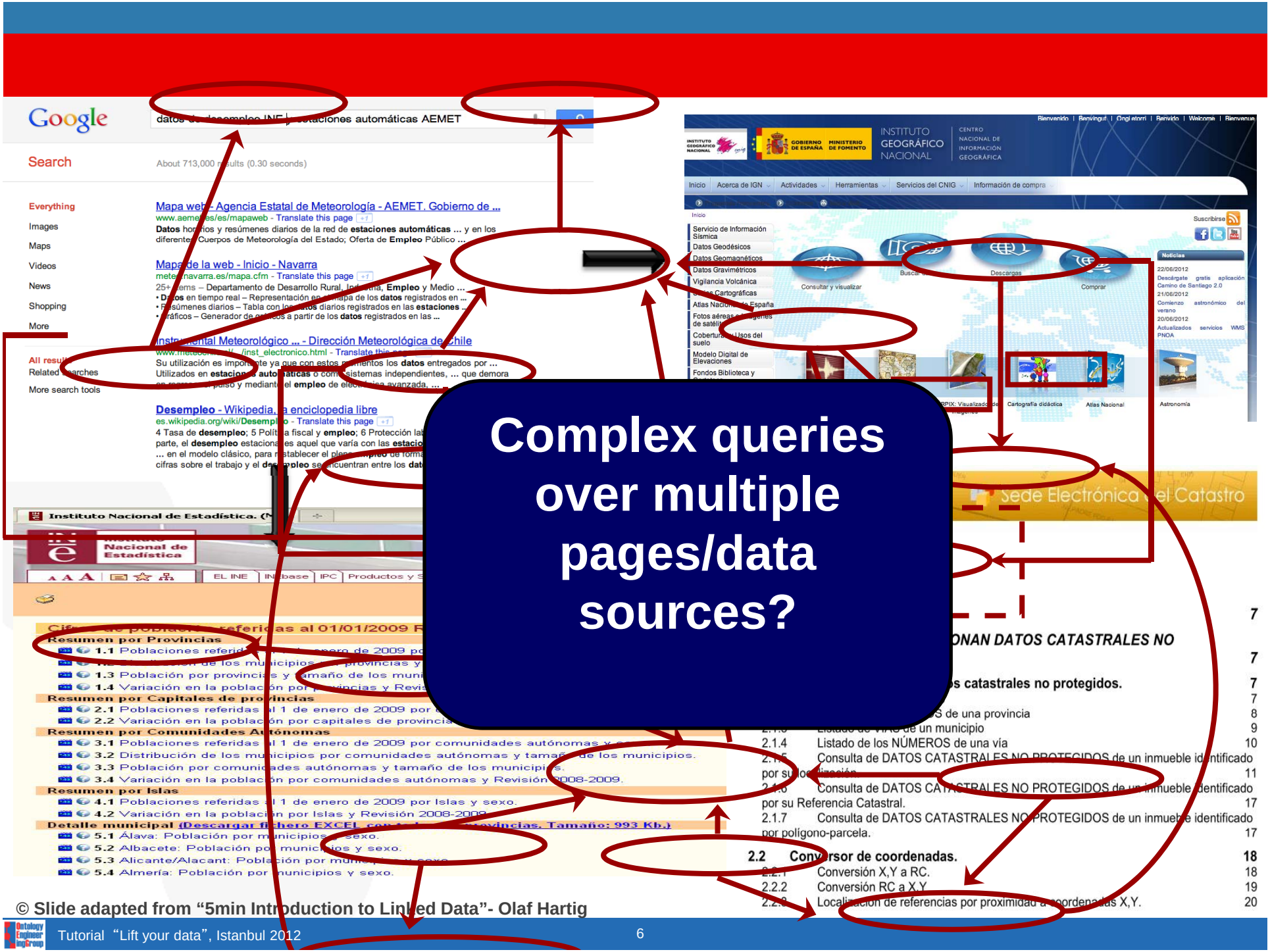
© Slide adapted from "5min Introduction to Linked Data"- Olaf Hartig

Information from *single pages* can be found via search engines





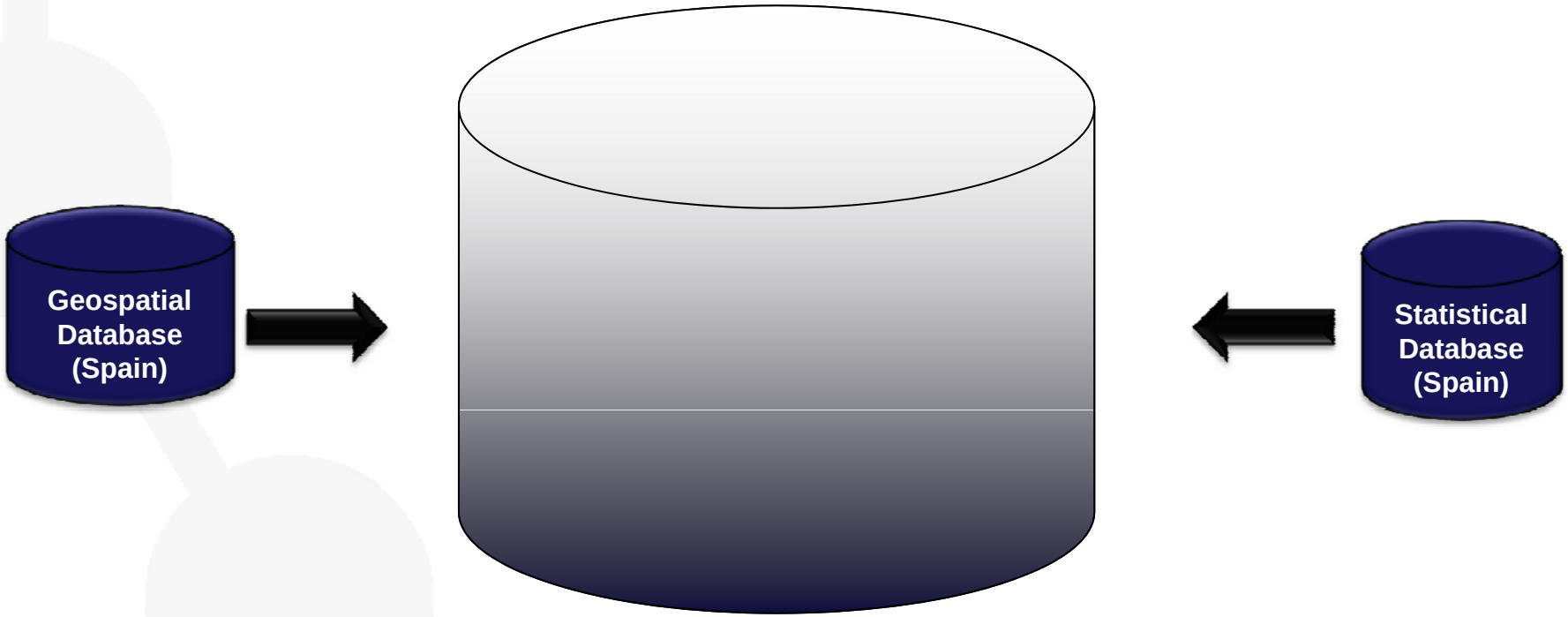
© Slide adapted from "5min Introduction to Linked Data"- Olaf Hartig



© Slide adapted from "5min Introduction to Linked Data"- Olaf Hartig

What do we actually want?

- Use the Web like a single global database
  - web of documents -> web of data



© Slide adapted from “5min Introduction to Linked Data”- Olaf Hartig

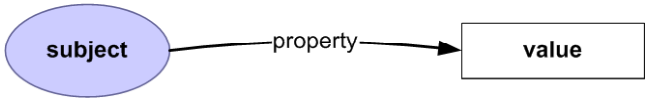
Linked Data enables such Web of Data

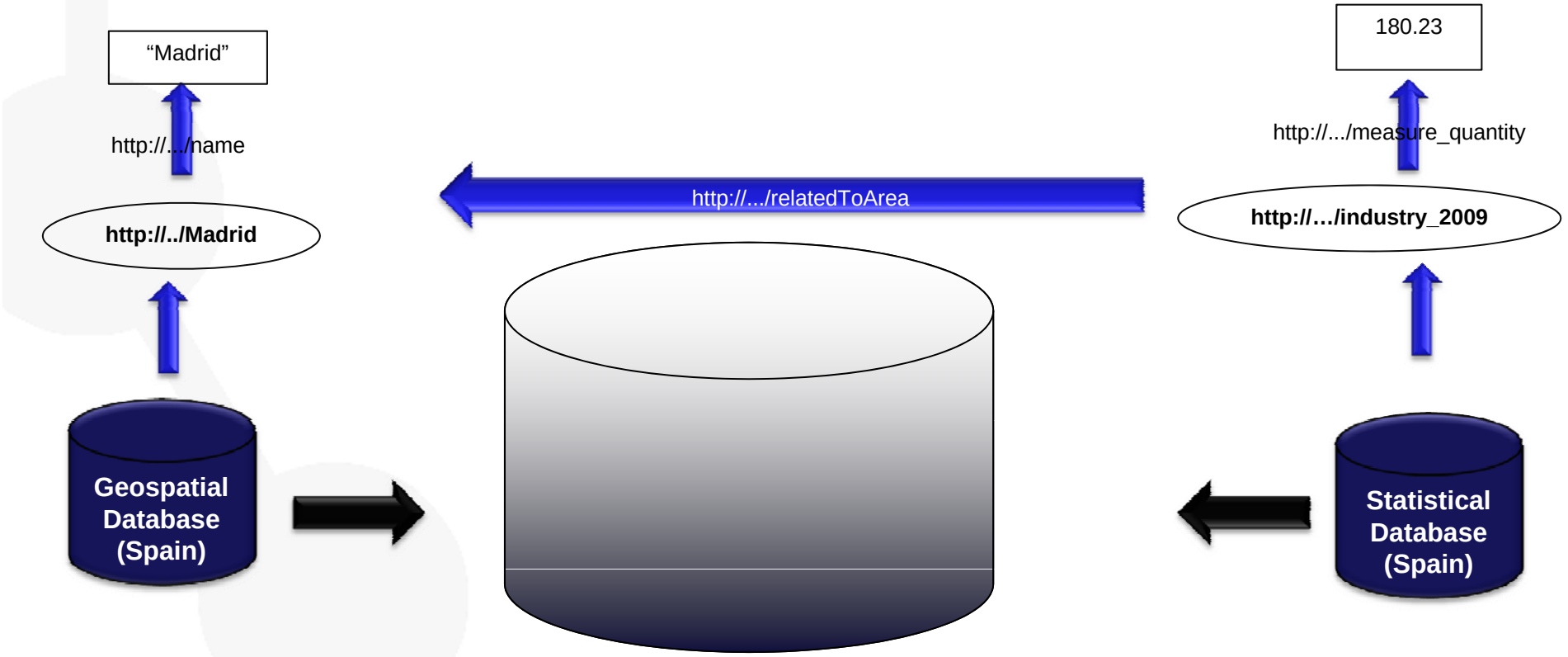
Global Identifier: URI (Uniform Resource Identifier), which is a string of characters used to identify a name or a resource on the Internet.

Data Model: RDF (Resource Description Framework), which is a standard model for data interchange on the Web

Access Mechanism: HTTP

Connection: Typed Links



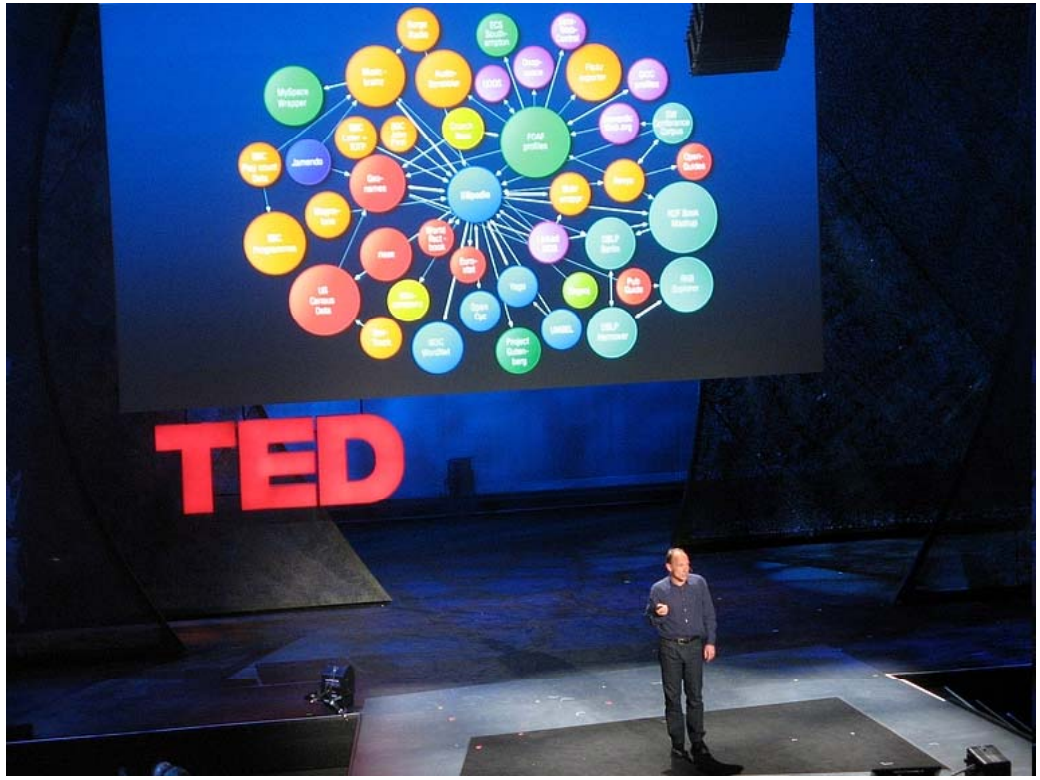


© Slide adapted from "5min Introduction to Linked Data"- Olaf Hartig

The four principles (Tim Berners Lee, 2006)

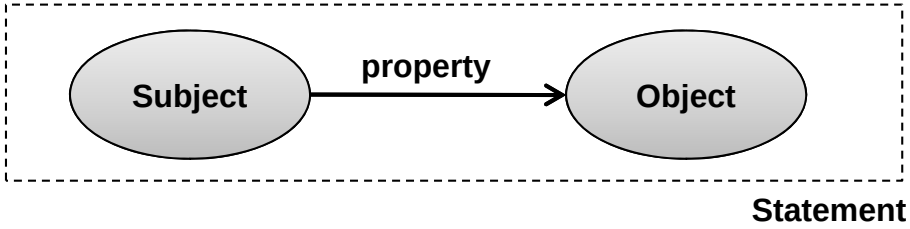
1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
4. Include links to other URIs, so that they can discover more things.

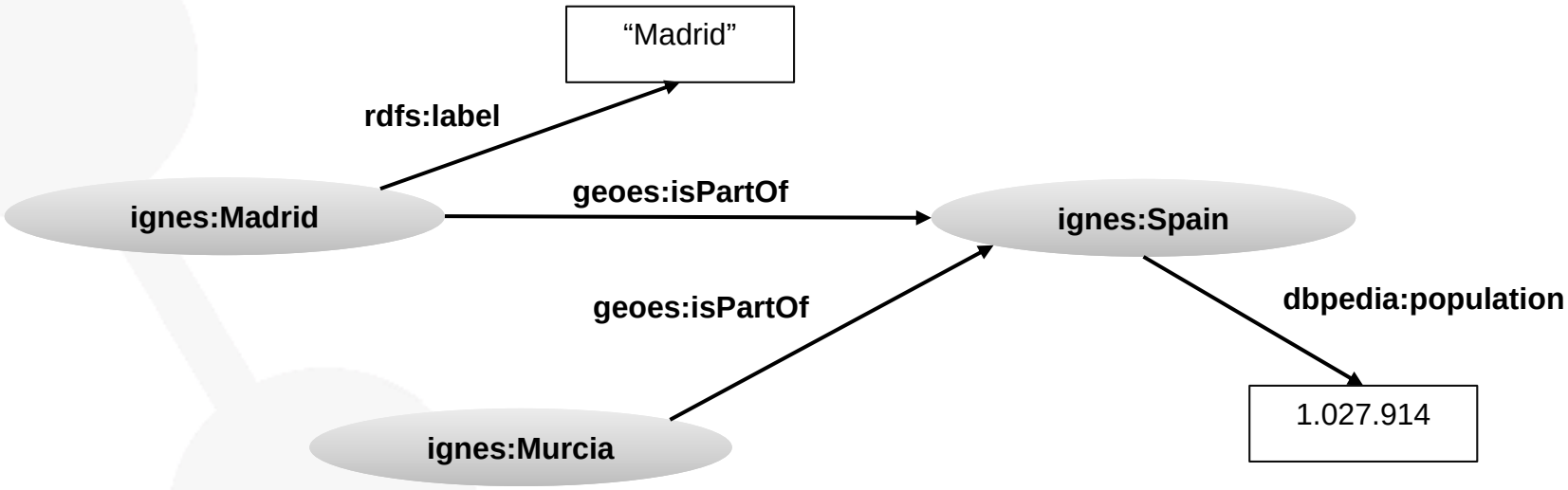
- **http://www.w3.org/DesignIssues/LinkedData.html**
- **http://www.ted.com/talks/tim_berners_lee_on_the_next_web.html**



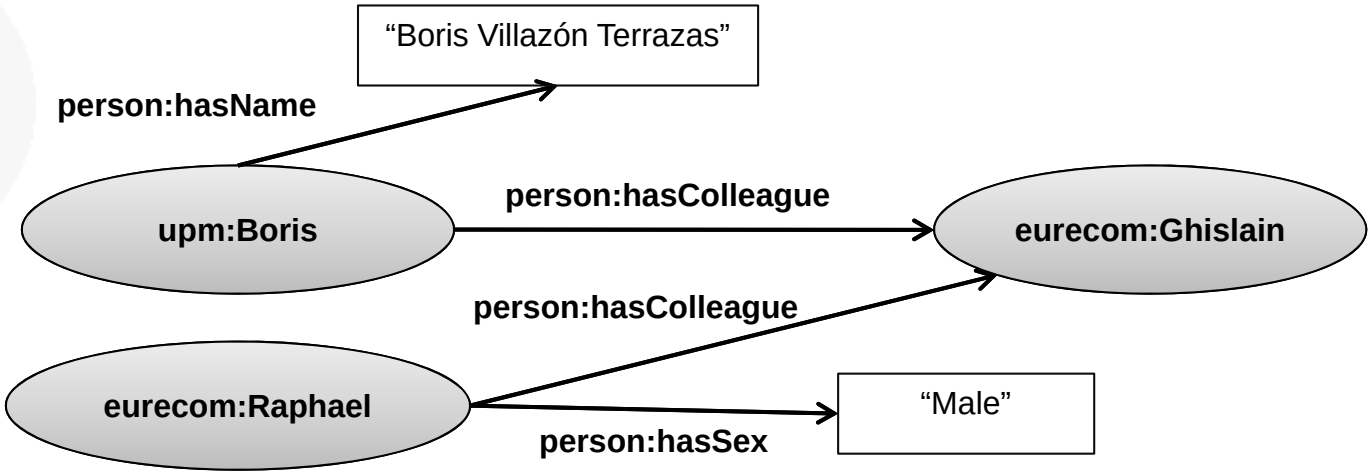
Resource Description Framework (RDF)

- RDF is a basic KR language based on semantic networks
  - Useful to represent metadata and describe any type of information in a machine-accesible way.

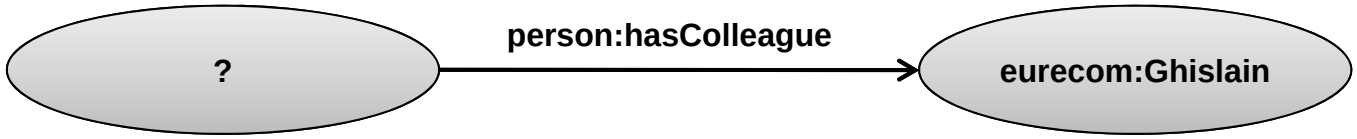




RDF - SPARQL



- Query: "Tell me who are the persons who have Ghislain as colleague"

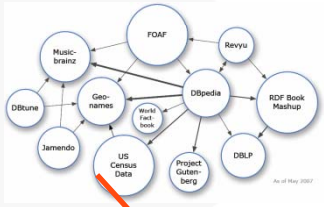


- Result: **upm:Boris** and **eurecom:Raphael**

- SPARQL query language for RDF. W3C recommendation

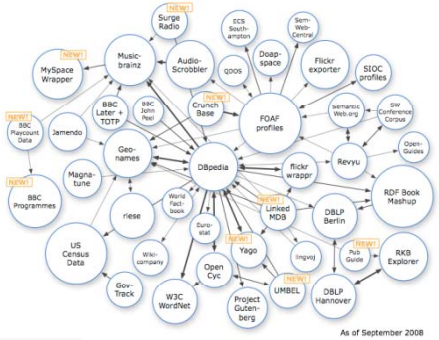
```
SELECT ?s
WHERE { ?s person:hasColleague eurecom:Ghislain . }
```

Evolution of the *Linked Open Data*

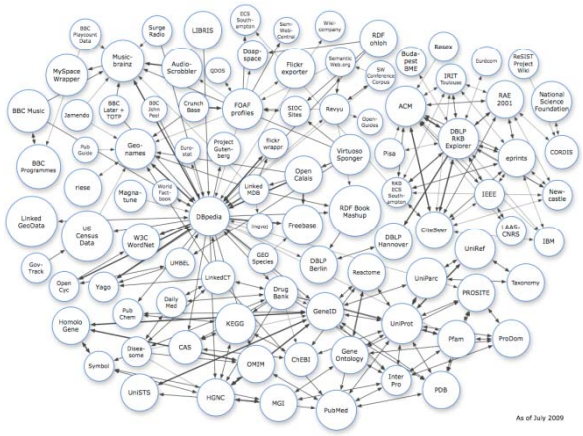


- 2007

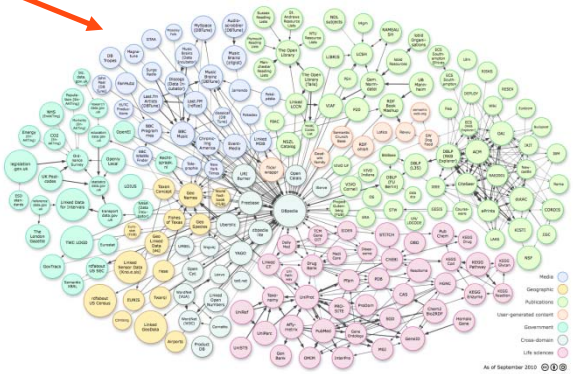
- 2008



- 2009

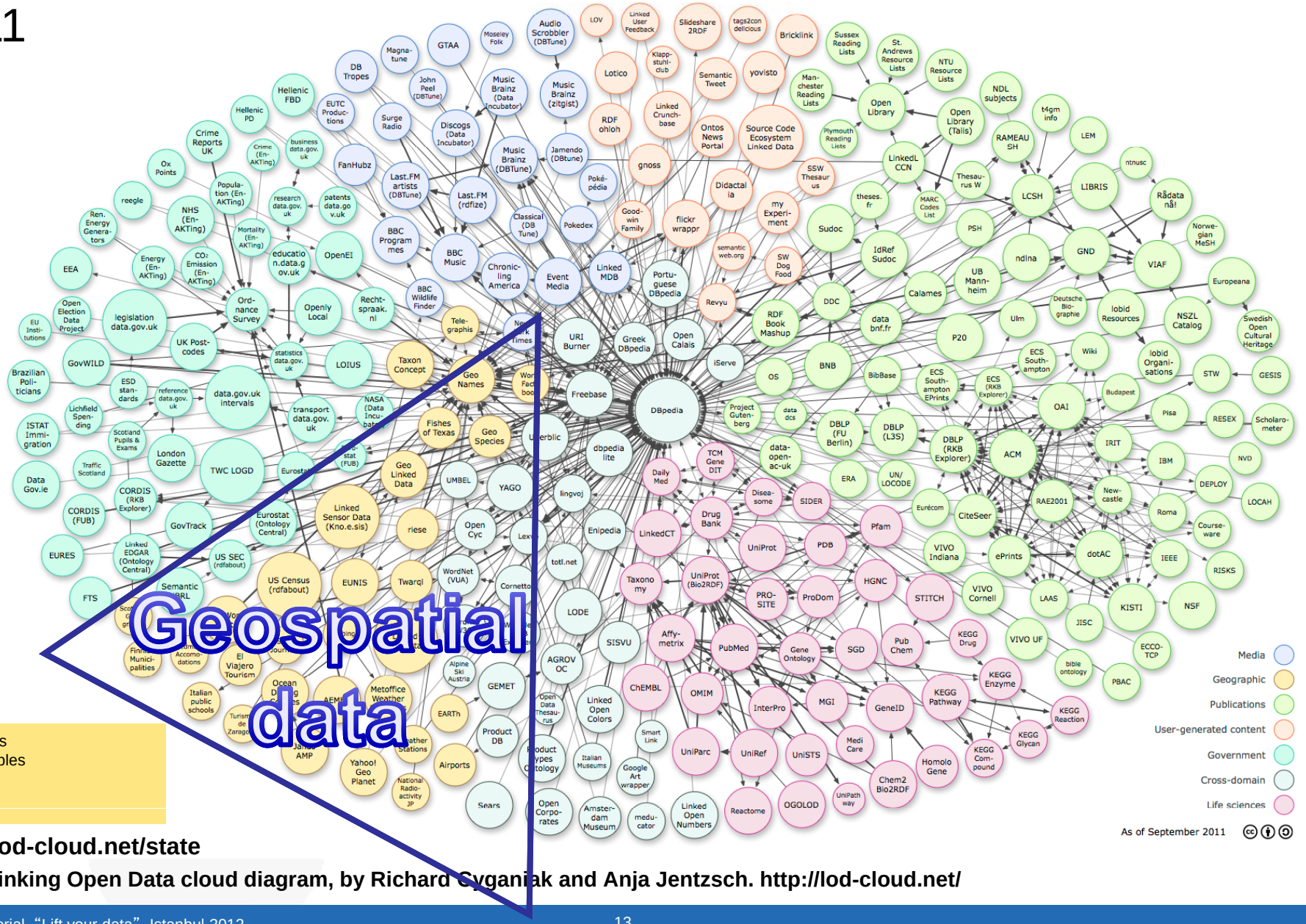


- 2010



Linked Open Data

2011



31 datasets
19.43% triples

http://lod-cloud.net/state

Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. http://lod-cloud.net/

Agenda

- Linked Data
- Geospatial LD Datasets
- 5-star deployment scheme for Linked Open Data

GeoData is becoming increasingly relevant

- Datasets in the geospatial domain contribute more than one fifth of the RDF triples in the Web of Data.

- Well-know Geo LD datasets
  - Ordnance Survey
  - GeoLinkedData
  - LinkedGeoData
  - GeoNames
  - DBPedia
  - GDAM-RDF

Ordnance Survey http://data.ordnancesurvey.co.uk

- Topic: Administrative units
- Datasets: The administrative gazeeter for Great Britain.
- URI pattern: http://data.ordnancesurvey.co.uk/id/{id}
- Vocabulary: Spatial Relations Ontology, Administrative Geography Ontology, WGS84, FOAF, and Gazeeter Ontology

GeoLinkedData http://geo.linkeddata.es/

- Topic: Hydrography, Administrative units, statistical information, meteorological features
- Datasets: Statistical, geospatial, and meteorological data.
- URI pattern:
  - http://geo.linkeddata.es/ontology/{concept|property}
  - http://geo.linkeddata.es/resource/{type}/{name}
- Vocabulary: SCOVO, FAO Geopolitical, hydrOntology, WSG84, GeoLinkedData Geometry Model, and Time Ontology.

LinkedGeoData http://linkedgeodata.org/

- Topic: Points of interest
- Datasets: OpenStreetMap database
- URI pattern: http://linkedgeodata.org/triplify/{id}
- Vocabulary: LGD Ontology, WGS84, NeoGeo

GeoNames http://www.geonames.org/

- Topic: Toponyms
- Datasets: Datasets used by geonames
- URI pattern: http://sws.geonames.org/{id}
- Vocabulary: GeoNames Ontology, WGS84

DBPedia http://dbpedia.org/

- Topic: General knowledge
- Datasets: Wikipedia
- URI pattern: http://dbpedia.org/resource/{name}
- Vocabulary: DBPedia ontology, WGS84

GADM-RDF http://gadm.geovocab.org/

- Topic: Global administrative areas
- Datasets: countries and lower level subdivisions.
- URI pattern: http://gadm.geovocab.org/id/{id}
- Vocabulary: gadm ontology, NeoGeo Ontology, DBPedia ontology, WGS84

Agenda

- Linked Data
- Geospatial LD Datasets
- 5-star deployment scheme for Linked Open Data

The Five Stars deploymet scheme

Is your data five stars?

★ make your stuff available on the Web (whatever format) under an open license [1]

★★ make it available as structured data (e.g., Excel instead of image scan of a table) [2]

★★★ use non-proprietary formats (e.g., CSV instead of Excel) [3]

★★★★ use URIs to identify things, so that people can point at your stuff [4]

★★★★★ link your data to other data to provide context [5]

http://lab.linkeddata.deri.ie/2010/star-scheme-by-example/

One Star: Open Data & Open License

data.gouv.fr BETA
INNOVATION TRANSPARENCE . OUVERTURE

data.gov.uk BETA
Opening up government

Proyecto aporta
Reutilización de la Información del Sector Público

nature.com linked data

License	Domain	By	SA
Creative Commons Attribution	Content	Y	N
Creative Commons Attribution Share-Alike	Content	Y	Y
Creative Commons CCZero	Content, Data	N	N
GNU Free Documentation License	Content	Y	Y
UK PSI Public Sector Information	Content, Data	Y	N
Free Art License	Content	Y	Y
MirOS License	Code, Content	Y	N

http://opendefinition.org/licenses/

Two stars: Reusability

- Data are reusable
- Some formats are helpful
  - Excel, CSV, JSON, XML, GML
- Others not really
  - PDF, HTML, MS Word

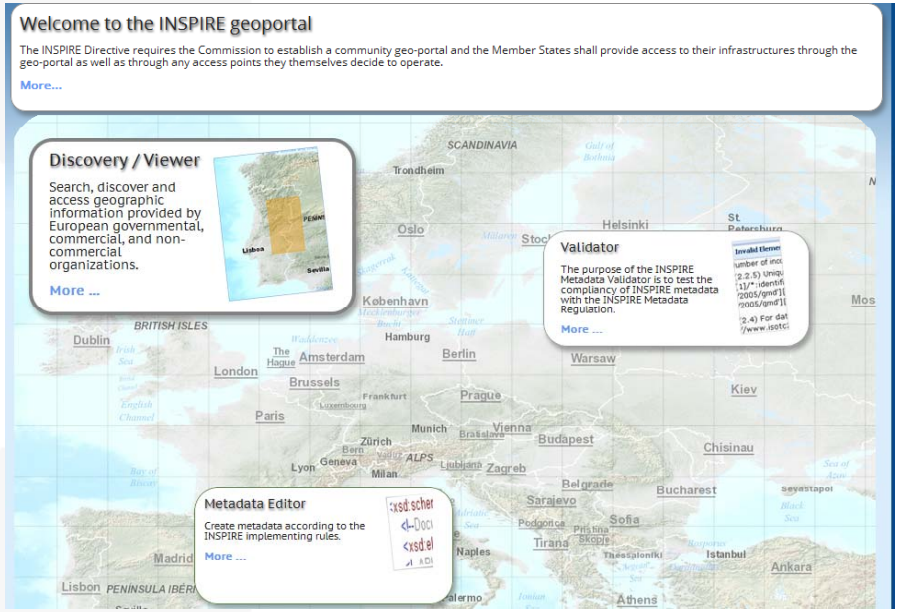
Credit: Bernadette Hyland : http://www.slideshare.net/3roundstones

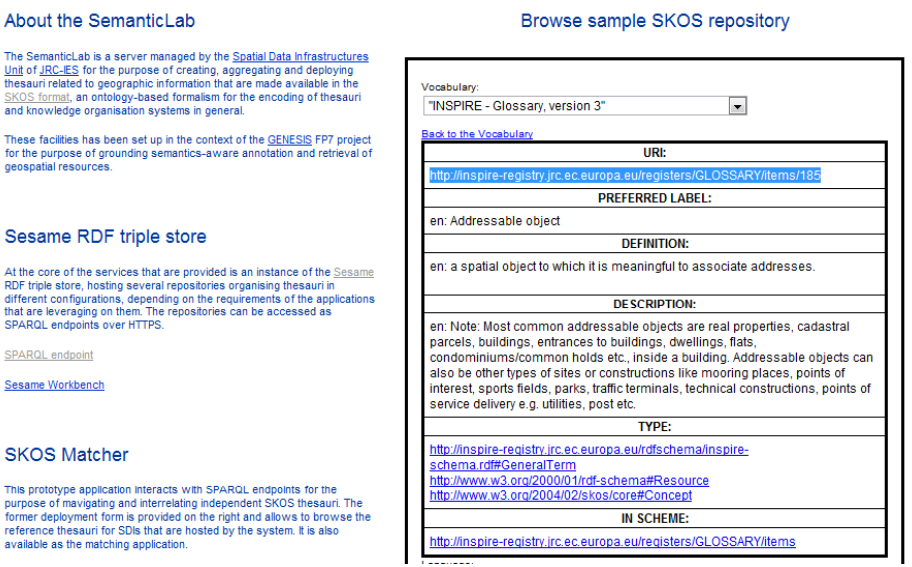
Three stars: Specialist formats

- Specialist tools often have specialist formats
  - Few people have the tools
  - Expensive
  - Difficult to re-use
  - Geospatial tools, statistical packages, etc..
- Use Open standards
  - CSV, JSON, XML, GML, OGC Web services, TSV, RDF

Credit: Richard Ciganiak: http://www.slideshare.net/cygri/edf2012-the-web-of-data-and-its-five-stars

Towards INSPIRE (5-Stars) Dataset?



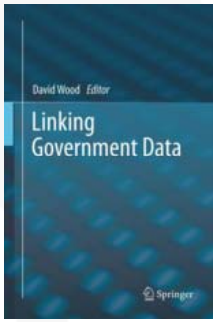


* INSPIRE GeoPortal (to find datasets)
http://inspire-geoportal.ec.europa.eu/discovery/

SKOS Glossary for Spatial Data Infrastructure:
http://semanticlab.jrc.ec.europa.eu/

http://www.flickr.com/photos/wwworks/4759535950/

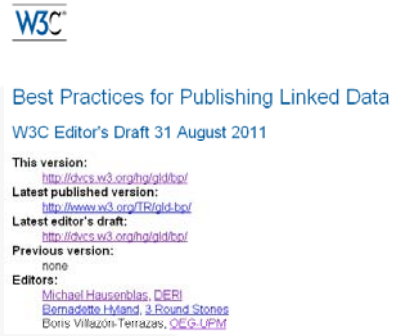
Some references



Wood, David (Ed) Linking Government Data - 2011

Methodological Guidelines for Publishing Government Linked Data

Boris Villazón-Terrazas, Luis M. Vilches, Oscar Corcho, Asunción Gómez-Pérez


Best Practices for Publishing Linked Data

W3C Editor's Draft 31 August 2011

This version:
http://dvcs.w3.org/hg/gld/bp/
Latest published version:
http://www.w3.org/TR/gld-bp/
Latest editor's draft:
http://dvcs.w3.org/hg/gld/bp/
Previous version:
none
Editors:
Michael Hausenblas, DERI
Bernadette Hyland, 3 Round Stones
Boris Villazón-Terrazas, OEG-UPM

Best Practices for Publishing Linked Data

W3C Editor's Draft – Government Linked Data Working Group

Bernadette Hyland, Boris Villazón-Terrazas, Michael Hausenblas,

https://dvcs.w3.org/hg/gld/raw-file/default/bp/index.html


Cookbook for Open Government Linked Data

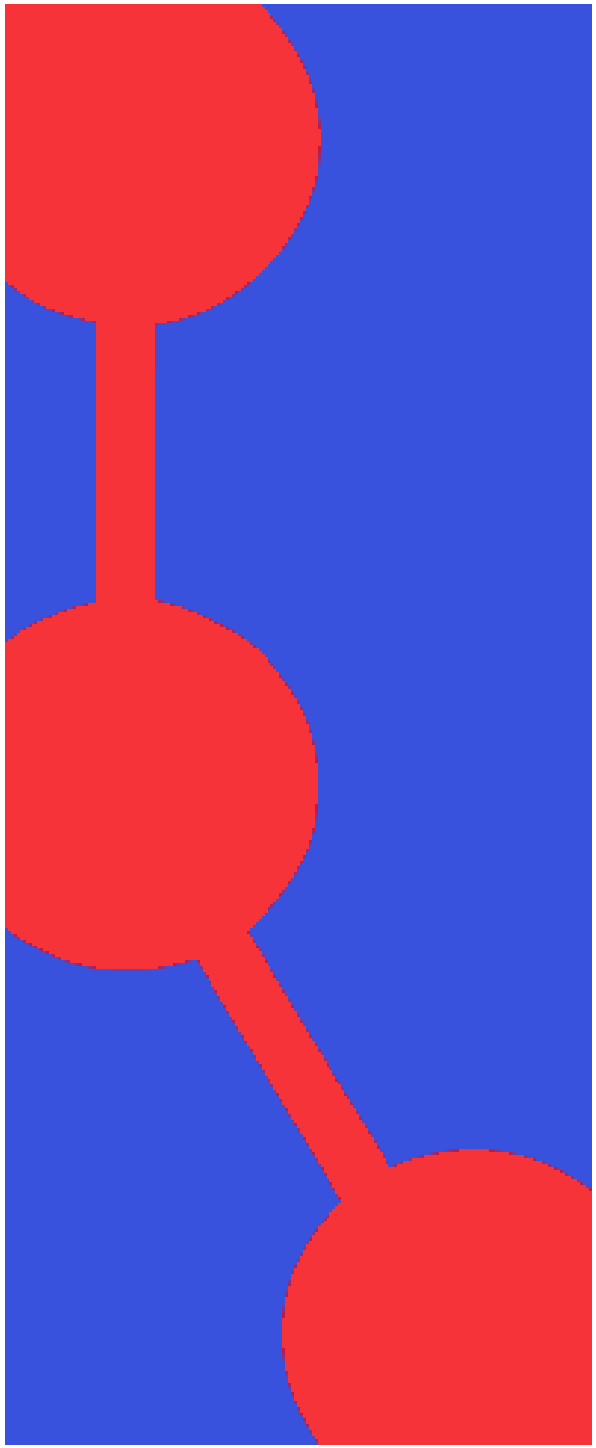
Editors:
Bernadette Hyland, (3 Round Stones)
Boris Villazón Terrazas (Universidad Politécnica de Madrid)
Sarven Capadisli (Digital Enterprise Research Institute)

Cookbook for Open Government Linked Data

W3C Editor's Draft – Government Linked Data Working Group

Bernadette Hyland, Boris Villazón-Terrazas, Sarven Capadisli

http://www.w3.org/2011/gld/wiki/Linked_Data_Cookbook



Lift your data

Introduction to Linked Data

Ghislain Atemezing[1] , Boris Villazón-Terrazas[2]

1 EURECOM, France

auguste.atemezing@eurecom.fr

2 Ontology Engineering Group, FI, UPM

bvillazon@fi.upm.es

Slides available at: http://www.slideshare.net/boricles/

INSPIRE Conference 2012, Istanbul

Acknowledgements: Oscar Corcho, Asunción Gómez-Pérez, Luis Vilches, Raphaël Troncy, Olaf Hartig, Luis Vilches and many others that we may have omitted.