

Linked Data I

Oscar Corcho, Asunción Gómez Pérez ({ocorcho, asun}@fi.upm.es)
Universidad Politécnica de Madrid

Credits: Raúl García Castro, Guillermo Alvaro, Oscar Muñoz, Jose Angel Ramos Gargantilla, María del Carmen Suárez de Figueroa, Boris Villazón-Terrazas, Alex de León, Víctor Saquicela, Luis Vilches, Miguel Angel García, Manuel Salvadores, Juan Sequeda, Carlos Ruiz Moreno and many others

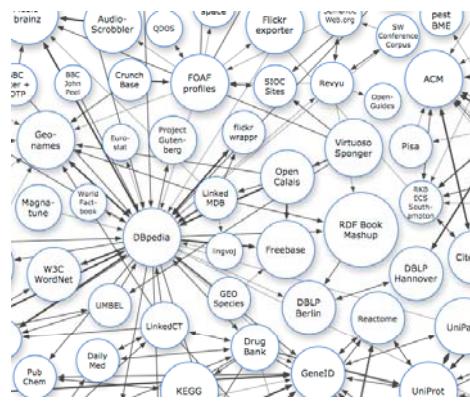
*Work distributed under the license Creative Commons
Attribution-Noncommercial-Share Alike 3.0*

Contents

- [Introduction to Linked Data](#)
- [Linked Data publication](#)
 - [Methodological guidelines for Linked Data publication](#)
 - [RDB2RDF tools](#)
 - [Technical aspects of Linked Data publication](#)
- [\[Linked Data consumption\]](#)

What is the Web of Linked Data?

- 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

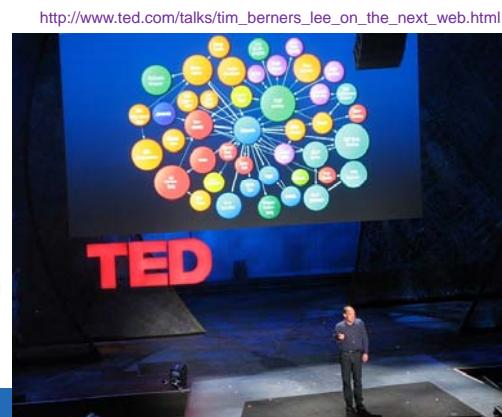


What is Linked Data?

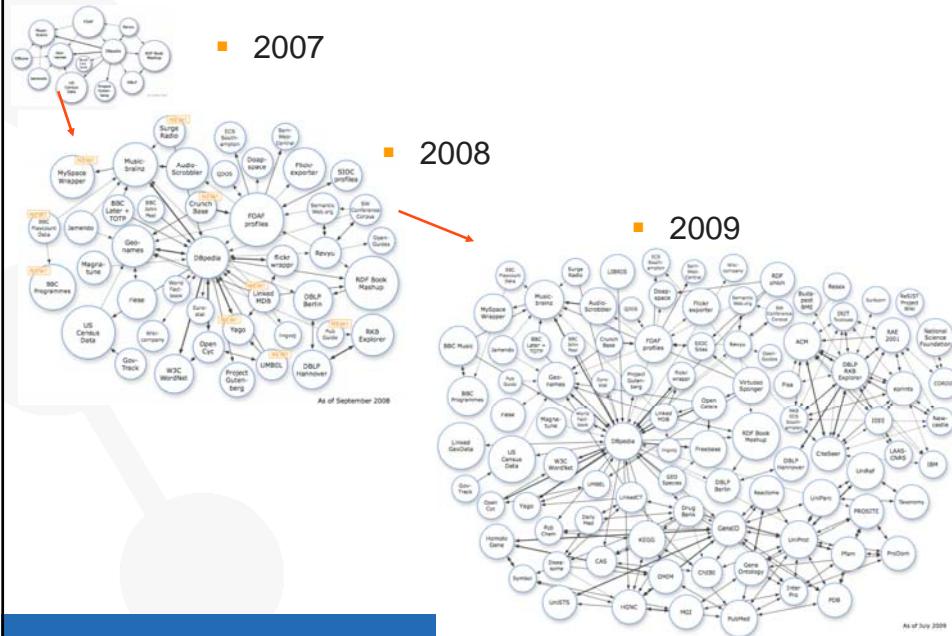
- Linked Data is a term used to describe a recommended **best practice** for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF.
 - Part of the Semantic Web
 - Exposing, sharing and connecting data
 - Technologies: URIs and RDF (although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

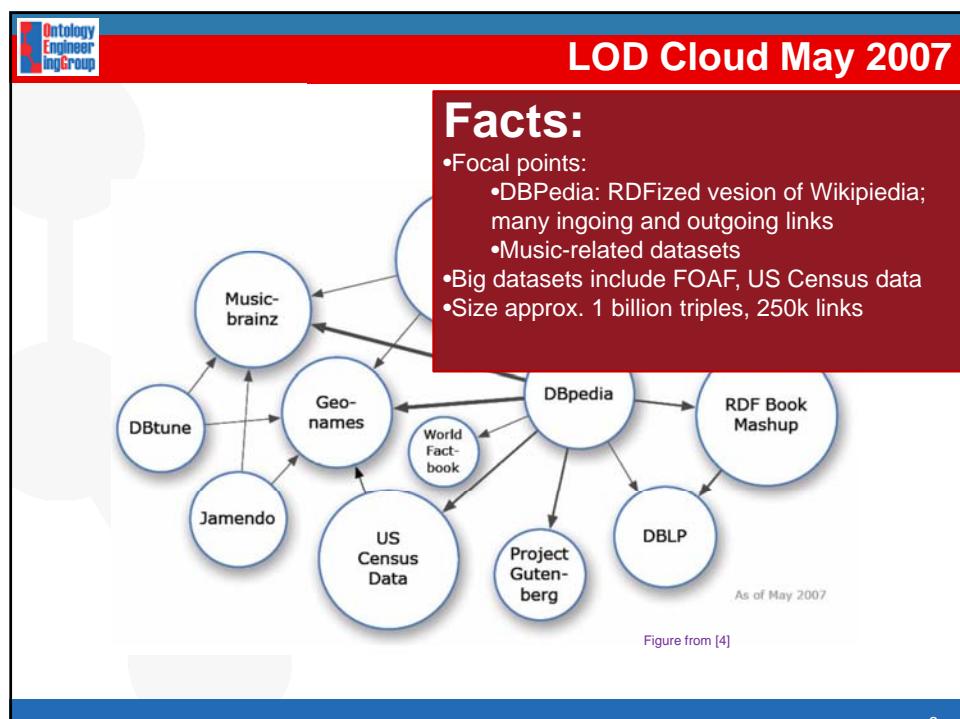
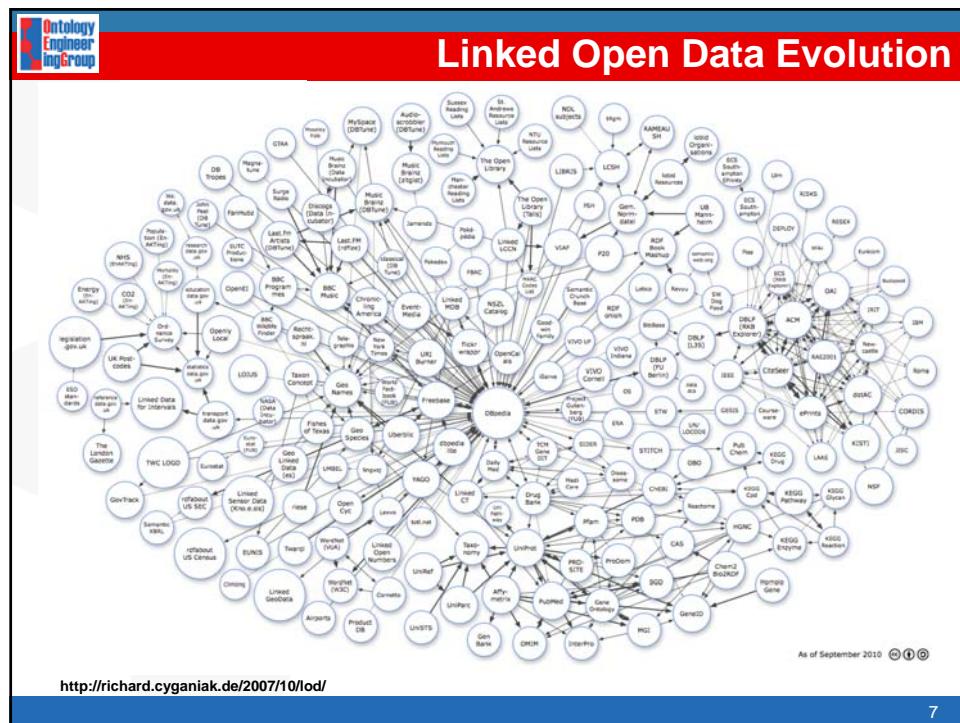
The four principles (Tim Berners Lee, 2006)

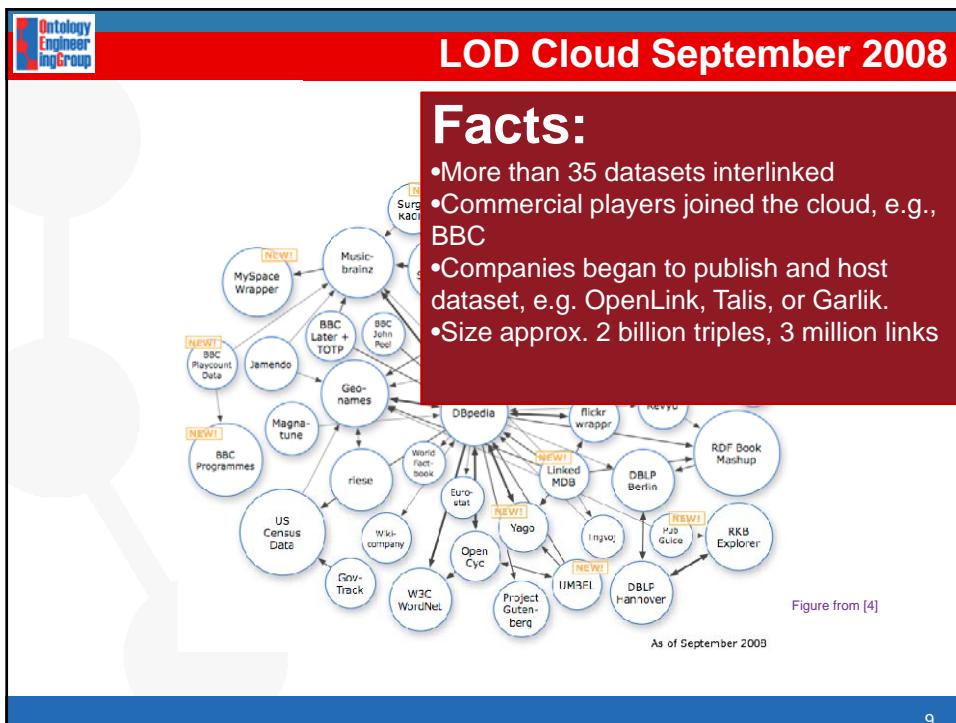
1. Use URIs as names for things
 2. Use HTTP URIs so that people can look up those names.
 3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
 4. Include links to other URIs, so that they can discover more things.
- <http://www.w3.org/DesignIssues/LinkedData.html>



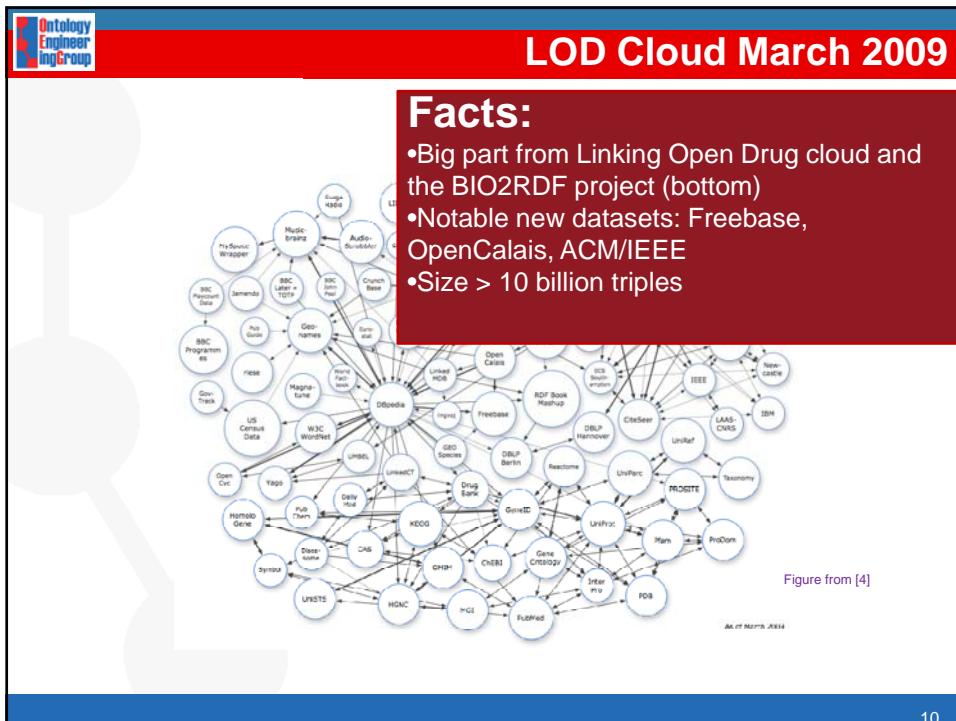
Linked Open Data evolution



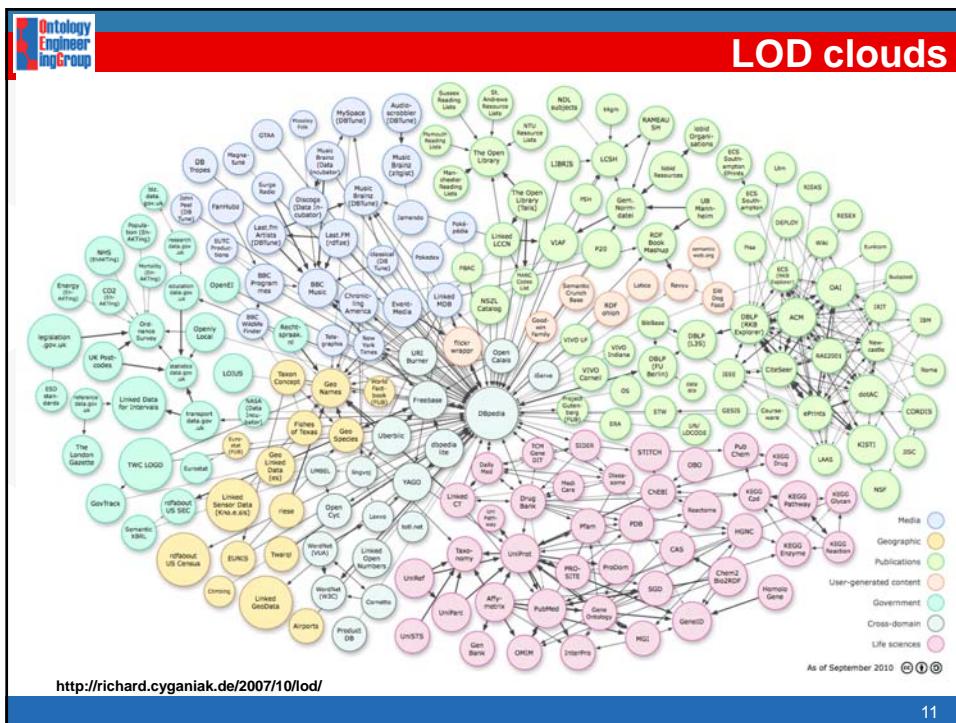




9



10



How should we publish data?

- Formats in which data is published nowadays...
 - XML
 - HTML
 - DBs
 - APIs
 - CSV
 - XLS
 - ...
- However, main limitations from a Web of Data point of view
 - Difficult to integrate
 - Data is not linked to each other, as it happens with Web documents.

How do we publish Linked Data?

1. Exposing Relational Databases or other similar formats into Linked Data
 - D2R
 - Triplify
 - R2O & ODEMapster
 - Virtuoso
 - Ultrawrap
 - NOR2O
 - ...
2. Using native RDF triplestores
 - Sesame
 - Jena
 - OwlIM
 - Talis platform
 - ...
3. Incorporating it in the form of RDFa in CMSs like Drupal

13

How do we consume Linked Data?

- Linked Data browsers
 - To explore things and datasets and to navigate between them.
 - Tabulator Browser (MIT, USA), Marbles (FU Berlin, DE), OpenLink RDF Browser (OpenLink, UK), Zitgist RDF Browser (Zitgist, USA), Disco Hyperdata Browser (FU Berlin, DE), Fenfire (DERI, Ireland)
- Linked Data mashups
 - Sites that mash up (thus combine Linked data)
 - Revyu.com (KMI, UK), DBtune Slashfacet (Queen Mary, UK), DBpedia Mobile (FU Berlin, DE), Semantic Web Pipes (DERI, Ireland)
- Search engines
 - To search for Linked Data.
 - Falcons (IWS, China), Sindice (DERI, Ireland), MicroSearch (Yahoo, Spain), Watson (Open University, UK), SWSE (DERI, Ireland), Swoogle (UMBC, USA)

Linked Data browsers (Disco)

The screenshot shows a Windows Internet Explorer window displaying a resource page from the Disco linked data browser. The URL is <http://www4.wiwiss.fu-berlin.de/dbp/resource/person/315759>. The page title is "Christian Bizer". The interface includes a navigation box, a resource description table, a list of source graphs, and a session cache link.

Property	Value	Sources
more data	List of all instances: People	G2
type	http://mmsn.com#t0f1#Person	G2
label	Christian Bizer	G2
sourceURL	Christian Bizer	G2
name	Christian Bizer	G2
is Creator of	Business to Consumer Markets on the Semantic Web	G2, G7
is Creator of	Semantic Web Publishing using Named Graphs	G2, G7
is Creator of	The Impact of Semantic Web Technologies on Job Recruitment Processes	G2, G4
is Creator of	DR MAP - A Database to RDF Mapping Language	G2, G7
is Creator of	Using context- and content-based trust policies on the semantic web	G2, G7
is Creator of	Named graphs, provenance and trust	G2, G7
is Creator of	A Web Service Market Model based on Dependencies	G2, G7
is Creator of	Named graphs	G2, G5
is sameAs	Chris Bizer	G3
is sourceURL of	Christian Bizer	G2

Labels and Boxes:

- Label of the displayed resource:** Christian Bizer
- Navigation box:** Includes links for "Home", "About", "Logout", and "Help".
- Resource description:** A table showing properties, values, and sources for the resource.
- List of all source graphs:** A list of RDF graphs used to display the information.
- Link for displaying the session cache:** A link to view the session cache.
- Go to URI button:** A button to go to the URI of the displayed resource.
- Sources of each piece of information:** A list of source URLs for each piece of information.

Linked Data Mashup (LinkedGeoData)

The screenshot shows a mashup application for LinkedGeoData. It features a map of Berlin with various geographical features and data overlays. On the left, there is a sidebar with a class hierarchy tree, a search bar, and a results table. The results table lists 11 highway traffic signals found in Berlin.

Index	Location
1	highway:traffic_signals
2	highway:traffic_signals
3	highway:traffic_signals
4	ReichskanzlerAllee
5	highway:traffic_signals
6	highway:traffic_signals
7	highway:traffic_signals
8	highway:traffic_signals
9	highway:traffic_signals
10	highway:turning_circle
11	highway:traffic_signals

Labels and Icons:

- Class Hierarchy:** Shows the class hierarchy for highway traffic signals.
- Instances:** Shows the count of instances (330) and the fact that some results have been omitted.
- Search:** A search bar with the placeholder "Search" and a "Powered by Planlayer" logo.
- Results:** A table showing the index, location, and name for each found highway traffic signal.
- Map:** A map of Berlin with red markers indicating the locations of the found highway traffic signals.
- View:** A modal dialog showing details for a specific location, including Name, Description, and Source ref.
- Links:** A magnifying glass icon over a map, and a link to <http://linkedgeo.org/>.

Ontology Engineer IngGroup

Linked Data Mashup (DBpedia Mobile)

<http://wiki.dbpedia.org/DBpediaMobile>

- Displays Wikipedia data on a map
- Smushes the data with data from other sources

Ontology Engineer IngGroup

Linked Data Search Engines (Sindice and SIG.MA)

- Entity lookup service. Find a document that mentions a URI or a keyword.

<http://sindice.com/>

<http://sig.ma/>

Ontology Engineer Group

Linked Data Search Engines (NYT)

- The New York Times: Alumni In The News
 - <http://data.nytimes.com/schools/schools.html>

Stanford University
ANALYSIS; Kelly Rushes Up The Irish Wish List - December 08, 2009 Attracting Valuable Coaches to the Prettiest College Town - November 10, 2007 As Usual, U.S.C. Is Loaded, But Now It's on Defense - July 27, 2007 Sports Briefing - December 19, 2006 PLUS: PRO FOOTBALL -- SAN DIEGO; Chargers Land Harbaugh in Deal - March 18, 1999 PRO FOOTBALL; Harbaugh Says His Injury Won't Sideline Him for Jets - September 08, 1998 PRO FOOTBALL; Jets Rely On Quiet Newcomer - August 16, 1998 PRO FOOTBALL; The New Line On Harbaugh: He Has a Line - August 05, 1998 PRO FOOTBALL: NOTEBOOK; Quarterbacks on Merry-Go-Round in A.F.C. East - February 15, 1998 PRO FOOTBALL; Harbaugh Dealt Baltimore Ravens - February 15, 1998
John Elway Born: June 28, 1960
PRO FOOTBALL; Elway Was a God. Plummer? He's Trying. - January 21, 2006

Ontology Engineer Group

Linked Data Search Engines (NYT)

- The New York Times: Source code is available


```
// SPARQL query that retrieves the Freebase ID(s) associated with an item
// in DBpedia
var sparql =
  "SELECT ?freebaseUri WHERE {" +
    "?s" + schoolUri + "> owl:sameAs ?freebaseUri FILTER regex(?freebaseUri, 'http://rdf\\\\\\\.freebase\\\\\\\.com/.+') " +
  "}"';

// DBpedia URI for executing the sparql query.
var dbpediaUrl =
  "http://dbpedia.org/sparql?default-graph-uri=http%3A%2F%2Fdbpedia.org&query=" +
  escape(sparql) +
  "&format=json";
```
- ... and is based on SPARQL queries



One additional motivation: Open Government

- Government and state administration should be opened at all levels to effective public scrutiny and oversight
- Objectives:
 - Transparency
 - Participation
 - Collaboration
 - Inclusion
- Cost reduction
 - Interoperability
 - Reusability
- Leadership
 - Market & Value

Some Links:

- B. Obama - [Transparency and Open Government](#)
- T. Berners-Lee - [Raw data now!](#)
- J. Manuel Alonso - [¿Qué es Open Data?](#)
- [Open Government Data](#)
- [8 Principles of Open Government Data](#)

21



Open Government. USA and UK

HM Government

Unlocking innovation
Working with UK Public Sector Information and data

Latest datasets

3 July Public sector savings over £100,000 now also covers NDPRs
25 June Costs and other data about central government websites
18 June Estimated government workforce including consultants
14 June More COINS public spending data now covers 2005 to 2010

What we do

data.gov.uk is a key part of the Government Transparency programme. It is the public-facing side of a wider policy, with Sir Tim Berners-Lee, Professor Ian Goldin and Tim Berners-Lee and others involved. The Cabinet Office Ministers new Public Sector Transparency Board will also seek to give a very into the wealth of government data becoming available. It's under constant development and we want to work with you to make it better.

See also

TOP-DOWN

DATA.GOV

HAPPY 1ST ANNIVERSARY DATA.GOV data.gov.uk

Community Log in / Sign up Local Data Panel What is the semantic Web

SEARCH OUR CATALOGS

Most Popular Datasets

- 1. U.S. Overseas Loans and Grants (Greenbook)
- 2. Worldwide Min+Earthquakes, Past 7 Days
- 3. Latest Volumes of Foreign Relations of the...
- 4. Local Area Unemployment Statistics
- 5. Personal Trends by Gender/Race

As the Web of linked documents moves to the next stage of limited data, we're working to maximize the potential of the Semantic Web technologies to realize the promise of

BOTTOM-UP

22

Linked Data Mashup (data.gov)

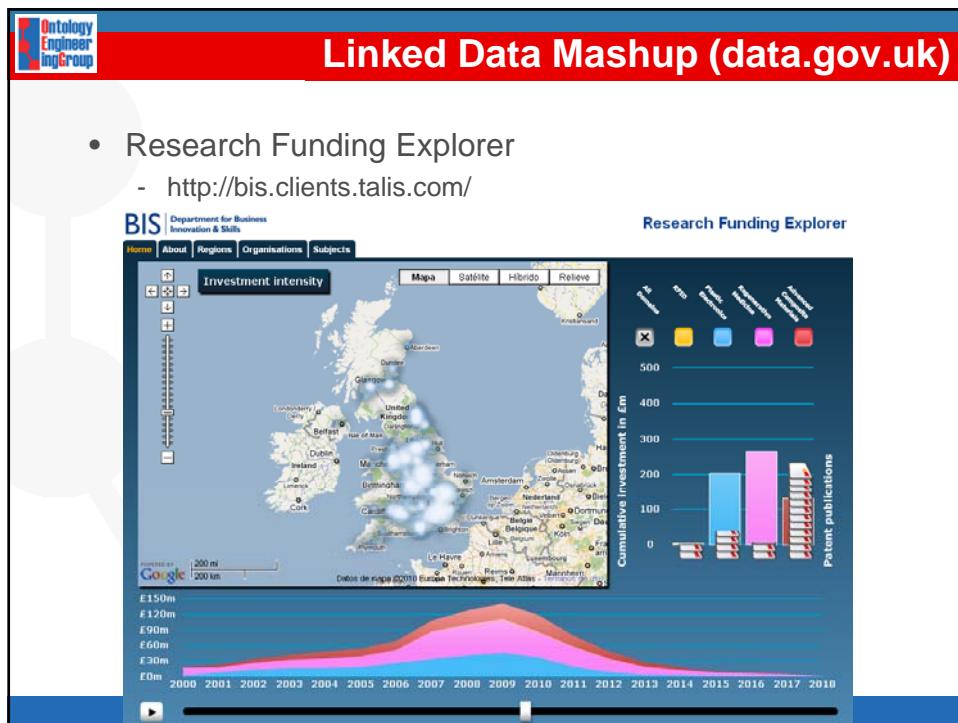
- Clean Air Status and Trends (CASTNET)
 - <http://data-gov.tw.rpi.edu/demo/exhibit/demo-8-castnet.php>

The screenshot displays three distinct web applications side-by-side:

- Clean Air Status and Trends Network (CASTNET):** A map of the western United States showing various locations with orange circles of varying sizes, representing different data points. A legend indicates Land Use (Agric, Coastal/Marsh, Desert, Forest) and Terrain (Ridge, MountainTop, Flat/Water, Flat). A color key shows Agency (EPA, NPS).
- employment market explorer:** A map of San Francisco Bay Area showing unemployment rates across different neighborhoods.
- National obesity comparison tool:** A map of California showing obesity rates by county. A legend indicates Obesity (% of pop) from 0% to 40%. A scatter plot below shows Obesity (% of pop) on the Y-axis versus another metric on the X-axis, with data points clustered around a diagonal line.

Linked Data in the UK

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



Ontology Engineer IngGroup

Open Government Spain. Abredatos

The image shows a smartphone displaying the Infocarretera app's user interface. The screen shows a map with road segments labeled 'Altura 2 (Baja)' and 'Altura 3 (Alta)'. To the right of the phone, there is a stylized illustration of a road with cars, a bus, and a traffic cone under a blue sky with clouds. The Infocarretera logo is prominently displayed.

infocarretera
toda la información de la red carreteras de euskadi en tu bolígrafo

PRIMER PREMIO
AbreDatos
DATAHACKS

infocarretera te acerca el estado de las

<http://www.infocarretera.com/>

27

Ontology Engineer IngGroup

Open Government Spain. Zaragoza

The screenshot shows the official website of the City of Zaragoza. At the top, there is a navigation bar with links for 'AYUNTAMIENTO', 'LA CIUDAD', 'CULTURA', 'PASAR AL DÍA', 'NOTICIAS', and 'TURISMO'. Below the navigation, there is a section titled 'SERVICIOS' and 'CONJUNTO DE DATOS'. A prominent call-to-action button says 'Descargar datos'. Further down, there is a table titled 'Contenido' with two rows:

Servicio	Formato
Agenda de actividades	+ PDF + RSS + HTML
Oficina de Empleo	+ PDF + RSS

At the bottom of the page, there is a footer with links for 'Ayuntamiento de Zaragoza', 'Política de Privacidad', 'Aviso Legal', 'Mapa del Sitio', and 'Contacto'.

28

Open Government Spain. Asturias

Catálogo de Datos de Asturias

En la siguiente tabla se enumeran los conjuntos de datos de Asturias, los cuales están representados por filas. Cada conjunto de datos puede estar representado en los diversos formatos que aparecen en la tabla y se puede obtener más información sobre cada uno de ellos (detalles técnicos, licencia, etc.) en las fichas asociadas a cada uno.

Consulta más sobre los [formatos disponibles](#) o sobre los [términos de uso](#) de los datos.

Conjuntos de datos y los formatos en los que aparecen

Nombre	Actualización	HTML	SPARQL/XML	SPARQL/JSON	RDF/XML	RDF/N3	TURTLE
Oferta Formativa del Servicio Público de Empleo del Principado de Asturias	17/02/2010	✓	✓	✓	✓	✓	✓
Organograma de las Consejerías del Principado de Asturias	17/02/2010	✓	✓	✓	✓	✓	✓
Plataforma de acceso a la documentación del Principado de Asturias	17/02/2010	✓	✓	✓	✓	✓	✓

<http://risp.asturias.es/catalogo/index.html>

29

Linked Data Mashup (Water quality)

- Water quality in Asturias' beaches
 - <http://datos.fundacionctic.org/sandbox/asturias/playas/>

Playa de Borizo

Calidad de las aguas en las playas asturianas

Mapa | Satélite | Hibrido

Datos de mapa ©2010 Tele Atlas - 10mms de resolución

- Introduction to Linked Data
- **Linked Data publication**
 - Methodological guidelines for Linked Data publication
 - RDB2RDF tools
 - Technical aspects of Linked Data publication
- [Linked Data consumption]

- Motivation
- Related Work
- GeoLinkedData
 - Identification of the data sources
 - Vocabulary Development
 - Generation of the RDF data
 - Publication of the RDF data
 - Data cleansing
 - Linking the RDF data
 - Enable effective discovery
- Future Work

GeoLinkedData

- It is an open initiative whose aim is to enrich the Web of Data with Spanish geospatial data.
- This initiative has started off by publishing diverse information sources, such as National Geographic Institute of Spain (IGN-E) and National Statistics Institute (INE)

GeoLinked Data

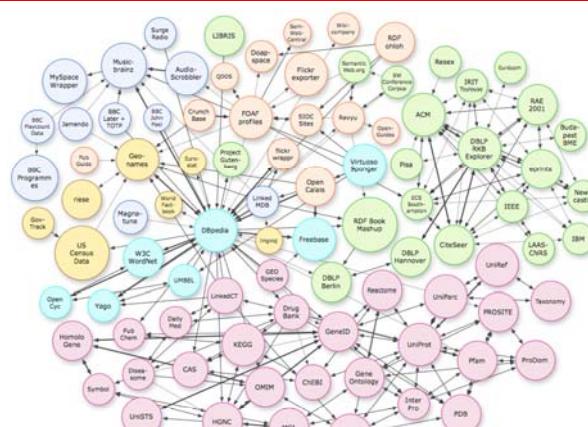


- <http://geo.linkeddata.es>

Motivation

The Web of Data is mainly for English speakers
Poor presence of Spanish

Total	3182428709
Non-literals	1494635231
Literals	1687793478
Literals without Lang tag	1282062417
Literals - en	99.171 %
Literals - ja	0.463 %
Literals - fr	0.054 %
Literals - de	0.034 %
Literals - pl	0.025 %
Literals - it	0.021 %
Literals - es	0.019 %
Literals - ru	0.018 %
Literals - nl	0.018 %



» 99.171 % English
» 0.019 % Spanish

Source: Billion Triples dataset at <http://km.aifb.kit.edu/projects/btc-2010/>
Thanks to Aidan, Richard, Andreas

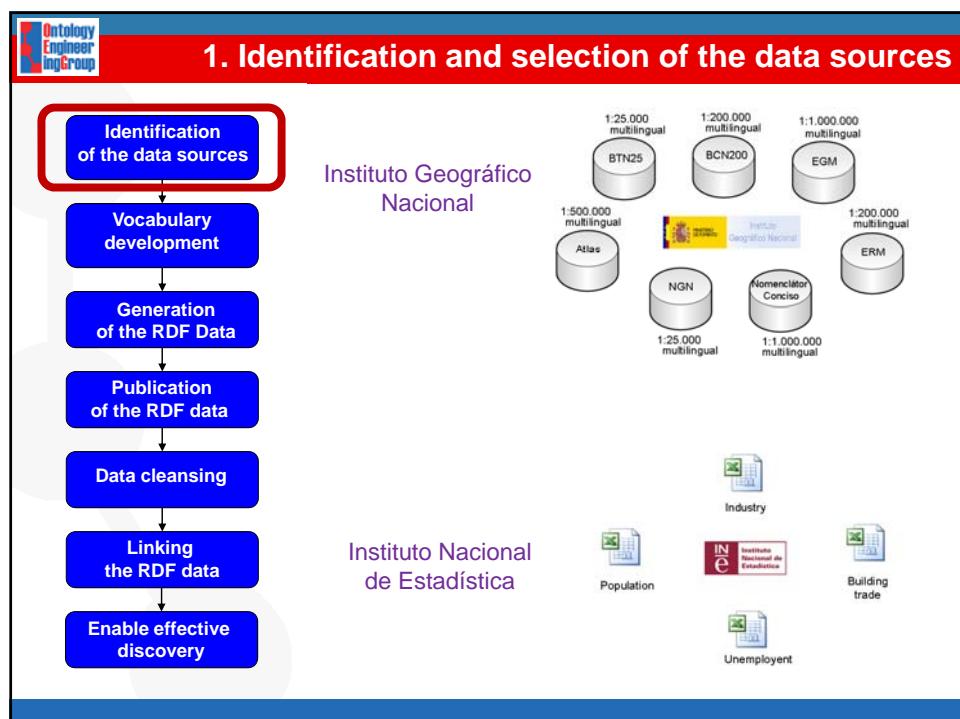
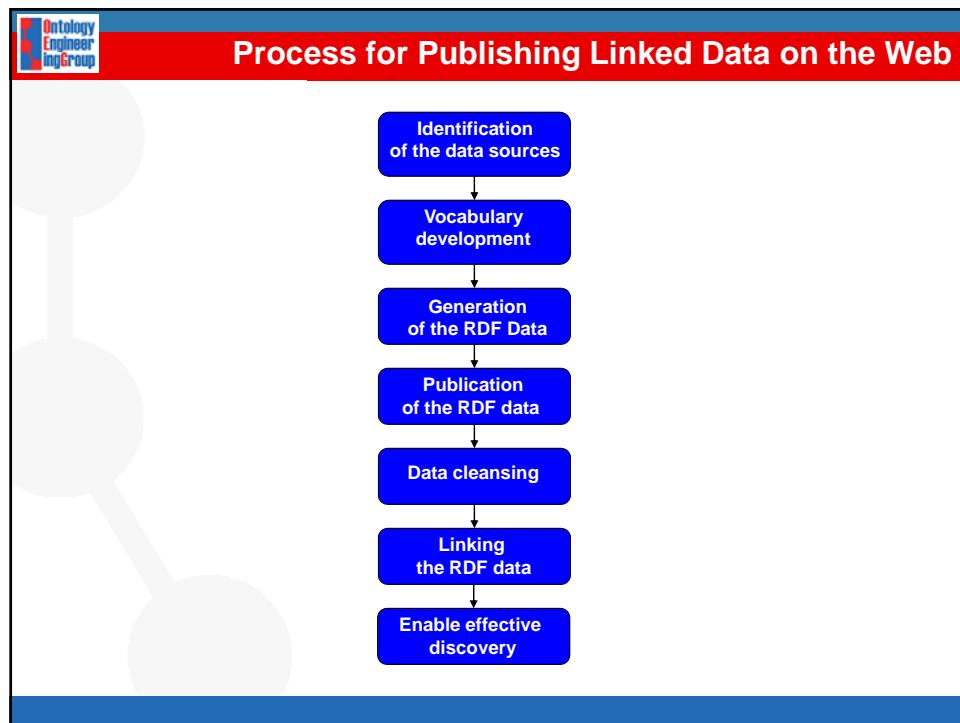
Related Work

Provider	Provenance data	Triple store	Topics	Datasets	RDF generation	URI pattern	Vocabulary	Geometry	Lang.
Ordnance survey	Ordnance survey	Talis	Administrative units	A Gazetteer (1:50,000) and administrative gazetteer for Great Britain	-	http://data.ordnancesurvey.co.uk/id/+	Spatial Relations Ontology, Administrative Geography Ontology, WGS84 Geo Positioning FOAF, OWL, and Gazetteer Ontology	Point	en
Linked GeoData	OpenStreet Map	Virtuoso	Points-of-interest	OpenStreetMap database (post offices, traffic lights, bus stops)	Triplify	http://linkedgeodata.org/triplify/+	LGD ontology, WGS84 Geo Positioning	Point	en
GeoNames	GeoNames	RDF dump	Toponyms	Datasources used by GeoNames	-	http://sws.geonames.org/id/	GeoNames Ontology, WGS84 Geo Positioning	Point	en, es, de, ca, nb, it, da, fr.
DBpedia	Wikipedia	Virtuoso	General knowledge	Wikipedia	D2R Server	http://dbpedia.org/resource/+	Dbpedia ontology, WGS84 Geo Positioning	Point	92 lang.

Impact of Geo.linkeddata.es

- Número de tripletes en Español (July): 1.412.248
- Número de tripletas en Español (End august): 21.463.088

Before	geo.linkeddata.es	After	geo.linkeddata.es
en	99,1712875	en	94,18744941
ja	0,463849377	es	5,044085342
fr	0,05447229	ja	0,440538697
de	0,034225134	fr	0,051734793
pl	0,02532934	de	0,032505155
it	0,021982542	pl	0,024056418
es	0,019584648	it	0,020877812





1. Identification and selection of the data sources

- Instituto Geográfico Nacional (Geographic Spanish Institute)
 - Multilingual (Spanish, Vasc, Galician, Catalan)
 - Conceptualization mismatches
 - Granularity (scale concept)
 - Textual information
 - Particularities
 - Longitude
 - latitude
- Instituto Nacional de Estadística (Statistic Spanish Institute)
 - Monolingual
 - Numerical information
 - Particularities
 - Geo (textual level)
 - Temporal

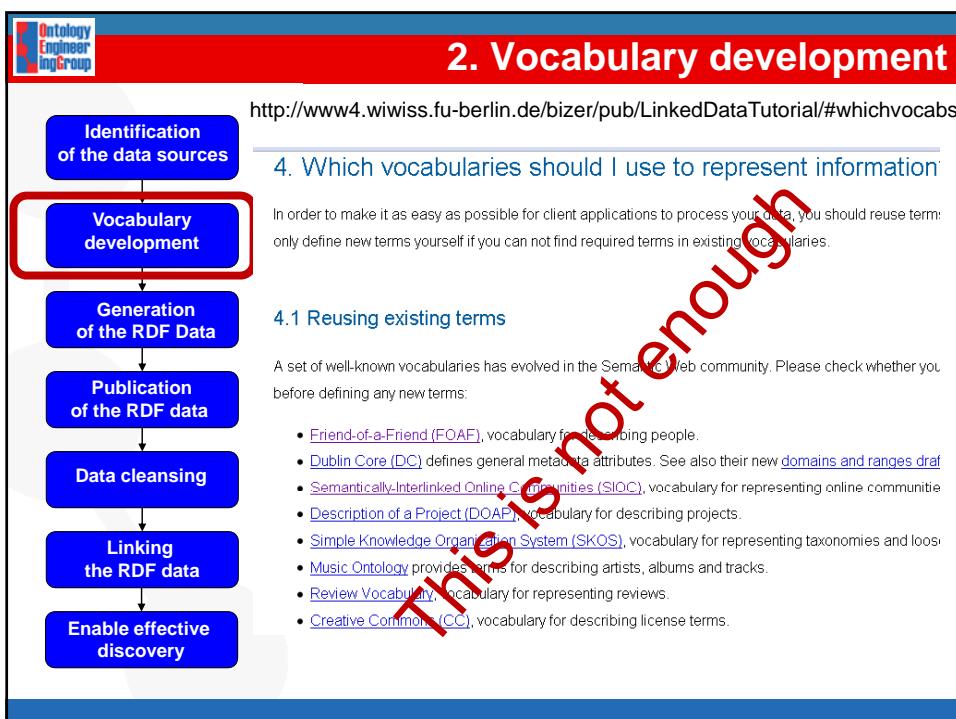
39

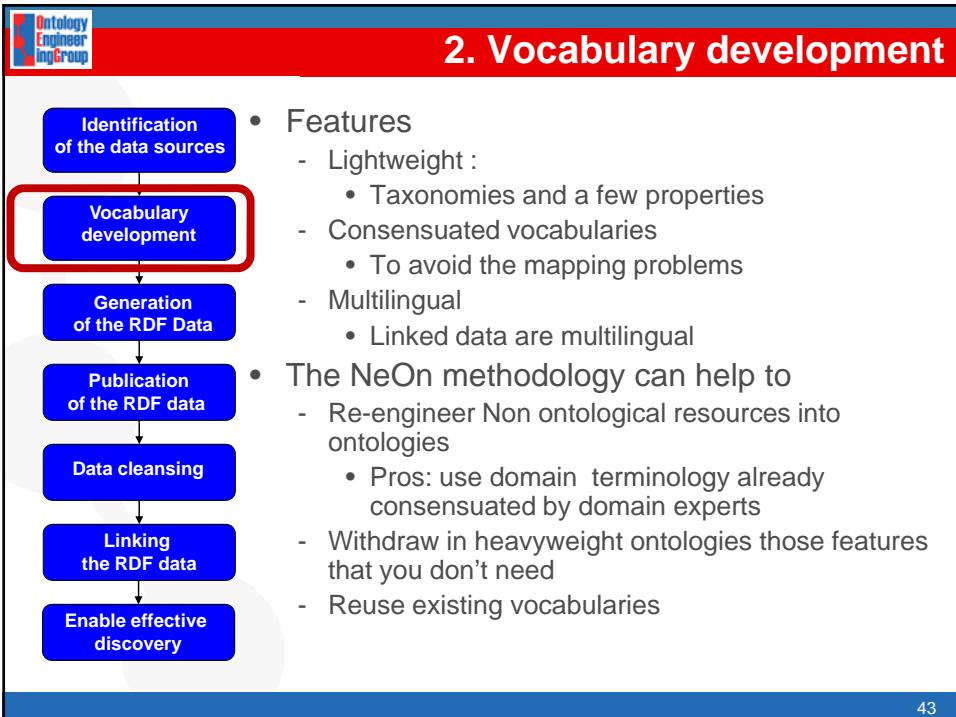
1. Identification and selection of the data sources												
NOMBRE	ENTIDAD	CON	PROV	CODINE	XUTM	YUTM	HUSO	HOJA25	LongitudG	LatitudG	LongitudGMS	LatitudGMS
Abejuela, Olivar de / Oli	Lugar/Paraje	0	02	02042	584300	246500	30	0867-3	-2,03512445814873	38,3620693572292	-02°02'06.44	38°21'
Abejuela, Rambla de	Corriente fluvial	0	04	04053	397560	152000	30	0974-4	-1,89618811907142	37,509177056304	-01°53'46.37	37°30'
Abejuela, Rambla de la	Corriente fluvial	1	46	46262	379190	123000	30	0638-2	-0,9028553170738	39,9374839997029	-00°54'10.27	39°56'
Abejuela, Rambla de la	Corriente fluvial	1	44	44002	378800	742000	30	0638-2	-0,907142388290964	39,9465700679236	-00°54'25.71	39°56'
Abelá, A	Población	0	27	27025	521500	336500	29	0003-3	-7,49293473672413	43,670782491816	-07°29'34.56	43°40'
Abeladaira	Población	0	27	27039	522000	774030	29	0073-1	-7,5006276897665	43,1084031553505	-07°30'02.38	43°06'
Abeladaira	Lugar/Paraje	0	27	27039	521550	774290	29	0073-1	-7,50615694547912	43,1100057447368	-07°30'22.15	43°06'
Abelaedo, Lugar	Paraje	0	27	27005	548100	321300	29	0010-3	-7,16730269277051	43,5291689758955	-07°10'02.28	43°31'
Abelaedo, Chao do	Llanura/Raso	0	27	27013	525500	330200	29	0009-1	-7,4448008623332	43,613141717808	-07°26'41.28	43°36'
Abelaedo, Monte do	Lugar/Paraje	0	27	27064	509000	341000	29	0002-4	-7,647027151208	43,7132267373192	-07°38'49.29	43°42'
Abelaedo, O	Población	0	27	27064	508500	340500	29	0002-4	-7,65333254552948	43,708799463333	-07°39'11.99	43°42'
Abelaedo, Rego do	Corriente fluvial	0	27	27064	509000	340540	29	0002-4	-7,64712030731384	43,7090862344245	-07°38'49.63	43°42'
NOMBRE									GML/GEOMETRY			
Abengibre									(HUGECLOB)			
Alatoz									(HUGECLOB)			
Albatana												
Balsa de Ves									 <gml:Polygon srsName="#SDO:8223" xmlns:gml="http://www.opengis.net/gml">			
Ralletern, FI									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,48374108,39,23127677,0-1,48419605,39,22531695,0-1,48544493,39,2231538,0-1,48858725,39,22212319,0-1,48680465,39,22141415,0-1,486050			
Albreca									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,48421128,39,21319056,0-1,48390069,39,21064575,0-1,48421965,39,20454254,0-1,48592926,39,20406290,0-1,49250			
Alcadizo									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,49590016,39,20836339,0-1,49005691,39,20405732,0-1,50256676,39,20446292,0-1,50338941,39,20465364,0-1,5040			
Alcalá del Júcar									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,51378498,39,2047147,0-1,51398161,39,20472622,0-1,5171434,39,20471246,0-1,51894734,39,20373528,0-1,51962			
									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,52344981,39,20108038,0-1,52693907,39,19901681,0-1,52849033,39,19815278,0-1,53151667,39,19819073,0-1,5320			
									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,5329951,39,19477493,0-1,53388695,39,19603246,0-1,53537021,39,19528414,0-1,5367317,39,19544615,0-1,5369			
									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,53953498,39,19602121,0-1,54099994,39,1962185,0-1,54110394,39,19630216,0-1,54231647,39,19646151,0-1,5441			
									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,54597924,39,19635592,0-1,55131402,39,1964386,0-1,55382203,39,19667799,0-1,55798564,39,19690882,0-1,5598			
									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,56633376,39,19750473,0-1,5703604,39,19766503,0-1,57231502,39,19777111,0-1,5728711,39,19790204,0-1,5736			
									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,576333787,39,1982863,0-1,57807687,39,20087703,0-1,57814542,39,20092292,0-1,57871557,39,20132037,0-1,5791			
									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,5805024,39,20233924,0-1,58088715,39,2027543,0-1,58107794,39,20313014,0-1,58122865,39,20365467,0-1,5814			
									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,58195139,39,20563864,0-1,58191475,39,20581637,0-1,58194513,39,20630299,0-1,58256235,39,20567667,0-1,5842			
									<gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1,58516216,39,21675162,0-1,58549891,39,21657712,0-1,58739734,29,21428,0-1,58756889,29,21417703,0-1,58924			

Ontology Engineer IngGroup

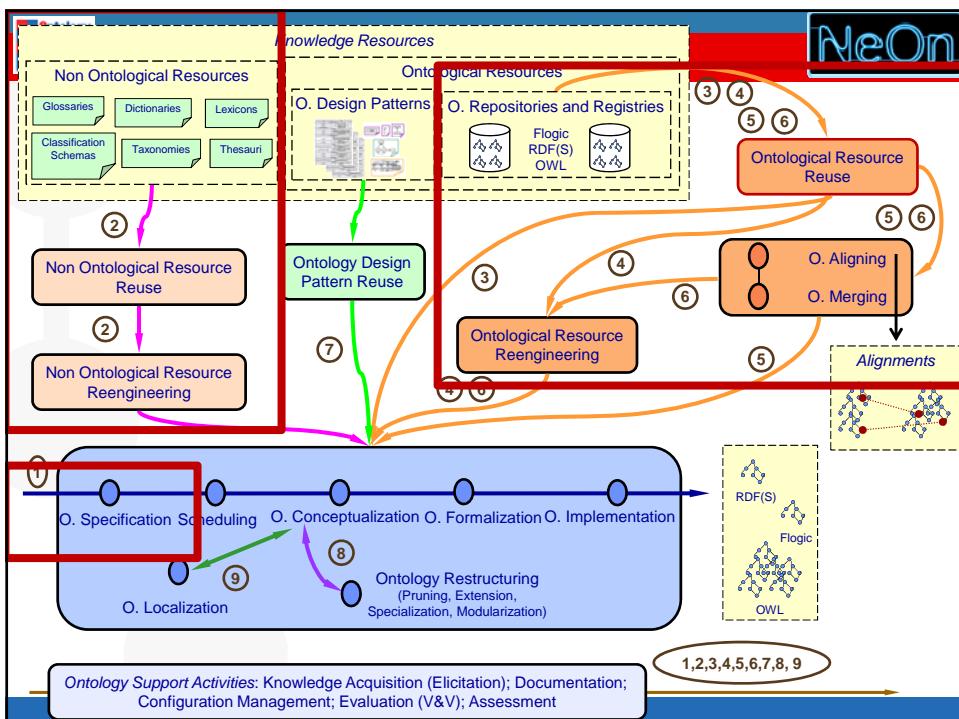
1. Identification and selection of the data sources

Province		Year							Industry Production Index	
		2009	2008	2007	2006	2005	2004	2003		
Total Nacional		3355830	342239	3338657	3174393	3064129	2942583	2813159		
Alava		21988	22318	20676	20349	19838	19779	19638		
Albacete		27380	27647	27068	25531	24685	23550	22547		
Alicante		136239	142307	140145	133016	123333	113852	111805		
Almería		43601	45130	43970	40871	38768	36260	33947		
Asturias		71863	73124	72276	70116	69178	67039	65082		
Ávila		11455	11708	11434	10590	10611	10319	10211		
Badajoz		40674	41358	40168	38645	37052	34972	34866		
Illes Balears		91826	93335	91254	88027	87024	85425	75951		
Barcelona		467385	477942	469432	444410	436294	417425	397693		
Burgos		25567	25891	25372	24504	23733	22882	22159		
Cáceres		26307	26494	26064	25039	24465	20598	23440		
Cádiz		62817	64505	63338	61691	58868	57138	54462		
Cantabria		39611	40333	39560	37690	36561	35649	34017		
Castellón		42122	43855	42476	39749	37685	37214	34213		
Ciudad Real		32046	33011	31881	30446	29521	29011	26778		
Córdoba		48979	50057	49302	47155	45405	43394	41964		
Coruña, A		83748	84220	82873	79170	77023	74809	71748		
Cuenca		14747	14928	14741	13822	13336	12829	12546		
Girona		58404	51467	50108	47169	46827	45145	52482		
Granada		60016	62269	61055	57223	54341	50508	49662		
Guadalajara		13507	13735	12874	11825	10438	10120	9422		
Guipúzcoa		62034	63569	59546	58486	57193	56498	55983		
Huelva		26783	27463	27063	25487	24777	24270	22547		
Huesca		18837	17109	16694	16025	15390	15078	14283		
Jaén		36557	37368	36962	35383	34675	33157	32444		
León		33564	34012	33563	32359	31664	30992	30256		
Lleida		36920	37638	36065	33956	32739	31515	29605		
Lugo		24861	25035	24609	23780	23122	22479	22396		
Madrid		511804	519307	503000	478202	456175	436074	407655		
Málaga		113362	116683	114547	108713	102382	96587	88257		
Murcia		95636	100075	97374	90698	85110	82484	75973		
Navarra		43282	43847	43142	41083	40730	39679	38936		
Ourense		23304	23711	23520	22843	22452	22118	21560		
Palencia		10954	11111	11060	10594	10575	10399	10297		





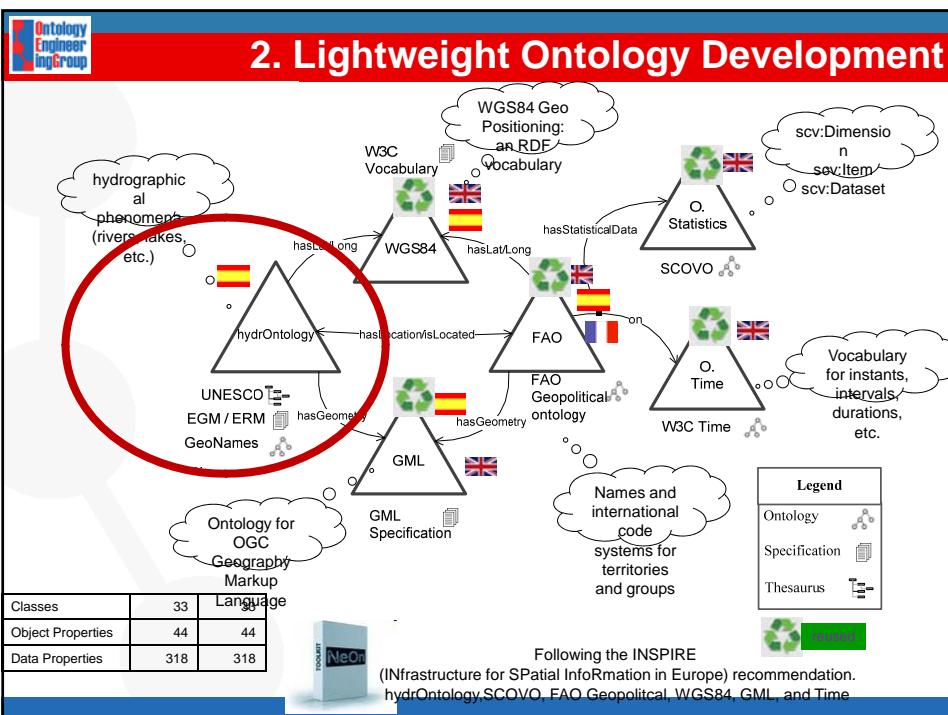
43



Vocabulary development: Specification

- Content requirements: Identify the set of questions that the ontology should answer
 - Which one are the provinces in Spain?
 - Where are the beaches?
 - Where are the reservoirs?
 - Identify the production index in Madrid
 - Which one is the city with higher production index?
 - Give me Madrid latitude and altitude
 -
- Non-content requirements
 - The ontology must be in the four official Spanish languages

45



Contexto – Directiva INSPIRE

- **Objetivos:**

- INSPIRE intenta conseguir fuentes armonizadas de Información Geográfica para dar soporte a la formulación, implementación y evaluación de políticas comunitarias (Medio Ambiente, etc).
- Fuentes de Información Geográfica: Bases de datos de los Estados Miembros (UE) a nivel local, regional, nacional e internacional.



European Environment Agency

INSPIRE - Anexos

Datos de Referencia

(Armonizados y Compartidos)

Consistencia
geométrica,
topológica,
semántica y lógica

Anexo I (INSPIRE)

Sistema de Ref. de Coordenadas
Cuadrículas geográficas
Nombres Geográficos
Unidades Administrativas
Redes de Transporte
Hidrografía
Lugares protegidos

Modelos de Elevación

Direcciones y Áreas Postales

Parcelas Catastrales

Ocupación del suelo

Ortofotos

Anexo II (INSPIRE)



- Existencia de gran diversidad de problemas (múltiples fuentes, heterogeneidad de contenido y estructuración, ambigüedad del lenguaje natural, etc.) en la información geográfica.
- Necesidad de un modelo compartido para solventar los **problemas de armonización y estructuración** de la información hidrográfica.
- *hydrOntology* es una ontología global de dominio desarrollada conforme a un acercamiento *top-down*.
 - Recubrir la mayoría de los fenómenos representables cartográficamente asociados al dominio hidrográfico.
 - Servir como marco de armonización entre los diferentes productores de información geo-espacial en el entorno nacional e internacional.
 - Comenzar con los pasos necesarios para obtener una **mejor organización y gestión** de la información geográfica (hidrográfica).

Fuentes



Catálogos de fenómenos



BCN25



EGM & ERM



BCN200



CC.AA.

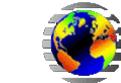


Nomenclátor Geográfico Nacional

Nomenclátor Conciso



Tesauros y Bibliografía



GEMET



WFD



Diccionarios y Monografías

Criterios de estructuración

- Directiva Marco del Agua
 - Propuesta por Parlamento y Consejo de la UE
 - Lista de definiciones de fenómenos hidrográficos
- Proyecto SDIGER
 - Proyecto piloto INSPIRE
 - Dos cuencas, países e idiomas
- Criterios semánticos
 - Diccionarios geográficos
 - Diccionario de la Real Academia de la Lengua
 - WordNet
 - Wikipedia
 - Bibliografía de varias áreas de conocimiento
- Herencia: Estructuración actual de catálogos
- Asesoramiento expertos en toponimia del IGN

Luis Manuel Vilches Blázquez

Nivel su

Nivel inferior

Luis Manuel Vilches Blázquez

Implementación & Formalización

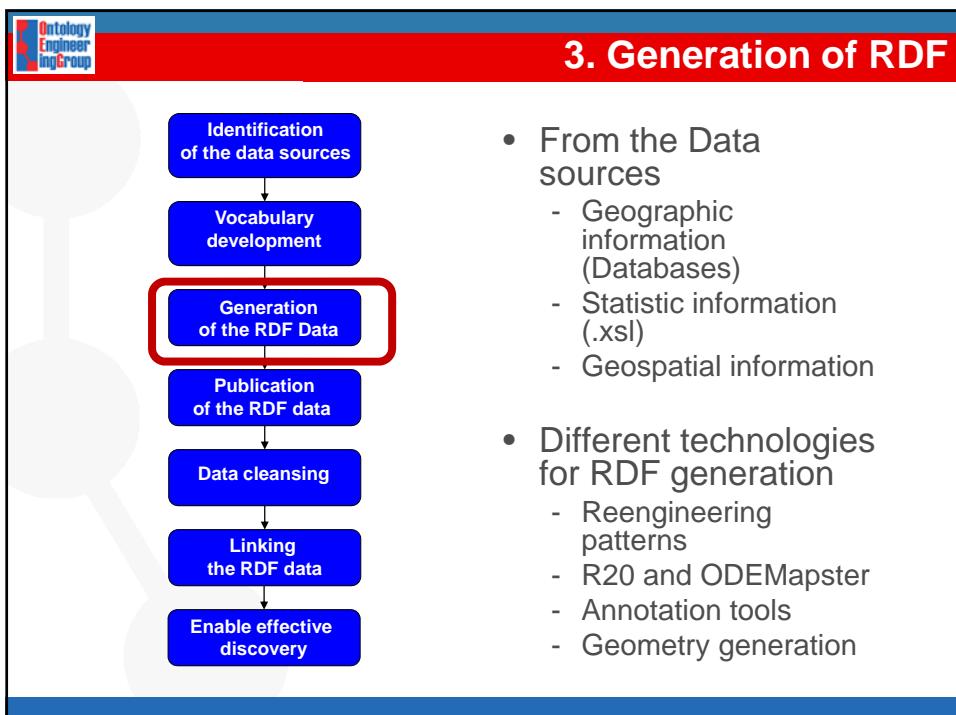
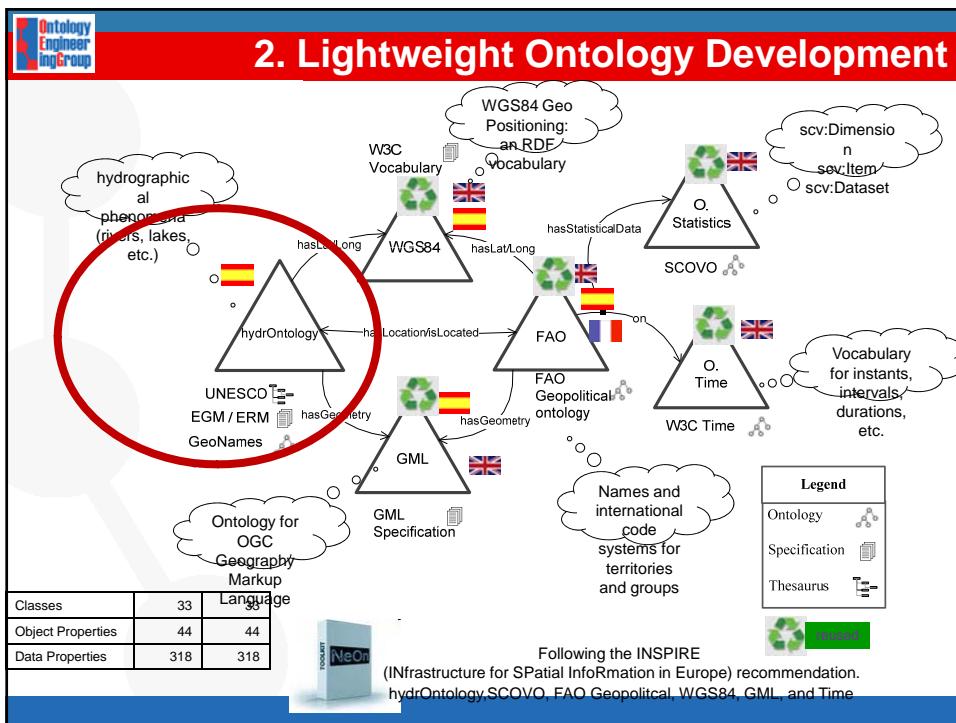
protégé + Pellet

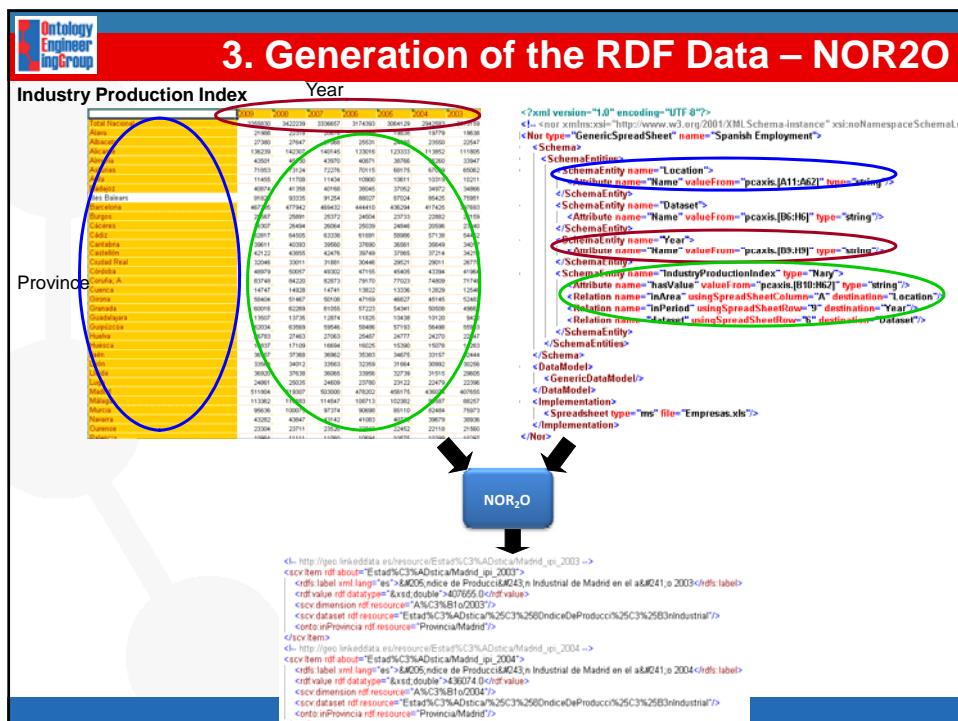
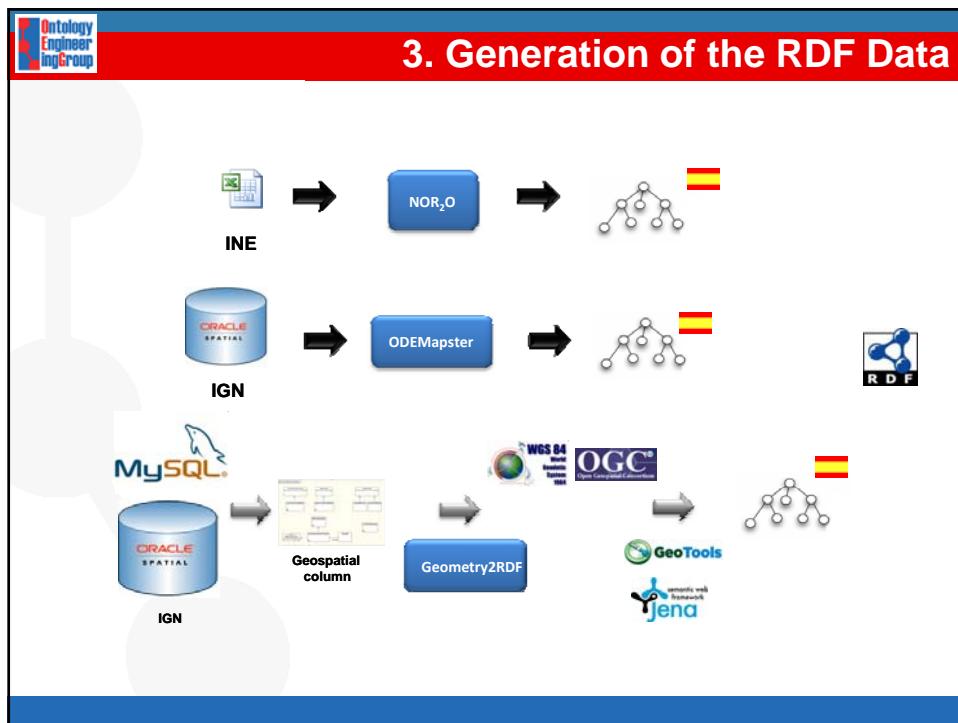
+ 150 conceptos (*classes*) , 47 tipos de relaciones (*properties*)
y 64 tipos de atributos (*attribute types*)

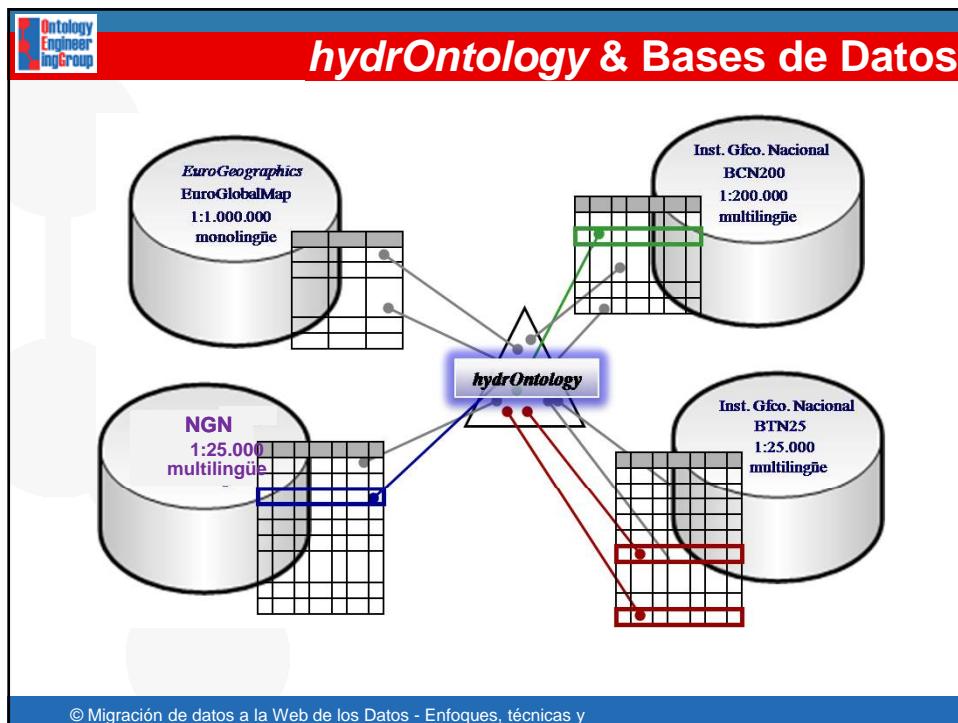
Luis Manuel Vilches Blazquez

2. Vocabulary development: HydrOntology

http://geo.linkeddata.es/ontology/Surgencia	http://geo.linkeddata.es/ontology/Lucio
http://geo.linkeddata.es/ontology/Marisma	http://geo.linkeddata.es/ontology/Rambla
http://geo.linkeddata.es/ontology/Desembocadura	http://geo.linkeddata.es/ontology/Esterero
http://geo.linkeddata.es/ontology/Albufera	http://geo.linkeddata.es/ontology/Regato
http://geo.linkeddata.es/ontology/Brazo	http://geo.linkeddata.es/ontology/Balsa
http://geo.linkeddata.es/ontology/Oio	http://geo.linkeddata.es/ontology/Charca
http://geo.linkeddata.es/ontology/Caz	http://geo.linkeddata.es/ontology/Azarbe
http://geo.linkeddata.es/ontology/Chorrera	http://geo.linkeddata.es/ontology/Garganta
http://geo.linkeddata.es/ontology/Lavajo	http://geo.linkeddata.es/ontology/Madre
http://geo.linkeddata.es/ontology/Reguero	http://geo.linkeddata.es/ontology/Estanque
http://geo.linkeddata.es/ontology/Ribera	http://geo.linkeddata.es/ontology/R%C3%A9tulo
http://geo.linkeddata.es/ontology/Embalse	http://geo.linkeddata.es/ontology/Fa%C3%ADa
http://geo.linkeddata.es/ontology/Laguna	http://geo.linkeddata.es/ontology/Ca%C3%B3n
http://geo.linkeddata.es/ontology/Nava	http://geo.linkeddata.es/ontology/Ca%C3%B3nB1
http://geo.linkeddata.es/ontology/Arroyo	http://geo.linkeddata.es/ontology/Ca%C3%B3nB1da
http://geo.linkeddata.es/ontology/Canal	http://geo.linkeddata.es/ontology/Oc%C3%A1n
http://geo.linkeddata.es/ontology/Acequia	http://geo.linkeddata.es/ontology/Captaci%C3%B3n
http://geo.linkeddata.es/ontology/Rivera	http://geo.linkeddata.es/ontology/A%C3%81rea
http://geo.linkeddata.es/ontology/Torrente	http://geo.linkeddata.es/ontology/Organizaci%C3%B3n
http://geo.linkeddata.es/ontology/Colector	http://geo.linkeddata.es/ontology/Humedal

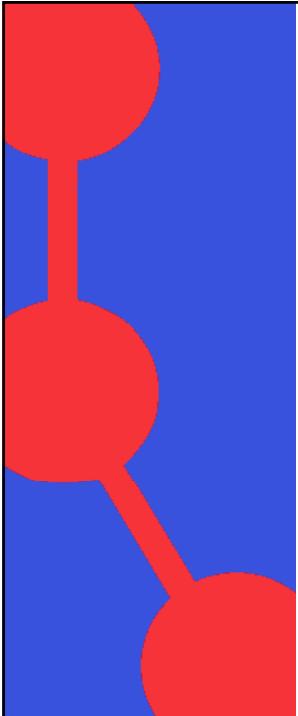






3. Generation of the RDF Data – R2O & ODEMapster

- Creation of the R2O Mappings



Linked Data Generation I

Oscar Corcho, Asunción Gómez Pérez ({ocorcho, asun}@fi.upm.es)
Universidad Politécnica de Madrid

Credits: Raúl García Castro, Guillermo Alvaro, Oscar Muñoz, Jose Angel Ramos Gargantilla, María del Carmen Suárez de Figueroa, Boris Villazón-Terrazas, Alex de León, Víctor Saquicela, Luis Vilches, Miguel Angel García, Manuel Salvadores, Juan Sequeda, Carlos Ruiz Moreno and many others

*Work distributed under the license Creative Commons
Attribution-Noncommercial-Share Alike 3.0*