



Towards high quality, industry-ready Linguistic Linked Licensed Data

*Víctor Rodríguez-Doncel¹, Daniel Vila-Suero¹, Asunción Gómez-Pérez¹,
Philipp Cimiano², John P. McCrae², and Guadalupe Aguado-de-Cea¹*

1 Ontology Engineering Group, Facultad de Informática, UPM. Madrid, Spain

{dvila, vrodriguez, asun, lupe}@fi.upm.es

2 Forschungsbau Intelligente Systeme (FBIS). Universität Bielefeld. Bielefeld, Germany

{cimiano, jmcrae}@cit-ec.uni-bielefeld.de



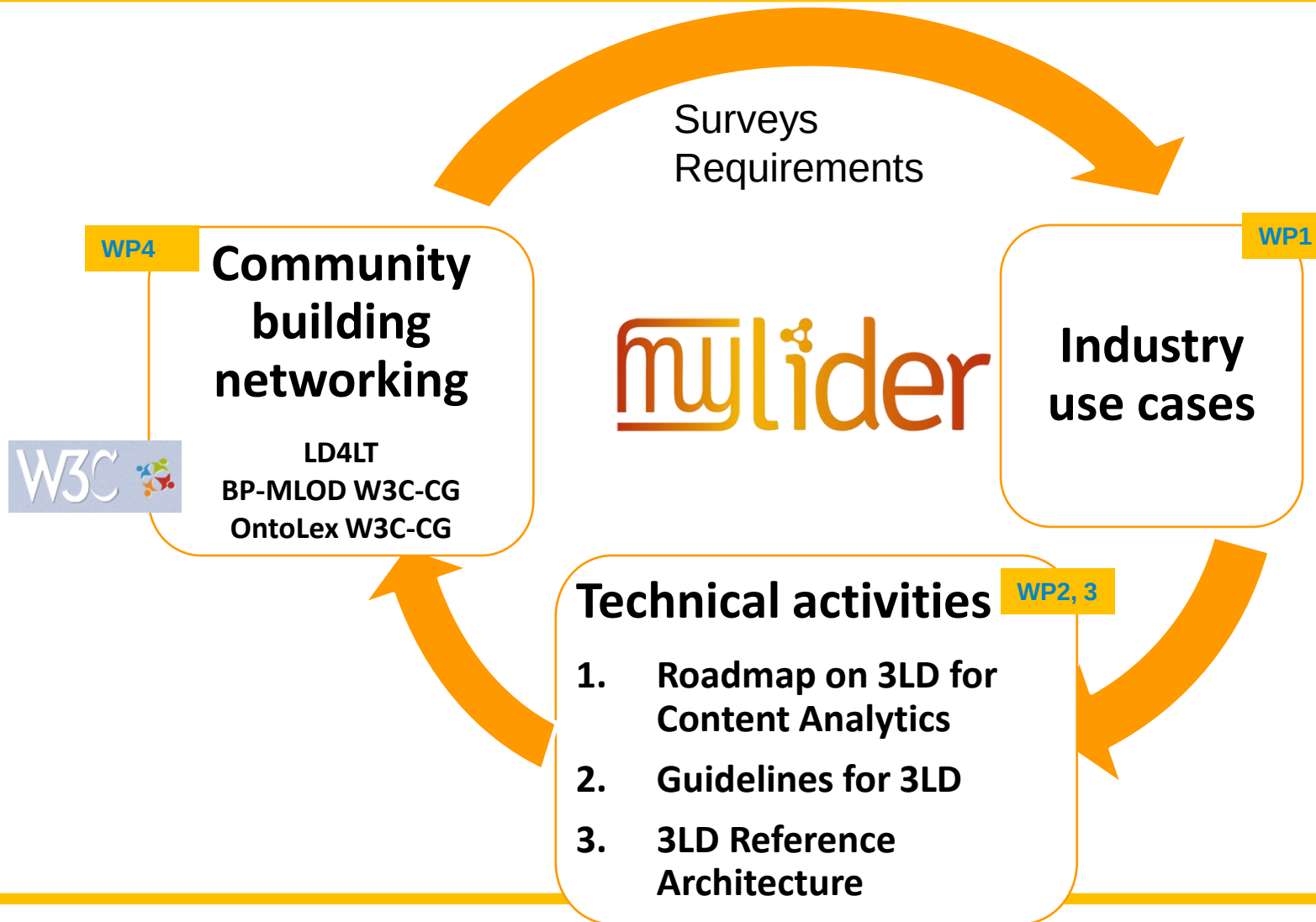
FEISGILTT 2014, Localization World Conference, Dublin 04 June 2014



NUI Galway
OÉ Gaillimh

UNIVERSITÄT LEIPZIG





3LD



Linguistic Linked Licensed Data

*Language resources
such as:*

- Lexica
- Corpora
- Dictionaries
- Grammars ..

*Using **RDF** and
standard data
models
(vocabularies):*



- Lexica 
- Corpora 
-

NLP Interchange Format

*Published along with
a **machine-readable
license.***

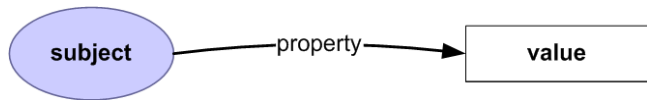
ODRL

Open Digital Rights Language





RDF(S) models

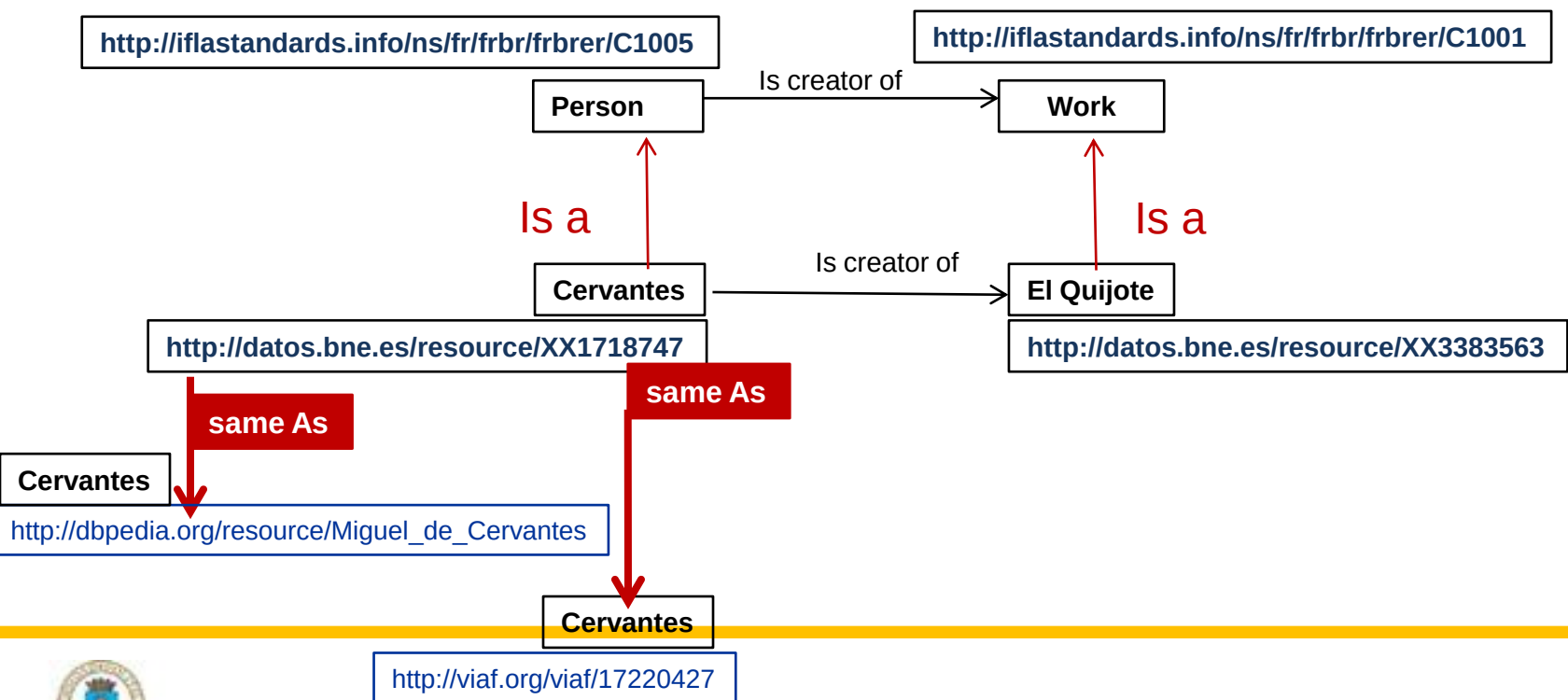


Unique identifiers: URI

identify or name a resource

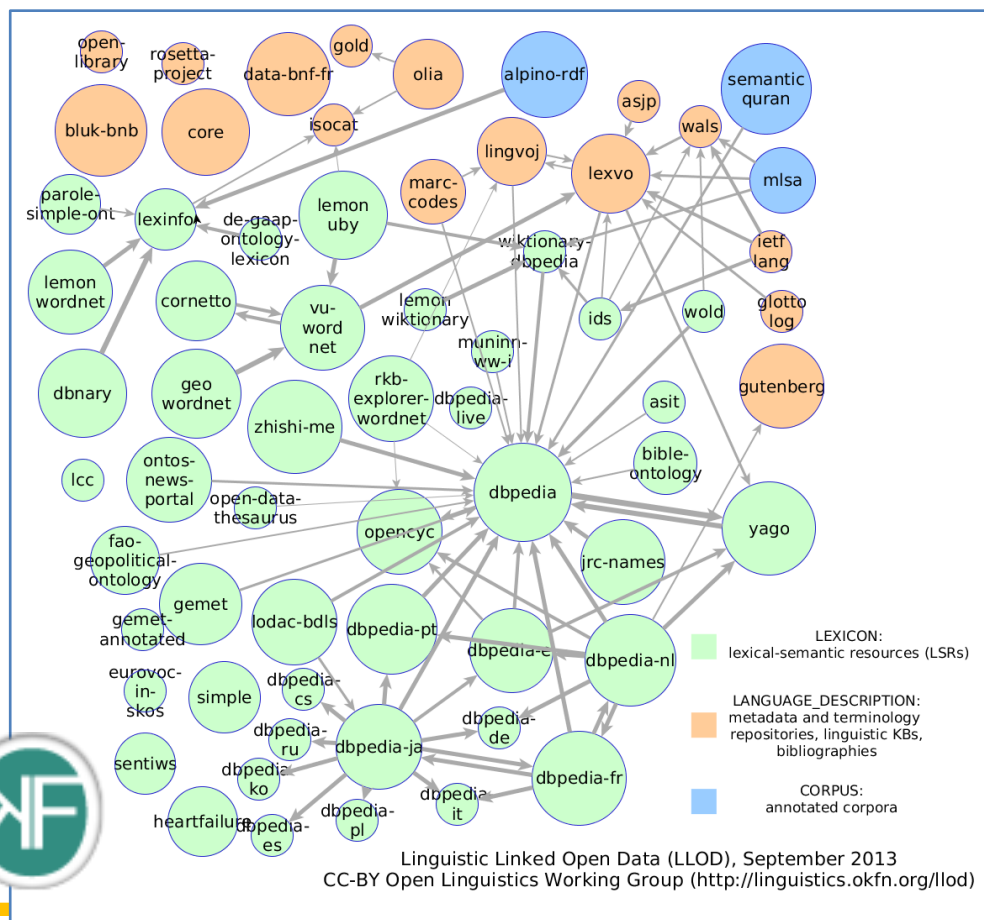
Equivalence links to other datasets

Owl:same As





Linguistic LOD (LLOD) cloud



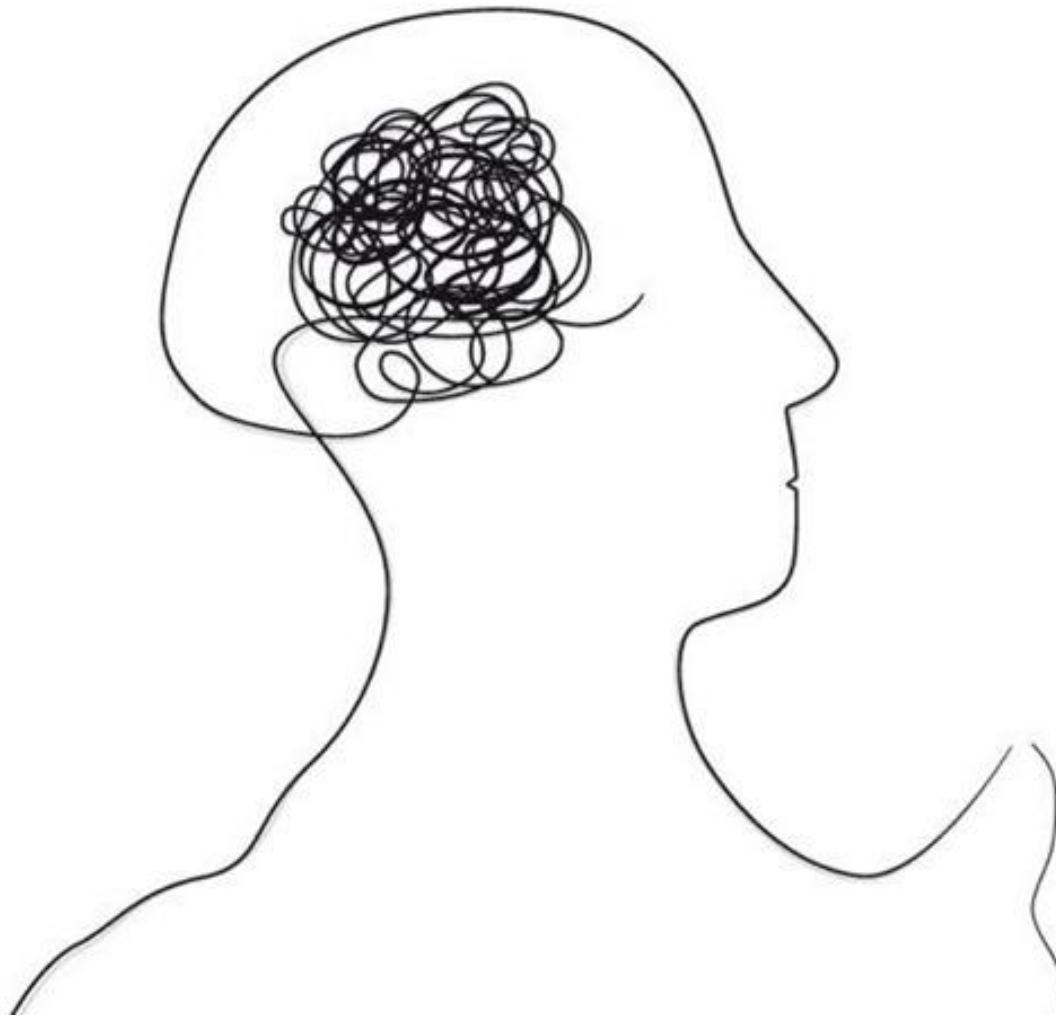
Language resources as Linked Data:

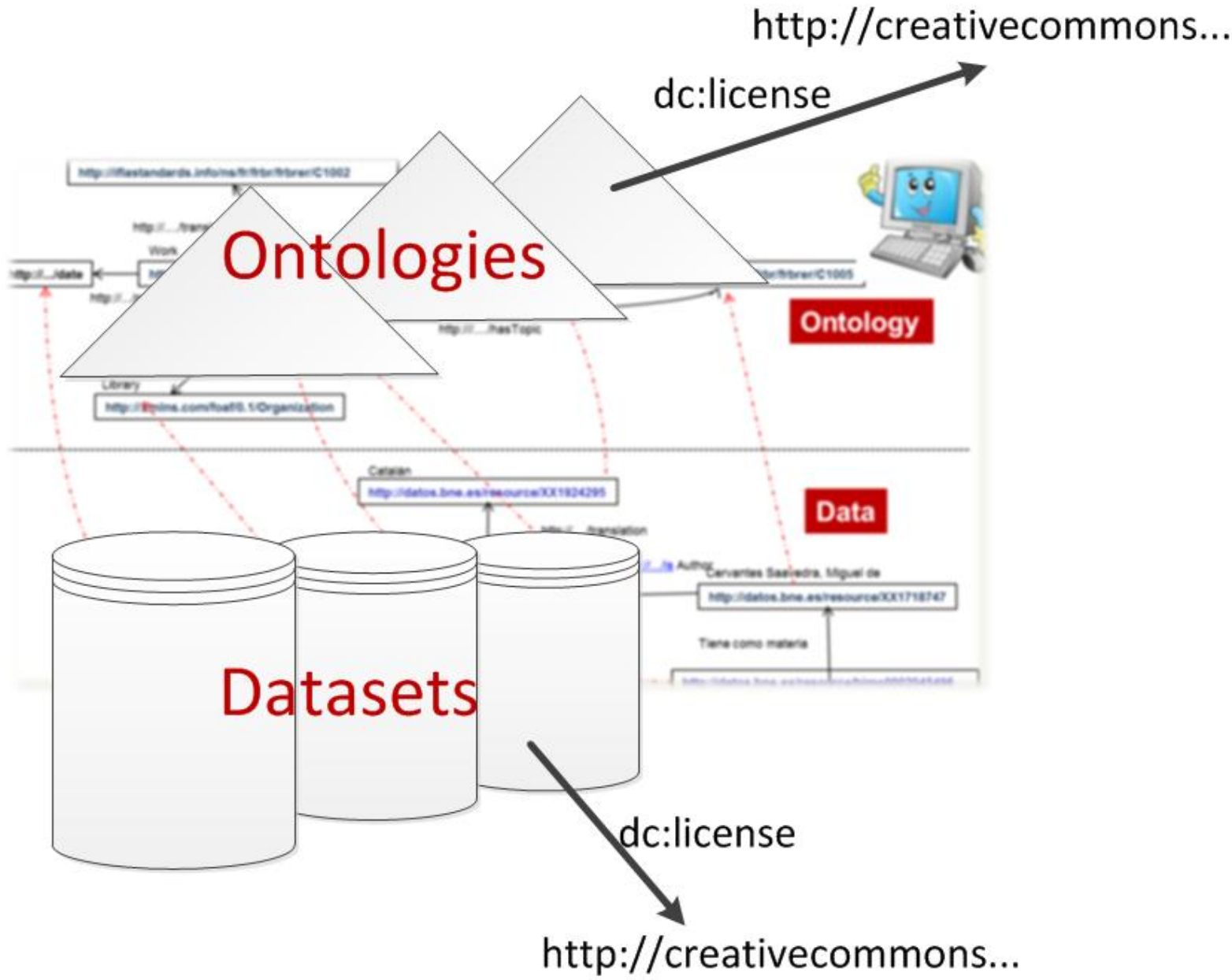
- Lexica
- Language descriptions
- Corpora

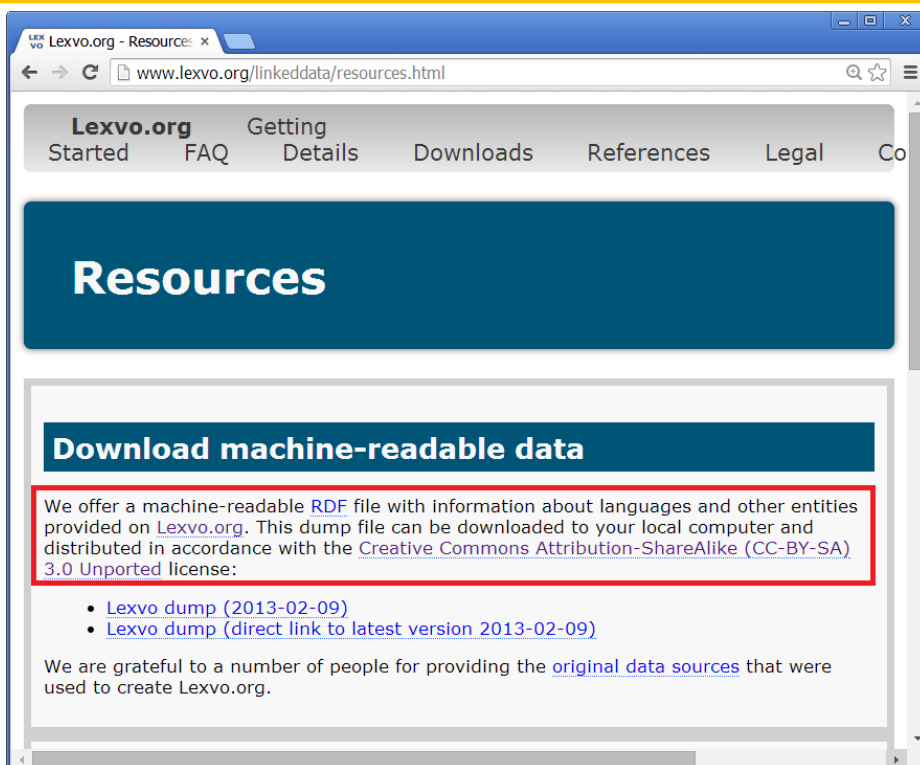
....

¹ "Open Data and Linguistics" working group, Open Knowledge Foundation, see more <http://linguistics.okfn.org/>









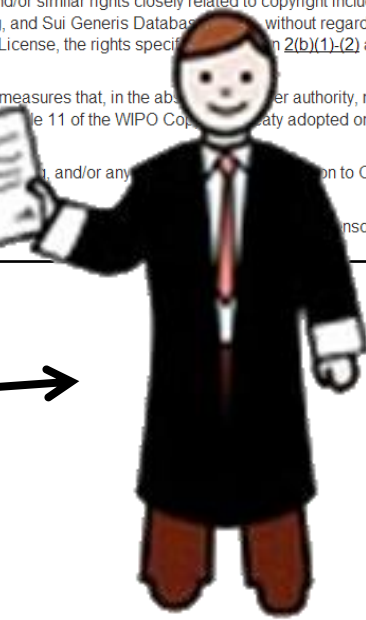
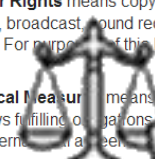
Creative Commons Attribution 4.0 International Public License

By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons Attribution 4.0 International Public License ("Public License"). To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in consideration of Your acceptance of these terms and conditions, and the Licensor grants You such rights in consideration of benefits to the Licensor received from making the Licensed Material available under these terms and conditions.

Section 1 – Definitions.

- Adapted Material** means material subject to Copyright and Similar Rights that is derived from or based upon the Licensed Material and in which the Licensed Material has been translated, altered, arranged, transformed, or otherwise modified in a manner requiring permission under the Copyright and Similar Rights held by the Licensor. For purposes of this Public License, where the Licensed Material is a musical work, performance, sound recording, or audio production, Adapted Material is a derivative work where the Licensed Material is synched in time to a moving image.
- Adapter's License** means the license You apply to Your Copyright and Similar Rights in Your contributions to Adapted Material in accordance with the terms and conditions of this Public License.
- Copyright and Similar Rights** means copyright and/or similar rights closely related to copyright including, without limitation, performance, broadcast, sound recording, and Sui Generis Database Rights, without regard to how the rights are labeled or categorized. For purposes of this Public License, the rights specified in article 2(b)(1)-(2) are not Copyright and Similar Rights.
- Effective Technological Measures** means those measures that, in the absence of applicable law, rights holder authority, may not be circumvented under laws fulfilling obligations under article 11 of the WIPO Copyright Treaty adopted on December 20, 1996, and/or similar international agreements.
- Exceptions and Limitations** means fair use, fair dealing, and/or any other exception or limitation to Copyright and Similar Rights that applies to Your use of the Licensed Material.
- Licensed Material** means the artistic or literary work, performance, sound recording, or audio production to which the Licensor applied this Public License.

URI



Resource



Contains a link to a linguistic resource (a RDF dump file)
Contains a link to a license

1

Add "rights" metadata in the dataset description
(e.g., VoID, DCAT)

DCAT 
Data catalog vocabulary



2

Use standard predicates to declare "rights" statements
(e.g., Dublin Core terms: `dc:rights`, `dct:license`)

Standard license available



Yes



No



 **creative commons**



3a

Use **URI of standard license** e.g., CC0

ODRL

Open Digital Rights Language

3b

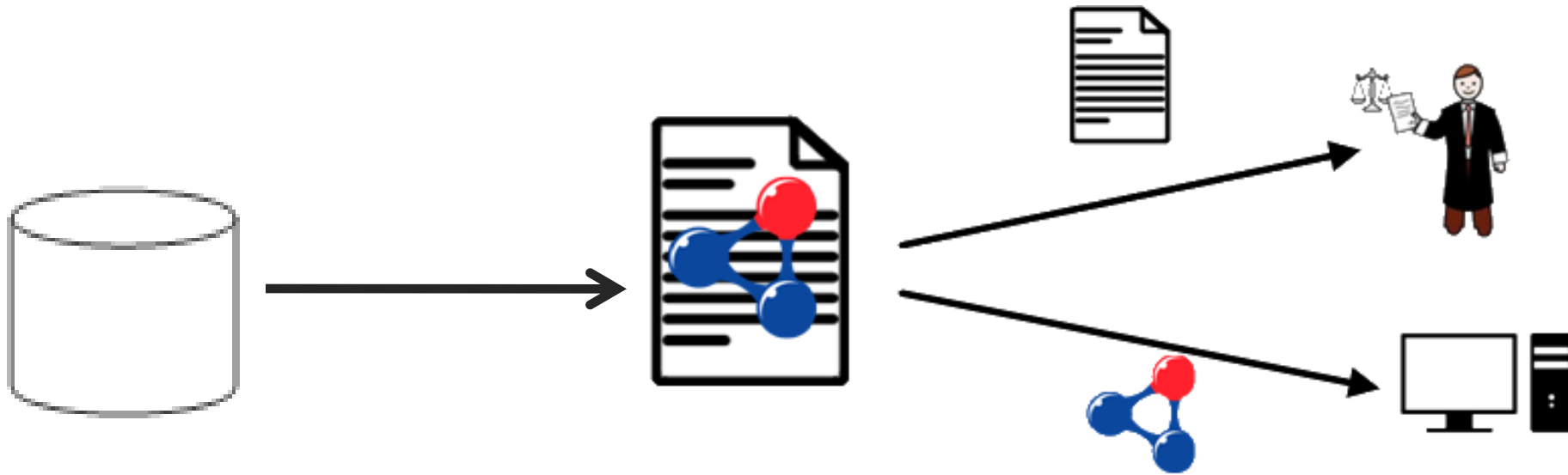
Use **rights declaration language**, e.g., ODRL

META-SHARE **NonCommercial** **NoRedistribution** **NoDerivatives** **For-a-Fee** Licence

```
@prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:   <http://www.w3.org/2000/01/rdf-schema#> .
@prefix odrl:   <http://www.w3.org/ns/odrl/2/> .

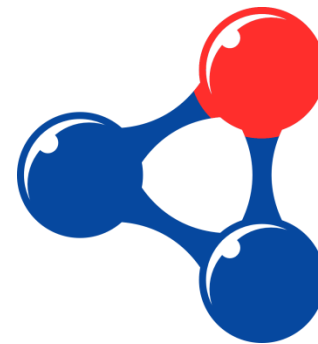
<http://example.com/nc-nored-nd-ff> a odrl:Policy ;
  rdfs:label      "NC-NoReD-ND-FF" ;
  rdfs:comment    "MetaShare NonCommercial, No Redistribution, No Derivatives, for a fee.
    Perpetual, worldwide, allowing no redistribution of the original. "@en ;
  rdfs:seeAlso    <http://www.meta-net.eu/meta-share....pdf> ;
  odrl:permission [ a          odrl:Permission ;
                    odrl:action odrl:reproduce;
                    odrl:duty   [ a          odrl:Duty ;
                                odrl:action  odrl:pay ;
                                odrl:target  "XXX EUR"
                              ]
                  ] ;
  odrl:prohibition [ a          odrl:Prohibition ;
                    odrl:action odrl:commercialize, odrl:distribute, odrl:derive
                  ] .
```

See demo at <http://conditional.linkeddata.es>



- **Prototype** developed at OEG-UPM
- A **licenses-aware Linked Data server** and a **data policies and licenses manager**
- Using Web standards (DCAT descriptions, SPARQL constructs, ODRL RDF policies, etc.)

Excerpt for a Demo taken from
Princeton Wordnet^{3.0}
in RDF



salonica.dia.fi.upm.es/ldr/resource/100006269-n



Conditional access to Linked Data

Resource info

Demo data

What is this?

Introduction

Technical details

User's view (demo)

Browse dataset

Offers

SPARQL

Account

Admin's view (demo)

Control panel

Accountability

life (Synset)

Open triples

Property	Value
sample	the oceans are teeming with life
part_of_speech	<u>noun</u>
translation	生命
translation	دڤڤ
translation	vie
lexical_domain	<u>noun.tops</u>
translation	takdir

Limited access triples

Property	Value
sameAs	Limited access (Offer --- Per RDF Statement)
sameAs	Limited access (Offer --- Per RDF Statement)





Conditional access to Linked Data Control panel

What is this?

Introduction

Technical details

User's view (demo)

Browse dataset

Offers

SPARQL

Account

Admin's view (demo)

Control panel

Accountability

Assignment of policies to resources

Choose licenses/policies for each pre-defined dataset

Datasets

[Mappings](#)

[Thing](#)

[General](#)

Dataset management

Mappings

Description: Mappings to terms in other resources

Number of triples: 519

Dataset license: [License1Cent]

License1Cent ▼

Add

Remove

View



Conditional access to Linked Data GeoLinkedData dataset

What is this?

Introduction

Technical details

User's view (demo)

Browse dataset

Offers

SPARQL

Account

Admin's view (demo)

Control panel

Accountability

Dataset RDF dump offers

Demo data

Different subsets of data can be offered under different conditions. These subsets can be portions of the dataset, mappings, specific resources or even individual statements.

Dataset	RDF dump	Licenses	Description
Mappings		1- License1Cent	Mappings to terms in other resources
Thing		1- License15euros	Synset for Thing
General	Download	1- Creative Commons CC-BY	General dataset



License: License1Cent

Rule (Permission)

Rights
reproduce,

Constraints
Pay (Pay 0.01 EUR) - per unit,
Target

To be acted on <http://localhost:99/ldr/resource/Provincia/Barcelona> - <http://www.w3.org/2004/02/skos/core#exactMatch>

 [View license in RDF](#)

 [purchase](#)

```
<http://localhost:99/ldr/policy/ee32f675-ccae-4ca9-a544-3c07abf0b16e>
  a      <http://www.w3.org/ns/odrl/2/Policy> , <http://www.w3.org/ns/odrl/2/Set>;
  <http://www.w3.org/2000/01/rdf-schema#comment>
    "Individual triples are available upon payment of 1 euro cent" ;
  <http://www.w3.org/ns/odrl/2/permission> ....
```



Conditional access to Linked Data Account

What is this?

Introduction

Technical details

User's view (demo)

Browse dataset

Offers

SPARQL

Account

Admin's view (demo)

Control panel

Accountability

Logged in as **vroddon@gmail.com**

User: **Victor Rodriguez Doncel** [logout](#)

You have purchased the following permissions:

Price	Asset	License
15,00 EUR	http://salonica.dia.fi.upm.es/ldr/datas	License15euros (view)

Reset portfolio

View portfolio

View signed portfolio



Conditional access to Linked Data

Introduction

What is this?

[Introduction](#)[Technical details](#)

User's view (demo)

[Browse dataset](#)[Offers](#)[SPARQL](#)[Account](#)

Admin's view (demo)

[Control panel](#)[Accountability](#)

SPARQL endpoint

This **SPARQL endpoint** provides access to the limited subset of RDF triples you are allowed to see

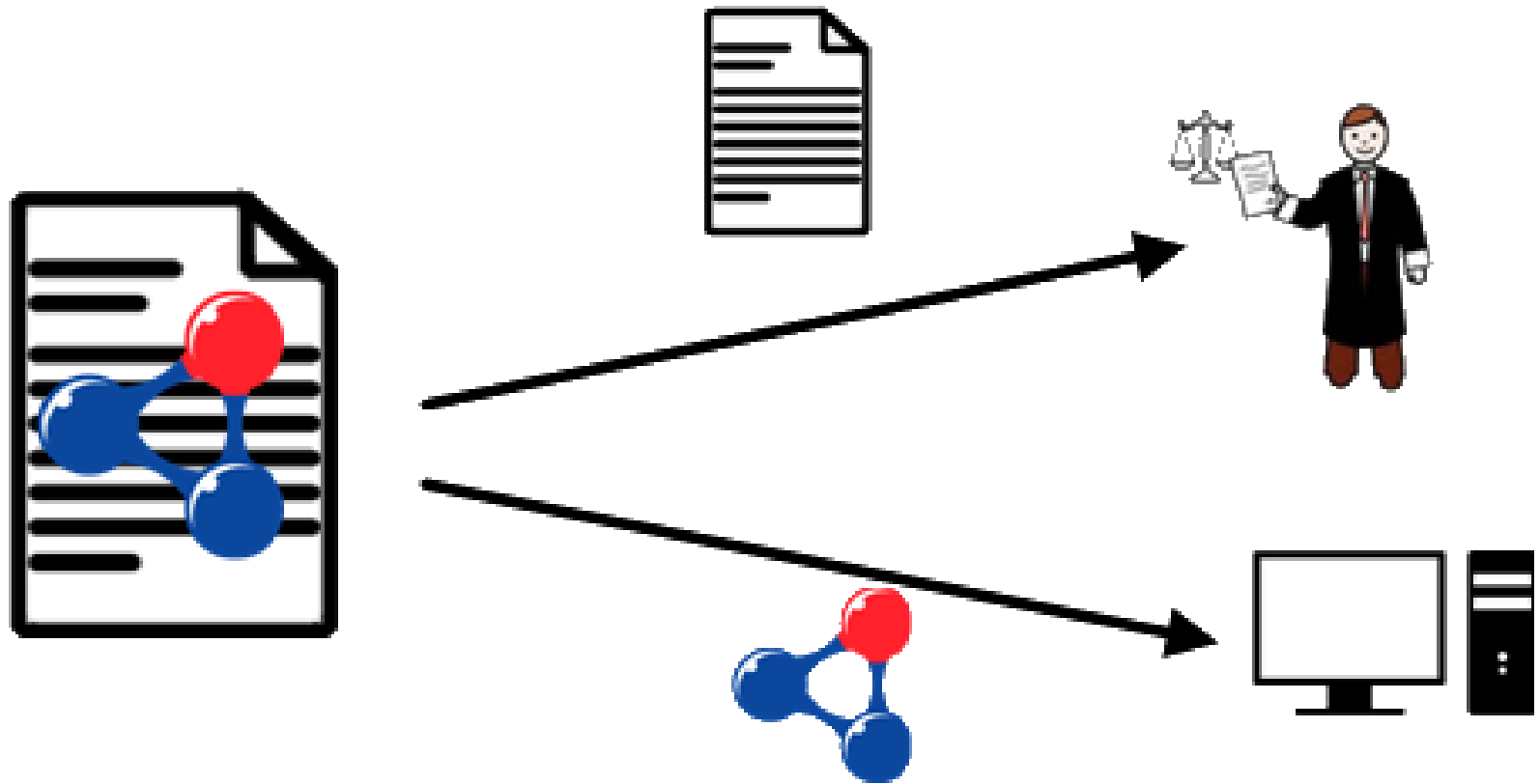
Output:

Text ▼

Superuser: ☒

Get Results





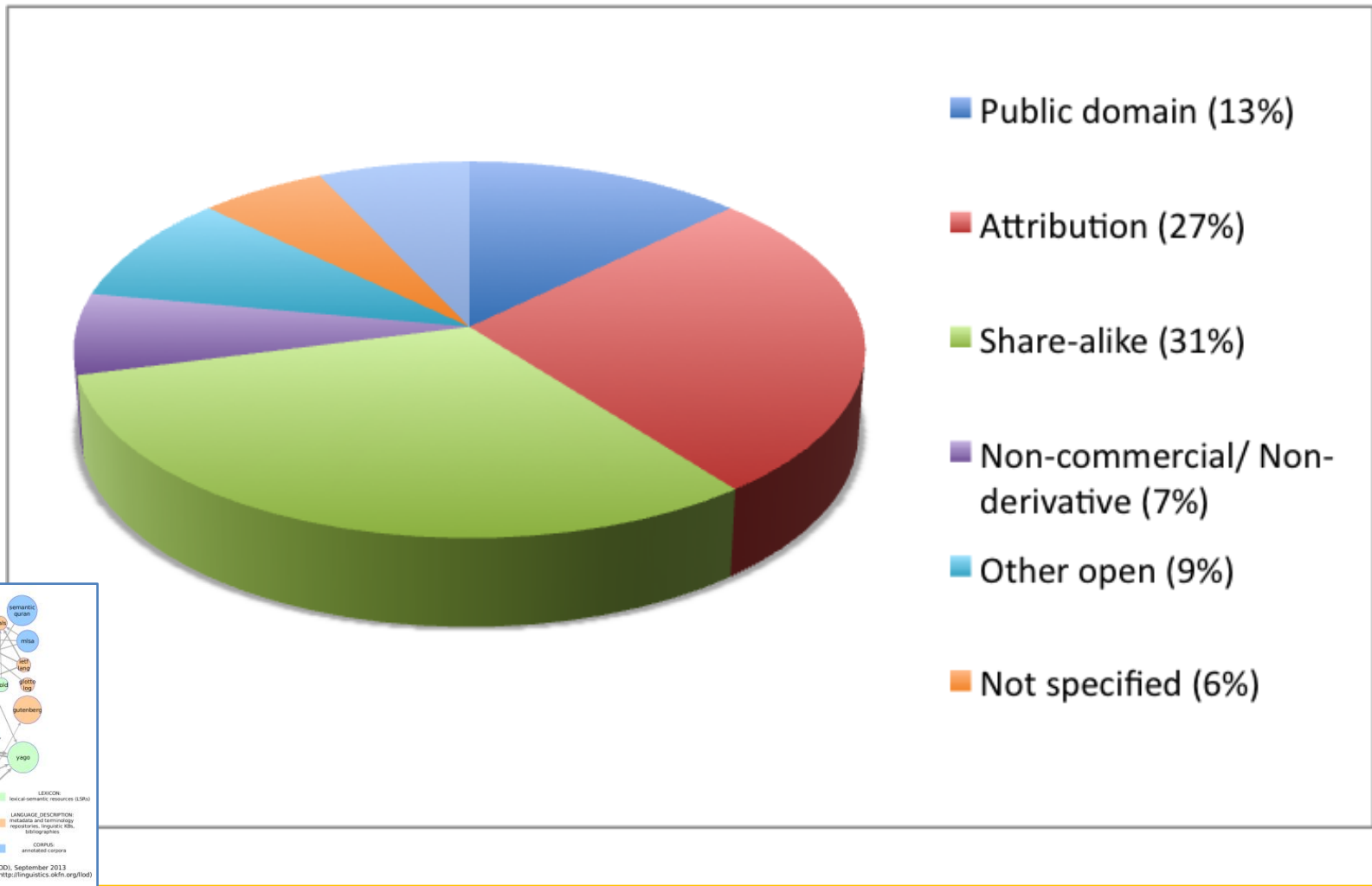
THANK YOU
FOR YOUR ATTENTION

QUESTIONS?

Demo available at:

<http://salonica.dia.fi.upm.es/ldr/en/>

<http://conditional.linkeddata.es>



- META-SHARE recommends a set of 21 licenses, classified in:
 - META-SHARE Non Redistribution
 - META-SHARE Commons («distribution towards META-SHARE members»)
 - Creative Commons
- ALL of them can be represented as RDF using extendedly used vocabularies such as ODRL (Open Digital Rights Licenses)
- Advantages of expressing licenses as RDF:
 - Unambiguous identification of well known licenses by their URIs
 - Enables conditional access to resources
 - Automatic license compatibility analysis when integrating resources

Lexvo.org - Resources

www.lexvo.org/linkeddata/resources.html

Lexvo.org Getting Started FAQ Details Downloads References Legal

Resources

Download machine-readable data

We offer a machine-readable [RDF](#) file with information about languages and other entities provided on [Lexvo.org](#). This dump file can be downloaded to your local computer and distributed in accordance with the [Creative Commons Attribution-ShareAlike \(CC-BY-SA\) 3.0 Unported](#) license:

- [Lexvo dump \(2013-02-09\)](#)
- [Lexvo dump \(direct link to latest version 2013-02-09\)](#)

We are grateful to a number of people for providing the [original data sources](#) that were used to create Lexvo.org.

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <rdf:RDF
3   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4   xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
5   xmlns:lvont="http://lexvo.org/ontology#"
6   xmlns:lexvo="http://lexvo.org/id/"
7   xmlns:owl="http://www.w3.org/2002/07/owl#"
8   xmlns:skos="http://www.w3.org/2008/05/skos#"
9   xmlns:skosxl="http://www.w3.org/2008/05/skos-xl#"
10  xmlns:foaf="http://xmlns.com/foaf/0.1/"
11  xmlns:xsd="http://www.w3.org/2001/XMLSchema#" >
12
13  <!--
14  This data file is a part of
15
16  Lexvo
17  http://www.lexvo.org/
18  Gerard de Melo, 2008-2013
19
20  License: Creative Commons Attribution-Share Alike 3.0 Unported
21  http://creativecommons.org/licenses/by-sa/3.0/
22
23  Portions copyright various contributors (e.g. article authors on Wikipedia).
24  For information about the data sources and the respective copyrights,
25  please see:
26  http://www.lexvo.org/linkeddata/sources.html
27  -->
28
29  <rdf:Description rdf:about="http://id.loc.gov/vocabulary/iso639-2/abk">
30    <foaf:focus rdf:resource="http://lexvo.org/id/iso639-3/abk" />
31  </rdf:Description>
32  <rdf:Description rdf:about="http://id.loc.gov/vocabulary/iso639-2/ace">
33    <foaf:focus rdf:resource="http://lexvo.org/id/iso639-3/ace" />
34  </rdf:Description>

```