# How to License Linked Data

Finding and declaring known licenses
A practical approach

***Tutorial: Rights and Licenses for Linked Data***
**within the European Semantic Web Conference 2014 ESWC14**

**Víctor Rodríguez-Doncel**

*Ontology Engineering Group*
*Universidad Politécnica de Madrid (Spain)*

Hersonissos (Greece)

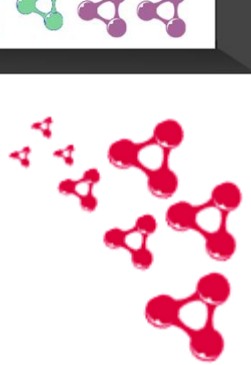
2014.05.26

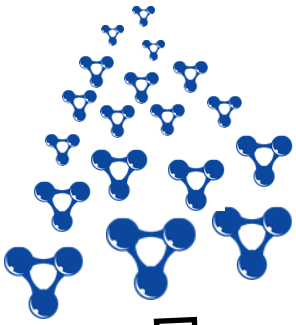
# Introduction: Objective

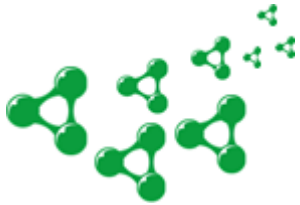
Data portals

Government

Individuals



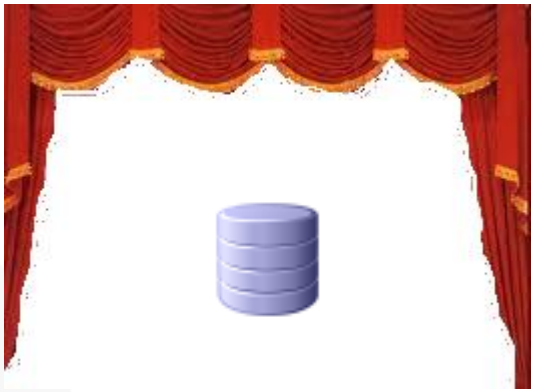




Create, consume, aggregate, derive and publish Linked Data being respectful with licenses

# Introduction: Licensed Linked Data

## Open Data



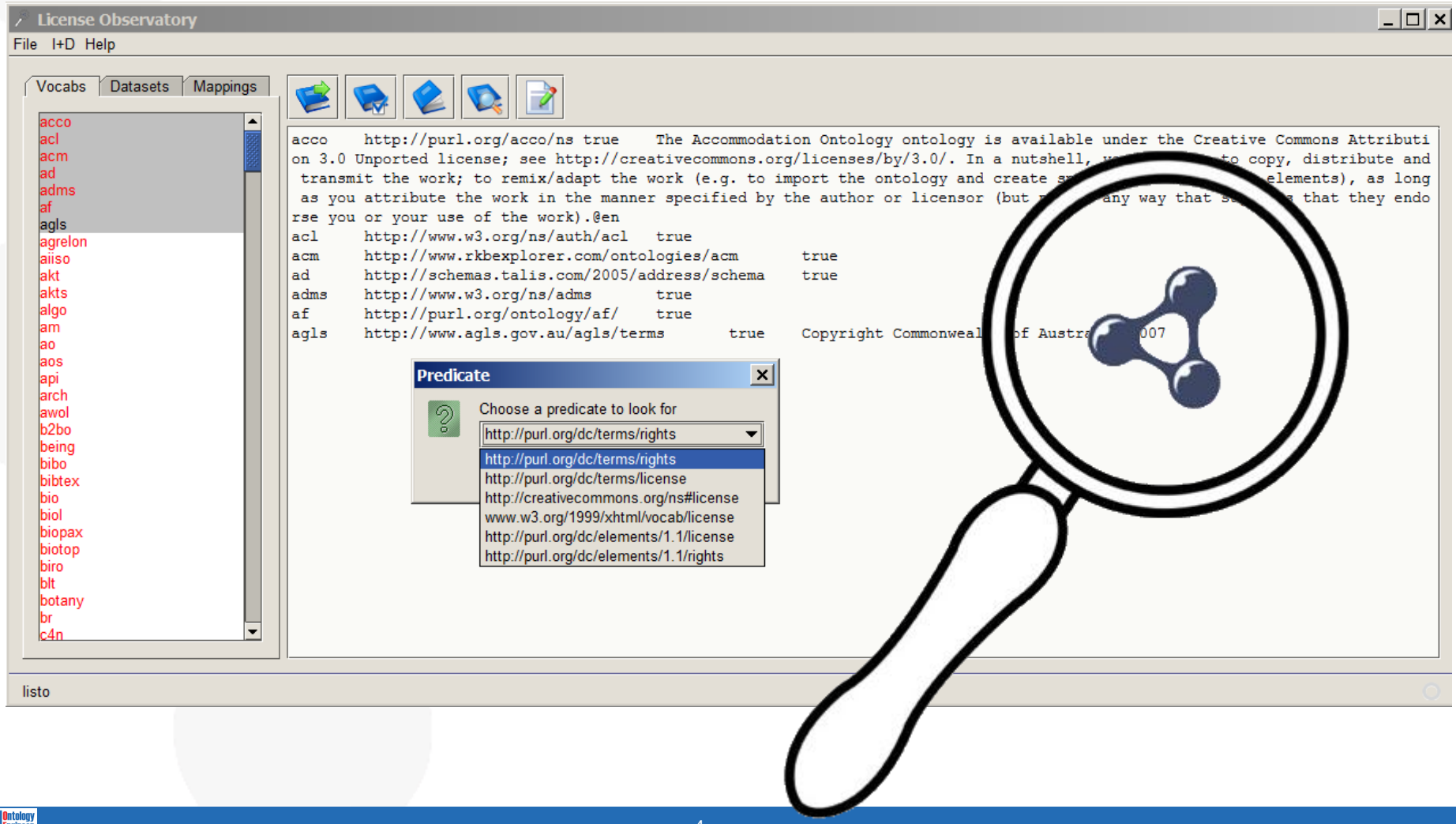
Gain visibility
Encourage re-use
Accesible, but in any case, **licensed**

## Propietary Data

Protect your data
Enable ways to track usage
New business models

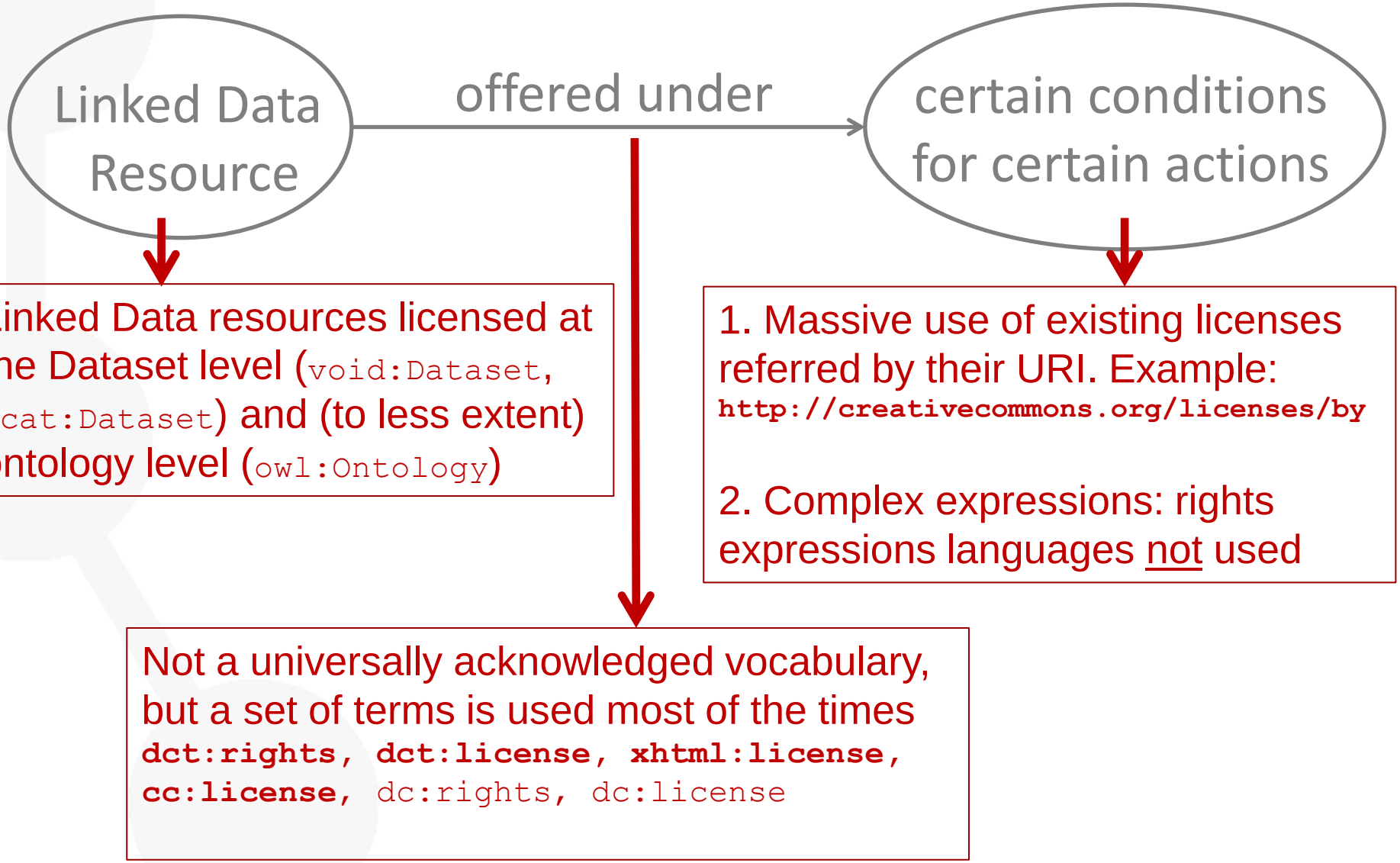
# Introduction: the current state of the affairs

- An observation on how Linked Data was being licensed…

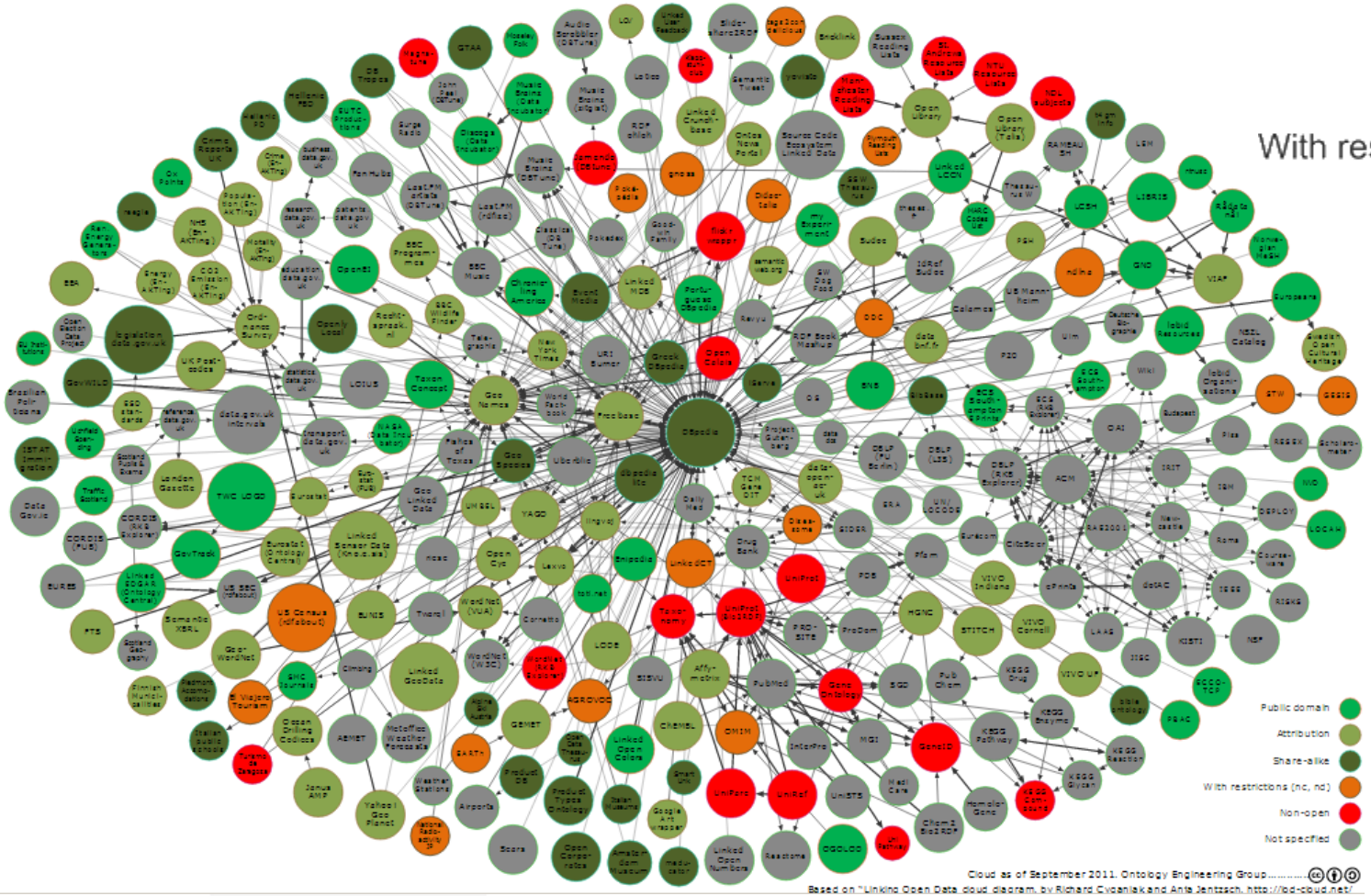


# Introduction: the current state of the affairs

- …revealing the following (qualitative results)

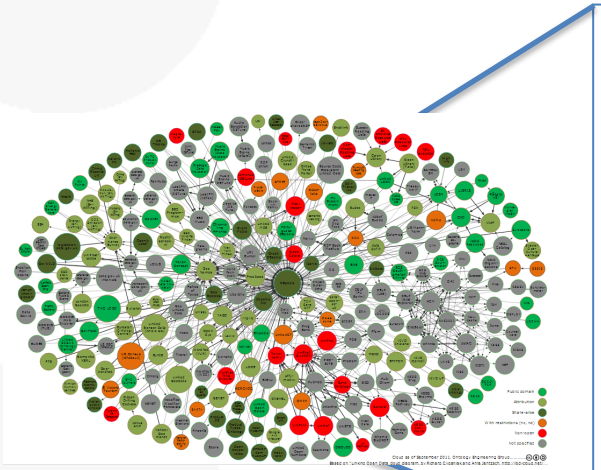


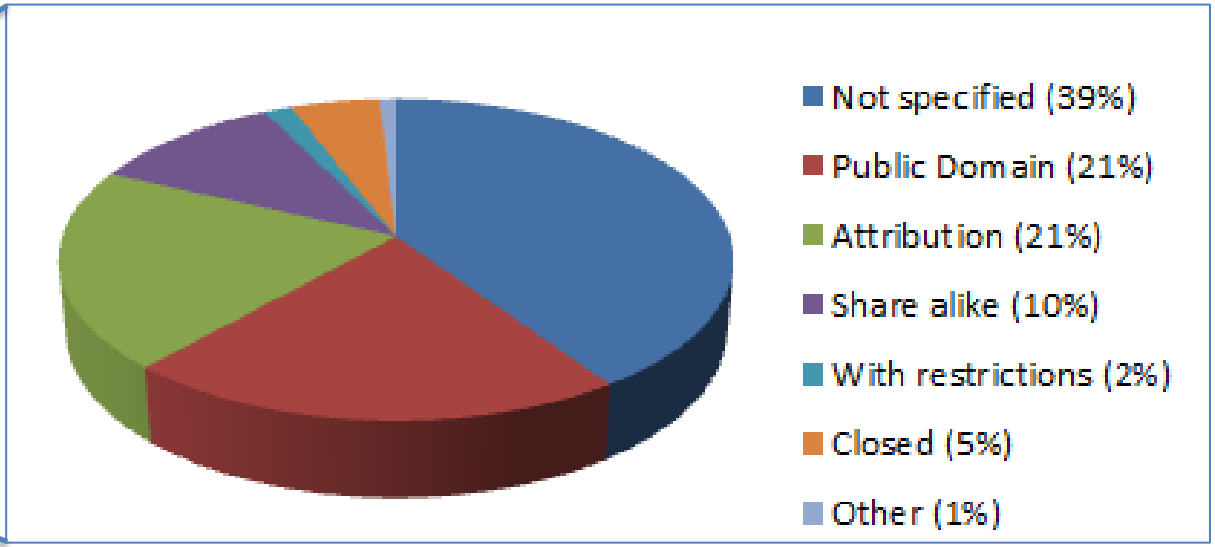
# Introduction: a coloured LOD cloud…



# Introduction: quantitative results

- 338 datasets in the official LOD cloud :






- Broader results

(vocabularies, non-LOD datasets)

datahub

| *Kind of license* | *No. of licenses in LOV vocabs* | *No. of licenses in CKAN datasets* |
|---|---|---|
| Not specified | 313 (86%) | 469 (26%) |
| Public Domain | 5 (1%) | 291 (16%) |
| Attribution | 18 (5%) | 440 (24%) |
| Share alike | 2 (2%) | 322 (18%) |
| With restrictions | 17 (5%) | 143 (8%) |
| Closed | 7 (2%) | 43(2%) |
| Other | 4 (1%) | 128(7%) |

[1] Ro
in Proc. of the 3rd

# Introduction: quantitative results

- Most used terms to declare the license

| *Element* | *LOV vocabs* | *LOD datasets* |
|---|---|---|
| `dct:license` | 34 (9%) | 62 (41%) |
| `cc:license` | 20 (5%) | 2 (1%) |
| `dct:rights` | 31 (8%) | 4 (3%) |

- But still not always introduced as predicates

```
  xmlns:skosxl="http://www.w3.org/2008/05/skos-xl#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#" >

<!--
This data file is a part of

Leavo
http://www.leavo.org/
Gerard de Melo, 2008-2013

License: Creative Commons Attribution-Share Alike 3.0 Unported
http://creativecommons.org/licenses/by-sa/3.0/

Portions copyright various contributors (e.g. article authors on Wikipedia).
For information about the data sources and the respective copyrights,
please see:
```

# Contents

- Be able to ….
  - **discover the license** of a Linked Data resource
  - **select** the appropriate license for your Linked Data resource
  - **declare the license** for your Linked Data resource

- Be able to ….
  - **discover the license** of a Linked Data resource
  - **select** the appropriate license for your Linked Data resource
  - **declare the license** for your Linked Data resource

# Discover the license

- Linked Data resources
  - Ontologies, vocabularies
  - Datasets as dump files (let's assume in RDF)
  - Linked Data from accesible endpoints (SPARQL etc.)
  - …

  →Each resource has different publication styles

# A full example

Eurostat RDF datasets ×

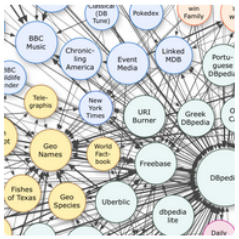
datahub.io/es/dataset/eurostat-rdf

/ Organizaciones / Linking Open Data Cloud / **Eurostat RDF datasets**

**Eurostat RDF datasets**

Seguidores
**1**

**Organización**



**Linking Open Data Cloud**

This group catalogs data sets that are available on the Web as Linked Data and contain data links pointing at other Linked Data sets. The descriptions of the data sets in this... leer más

**Social**

Google+

Twitter

Facebook

**Licencia**

Creative Commons Attribution

Conjunto de datos | Flujo de Actividad | Relacionados

## Eurostat RDF datasets

Three RDF datasets:

- Eurostat country codes
- Eurostat NUTS statistical regions (2008)
- Database of Eurostat-related legal acts

Each is available as an RDF dump. URIs resolve directly into the RDF dump. The country codes are linked to Geonames, to package:fu-berlin-eurostat, and to Telegraphis (40 links each).

### Datos y Recursos

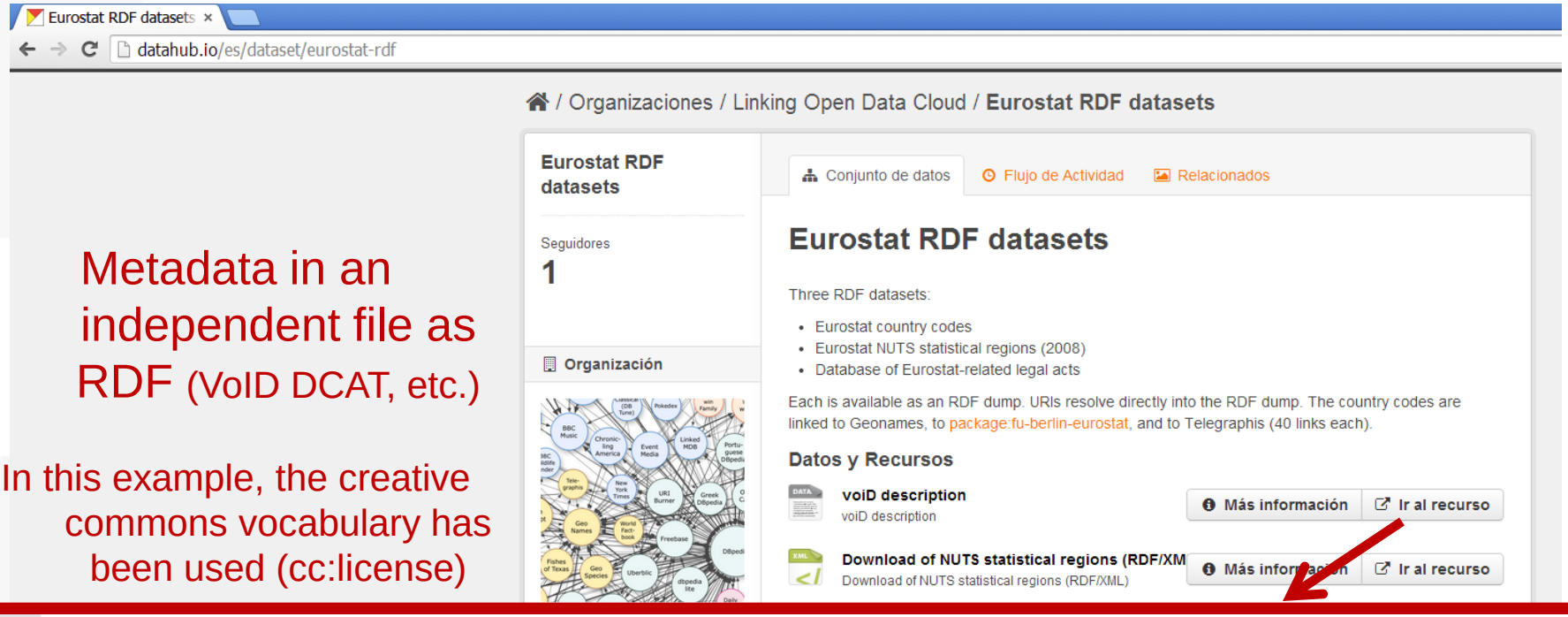
| | | |
|---|---|---|
| **voiD description**<br>voiD description | Más información | Ir al recurso |
| **Download of NUTS statistical regions (RDF/XM**<br>Download of NUTS statistical regions (RDF/XML) | Más información | Ir al recurso |
| **Download of country codes (RDF/XML)**<br>Download of country codes (RDF/XML) | Más información | Ir al recurso |
| **Download of legislation database (RDF/XML)**<br>Download of legislation database (RDF/XML) | Más información | Ir al recurso |
| **Download of NACE (Revision 2) codes (RDF/X**<br>Download of NACE (Revision 2) codes (RDF/XML) | Más información | Ir al recurso |
| **Example region (RDF/XML)**<br>Example region (RDF/XML) | Más información | Ir al recurso |
| **Example country (RDF/XML)**<br>Example country (RDF/XML) | Más información | Ir al recurso |
| **Example legislation record (RDF/XML)**<br>Example legislation record (RDF/XML) | Más información | Ir al recurso |
| **Download of NACE (Revision 2) translations ...**<br>Download of NACE (Revision 2) translations (RDF/XML) | Más información | Ir al recurso |

deref-vocab | eu | europe | eurostat | format-cc | format-rdf | geographic

# A full example

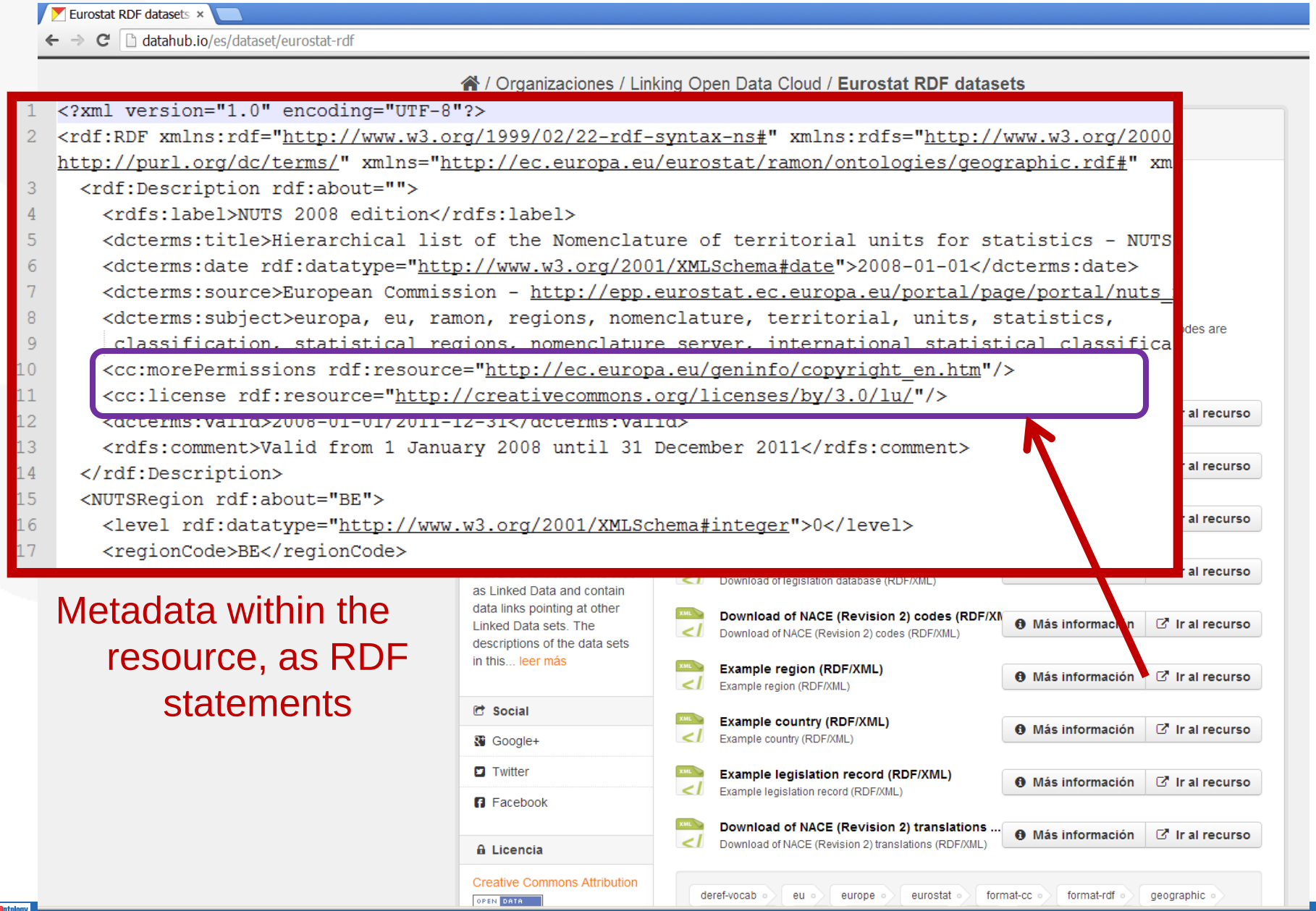
Metadata in an independent file as RDF (VoID DCAT, etc.)

In this example, the creative commons vocabulary has been used (cc:license)

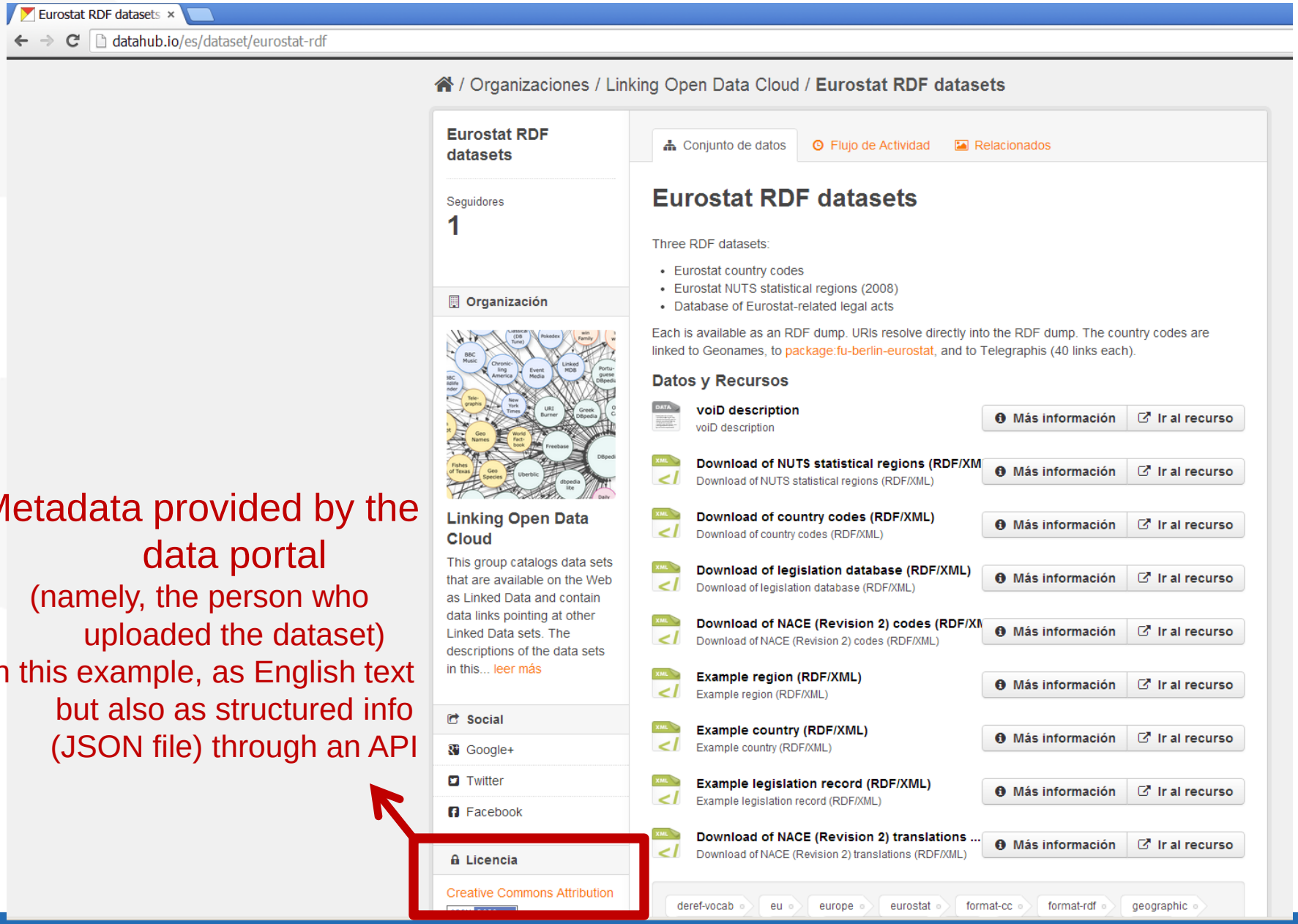



```
26  <void:Dataset rdf:ID="nuts2008">
27    <dcterms:title>NUTS 2008 edition</dcterms:title>
28    <dcterms:description>Hierarchical list of the Nomenclature of territorial units for statistics - NUTS
29    <dcterms:creator rdf:resource="#Eurostat"/>
30    <dcterms:subject rdf:resource="http://dbpedia.org/resource/Nomenclature_of_Territorial_Units_for_Stati
31    <dcterms:date>2008-01-01</dcterms:date>
32    <foaf:homepage rdf:resource="http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/int
33    <rdfs:seeAlso rdf:resource="&rdfdata;estat-legis/13000"/>
34    <cc:license rdf:resource="http://creativecommons.org/licenses/by/3.0/lu/"/>
35    <cc:morePermissions rdf:resource="http://ec.europa.eu/geninfo/copyright_en.htm"/>
36    <void:exampleResource rdf:resource="&rdfdata;nuts2008/LU"/>
37    <void:dataDump rdf:resource="&rdfdata;nuts2008.rdf"/>
38  </void:Dataset>
```

# A full example



# A full example

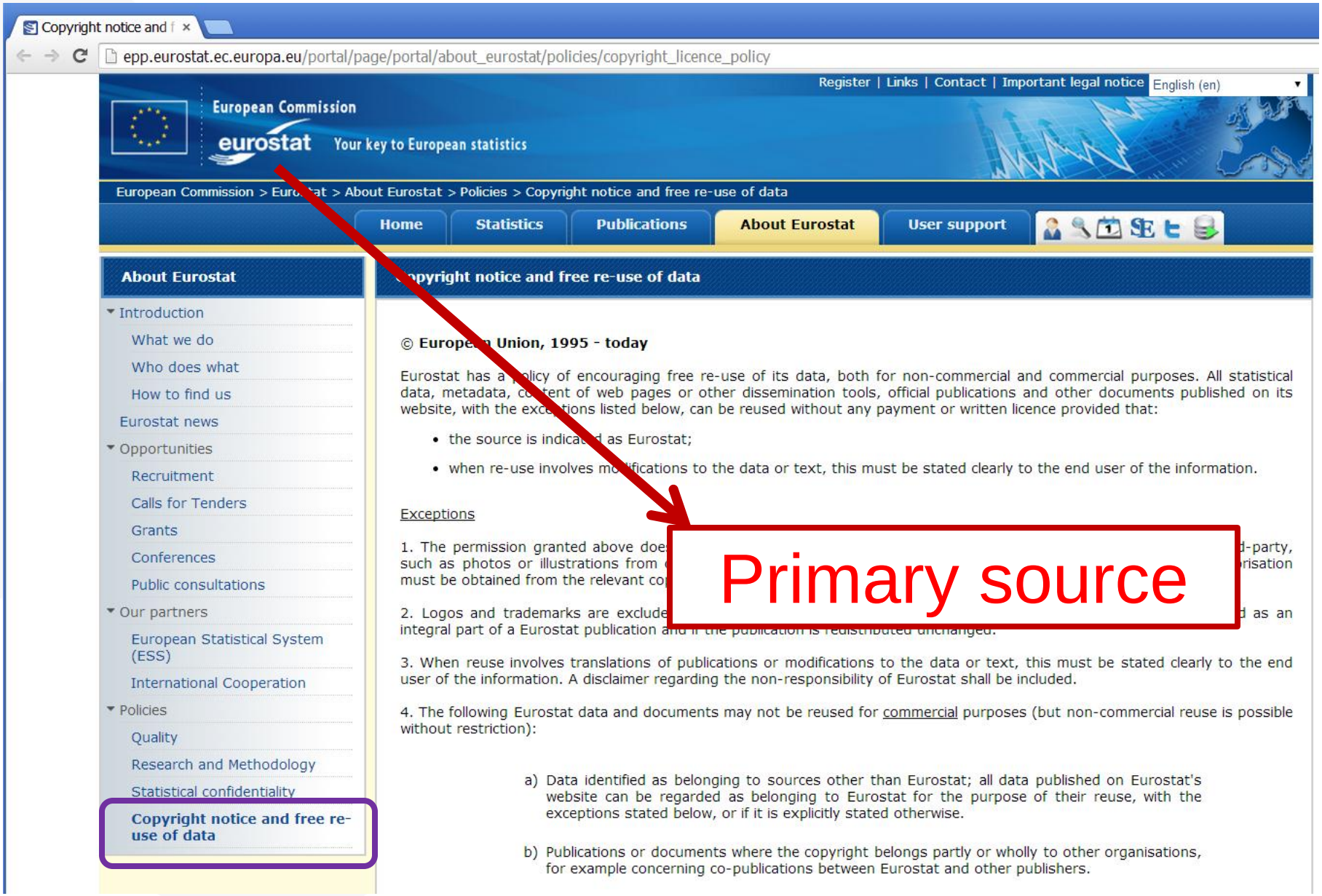



Metadata provided by the data portal

(namely, the person who uploaded the dataset)

In this example, as English text but also as structured info (JSON file) through an API

# A full example



# A full example

## Different targets

Machines (RDF, JSON via API)

Humans (Plain English text, icons)

## Different vocabs & terms

CC-REL, DublinCore,…

dc:license, dc:rights, etc.

## Different detail levels

Creative Commons Attribution

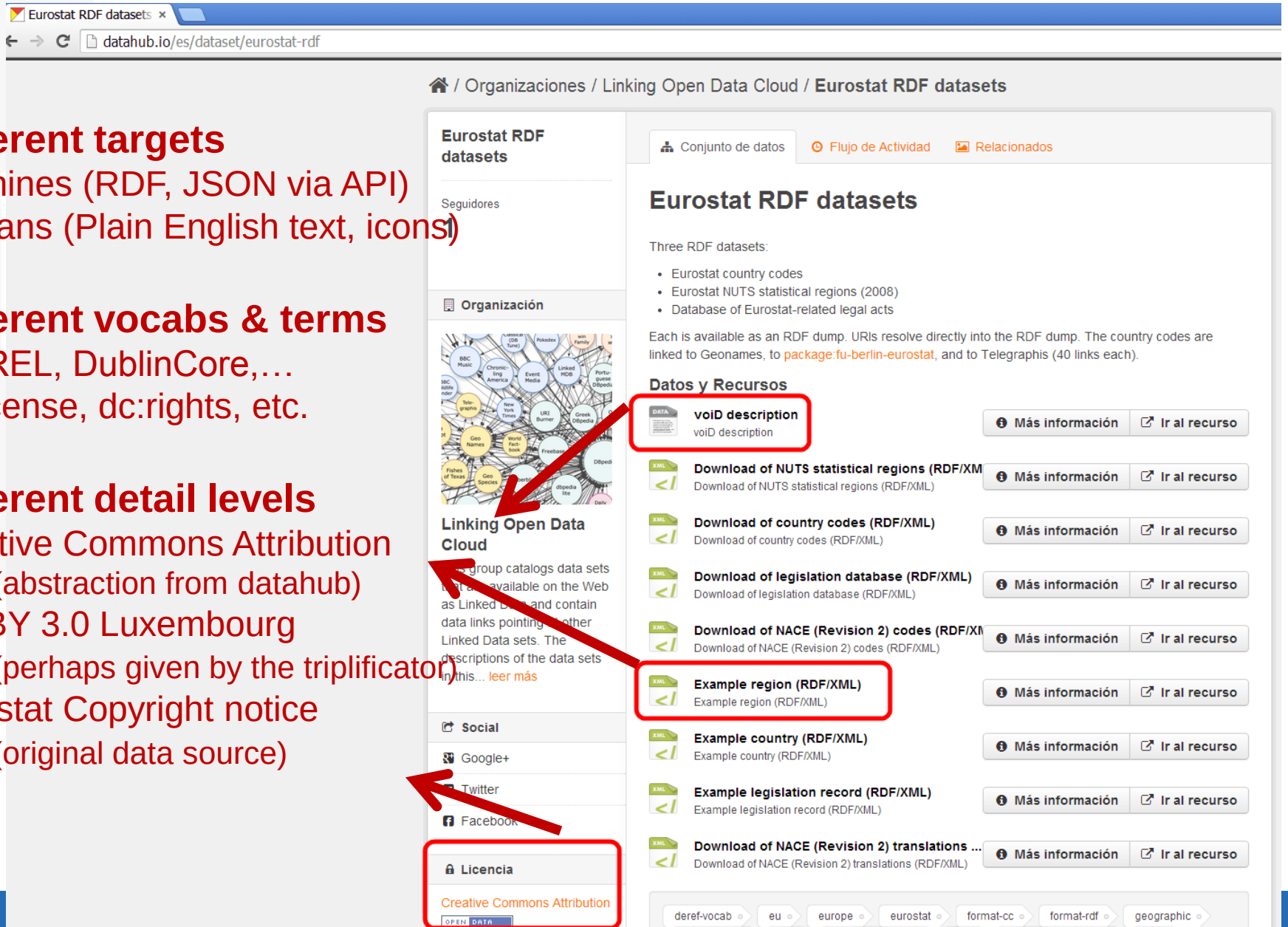
(abstraction from datahub)

CC-BY 3.0 Luxembourg

(perhaps given by the triplificator)

Eurostat Copyright notice

(original data source)



## Some common sense ideas

- There is no systematic procedure
  - Different targets, vocabularies, detail levels…

  →The process is a manual process (as of today)
- It is not always easy to find the primary source
  - Primary font. Through the creator (or rightsholder) website
  - Secondary font. Through a third party (aggregator, data portal, etc.)

  → Mind who is saying “this is the license”. Always check the primary font
- Licensing information is replicated
  - Sometimes different licensing info (or even contradictory)
  - Sometimes outdated licensing info

  →Do not rely on the first info

- Be able to ….
  - **discover the license** of a Linked Data resource
  - **select the appropriate license for your Linked Data resource**
  - **declare the license** for your Linked Data resource

# General

- Is your material original?
  - Yes: You can choose any license you like
  - No: You may be restricted by the license your resource is based upon
- Do you have peculiar restrictions?
  - Yes: Express the text of your choice
  - No: Use a well-known license (Open Data Commons, Creative Commons)
- Is your resource a general work, or plain data?
  - Yes: Use any license for works
  - No: Use data licenses (or CreativeCommons version > 4.0)

# Well-known licenses

- Creative Commons corporation
  - CC0*
  - CC-BY
  - CC-BY-SA
  - CC-BY-ND
  - CC-BY-NC
  - CC-BY-NC-SA
  - CC-BY-NC-ND
- Open Knowledge Foundation (data)
  - Open Data Commons Attribution License
  - Open Data Commons Open Database License (ODbL)
  - Open Data Commons Public Domain Dedication and License (PDDL)
- Others
  - UK-OGL, GFDL, etc.

•Mind the version
•Mind the jurisdiction port

- Be able to ….
  - **discover the license** of a Linked Data resource
  - **select** the appropriate license for your Linked Data resource
  - **declare the license for your Linked Data resource**

# Good practices

- Rights declaration for humans
  - Provide different, non-contradictory levels of information (icon, summary, full license text)
- Rights declaration for machines
  - Declaring metadata in a separate VoID/DCAT file is advisable for LD
    - It is also the best placeholder for the RDF declaration of the license
  - Use RDFa for declaring the license within the HTML website
  - Use the best acknowledged terms: dc:license, additional terms only if needed
- Advanced rights declaration
  - Next session with ODRL2.0

# Example

(see the live example)

Thanks!