

# Methodological Guidelines for Publishing Linked Data

Boris Villazón-Terrazas, Oscar Corcho

Facultad de Informática, Universidad Politécnica de Madrid

Campus de Montegancedo sn, 28660 Boadilla del Monte, Madrid

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

[bvillazon@fi.upm.es](mailto:bvillazon@fi.upm.es)

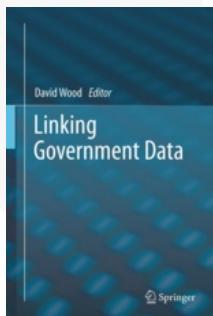
Phone: 34.91.3366605, Fax: 34.91.3524819

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

**Acknowledgements:** Oscar Corcho, Asunción Gómez-Pérez, Luis M. Vilches, Victor Saquicela, Alexander de León

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

# Main 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, DERC](#)

[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

Michael Hausenblas, Bernadette Hyland, Boris Villazón-Terrazas

<https://dvcs.w3.org/hg/gld/raw-file/bcb72f87b5cc/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

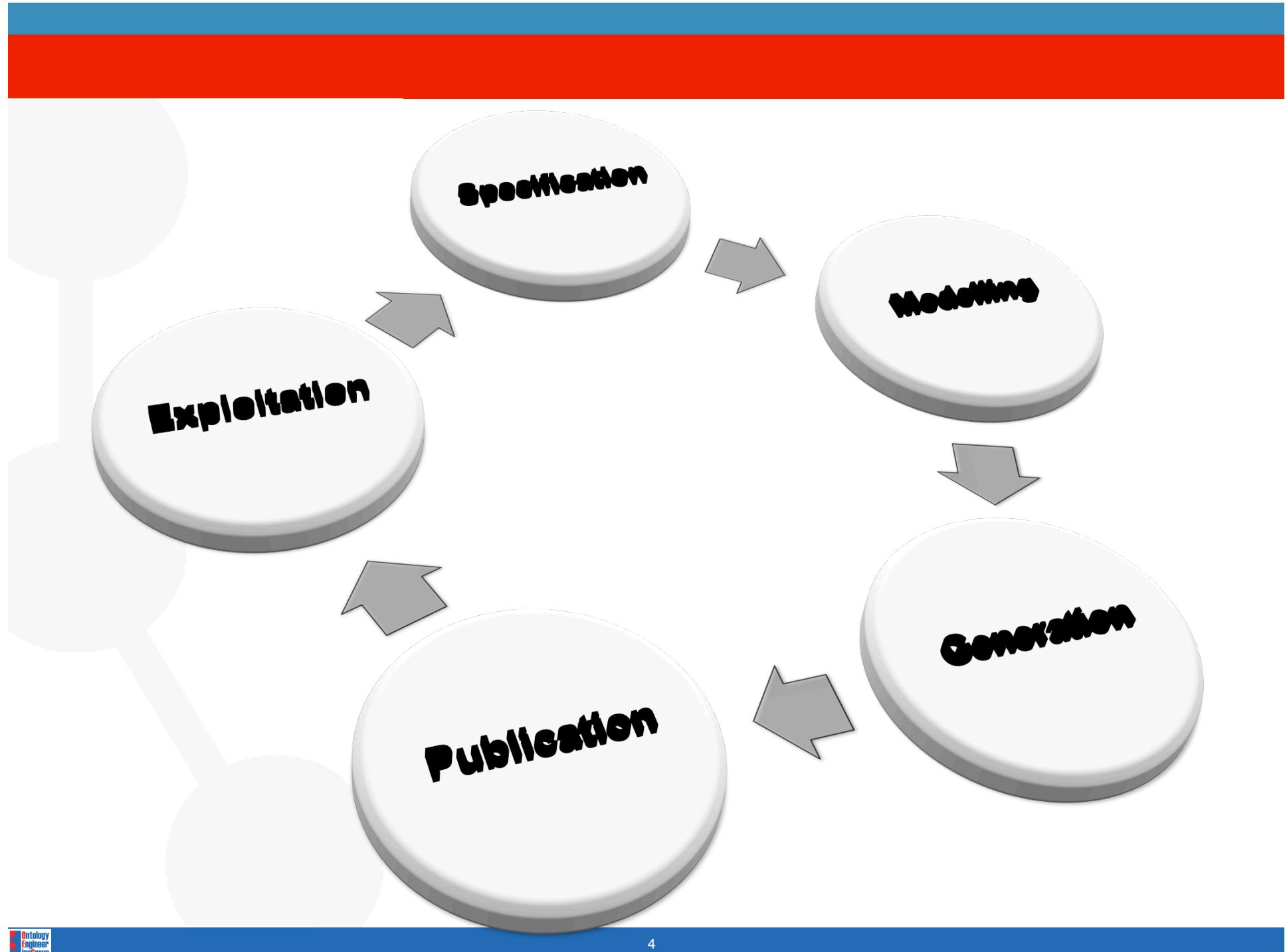
W3C Editor's Draft – Government Linked Data Working Group

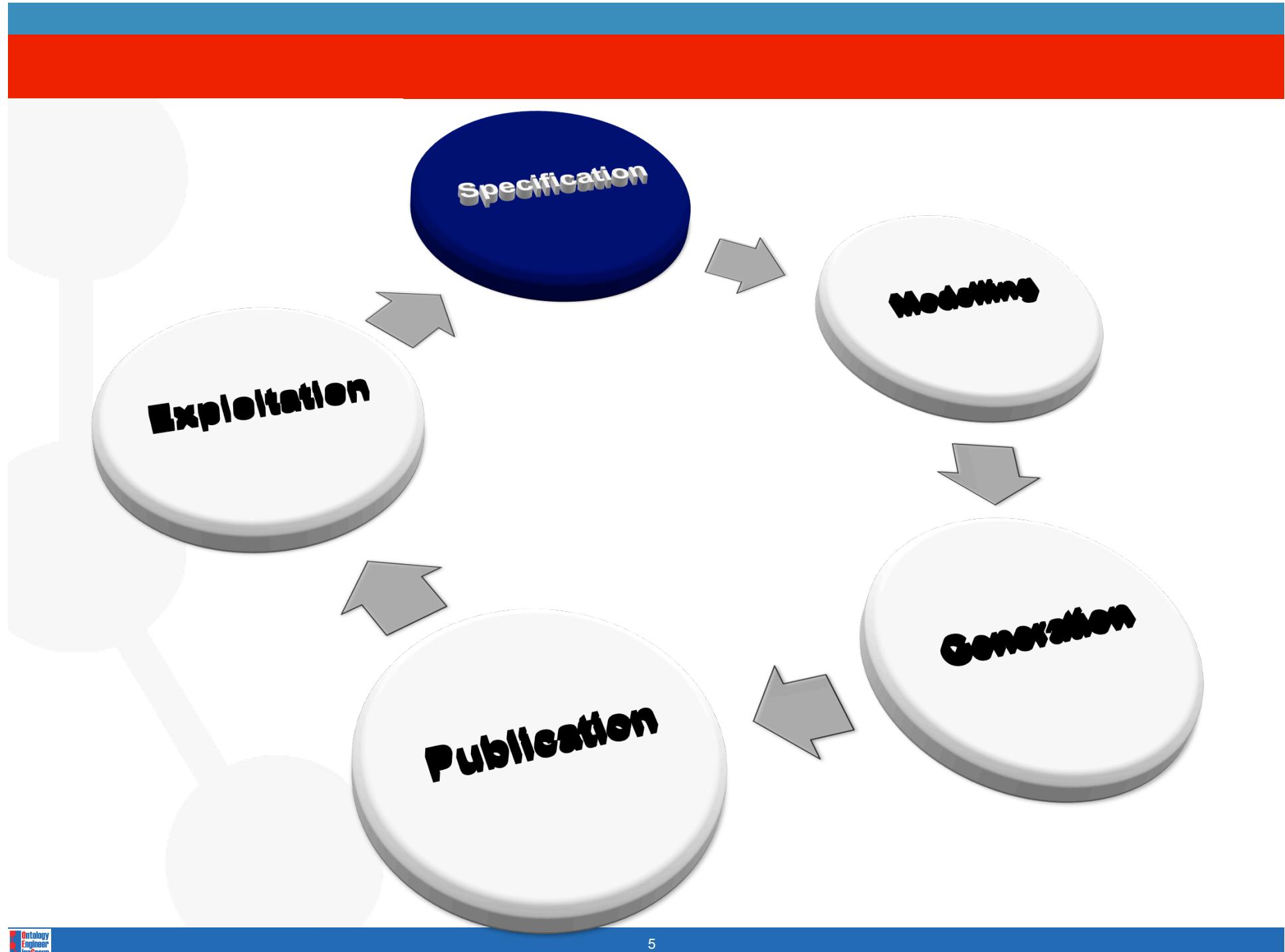
Bernadette Hyland, Boris Villazón-Terrazas, Sarven Capadisli

[http://www.w3.org/2011/gld/wiki/Linked\\_Data\\_Cookbook](http://www.w3.org/2011/gld/wiki/Linked_Data_Cookbook)

# Guidelines for Publishing Linked Data

- The process of publishing Linked Data has an iterative incremental life cycle model.
- Based on our experience in the production of Linked Data in several Governmental Contexts, have been applied in real case scenarios.





# Specification

- Identification and analysis of the data sources
- URI design
- Definition of the license

## Identification and analysis of the data sources

We have to distinguish

- Open and publish data that government agencies have not yet opened up and published
  - Task that may require contacting to specific government data owners to get access to their legacy data
- Reuse and leverage on data already opened up and published by government agencies
  - Task to look for these data in public government catalogs
    - Open Government Data
    - datacatalogs.org
    - Open Government Catalog

## Identification and analysis of the data sources

After we have identified and selected the government data sources

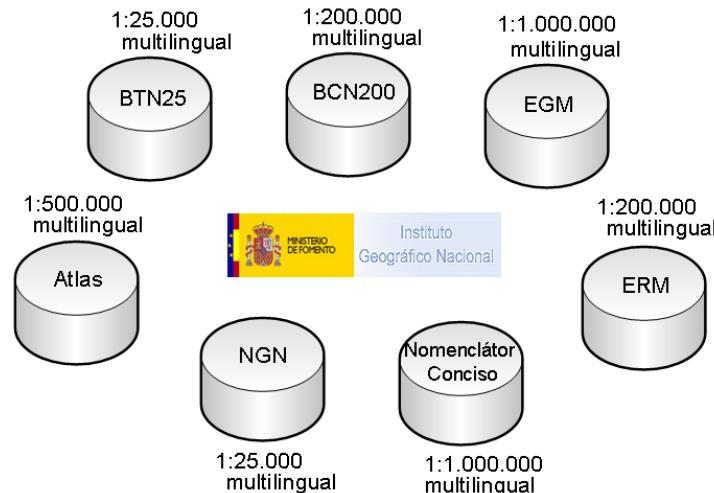
- Search and compile all the available data and documentation about those resources
- Identify the schema of those resources including conceptual components and their relationships
- Identify the items in the domain, i.e., things whose properties and relations are described in the data sources

# GeoLinkedData – Identification of the data sources

**IGN**

National Geographic Institute of Spain

Oracle & MySQL



Agreement with the IGN

**INE**

National Statistic Institute of Spain



Data sources available  
in a public data catalog

# GeoLinkedData – Analysis of the data sources

NOMBRE	ENTIDAD	CON	PROV	CODINE	XUTM	YUTM	HUSO	HOJA25	LongitudG	LatitudG	LongitudGMS	LatitudG
Abejuela, Olivar de / Oliv	Lugar/Paraje	0	02	02042	584300	246500	30	0867-3	-2,03512445814873	38,3620693572292	-02°02'06.44	38°21'43
Abejuela, Rambla de	Corriente fluvial	0	04	04053	597560	152000	30	0974-4	-1,89618811907142	37,509177056304	-01°53'46.27	37°30'33
Abejuela, Rambla de la	Corriente fluvial	1	46	46262	579190	123000	30	0638-2	-0,90285531710738	39,9374839997029	-00°54'10.27	39°56'14
Abejuela, Rambla de la	Corriente fluvial	1	44	44002	578800	124000	30	0638-2	-0,907142388290964	39,9465700679236	-00°54'25.71	39°56'47
Abelá, A	Población	0	27	27025	521500	336500	29	0003-3	-7,49293473672413	43,670782491816	-07°29'34.56	43°40'14
Abeladoira	Población	0	27	27039	522000	374030	29	0073-1	-7,50066276897865	43,1084031553505	-07°30'02.38	43°06'30
Abeladoira	Lugar/Paraje	0	27	27039	521550	374200	29	0073-1	-7,50615404547912	43,1100057447368	-07°30'22.15	43°06'36
Abelaedo	Lugar/Paraje	0	27	27005	548100	321300	29	0010-3	-7,16730269277051	43,5291689758955	-07°10'02.28	43°31'45
Abelaedo, Chao do	Llanura/Raso	0	27	27013	525500	330200	29	0009-1	-7,4448008623332	43,613414717808	-07°26'41.28	43°36'48
Abelaedo, Monte do	Lugar/Paraje	0	27	27064	509000	341000	29	0002-4	-7,6470271512085	43,7132267731492	-07°38'49.29	43°42'47
Abelaedo, O	Población	0	27	27064	508500	340500	29	0002-4	-7,65333254552948	43,708799463333	-07°39'11.99	43°42'31
Abelaedo, Rego do	Corriente fluvial	0	27	27064	509000	340540	29	0002-4	-7,64712030731384	43,7090862344245	-07°38'49.63	43°42'32

Year

	2009	2008	2007	2006	2005	2004	2003
Total Nacional	3355930	3422239	3336657	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	43501	45130	43970	40871	38766	36260	33947
Asturias	71853	73124	72276	70115	68175	67039	65062
Ávila	11455	11708	11434	10900	10611	10319	10211
Badajoz	40874	41358	40168	38045	37052	34972	34866
Illes Balears	91826	93335	91254	88027	87024	85425	75951
Barcelona	467385	477942	469432	444410	436294	417425	397693
Burgos	25957	25891	25372	24504	23733	22882	22159
Cáceres	26307	26494	26064	25039	24846	20598	23440
Cádiz	62817	64505	63338	61691	58986	57138	54462
Cantabria	39611	40393	39560	37690	36561	35649	34017
Castellón	42122	43855	42476	39749	37865	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	16837	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	102362	96567	86257
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	10964	11114	11060	10694	10575	10390	10207

Province

Industry Production Index

- Use meaningful URIs, instead of opaque URIs, when possible
- Separate TBox (ontology model) from ABox (instances) URIs.
  - Base URI
    - `http://data.gov.bo/`
    - `http://health.data.gov.bo/`
  - TBox URIs
    - `http://data.gov.bo/ontology/{class|property}`
  - ABox URIs
    - `http://data.gov.bo/resource/`
    - `http://data.gov.bo/resource/province/Tiraque`

- Base URI

`http://linkeddata.es/`

`http://geo.linkeddata.es/`

- TBox URIs

`http://geo.linkeddata.es/ontology/{concept|property}`

`http://geo.linkeddata.es/ontology/Provincia`

- ABox URIs

`http://geo.linkeddata.es/resource/{r. type}/{r. name}`

`http://geo.linkeddata.es/resource/Provincia/Madrid`

- Several possibilities
  - The UK Open Government License
  - Open Database License
  - Public Domain Dedication and License
  - Open Data Commons Attribution License
  - The Creative Commons Licenses

It is also possible to reuse and apply an existing license of the government data sources.

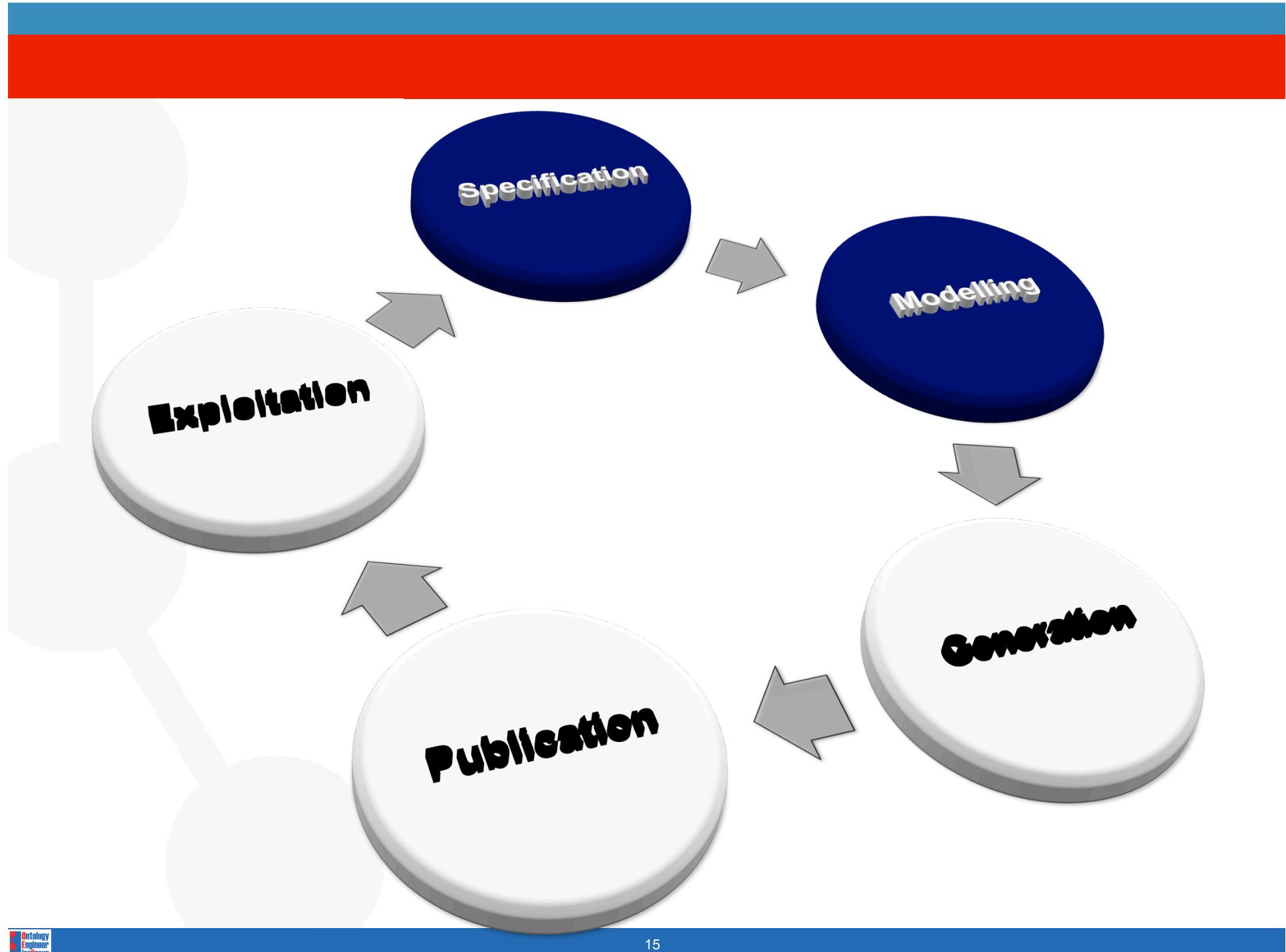
# GeoLinkedData - Definition of the license

- Reusing the original license of the government data sources. IGN and INE data sources have their own license, similar to Attribution-Share Alike 2.5 Generic License



<http://creativecommons.org/licenses/by-sa/2.5/>

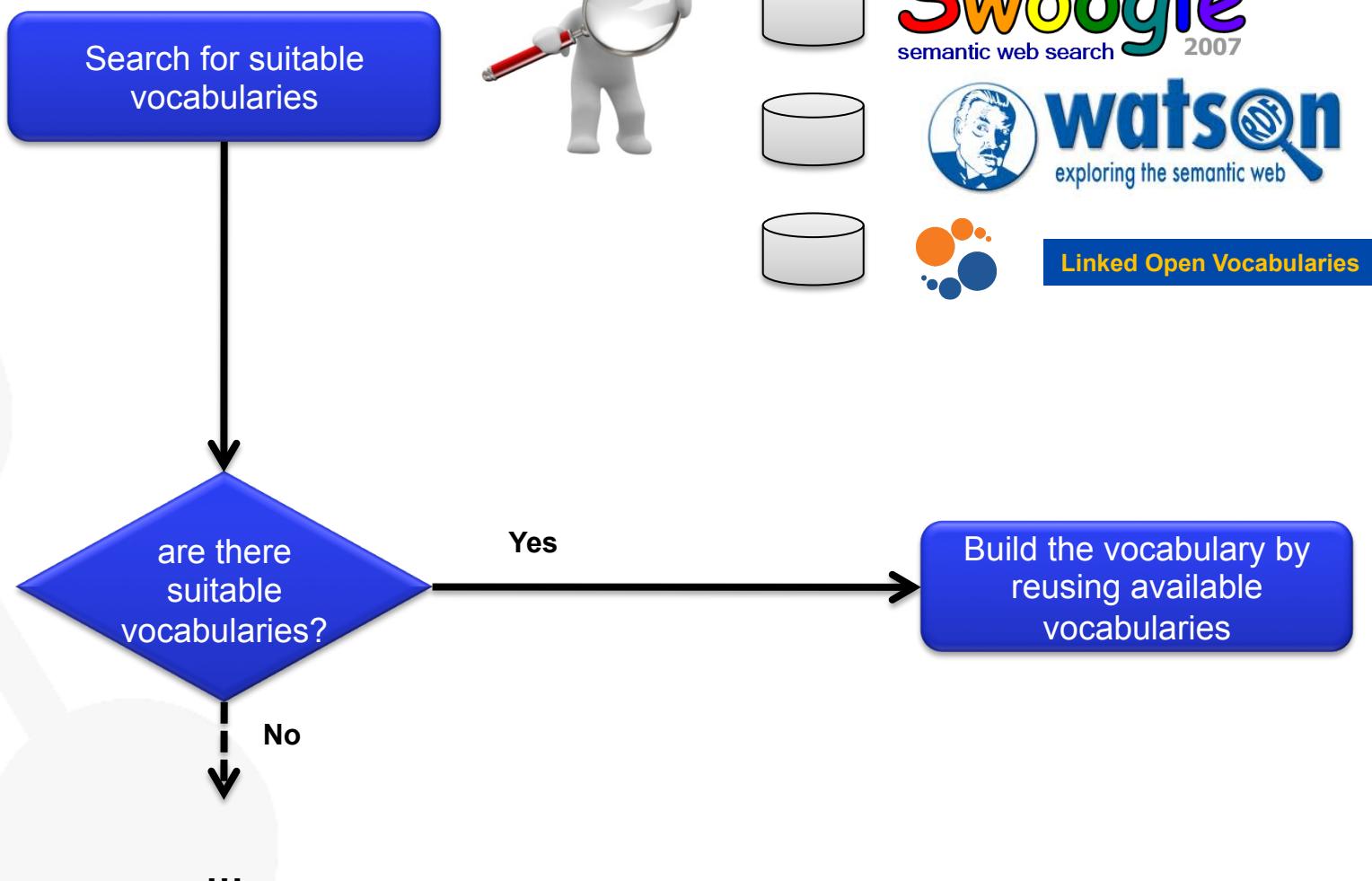




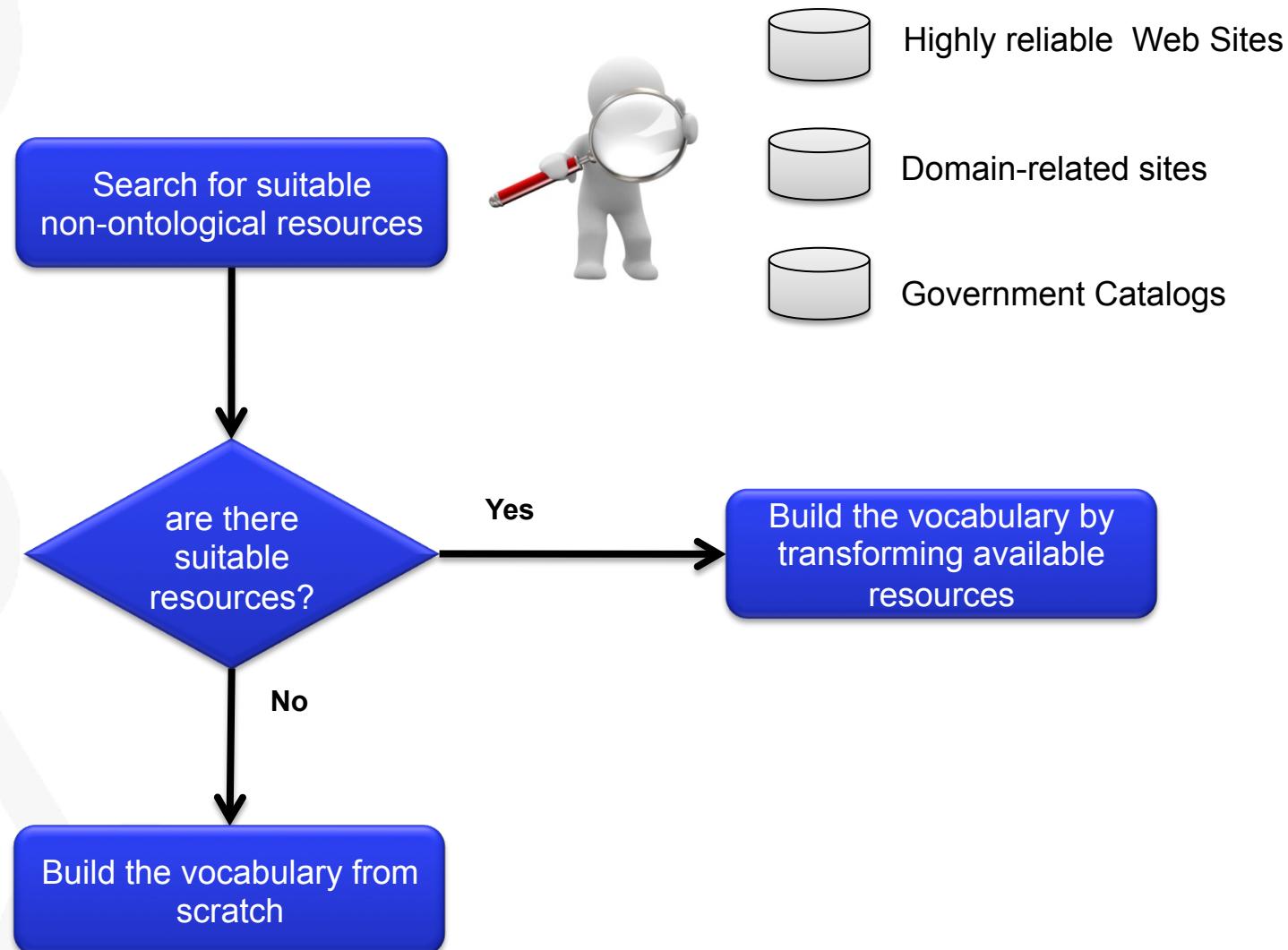
- An ontology is an engineering artifact, which provides:
  - A set of terms
  - A set of explicit assumptions regarding the intended meaning of the terms.
    - Almost always including concepts and their classification
    - Almost always including properties between concepts
- Shared understanding of a domain of interest
- Ontologies expressed in OWL or RDF(S), both based on RDF



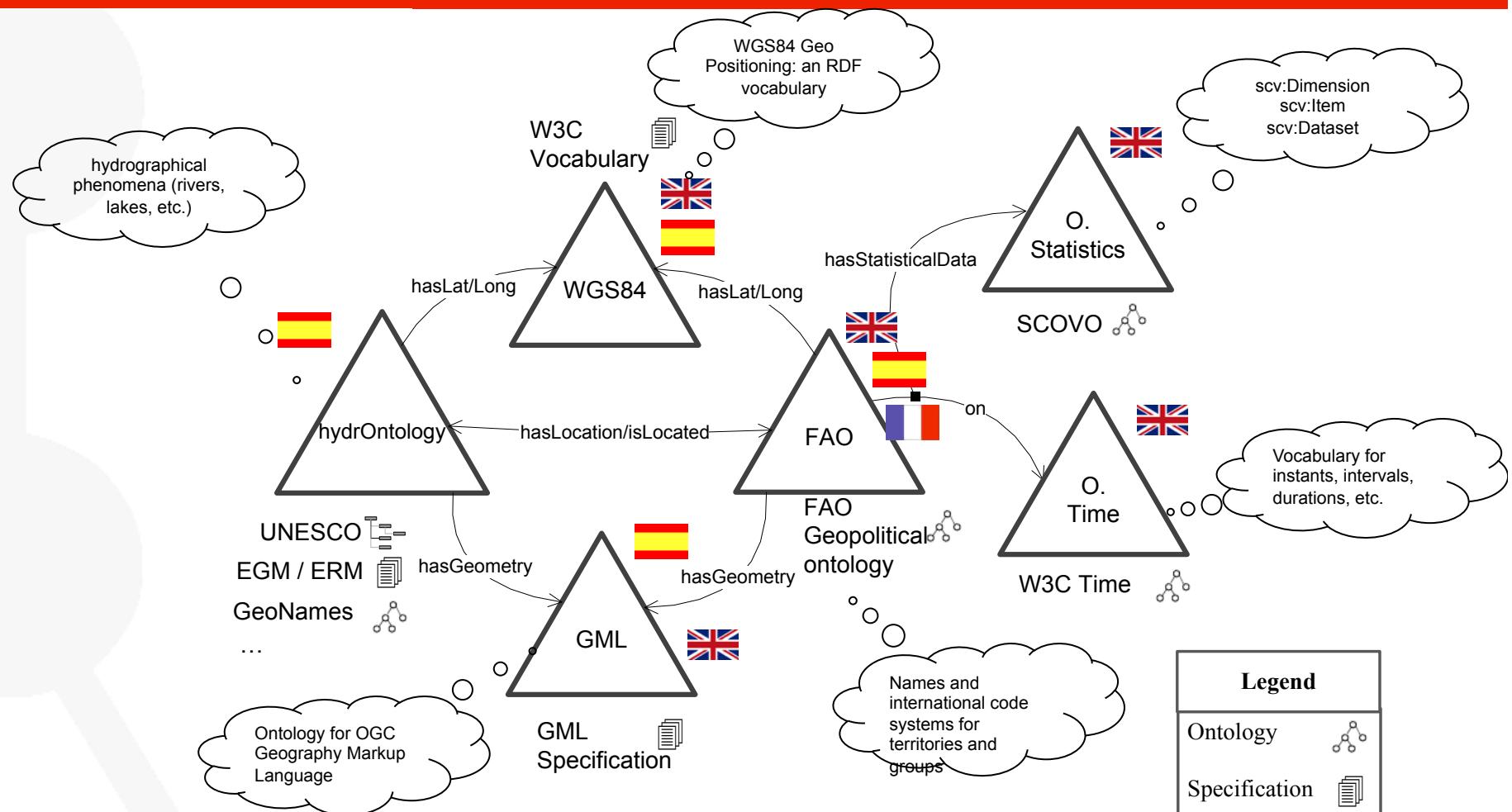
# Reuse available vocabularies



# Reuse available non-ontological resources



# Modelling GeoLinkedData



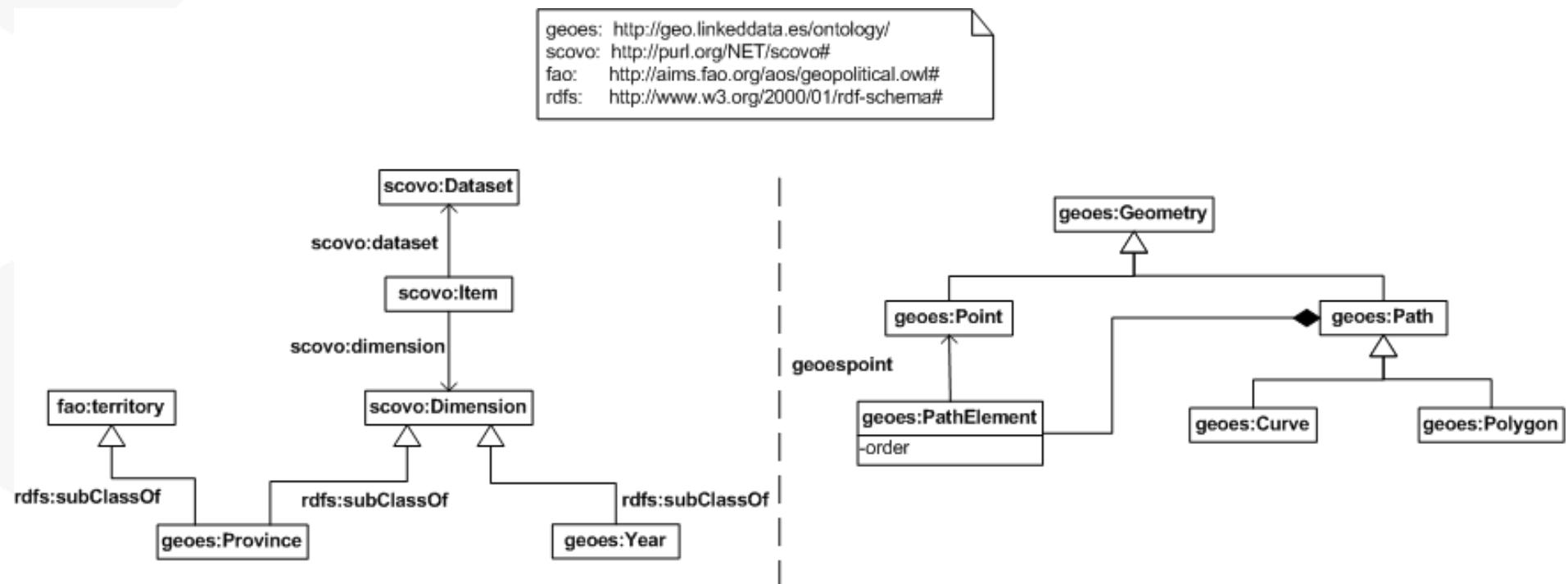
Classes	33	33
Object Properties	44	44
Data Properties	318	318

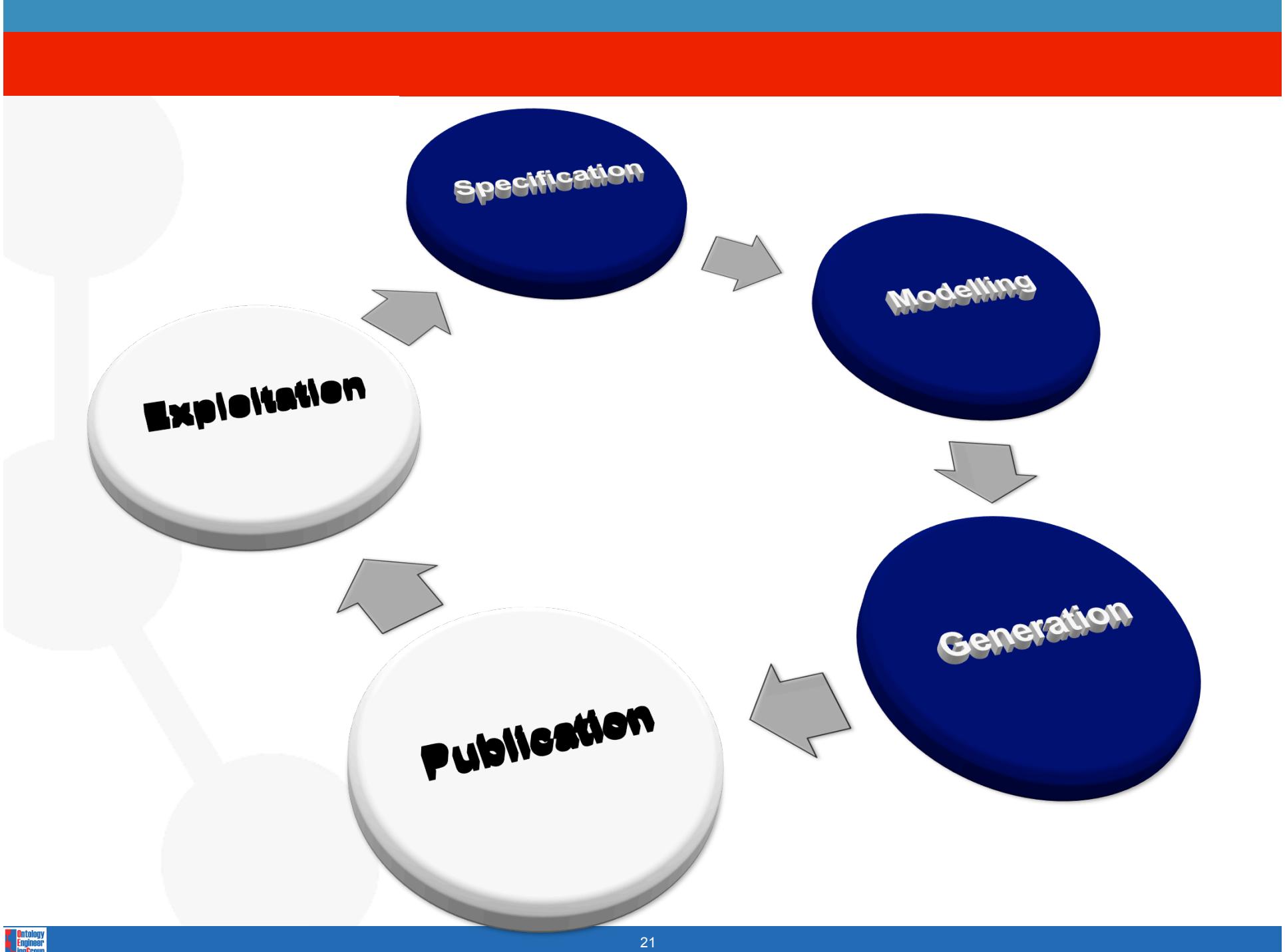


<http://neon-toolkit.org/>

Legend	
Ontology	
Specification	
Thesaurus	

# Modelling GeoLinkedData



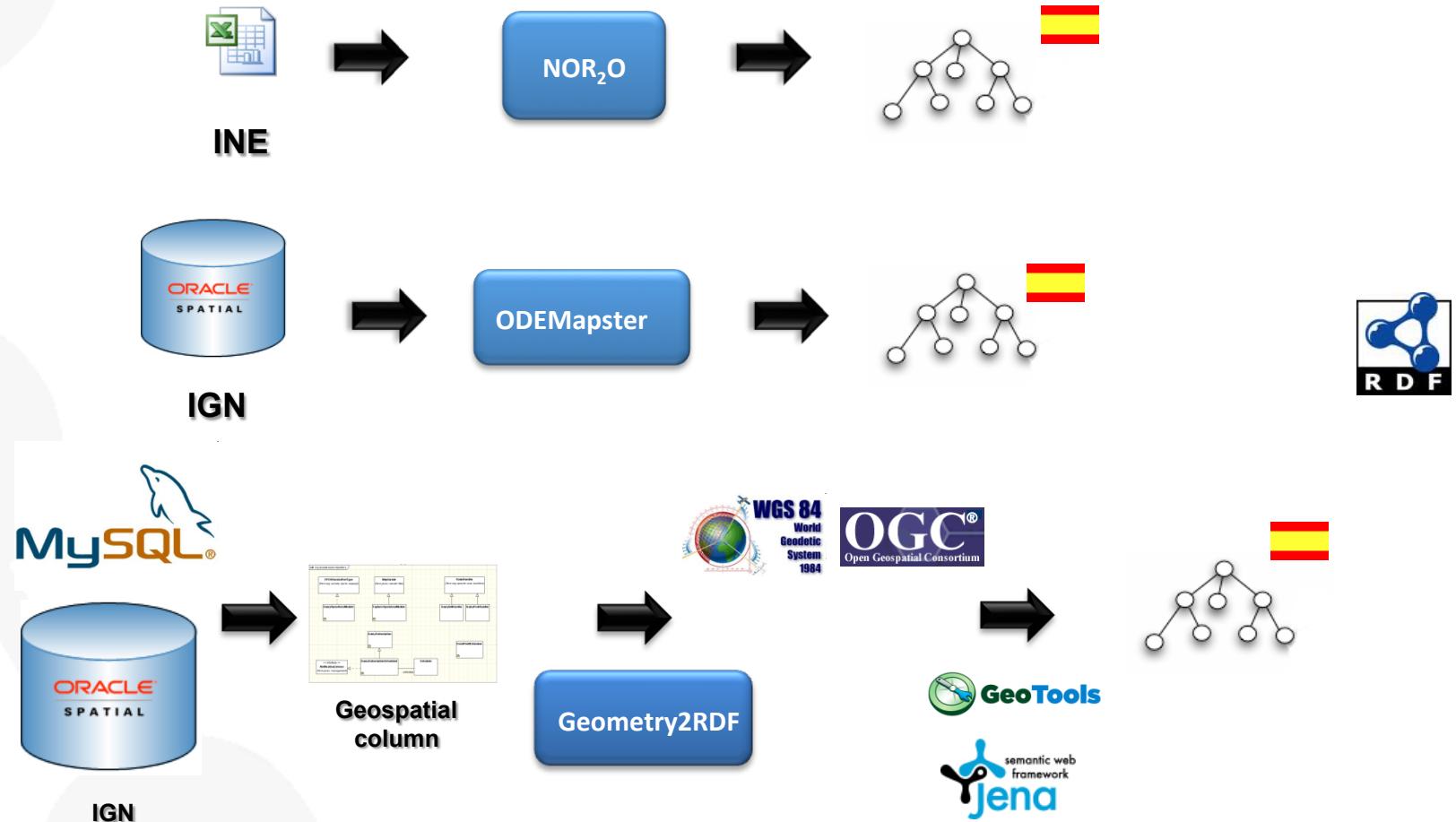


# Generation

- Transformation
- Data cleansing
- Linking

- Take the data sources selected in the specification activity and transform them to RDF according to the vocabulary created in the modelling activity
- Some tools
  - CSV and spreadsheets
    - RDF extension of Google Refine, XLWrap, RDF123, NOR<sub>2</sub>O
  - RDB
    - D2R Server, ODEMapster, W3C RDB2RDF WG – R2RML
  - XML
    - GRDDL, ReDeFer

# GeoLinkedData - Transformation



# GeoLinkedData - Transformation

## Industry Production Index

Year

	2009	2008	2007	2006	2005	2004	2003
Total Nacional	3355830	3422239	3336657	3174393	3064129	2942583	2915159
Alava	21988	22318	20570	19688	19779	19638	
Albacete	27380	27647	20668	25531	24285	23550	22547
Alicante	136239	142307	140145	133016	123333	113852	111805
Almería	43501	45100	43970	40871	38766	36260	33947
Ávila	71653	73124	72276	70115	68175	67049	65062
Badajoz	11455	11708	11434	10900	10611	10319	10211
Illes Balears	40874	41358	40168	38045	37052	34972	34866
Barcelona	91861	93335	91254	88027	87024	85425	75951
Burgos	467365	477942	459432	444410	436294	417425	37693
Cáceres	25367	25891	25372	24504	23733	22882	2159
Cádiz	6307	26494	26064	25039	24846	20596	2340
Cantabria	52817	64505	63338	61691	58986	57138	5442
Castellón	39611	40393	39560	37690	36561	35649	34017
Castellón	42122	43855	42476	39749	37865	37214	3421
Ciudad Real	32046	33011	31881	30446	29521	29011	26778
Córdoba	48979	50057	49302	47155	45405	43394	41964
Conuría, 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	4966
Guadalajara	13507	13735	12874	11825	10438	10120	942
Guipúzcoa	62034	63569	59546	56486	57193	56498	5593
Huelva	6783	27463	27063	25487	24777	24270	2237
Huesca	1.337	17109	16694	16025	15390	15078	14283
Jaén	36167	37368	36962	35383	34675	33157	2444
Léon	3356	34012	33563	32359	31664	30992	30256
Lleida	36920	37638	36065	33956	32739	31515	29605
Lugo	24981	25035	24609	23780	23122	22479	22396
Madrid	511804	519307	503000	478202	456175	43607	407655
Málaga	113362	117383	114547	108713	102382	96587	88257
Murcia	95636	100073	97374	90698	85110	82494	75973
Navarra	43282	43847	43142	41083	40728	39679	38936
Ourense	23304	23711	23520	22843	22452	22118	21560
Palencia	10664	11111	10660	10694	10575	10399	10297

Province

```

<?xml version="1.0" encoding="UTF-8"?>
<!--<nor xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="NOR.xsd">
<Nor type="GenericSpreadSheet" name="Spanish Employment">
  <Schema>
    <SchemaEntities>
      <SchemaEntity name="Location">
        <Attribute name="Name" valueFrom="pcaxis.[A11:A62]" type="string"/>
      </SchemaEntity>
      <SchemaEntity name="Dataset">
        <Attribute name="Name" valueFrom="pcaxis.[B6:H6]" type="string"/>
      </SchemaEntity>
      <SchemaEntity name="Year">
        <Attribute name="Name" valueFrom="pcaxis.[B9:H9]" type="string"/>
      </SchemaEntity>
    </SchemaEntities>
    <SchemaEntity name="IndustryProductionIndex" type="Nary">
      <Attribute name="hasValue" valueFrom="pcaxis.[B10:H62]" type="string"/>
      <Relation name="inArea" usingSpreadSheetColumn="A" destination="Location"/>
      <Relation name="inPeriod" usingSpreadSheetRow="9" destination="Year"/>
      <Relation name="dataset" usingSpreadSheetRow="6" destination="Dataset"/>
    </SchemaEntity>
  </Schema>
</Schema>
<DataModel>
  <GenericDataModel/>
</DataModel>
<Implementation>
  <Spreadsheet type="ms" file="Empresas.xls"/>
</Implementation>
</Nor>

```

NOR<sub>2</sub>O

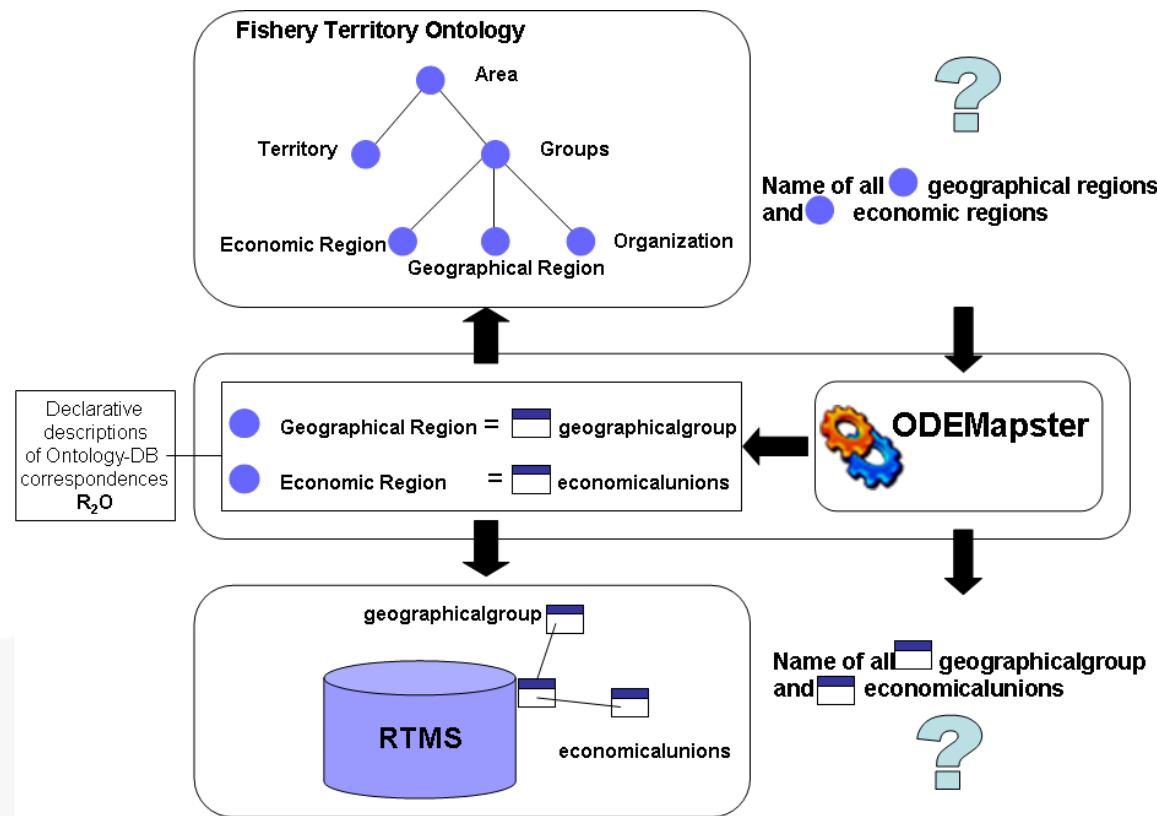
```

<!-- http://geo.linkeddata.es/resource/Estad%C3%ADstica/Madrid_ipi_2003 -->
<scv:Item rdf:about="Estat%C3%ADstica/Madrid_ipi_2003">
  <rdfs:label xml:lang="es">indice de Producci&#243;n Industrial de Madrid en el a&#241;o 2003</rdfs:label>
  <rdf:value rdf:datatype="&xsd;double">407655.0</rdf:value>
  <scv:dimension rdf:resource="A%C3%B1o/2003"/>
  <scv:dataset rdf:resource="Estat%C3%ADstica/%25C3%258DndiceDeProducci%25C3%25B3nIndustrial"/>
  <onto:inProvincia rdf:resource="Provincia/Madrid"/>
</scv:Item>
<!-- http://geo.linkeddata.es/resource/Estad%C3%ADstica/Madrid_ipi_2004 -->
<scv:Item rdf:about="Estat%C3%ADstica/Madrid_ipi_2004">
  <rdfs:label xml:lang="es">indice de Producci&#243;n Industrial de Madrid en el a&#241;o 2004</rdfs:label>
  <rdf:value rdf:datatype="&xsd;double">436074.0</rdf:value>
  <scv:dimension rdf:resource="A%C3%B1o/2004"/>
  <scv:dataset rdf:resource="Estat%C3%ADstica/%25C3%258DndiceDeProducci%25C3%25B3nIndustrial"/>
  <onto:inProvincia rdf:resource="Provincia/Madrid"/>
</scv:Item>

```

# GeoLinkedData - Transformation

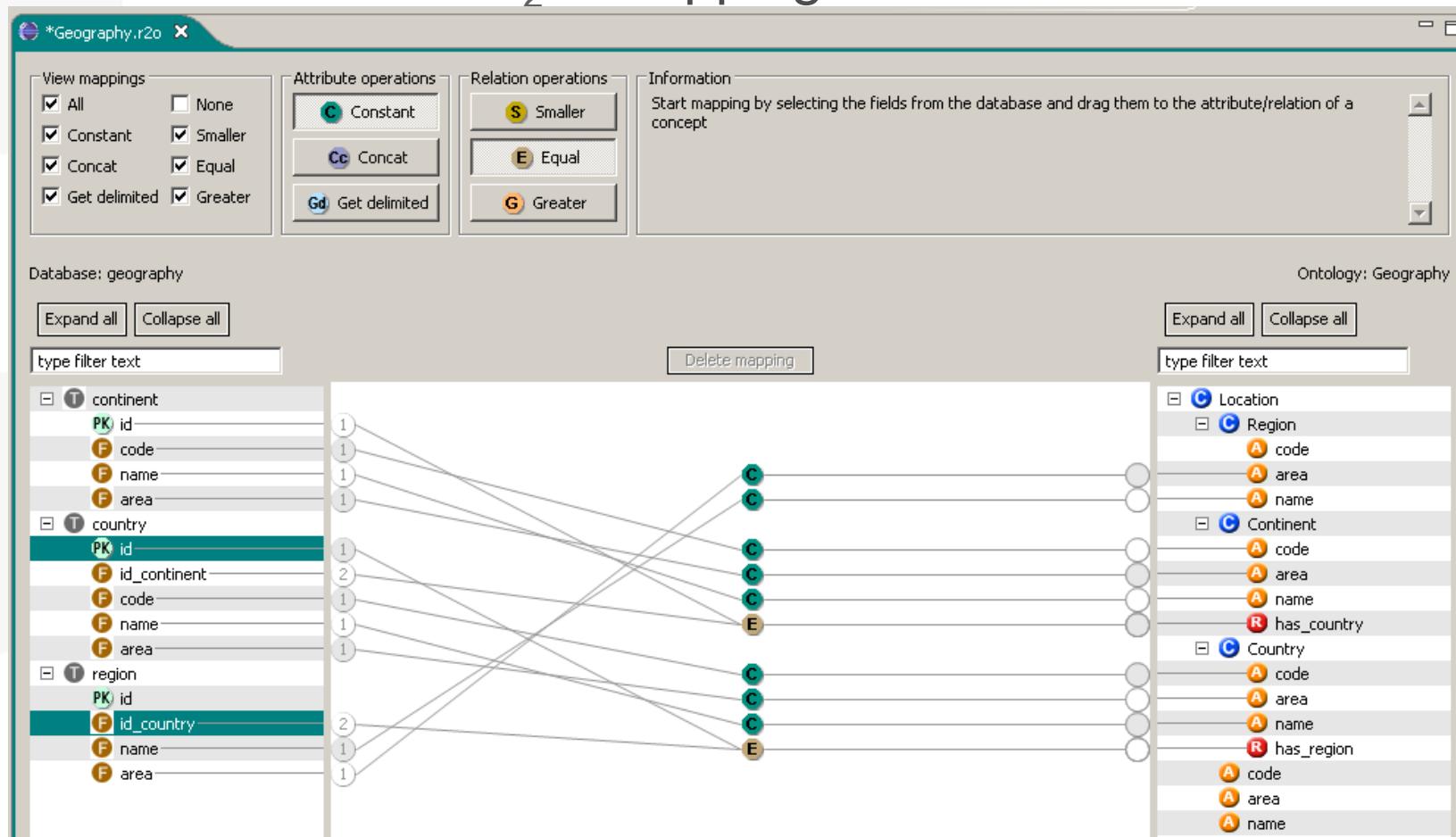
- R<sub>2</sub>O is an extensible, fully declarative language to describe mappings between relational database schemas and ontologies.
- The ODEMapster processor generates RDF instances from relational instances based on the mapping description expressed in the R<sub>2</sub>O document



[www.oeg-upm.net/index.php/en/downloads/9-r2o-odemaster](http://www.oeg-upm.net/index.php/en/downloads/9-r2o-odemaster)

# GeoLinkedData - Transformation

- Creation of the R<sub>2</sub>O Mappings

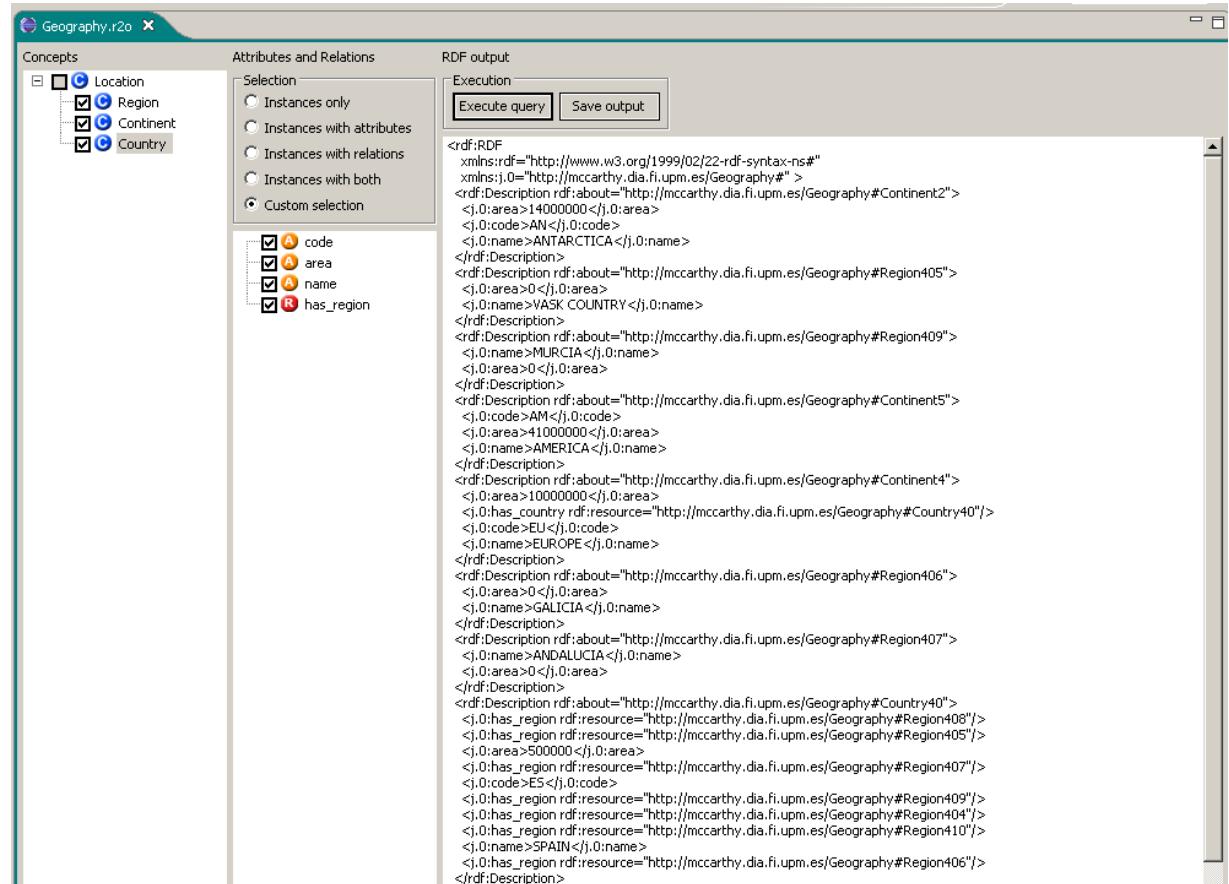


# GeoLinkedData - Transformation

```

<conceptmap-def name="http://mccarthy.dia.fi.upm.es/Geography#Region">
  <uri-as>
    <operation oper-id="concat">
      <arg-restriction on-param="string1">
        <has-value>http://mccarthy.dia.fi.upm.es/Geography#Region</has-value>
      </arg-restriction>
      <arg-restriction on-param="string2">
        <has-column>region.id</has-column>
      </arg-restriction>
    </operation>
  </uri-as>
  <described-by>
    </conceptmap-def>
  
```

## Excerpt of the R<sub>2</sub>O document



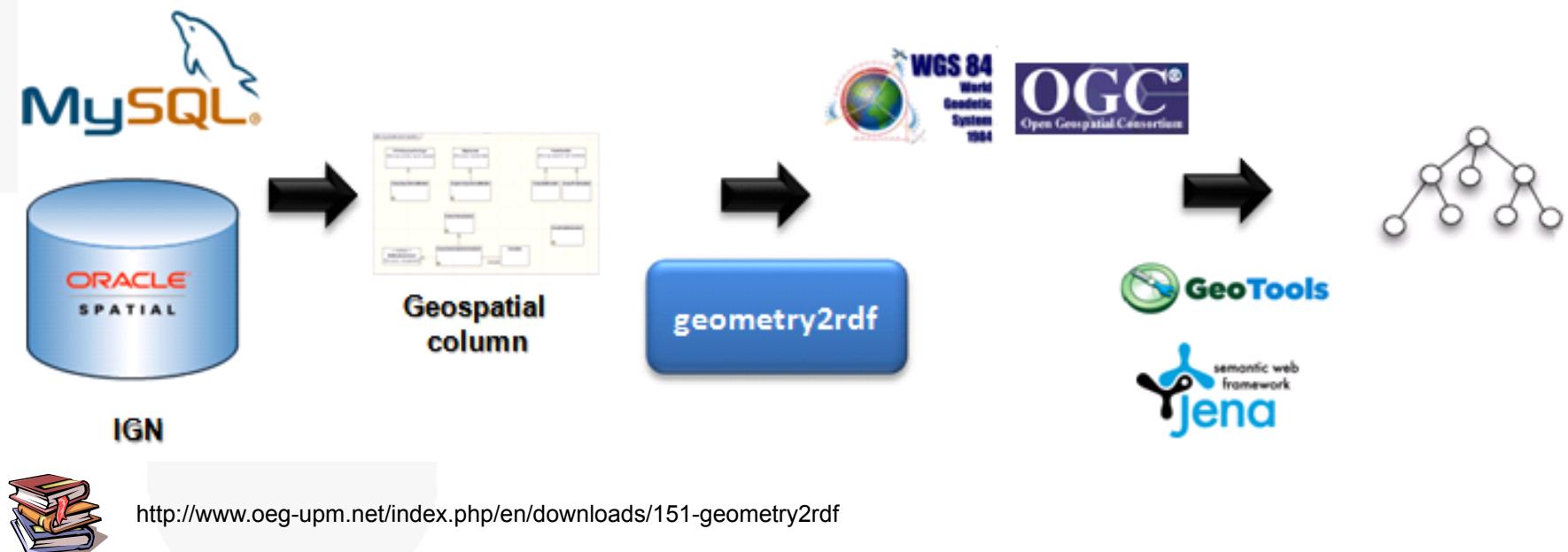
The screenshot shows the 'Geography.r2o' application window. On the left, there's a tree view under 'Concepts' with nodes for Location, Region, Continent, and Country, all of which are selected. To the right of the tree are sections for 'Attributes and Relations' and 'RDF output'. Under 'Attributes and Relations', there are checkboxes for 'Selection' (with 'Custom selection' checked) and several attributes: code (checked), area (checked), name (checked), and has\_region (checked). Below these is a large text area displaying an RDF query result. The result includes multiple descriptions for regions like ANTRARCTICA, MURCIA, VASK COUNTRY, AMERICA, and EUROPE, each with its code, area, name, and continent information.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:j_0="http://mccarthy.dia.fi.upm.es/Geography#"
  <j_0:area>14000000</j_0:area>
  <j_0:code>AN</j_0:code>
  <j_0:name>ANTARCTICA</j_0:name>
</rdf:Description>
<rdf:Description rdf:about="http://mccarthy.dia.fi.upm.es/Geography#Region405">
  <j_0:area></j_0:area>
  <j_0:name>VASK COUNTRY</j_0:name>
</rdf:Description>
<rdf:Description rdf:about="http://mccarthy.dia.fi.upm.es/Geography#Continent5">
  <j_0:code>AM</j_0:code>
  <j_0:area>41000000</j_0:area>
  <j_0:name>AMERICA</j_0:name>
</rdf:Description>
<rdf:Description rdf:about="http://mccarthy.dia.fi.upm.es/Geography#Continent4">
  <j_0:area>10000000</j_0:area>
  <j_0:has_country rdf:resource="http://mccarthy.dia.fi.upm.es/Geography#Country40"/>
  <j_0:code>EU</j_0:code>
  <j_0:name>EUROPE</j_0:name>
</rdf:Description>
<rdf:Description rdf:about="http://mccarthy.dia.fi.upm.es/Geography#Region406">
  <j_0:area></j_0:area>
  <j_0:name>GALICIA</j_0:name>
</rdf:Description>
<rdf:Description rdf:about="http://mccarthy.dia.fi.upm.es/Geography#Region407">
  <j_0:name>ANDALUCIA</j_0:name>
  <j_0:area></j_0:area>
</rdf:Description>
<rdf:Description rdf:about="http://mccarthy.dia.fi.upm.es/Geography#Country40">
  <j_0:has_region rdf:resource="http://mccarthy.dia.fi.upm.es/Geography#Region408"/>
  <j_0:has_region rdf:resource="http://mccarthy.dia.fi.upm.es/Geography#Region405"/>
  <j_0:area>500000</j_0:area>
  <j_0:has_region rdf:resource="http://mccarthy.dia.fi.upm.es/Geography#Region407"/>
  <j_0:code>ES</j_0:code>
  <j_0:has_region rdf:resource="http://mccarthy.dia.fi.upm.es/Geography#Region409"/>
  <j_0:has_region rdf:resource="http://mccarthy.dia.fi.upm.es/Geography#Region404"/>
  <j_0:has_region rdf:resource="http://mccarthy.dia.fi.upm.es/Geography#Region410"/>
  <j_0:name>SPAIN</j_0:name>
  <j_0:has_region rdf:resource="http://mccarthy.dia.fi.upm.es/Geography#Region406"/>
</rdf:Description>
  
```

# GeoLinkedData - Transformation

- Tool for generating RDF from geometrical information
- The geometry could be available in GML or WKT
- The RDF generated follows our Geometry Model



# GeoLinkedData - Transformation

NOMBRE	GMLGEOMETRY
Abengibre	(HUGELOB)
Alatoz	(HUGELOB)
Albatana	(HUGELOB)
Balsa de Ves	(HUGELOB)
Ballesteros, El	(HUGELOB)
Alborea	(HUGELOB)
Alcadozo	(HUGELOB)
Alcalá del Júcar	(HUGELOB)

## Oracle STO UTIL package

```
SELECT TO_CHAR(SDO_UTIL.TO_GML311GEOMETRY(geometry))
      AS Gml311Geometry
  FROM "BCN200"."BCN200_0301L_RIO" c
 WHERE c.Etiqueta='Arroyo'
```



```
<gml:Polygon srsName="SDO:8223" xmlns:gml="http://www.opengis.net/gml">
  <gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1.48374108,39.23127677,0 -1.48404
  -1.48480765,39.22531695,0 -1.48544493,39.2231538,0 -1.4858725,39.22212319,0 -1.48608465,39.22141415,0 -1.4860584:
  -1.48421128,39.21319056,0 -1.48390698,39.21064575,0 -1.48429166,39.2094524,0 -1.48529268,39.20490629,0 -1.4925061
  -1.49590016,39.20383639,0 -1.49905691,39.20405732,0 -1.50256876,39.20446292,0 -1.50338941,39.20453654,0 -1.504831
  -1.51378498,39.2047147,0 -1.51398161,39.20472622,0 -1.5171434,39.20471246,0 -1.51894734,39.20373528,0 -1.5196271:
  -1.52344981,39.20108038,0 -1.52699007,39.19901681,0 -1.5284903,39.19815278,0 -1.53151867,39.19619073,0 -1.5320861
  -1.5329951,39.19479473,0 -1.533988595,39.19503246,0 -1.53537021,39.19528414,0 -1.53607317,39.19544615,0 -1.536914
  -1.53953498,39.19602121,0 -1.54099994,39.19629185,0 -1.54110394,39.19630216,0 -1.54231647,39.19646151,0 -1.54461:
  -1.54957924,39.19633592,0 -1.55131402,39.1964386,0 -1.55382202,39.19666799,0 -1.55798564,39.19690882,0 -1.559800:
  -1.56633376,39.19750743,0 -1.5703604,39.19765603,0 -1.57231502,39.19777011,0 -1.57278711,39.19790204,0 -1.573636
  -1.57633787,39.19982863,0 -1.57807687,39.20087703,0 -1.57814542,39.20092292,0 -1.57871557,39.20136237,0 -1.57915
  -1.5805024,39.20233924,0 -1.58088715,39.20278543,0 -1.58107708,39.20313014,0 -1.58122865,39.20365457,0 -1.581447
  -1.5819531,39.20563664,0 -1.58191475,39.20581637,0 -1.58194513,39.20603299,0 -1.58285235,39.20756676,0 -1.584270:
  -1.5851004,39.2080704,0 -1.58540001,39.20907710,0 -1.58700074,39.21420,0 -1.58760009,39.21444070,0 -1.588004400
```

# GeoLinkedData - Transformation

```
<gml:Polygon srsName="SDO:8223" xmlns:gml="http://www.opengis.net/gml">
  <gml:outerBoundaryIs><gml:LinearRing><gml:coordinates decimal="." cs="" ts="">-1.48374108,39.23127677,0 -1.48404
  -1.4840765,39.22531695,0 -1.48544493,39.2231538,0 -1.4858725,39.22212319,0 -1.48608465,39.22141415,0 -1.4860584:
  -1.48421128,39.21319056,0 -1.48390698,39.21064575,0 -1.48429166,39.2094524,0 -1.48529268,39.20490629,0 -1.4925061
  -1.49890016,39.20383639,0 -1.49905691,39.20405732,0 -1.50256876,39.20446292,0 -1.50338941,39.20453654,0 -1.504831
  -1.51378498,39.2047147,0 -1.51398161,39.20472622,0 -1.5171434,39.20471246,0 -1.51894734,39.20373528,0 -1.5196271:
  -1.52344981,39.20108038,0 -1.52699007,39.19901681,0 -1.5284903,39.19815278,0 -1.53151867,39.19619073,0 -1.5320861
  -1.5329951,39.19479473,0 -1.53398595,39.19503246,0 -1.53537021,39.19528414,0 -1.53607317,39.19544615,0 -1.536914
  -1.53953498,39.19602121,0 -1.54099994,39.19629185,0 -1.54110394,39.19630216,0 -1.54231647,39.19646151,0 -1.54461:
  -1.54957924,39.19633592,0 -1.55131402,39.1964386,0 -1.55382202,39.19666799,0 -1.55798564,39.19690882,0 -1.559800:
  -1.56633376,39.19750743,0 -1.5703604,39.19765603,0 -1.57231502,39.19777011,0 -1.57278711,39.19790204,0 -1.573636
  -1.57633787,39.19982863,0 -1.57807687,39.20087703,0 -1.57814542,39.20092292,0 -1.57871557,39.20136237,0 -1.579151
  -1.5805024,39.20233924,0 -1.58088715,39.20278543,0 -1.58107708,39.20313014,0 -1.58122865,39.20365457,0 -1.581447
  -1.5819531,39.20563664,0 -1.58191475,39.20581637,0 -1.58194513,39.20603299,0 -1.58285235,39.20756676,0 -1.584270:
  -1.584471,39.2080364,0 -1.58449004,39.20807712,0 -1.58450701,39.2081000,0 -1.58452702,39.2081300,0 -1.58454703,39.2081400:
```

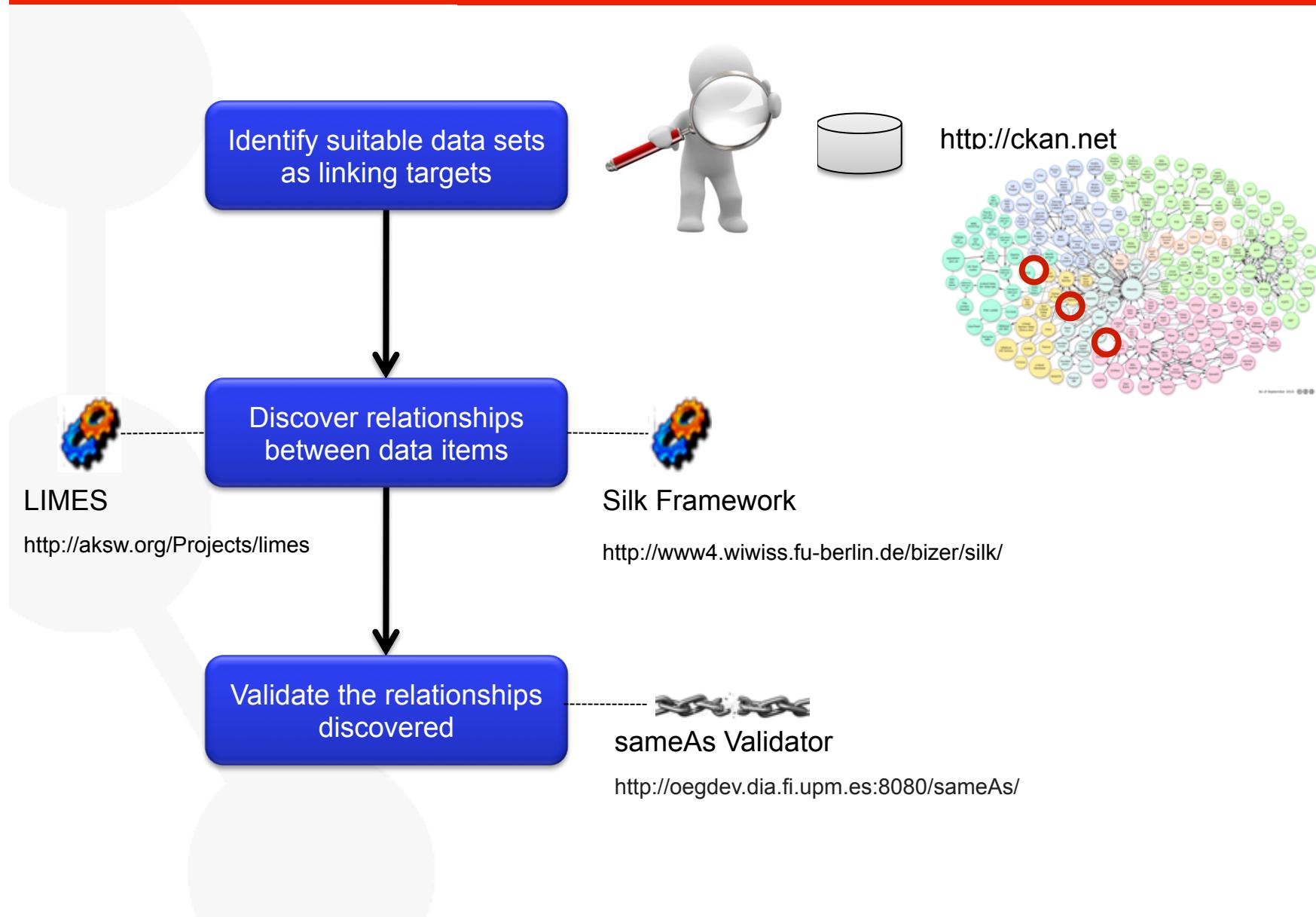


```
<rdf:Description rdf:about="http://geo.linkeddata.es/resource/wgs84/39.158523176194414_-1.6341427210305737">
  <geontology:order rdf:datatype="http://www.w3.org/2001/XMLSchema#int">205</geontology:order>
  <geo:long rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">-1.6341427210305737</geo:long>
  <geo:lat rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">39.158523176194414</geo:lat>
  <rdf:type rdf:resource="http://www.w3.org/2003/01/geo/wgs84_pos#Point"/>
</rdf:Description>
<rdf:Description rdf:about="http://geo.linkeddata.es/resource/wgs84/38.65501872007066_-2.6158910085778153">
  <geontology:order rdf:datatype="http://www.w3.org/2001/XMLSchema#int">6</geontology:order>
  <geo:long rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">-2.6158910085778153</geo:long>
  <geo:lat rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">38.65501872007066</geo:lat>
  <rdf:type rdf:resource="http://www.w3.org/2003/01/geo/wgs84_pos#Point"/>
</rdf:Description>
<rdf:Description rdf:about="http://geo.linkeddata.es/resource/wgs84/38.37015012539368_-1.6964507982671462">
  <geontology:order rdf:datatype="http://www.w3.org/2001/XMLSchema#int">141</geontology:order>
  <geo:long rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">-1.6964507982671462</geo:long>
  <geo:lat rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">38.37015012539368</geo:lat>
  <rdf:type rdf:resource="http://www.w3.org/2003/01/geo/wgs84_pos#Point"/>
</rdf:Description>
```

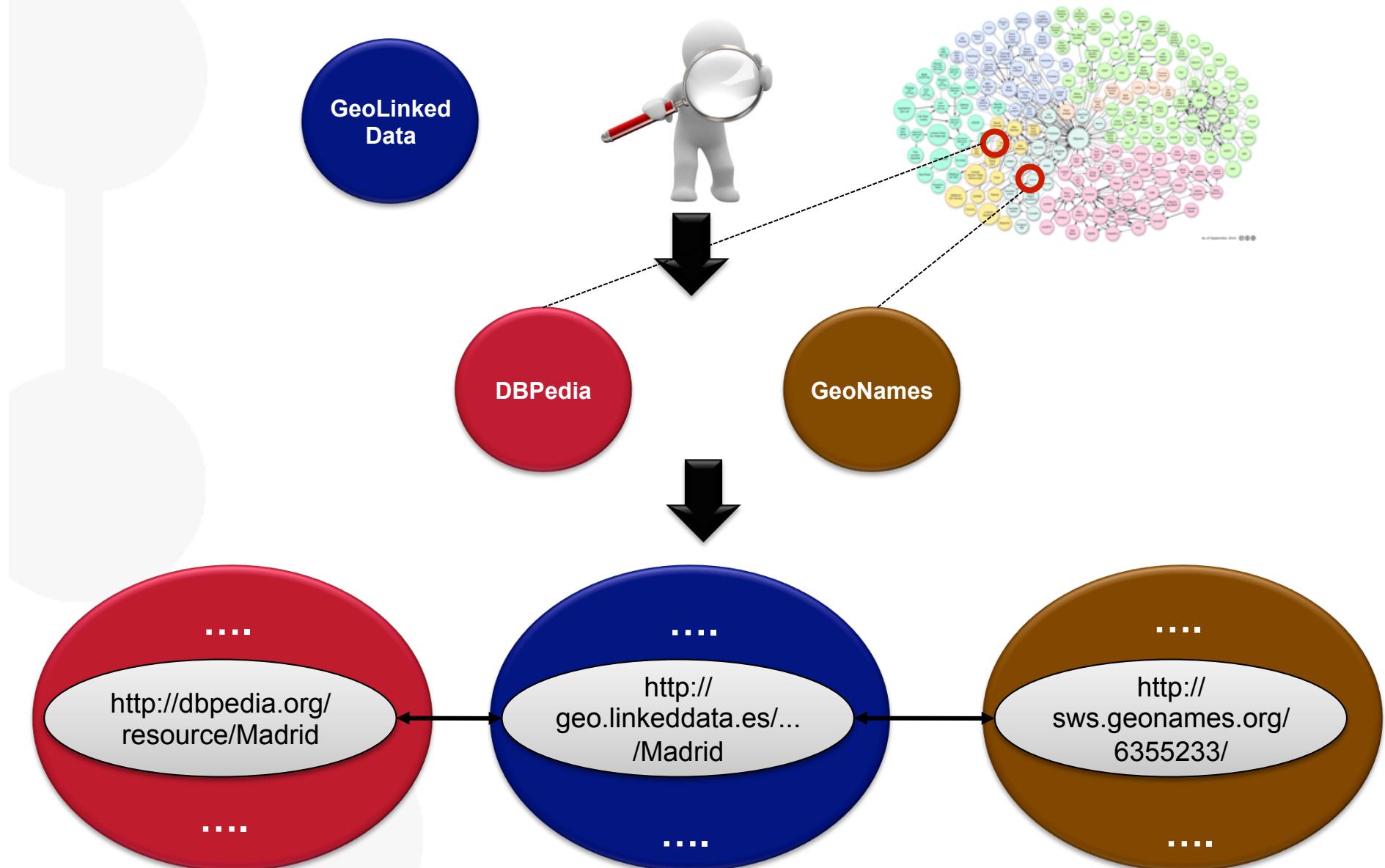
- To find possible errors, identified by Hogan et al.
  - http-level issues, such as accessibility and derefencability, e.g., HTTP URIs return 40x/50x errors
  - reasoning issues such as namespace without vocabulary, e.g., *rss:item* term invented
  - malformed/incompatible datatypes, e.g., “true” as xsd:int
- To fix the identified errors

# GeoLinkedData – Data Cleansing

- Errors
  - Some resources, with the same name, were mixed. For example, Granada municipality belongs to Granada province, and La Granada municipality belongs to Barcelona Province.
  - Autonomous communities that only have one province, e.g., Murcia Region, missed some municipalities, but their corresponding provinces, e.g., Murcia Province, have the correct number of municipalities.
  - Some hydrographical resources missed some parts of their geometrical information.



# GeoLinkedData - Linking



# GeoLinkedData - Linking

Are both resources equivalent?		
Yes: <input type="radio"/>	No: <input type="radio"/>	N/A: <input type="radio"/>
<input type="button" value="Submit"/>		

[http://dbpedia.org/resource/Province\\_of\\_M%C3%A1laga](http://dbpedia.org/resource/Province_of_M%C3%A1laga)

**About: Province of Málaga**

An Entity of Type : [Provinces of Spain](#), from Named Graph : <http://dbpedia.org>, within Data Space : [dbpedia.org](http://dbpedia.org)

The Province of Málaga (Spanish Provincia de Málaga) is located on the southern coast of Spain, in the Autonomous Community of Andalusia. It is bordered by the Mediterranean Sea to the South, and by the provinces of Cádiz, Sevilla, Córdoba and Granada. Its area is 7,308 km<sup>2</sup>. Its population is 1,330,010 (2002), of whom two-fifths live in the capital Málaga, and its population density is 181.99/km<sup>2</sup>.

Property	Value
dbpedia-owl:abstract	<ul style="list-style-type: none"> <li>Die Provinz Málaga (span. Provincia de Málaga) ist eine der si</li> <li>The Province of Málaga (Spanish Provincia de Málaga) is loca</li> <li>South, and by the provinces of Cádiz, Sevilla, Córdoba and Gr</li> <li>density is 181.99/km<sup>2</sup>. Its main industry and claim to fame is i</li> <li>European tourists. But besides the beaches, the mountainous</li> <li>composer Ernesto Lecuona, "Malagueña", is named for the m</li> <li>Besides the capital, its main cities are Marbella, Vélez-Málag</li> <li>The population density surpasses both the Andalusia and Spa</li> <li>located in the interior. The prevailing climate is a warm Medite</li> <li>the Eastern coastal zone has a subtropical Mediterranean clir</li> <li>Continental Mediterranean climate</li> <li>La provincia de Málaga es una de las ocho provincias español</li> <li>las provincias de Granada, al este, y Cádiz, al oeste. Al norte</li> <li>101 municipios, 9 comarcas y 11 partidos judiciales. Su pobla</li> <li>España por población. Quedó constituida como provincia en l</li> <li>Sevilla. El código postal de los municipios de Málaga empieza</li> <li>Málaga en su punto más meridional, en el término Femení</li> </ul>

<http://geo.linkeddata.es/resource/Provincia/M%C3%A1laga>

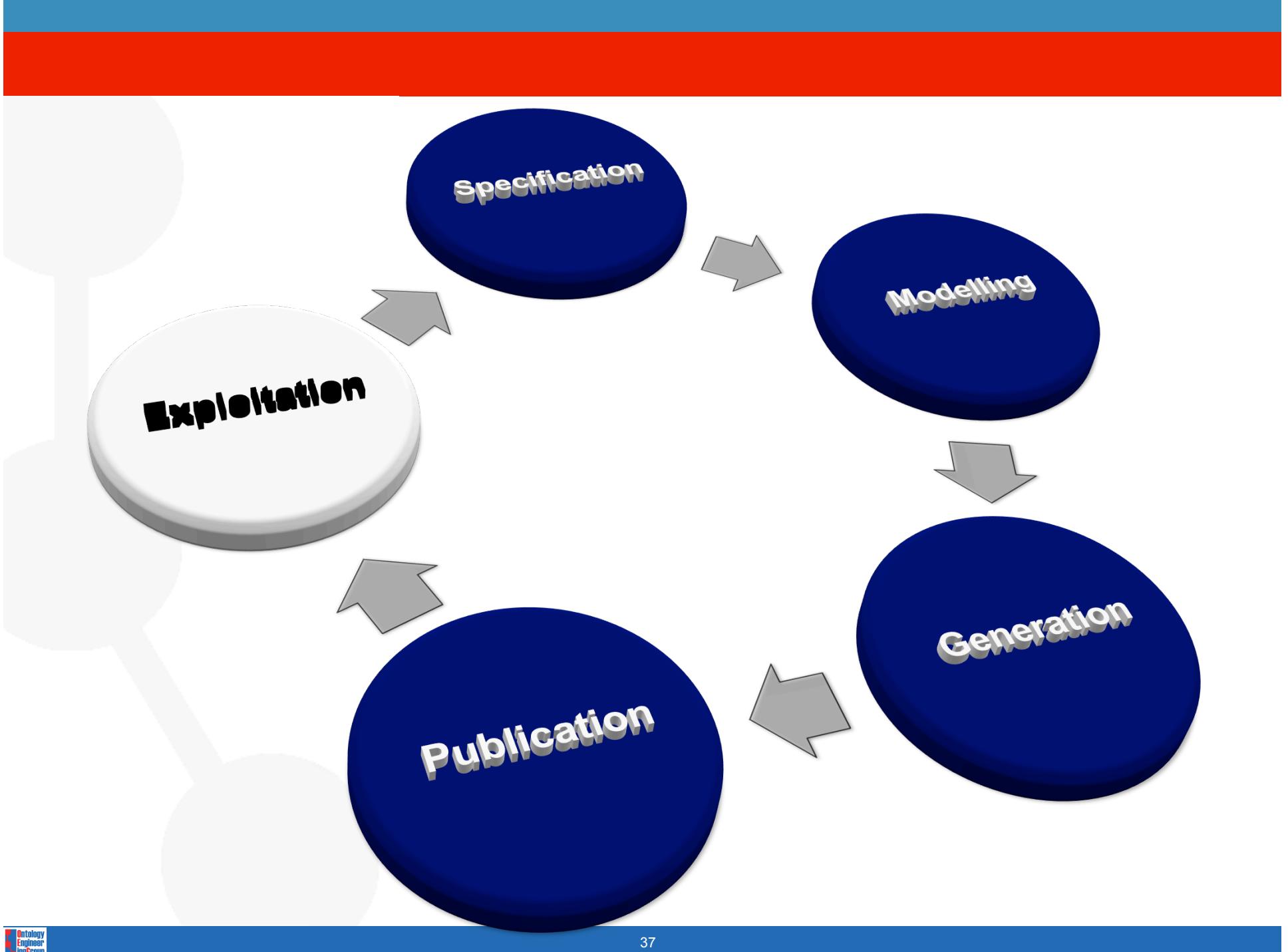
**Málaga** at geo.linkeddata.es

<http://geo.linkeddata.es/resource/Provincia/M%C3%A1laga>

Property	Value
geoes.formaParteDe	< <a href="http://geo.linkeddata.es/resource/ComunidadAut%C3%B3noma/Andaluc%C3%ADa">http://geo.linkeddata.es/resource/ComunidadAut%C3%B3noma/Andaluc%C3%ADa</a> >
geoes.formadoPor	<ul style="list-style-type: none"> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/%C3%81lora">http://geo.linkeddata.es/resource/Municipio/%C3%81lora</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/%C3%81rchez">http://geo.linkeddata.es/resource/Municipio/%C3%81rchez</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Alameda">http://geo.linkeddata.es/resource/Municipio/Alameda</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Alcauc%C3%ADn">http://geo.linkeddata.es/resource/Municipio/Alcauc%C3%ADn</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Alfarate">http://geo.linkeddata.es/resource/Municipio/Alfarate</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Alfaratejo">http://geo.linkeddata.es/resource/Municipio/Alfaratejo</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Algarrobo">http://geo.linkeddata.es/resource/Municipio/Algarrobo</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Algatoc%C3%ADn">http://geo.linkeddata.es/resource/Municipio/Algatoc%C3%ADn</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Alhaur%C3%ADn">http://geo.linkeddata.es/resource/Municipio/Alhaur%C3%ADn</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Alhaur%C3%ADn%20de%20la%20Gra">http://geo.linkeddata.es/resource/Municipio/Alhaur%C3%ADn%20de%20la%20Gra</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Alm%C3%A1char">http://geo.linkeddata.es/resource/Municipio/Alm%C3%A1char</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Almargen">http://geo.linkeddata.es/resource/Municipio/Almargen</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Almog%C3%ADa">http://geo.linkeddata.es/resource/Municipio/Almog%C3%ADa</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Alozaina">http://geo.linkeddata.es/resource/Municipio/Alozaina</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Alpandeire">http://geo.linkeddata.es/resource/Municipio/Alpandeire</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Artequer">http://geo.linkeddata.es/resource/Municipio/Artequer</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Archidona">http://geo.linkeddata.es/resource/Municipio/Archidona</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Ardales">http://geo.linkeddata.es/resource/Municipio/Ardales</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Arenas">http://geo.linkeddata.es/resource/Municipio/Arenas</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Ariate">http://geo.linkeddata.es/resource/Municipio/Ariate</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Atajate">http://geo.linkeddata.es/resource/Municipio/Atajate</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Benadalid">http://geo.linkeddata.es/resource/Municipio/Benadalid</a>&gt;</li> <li>&lt;<a href="http://geo.linkeddata.es/resource/Municipio/Benahav%C3%ADs">http://geo.linkeddata.es/resource/Municipio/Benahav%C3%ADs</a>&gt;</li> </ul>



<http://oegdev.dia.fi.upm.es:8080/sameAs/>



# Publication

- Dataset publication
- Metadata publication
- Dataset discovery

- Tools for storing RDF
  - Virtuoso Universal Server, Jena, Sesame, 4Store, YARS, OWLIM
- SPARQL endpoint and Linked Data frontend
  - Pubby, Talis Platform, Fuseki

- VoID allows to express metadata about RDF datasets



Describing Linked Datasets with the VoID Vocabulary

W3C Interest Group Note 03 March 2011

This version:

<http://www.w3.org/TR/2011/NOTE-void-20110303/>

Latest version:

<http://www.w3.org/TR/void/>

Authors:

[Keith Alexander](#) (Talis)

[Richard Cyganiak](#) (DERI, National University of Ireland, Galway)

[Michael Hausenblas](#) (DERI, National University of Ireland, Galway)

[Jun Zhao](#) (University of Oxford)

- Open Prov

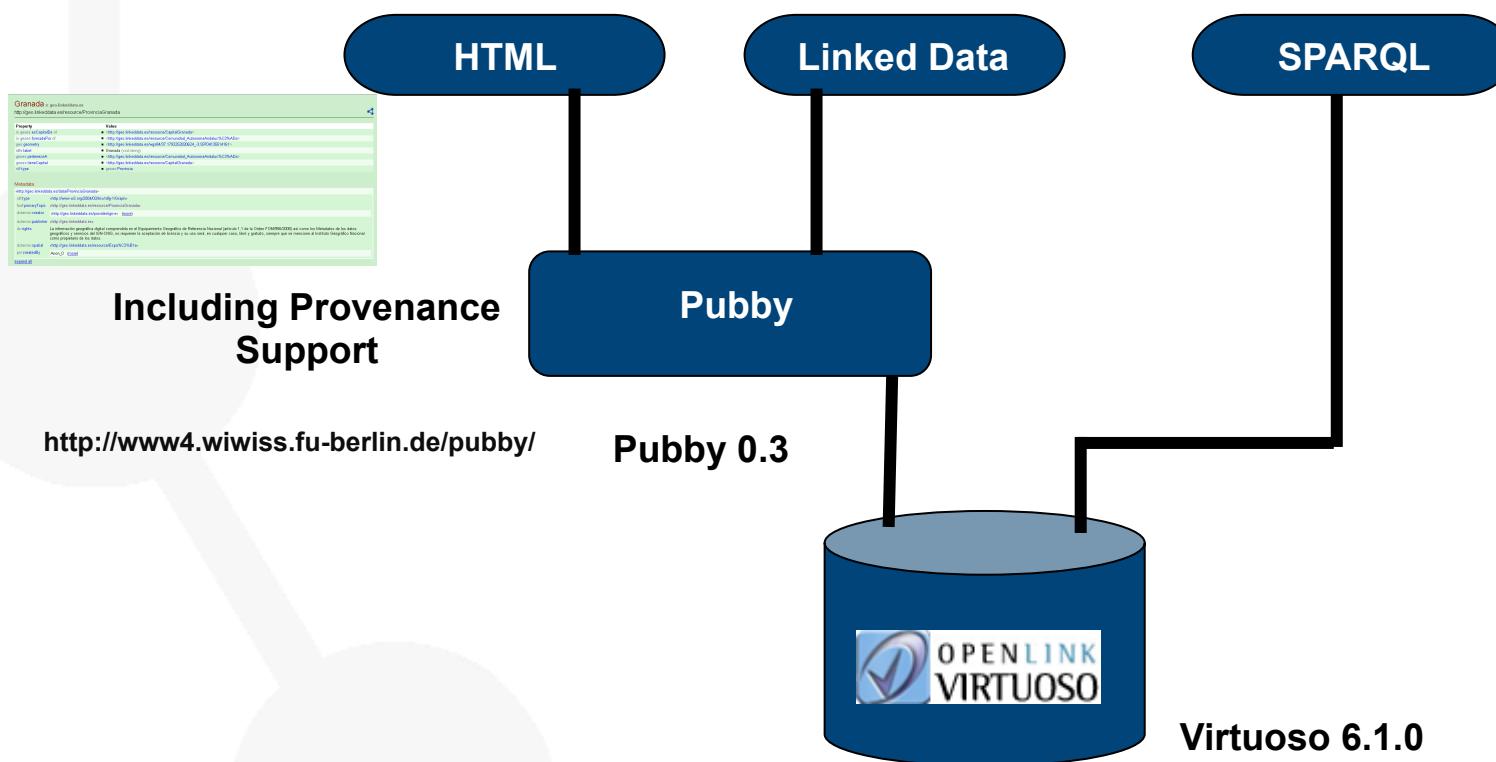


- Register the dataset into CKAN Registry
- Generate sitemap files for your dataset, by using sitemap4rdf
- Submit the sitemap location to Google and Sindice



<http://www.w3.org/wiki/TaskForces/CommunityProjects/LinkingOpenData/DataSets/CKANmetainformation>

# GeoLinkedData – Dataset publication



# GeoLinkedData – Dataset discovery

## GeoLinkedData

 View    Edit    History

GeoLinkedData (es) is an open initiative whose aim is to enrich the Web of Data with Spanish geospatial data. This initiative started off by publishing diverse information sources belonging to the National Geographic Institute of Spain. Such sources are made available as RDF (Resource Description Framework) knowledge bases according to the Linked Data principles. These data are interlinked with other knowledge bases belonging to the Linking Open Data Initiative.

### Resources

Resource	Format	Actions
SPARQL endpoint	api/sparql	
Download	application/x-ntriples	
RDF Schema	meta/rdf-schema	
Link to an example data item within the dataset (RDF/XML)	example/rdf+xml	

### Additional Information

Field	Value
links:dbpedia	51
links:geonames-semantic-web	51
namespace	<a href="http://geo.linkeddata.es/resource/">http://geo.linkeddata.es/resource/</a>
triples	21564199

First time at the Data Hub?

the Data Hub is a catalogue for data.  
[Click here to find out more ...](#)

#### Source

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

#### Author

Ontology Engineering Group, Facultad de Informática, Universidad Politécnica de Madrid.

#### Maintainer

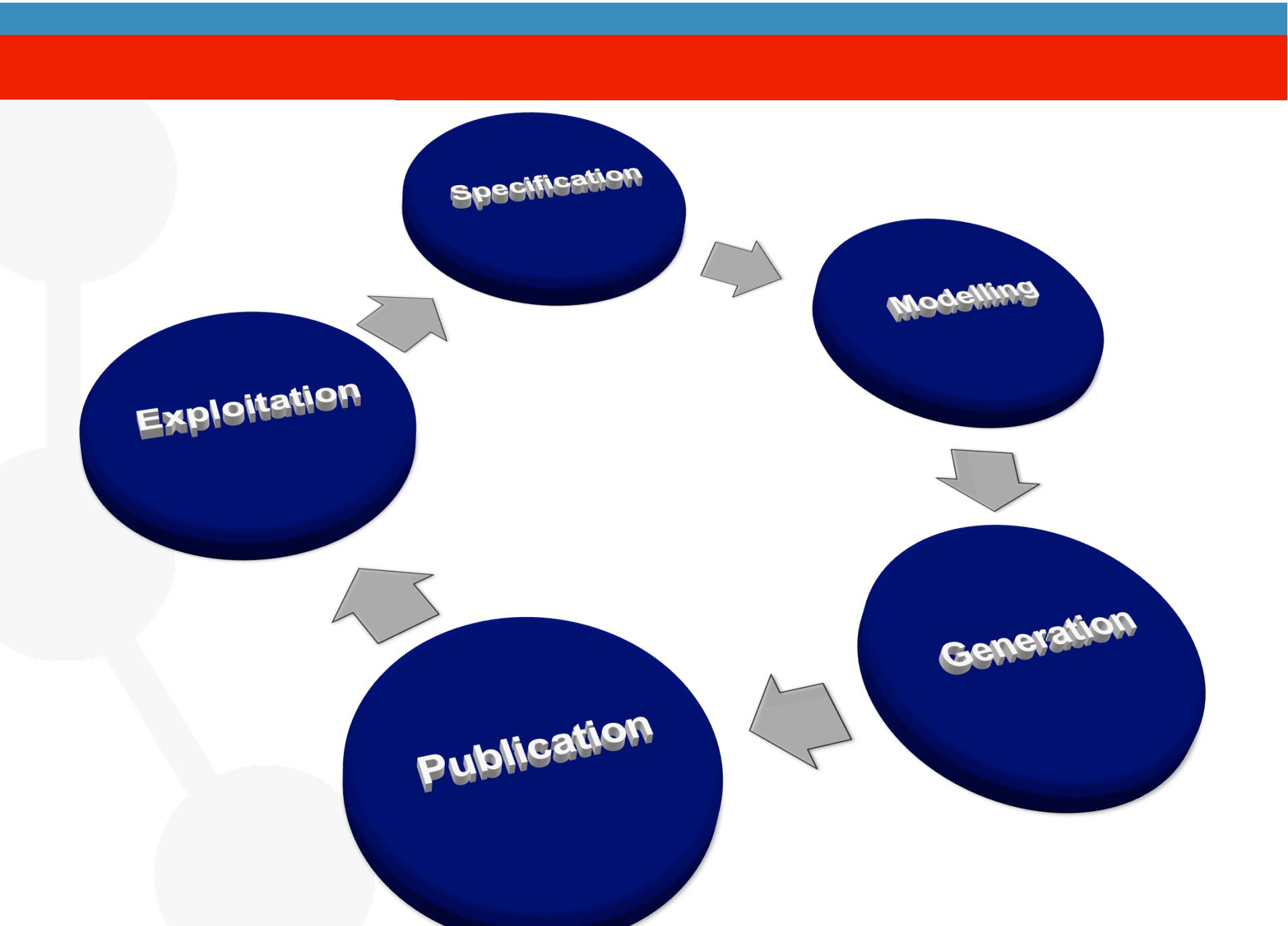
GeoLinkedData Team

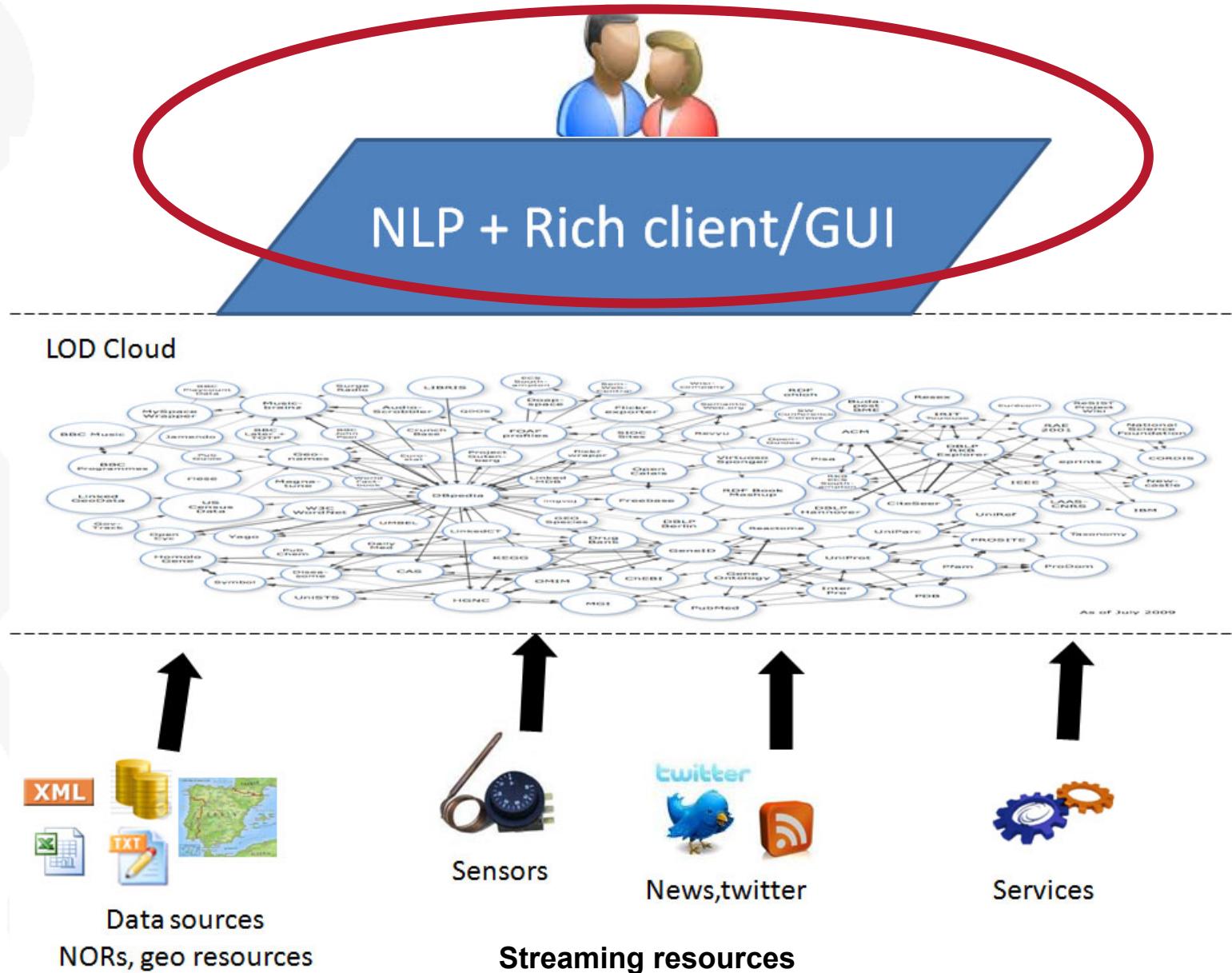
#### Version

2010-08

#### Tags

- country-spain
- deref-vocab
- format-dc
- format-rdf
- geodata
- geographic
- government
- lod
- no-license-metadata
- no-provenance-metadata





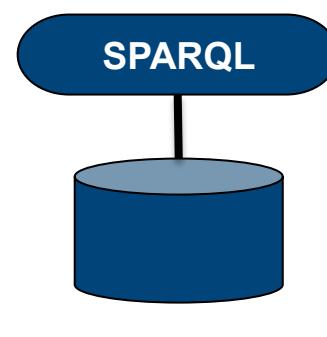
<http://oegdev.dia.fi.upm.es/projects/map4rdf/>



## map4rdf

### map4rdf:

- Google maps viewer of RDF resources
  - Resources with spatial information
- Extensible with google plugins
- Used in other applications like Aemet, Goodrelations



# DEMO

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

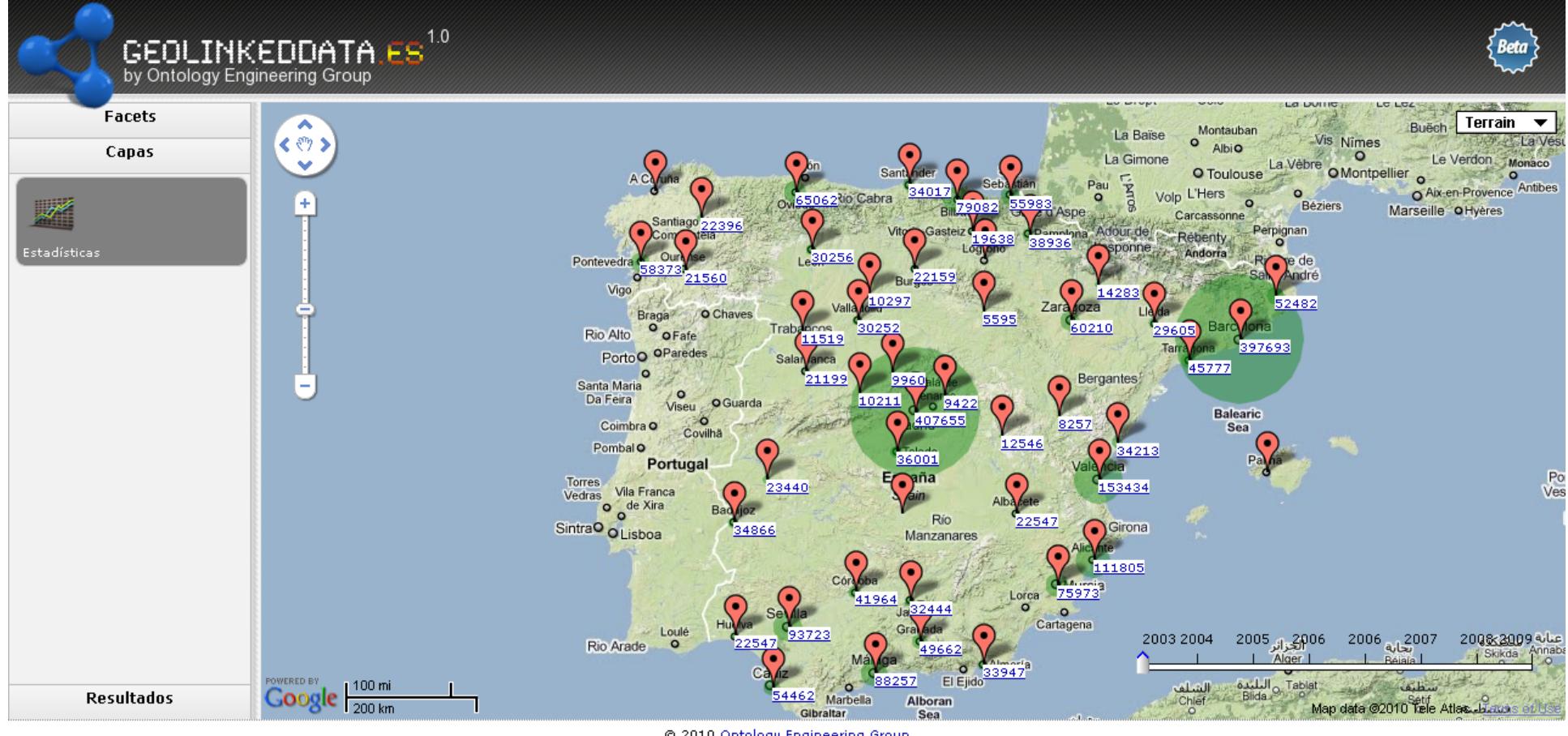
# Provinces



# Capital of Province



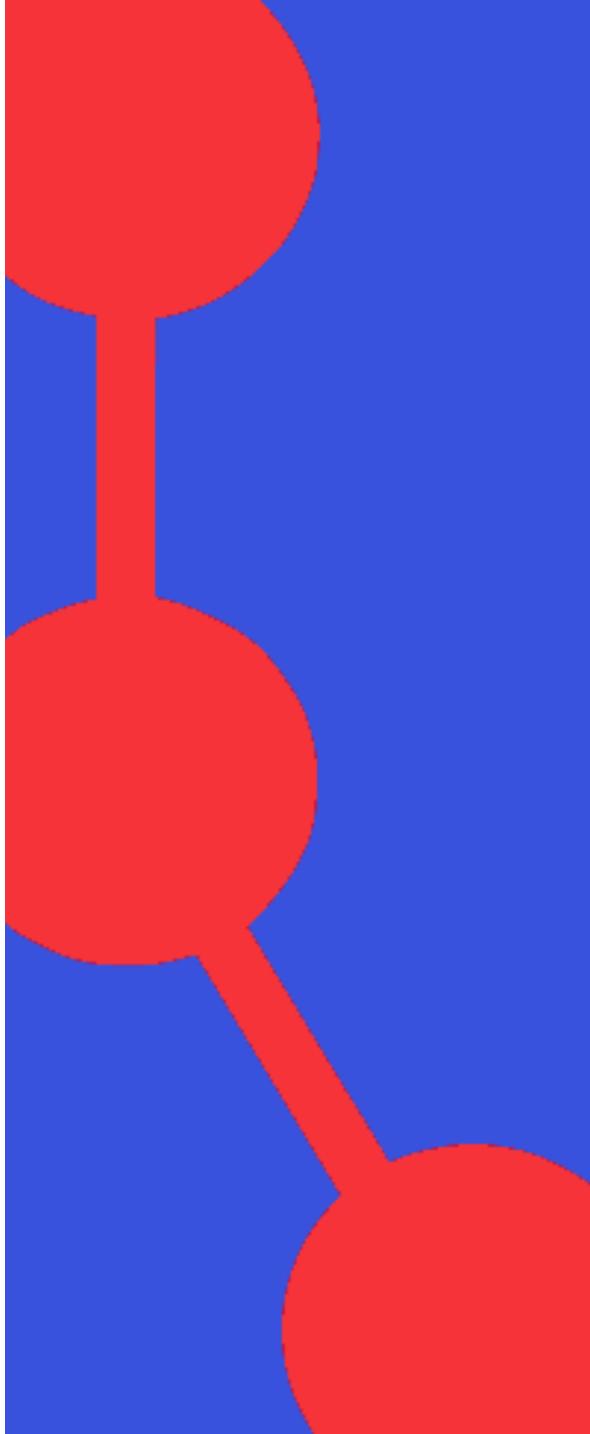
# Provinces – Industry Production Index



# Beaches







# Methodological Guidelines for Publishing Linked Data

Boris Villazón-Terrazas, Oscar Corcho

Facultad de Informática, Universidad Politécnica de Madrid

Campus de Montegancedo sn, 28660 Boadilla del Monte, Madrid

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

[bvillazon@fi.upm.es](mailto:bvillazon@fi.upm.es)

Phone: 34.91.3366605, Fax: 34.91.3524819

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

**Acknowledgements:** Asunción Gómez-Pérez, Luis M. Vilches,  
Victor Saquicela, Alexander de León

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