

The Web of (Linked) Data

Oscar Corcho, Boris Villazón-Terrazas, Asunción Gómez-Pérez

Facultad de Informática, Universidad Politécnica de Madrid

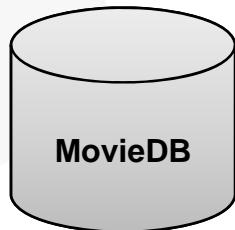
Campus de Montegancedo sn, 28660 Boadilla del Monte, Madrid

<http://www.oeg-upm.net>

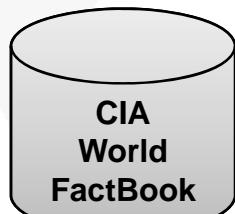
{ocorcho,bvillazon,asun}@fi.upm.es

Acknowledgements: Victor Saquicela, Guillermo Alvaro Rey, Olaf Hartig, Juan Sequeda, Luis M. Vilches and many others that we may have omitted.

Classic Web



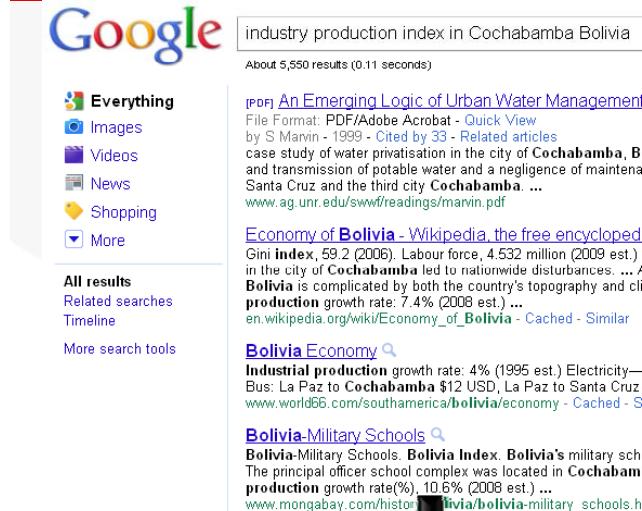
A screenshot of the IMDb website showing the movie "Even the Rain (2010)". The page includes the movie poster, title, release year, user rating (7.3/10), and director information. A large blue callout box overlaid on the page contains the text: "Data exposed to the Web via HTML, pdf, etc.".



A screenshot of the CIA World FactBook website for Bolivia. It shows a map of Bolivia, its flag, and some basic facts about the country. A large blue callout box overlaid on the page contains the text: "Data exposed to the Web via HTML, pdf, etc.".

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

Classic Web





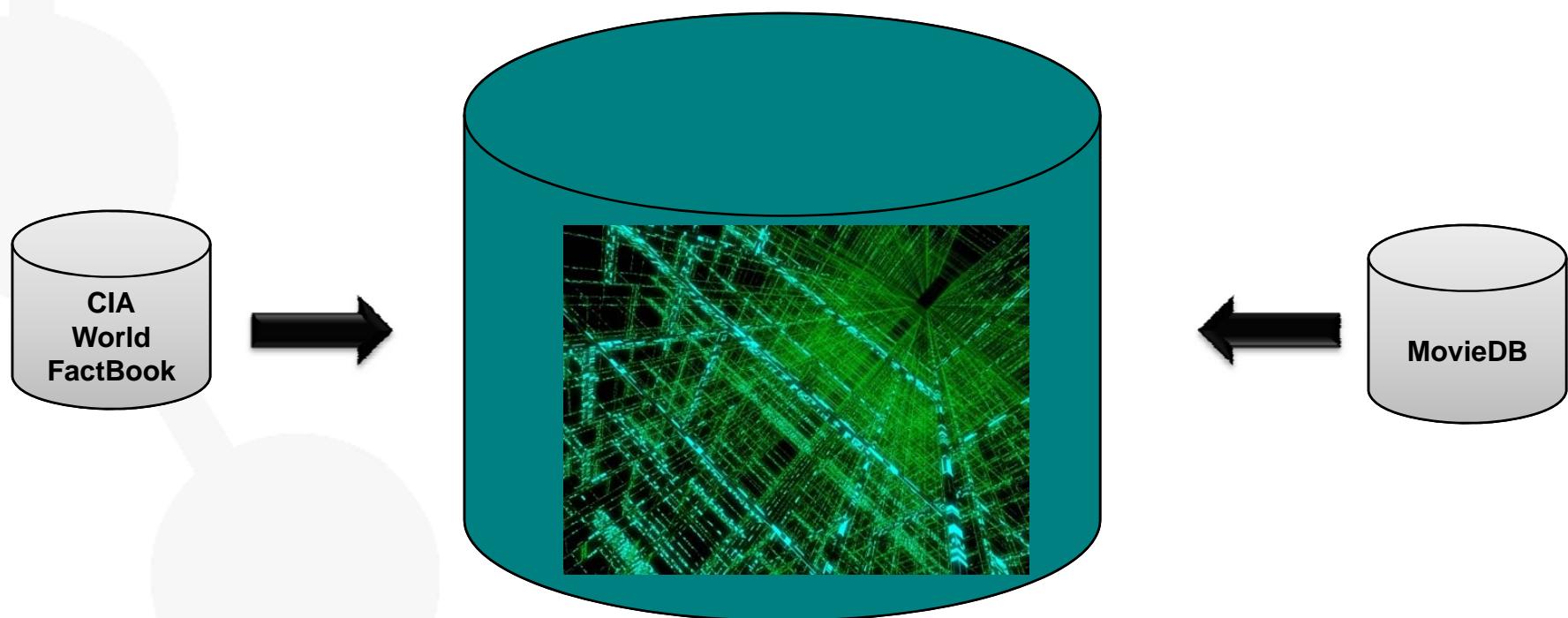
Complex queries over multiple pages / data sources?

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



What do we actually want?

- Use the Web like a single **global database**
 - Move from a Web of documents to a **Web of Data**



Let's try an example...

- Tell me which football players, born in the province of Albacete, in Spain, have scored a goal in the World Cup final

La Web Imágenes Vídeos Maps Libros Traductor Gmail Más ▾



jugadores de fútbol de la provincia de albacete

Aproximadamente 1,780 resultados (0.50 segundos)

Todo
Más

Mostrar herramientas de

[Mundial Sudáfrica 2010: Los nombres de la jornada](#)

Un gol que, como la pedazo de persona que también es, ... el del Mundial, además de haber sido, con cinco goles, ... ¿Pero FIFA que este hombre vale para arbitrar la final de un Mundial? Sido los últimos "nombres de la jornada" en ...

www.paraisofutbol.com/.../mundial-sudáfrica-2010-los-nombr

[Taringal - Delanteros argentinos Mundial 2010](#)

7 Abr 2010 ... Su habilidad y estilo de juego lo han comparado el 1 de mayo de la temporada 2004/2005 contra el Albacete B un jugador de fútbol argentino nacido el 2 de marzo de su máximo goleador de dicho torneo. ...

www.taringa.net/.../Delanteros-argentinos-Mundial-2010.html

[Ozil celebra con sus compañeros alemanes el gol m](#)

El jugador alemán Mesut Ozil celebra el gol encajado en el tiempo. Cerrar. Introduzca la ciudad o provincia ... El pulpo Pa Holanda en la final del Mundial de Sudáfrica. ... Los comentaristas intentaron de nuevo, pulsó la tecla F5 o el ...

www.que.es/.../fútbol/mundial/.../ozil-celebra-companeros-ale

En caché

[¿Puedes elegir a los héroes de 'La Roja'? -- Qué es](#)

13 Jul 2010 ... En la mayoría de los equipos de fútbol son los salvadores, muchas veces él solo contra el mundo. ... Los 2 títulos para describir sus cualidades. ... Mete el gol de la victoria. Mundial y va y se quita la zamarra para ...

www.que.es/.../fútbol/mundial/.../201007130924-puedes-elegir-a-los-héroes-de-la-roja

[Mostrar más resultados de www.que.es](#)

[Deportes | Mari Guerrero](#)

Hoy a las 12.30 gran final del mundial de basket: España VS pequeñas han jugado contra Albacete y han perdido digamos metido otra vez otros tantos goles, solo que esta vez las nuevitas por ejemplo por el físico de los futbolistas, ...

mari-21.blogia.com/temas/deportes.php - En caché



WIKIPEDIA
The Free Encyclopedia

Main page

Contents

Featured content

Current events

Random article

Interaction

About Wikipedia

Community portal

Recent changes

Contact Wikipedia

Donate to Wikipedia

Help

Toolbox

Print/export

Languages

العربية

Aragonés

Azərbaycan

Беларуская (тарашкевіца)

Български

Հայերեն

Historial web | Configuración de búsqueda | Acceder

New features Log in / create account

Article Discussion Read Edit View history Search

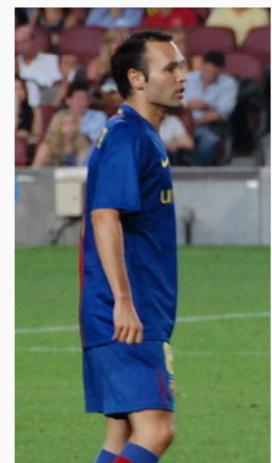
Andrés Iniesta

From Wikipedia, the free encyclopedia

This is a Spanish name; the first family name is Iniesta and the second is Luján.

Andrés Iniesta (Spanish pronunciation: [an'drés iñesta]; born 11 May 1984 in Fuentealbilla, Albacete, Castile-La Mancha) is a World Cup winning Spanish international football midfielder who currently plays for Spanish La Liga club FC Barcelona. His willingness to play anywhere on the pitch, coupled with a natural humility, has earned him the sobriquet *El Ilusionista* (The Dreamer), *El Anti-Galáctico* (The Anti-Galáctico), *Cerebro* (The Brain) and most recently *Don Andrés* from the Spanish press.^[1] After the 2009 UEFA Champions League Final, Manchester United striker Wayne Rooney asserted that he believed the midfielder to be the best player in the world.^[2] His current contract with Barcelona runs until 2015. He scored the winning goal for Spain in the 116th minute of the 2010 FIFA World Cup Final against the Netherlands and was the *Man of the Match* for the game.

Andrés Iniesta



Personal information

Full name	Andrés Iniesta Luján
Date of birth	11 May 1984 (age 26)
Place of birth	Fuentealbilla, Spain
Height	1.70 m (5 ft 7 in)
Playing position	Midfielder / Winger



Adapted from: Guillermo Alvaro Rey

It would be better to make a data query...

SPARQL Explorer for <http://dbpedia.org/sparql>

SPARQL:

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX : <http://dbpedia.org/resource/>
PREFIX dbpedia2: <http://dbpedia.org/property/>
PREFIX dbpedia: <http://dbpedia.org/>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT DISTINCT ?player ?village (
  ?s foaf:page ?player .
  ?s rdf:type <http://dbpedia.org/ontology/SoccerPlayer> .
  ?s skos:subject <http://dbpedia.org/resource/Category:UEFA_Euro_2008_players> .
  ?s <http://dbpedia.org/ontology/birthPlace> ?village .
  ?village skos:subject <http://dbpedia.org/resource/Category:Municipalities_in_Albacete> .
)
```

Results: [Browse](#) [Go!](#) [Reset](#)

SPARQL results:

player	village
http://en.wikipedia.org/wiki/Andr%C3%A9s_Iniesta	Fuentealbilla



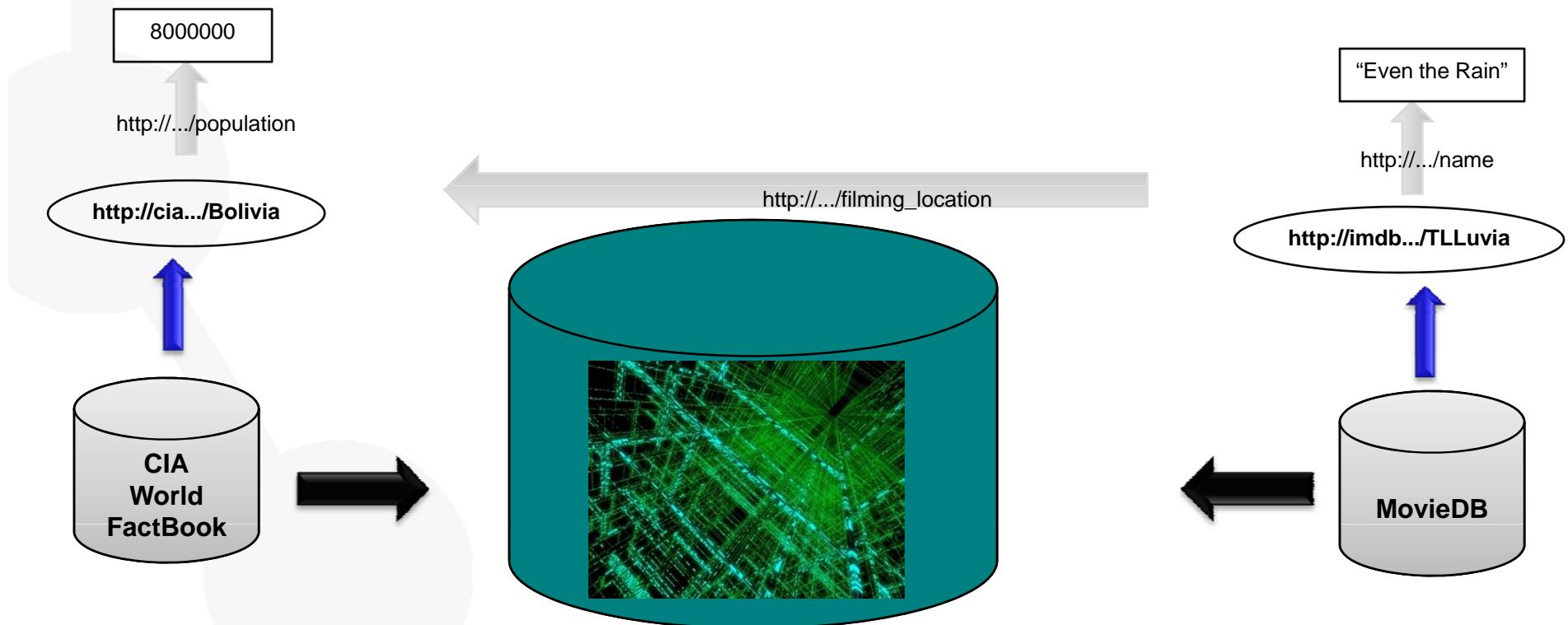
Linked Data enables such Web of Data

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**

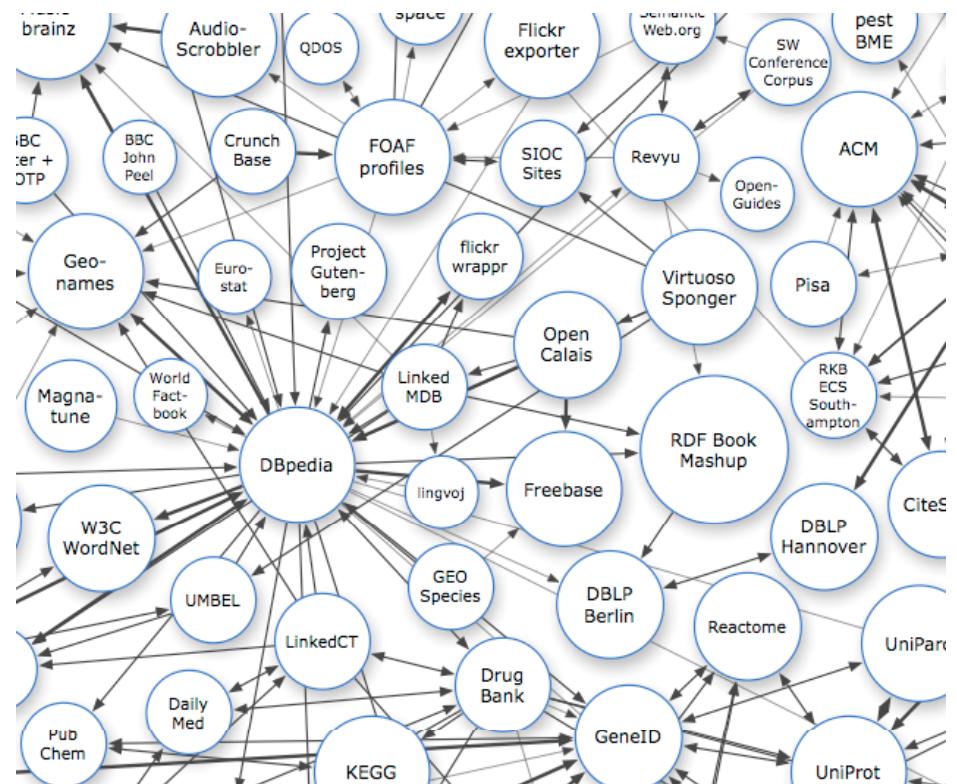
Connection: Typed Links



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

Linked Data in a nutshell

- An extension of the current Web...
 - ... where **data** are given **well-defined** and **explicitly represented** meaning,
 - ...
 - ... so that it can be **shared** and used by **humans and machines**, ...
 - ... better enabling them to work in cooperation
- How?
 - Promoting information exchange by **tagging web content** with machine processable descriptions of its meaning.
 - And **technologies** and **infrastructure** to do this
 - And **clear principles** on how to publish data



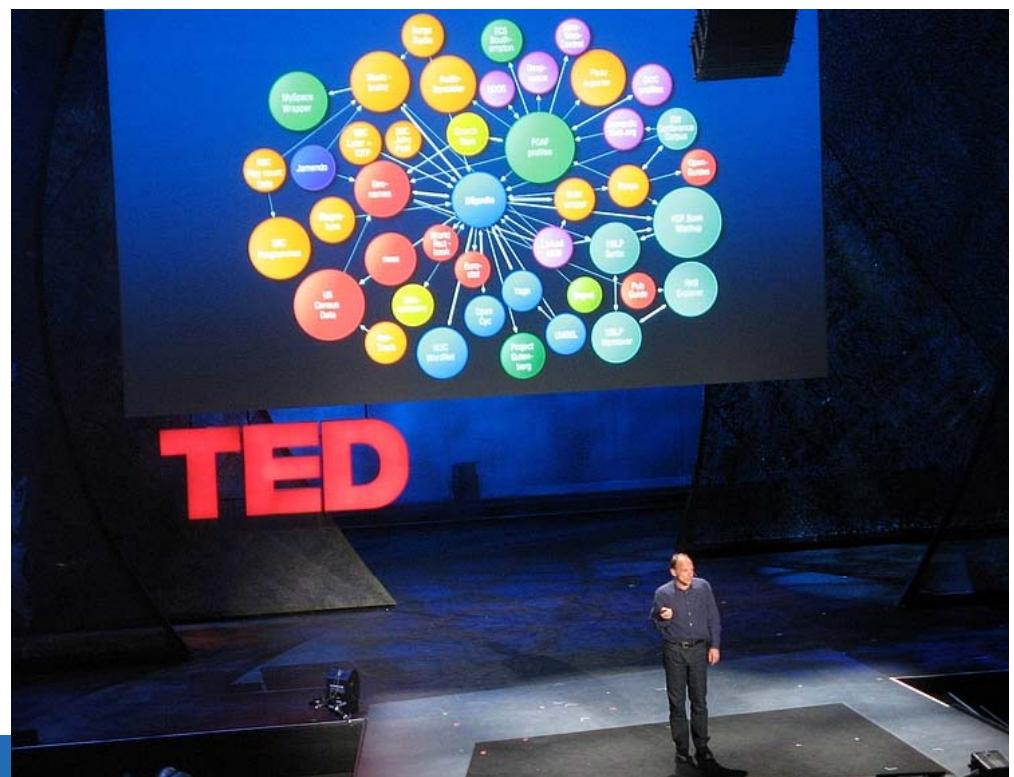
What is Linked Data?

- Linked Data is a term used to describe a recommended **best practice** for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF.
 - Part of the Semantic Web
 - Exposing, sharing and connecting data
 - Technologies: URIs and RDF (although others are also important)

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



So does that mean I have to publish my data as Linked Data, now?

- But, why?

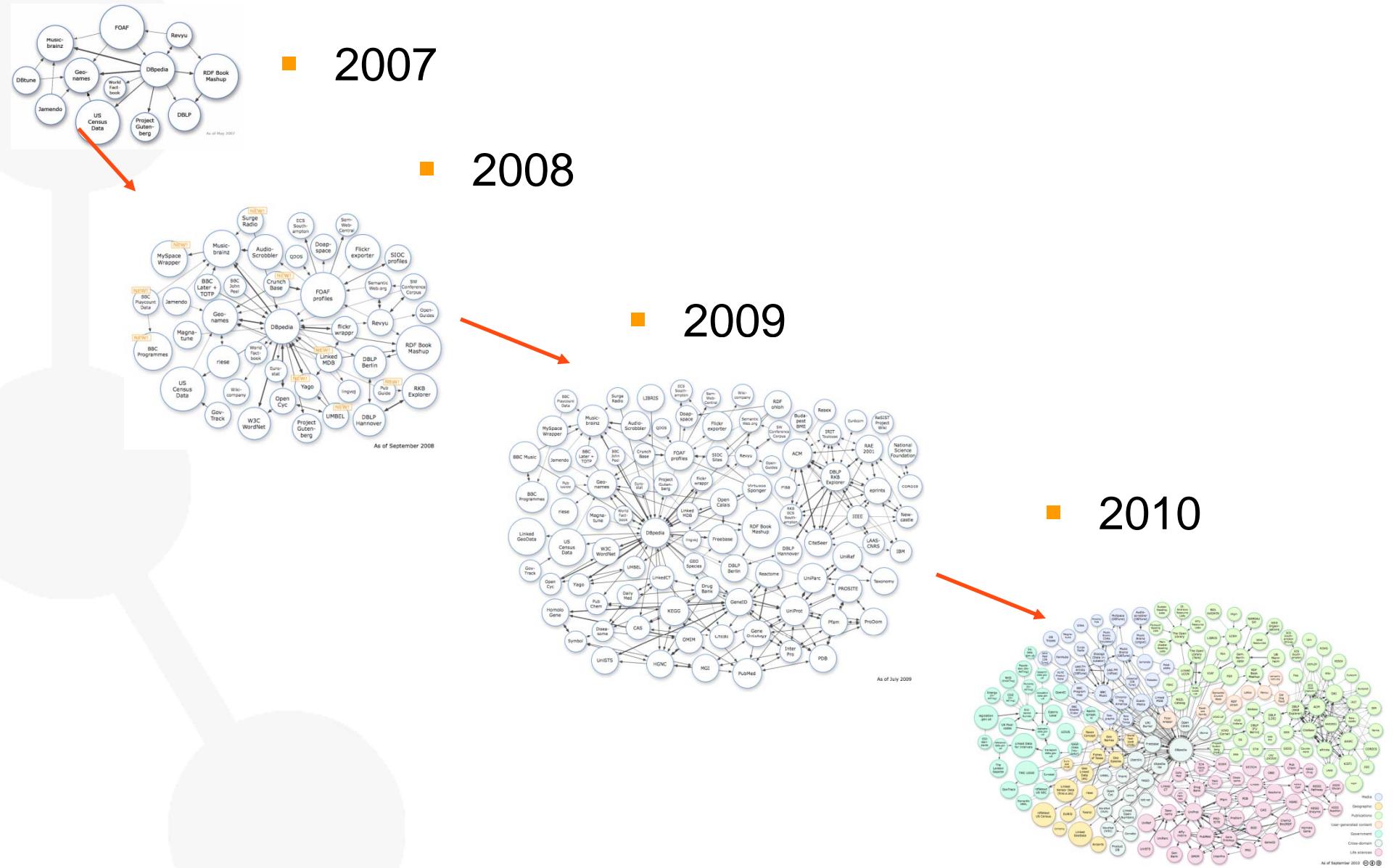


- What was your incentive to publish an HTML page in 1990?
 - Share data in documents and because your neighbor was doing it
- So, why should we publish Linked Data in 2011?
 - Share data as data and because your neighbor is doing it

And guess who is starting to publish Linked Data now?

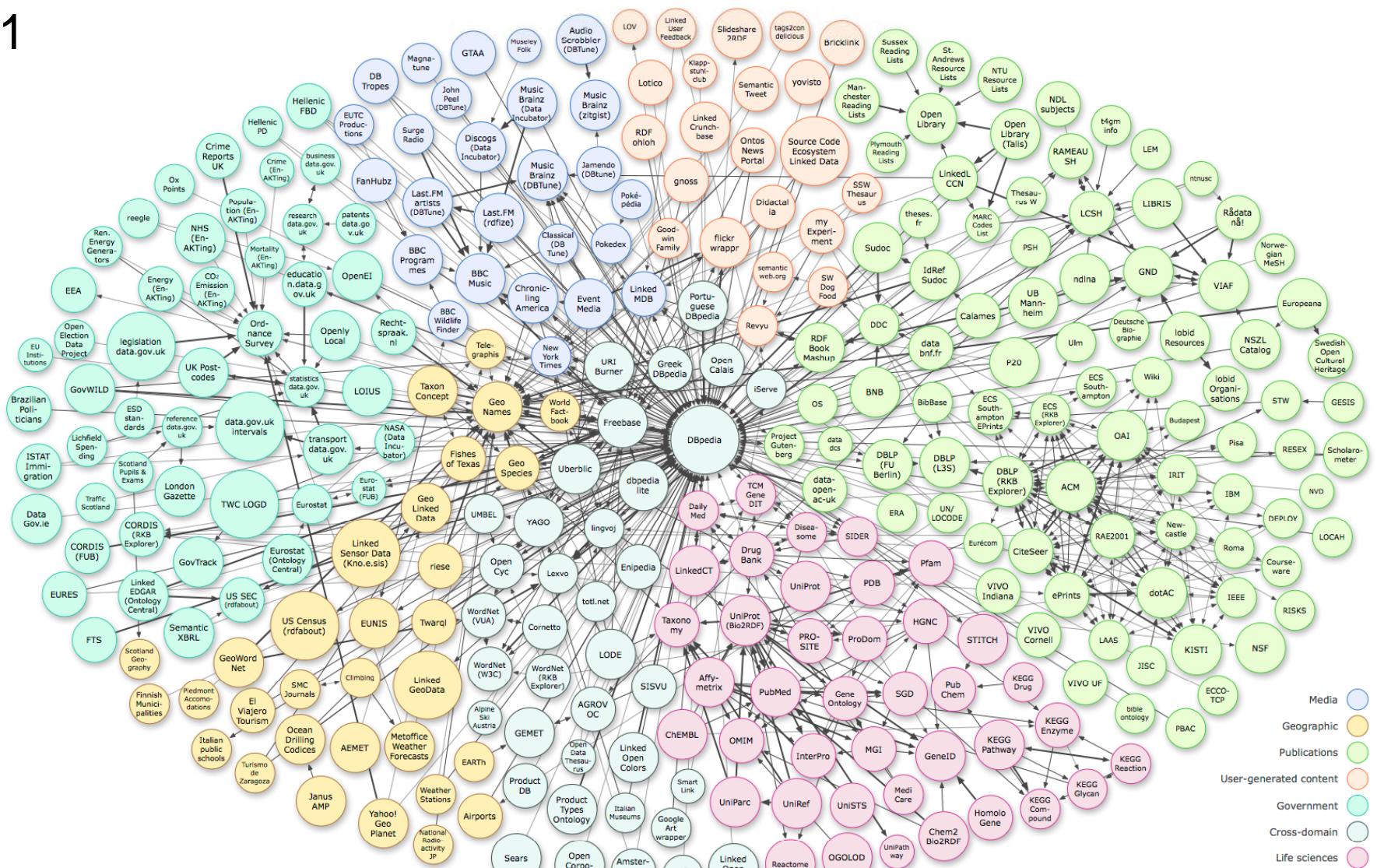
- UK Government
- US Government
- BBC
- Open Calais
- Freebase
- NY Times
- CNET
- Dbpedia
-

Linked Open Data evolution



Linked Open Data

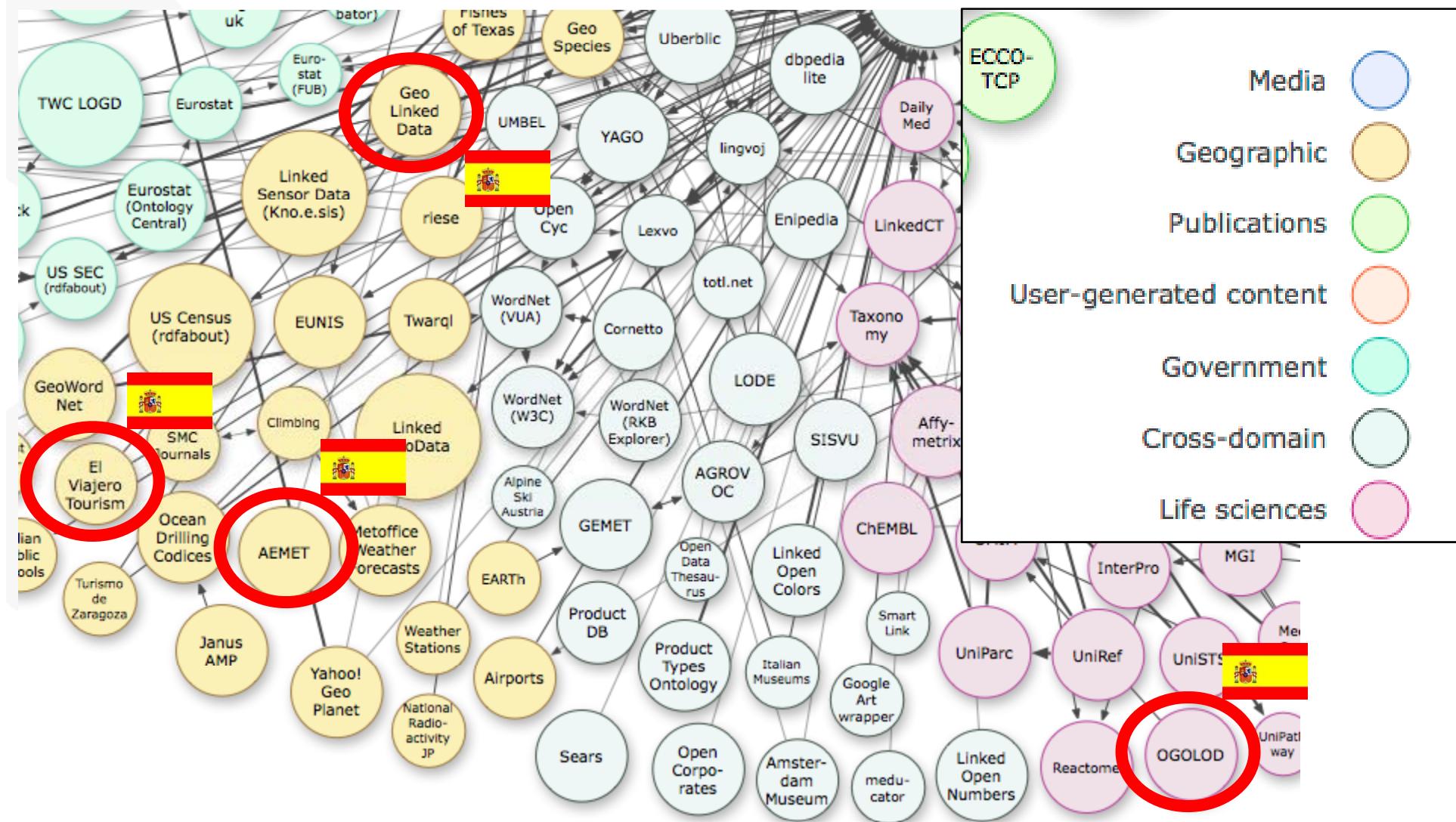
2011



As of September 2011 CC BY-SA

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

Linked Open Data – Some Spanish Datasets



Open Government Initiatives

- W3C eGovernment Activity
 - Improving Access to Government through Better Use of the Web
 - Publishing Open Government Data
 - W3C Government Linked Data WG
- Open Knowledge Foundation
 - Open Data Manual
- 5-star deployment scheme for Linked Open Data
 - ★ make your stuff available on the Web (whatever format) under an open license¹
 - ★★ 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 URLs to identify things, so that people can point at your stuff⁴
 - ★★★★★ link your data to other data to provide context⁵

Open Government. USA and UK

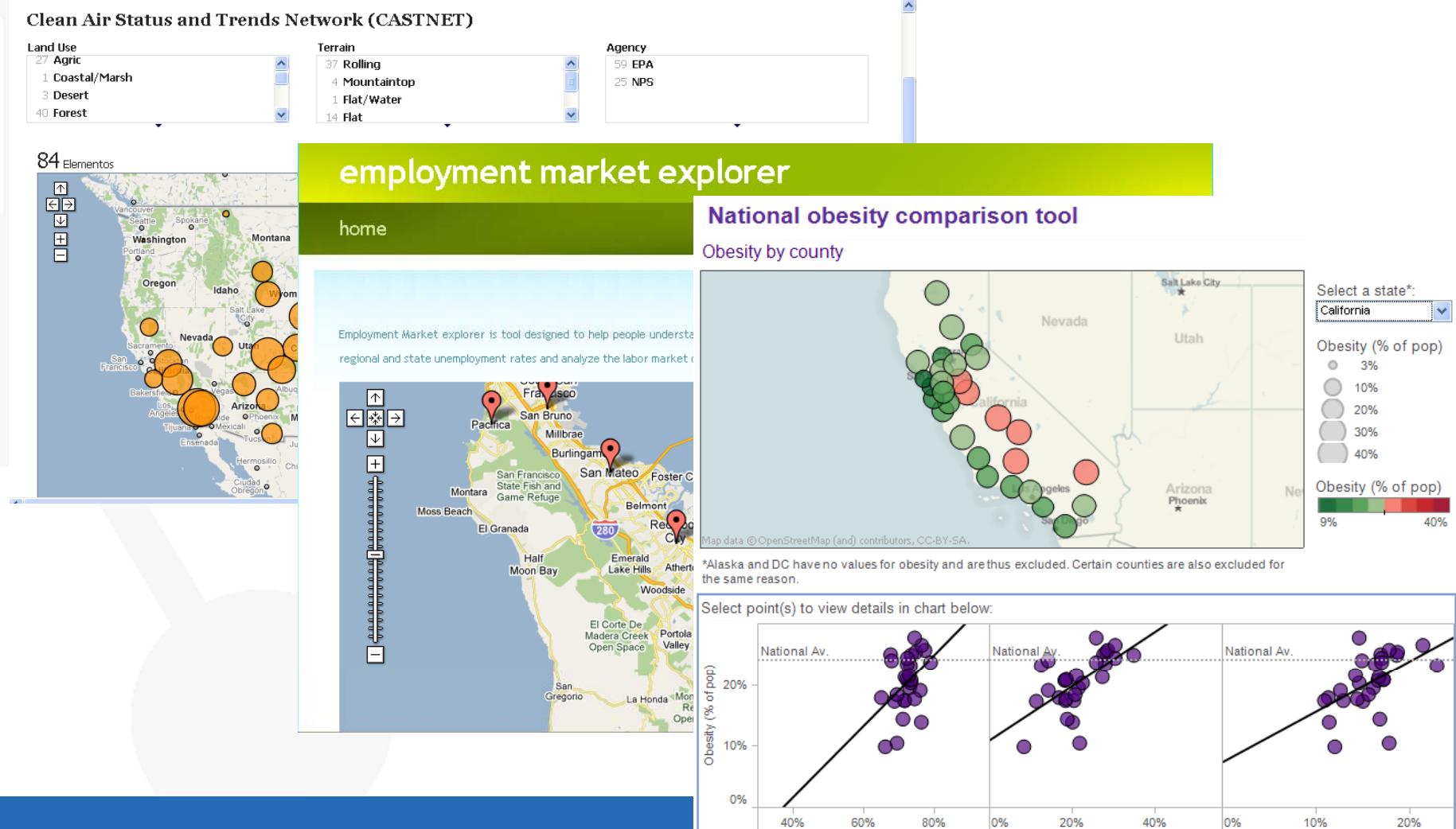
The screenshot shows the Data.gov.uk homepage. At the top, there's a banner for the "HAPPY 1ST ANNIVERSARY DATA.GOV" with the URL "data.gov.uk". To the right is a large image of a cupcake with one lit candle. Below the banner, there's a navigation bar with links for HOME, DATA, TOOLS, COMMUNITY, METRICS, and DIALOGUE. On the left, there's a sidebar with the HM Government logo and a "Unlocking innovation" section. The main content area has several sections: "Latest datasets" (listing dates like 2 July, 25 June, 18 June, and 14 June), "Community" (with links for "Log in / Sign up" and "Local Data Panel"), and "Semantic Web" (with a blue 3D cube icon). A large green banner at the bottom right reads "BOTTOM-UP". Another green banner at the bottom left reads "TOP-DOWN".

BOTTOM-UP

TOP-DOWN

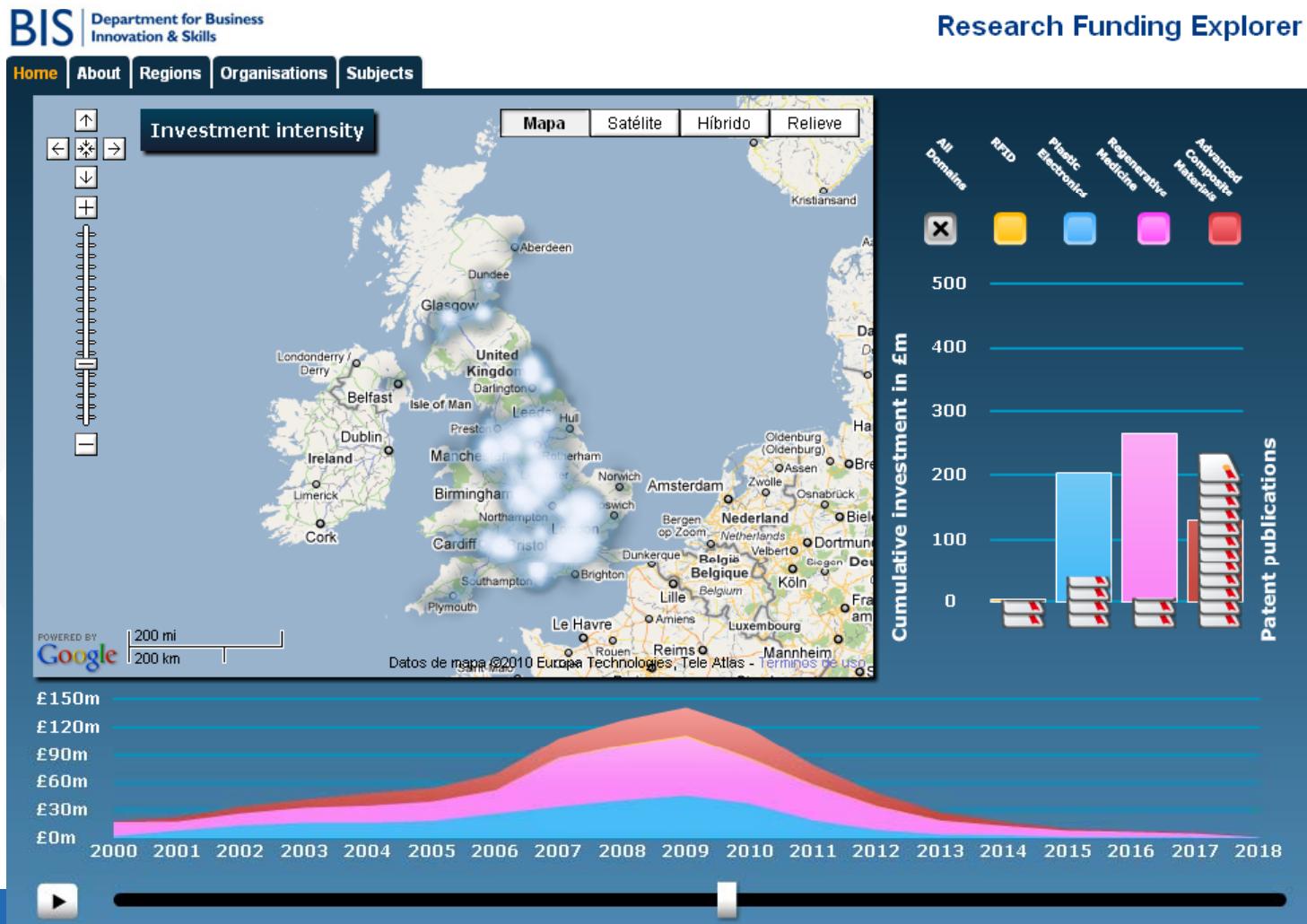
Linked Data Mashup (data.gov)

- Clean Air Status and Trends (CASTNET)
 - <http://data-gov.tw.rpi.edu/demo/exhibit/demo-8-castnet.php>



Linked Data Mashup (data.gov.uk)

- Research Funding Explorer
 - <http://bis.clients.talis.com/>



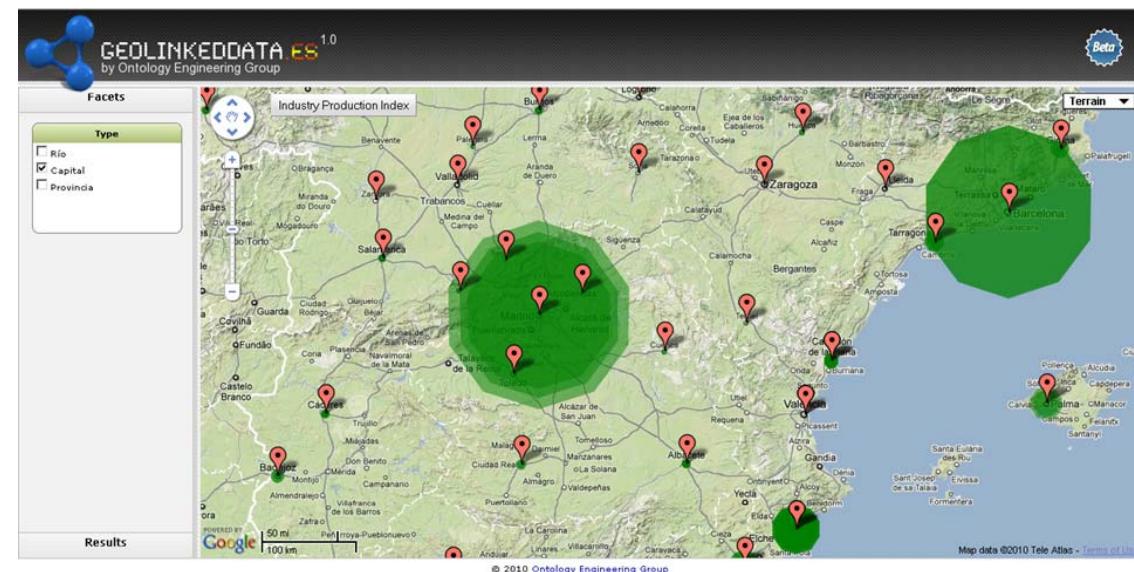
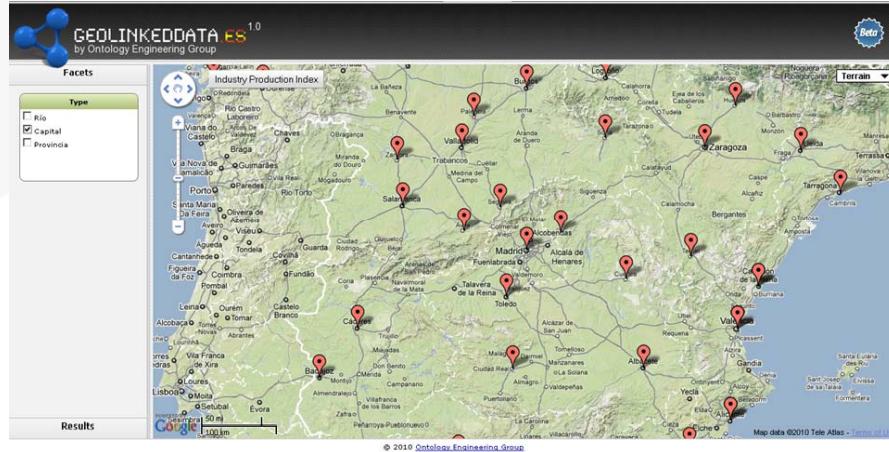
Designing URI sets for the Public Sector (UK)

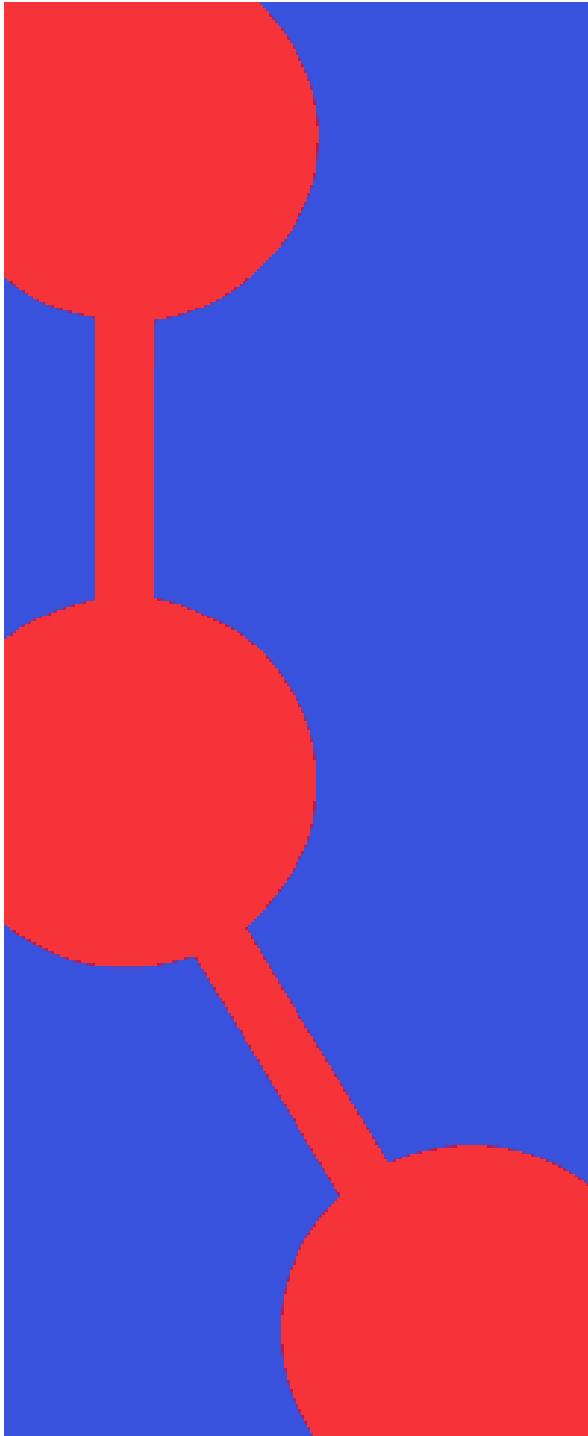
- http://www.cabinetoffice.gov.uk/media/301253/public_sector_uri.pdf

URI Type	URI structure	Examples
Identifier	http://{domain}/id/{concept}/{reference} or http://{domain}/{concept}/{reference}#{id}	http://education.data.gov.uk/id/school/78 http://education.data.gov.uk/school/78#id http://transport.data.gov.uk/id/road/M5/junction/24
Document	http://{domain}/doc/{concept}/{reference}	http://education.data.gov.uk/doc/school/78
Representation	http://{domain}/doc/{concept}/{reference}/{doc.file-extension}	http://education.data.gov.uk/doc/school/78/doc.rdf
Definition of the scheme concept	http://{domain}/def/{concept}	http://education.data.gov.uk/def/school
List of scheme identifiers	http://{domain}/doc/{concept}	http://education.data.gov.uk/doc/school
Set	http://{domain}/set/{concept}	http://education.data.gov.uk/set/school

- Education
 - <http://education.data.gov.uk/id/school/106661>
- Parliament
 - <http://parliament.psi.enakting.org/id/member/1227>
- Maps
 - E.g., London:
<http://data.ordnancesurvey.co.uk/id/7000000000041428>
 - <http://map.psi.enakting.org>
- Transport
 - <http://www.dft.gov.uk/naptan/>
- SameAs service
 - <http://www.sameas.org>
- Challenges
 - <http://gov.tso.co.uk/openup/sparql/gov-transport>

Linked Data in Spain. Example





The Web of (Linked) Data

Oscar Corcho, Boris Villazón-Terrazas, Asunción Gómez-Pérez

Facultad de Informática, Universidad Politécnica de Madrid

Campus de Montegancedo sn, 28660 Boadilla del Monte, Madrid

<http://www.oeg-upm.net>

{ocorcho,bvillazon,asun}@fi.upm.es

Acknowledgements: Luis M. Vilches, Victor Saquicela, Guillermo Alvaro Rey, Olaf Hartig, Juan Sequeda, and many others that we may have omitted.