

Presentation metadata

Open Data Support is funded by the European Commission under SMART 2012/0107 'Lot 2: Provision of services for the Publication, Access and Reuse of Open Public Data across the European Union, through existing open data portals' (Contract No. 30-CE-0530965/00-17).

© 2014 European Commission



OPEN DATA SUPPORT

Training Module 2.4

Designing and developing RDF vocabularies

Learning objectives

By the end of this training module you should have an understanding of:

- What the best practices are for creating an RDF vocabulary for modelling your data
- Where to find RDF vocabularies for reuse.
- How you can create your own RDF vocabulary.
- How to publish your RDF vocabulary.
- The process and methodology for developing semantic agreements developed by the ISA Programme of the European Commission.

Content

This module is about...

- The steps for modelling your data.
- How to reuse existing vocabularies to model your data.
- How to create new classes and properties in RDF.
- How and where to publish your RDF vocabulary so that it can be reused by others.

RDF Vocabulary

“A vocabulary is a data model comprising classes, properties and relationships which can be used for describing your data and metadata.”

- RDF Vocabularies are **sets of terms** used to describe things.
- A term is either a **class or a property**.
 - Object type properties (relationships)
 - Data type properties (attributes)

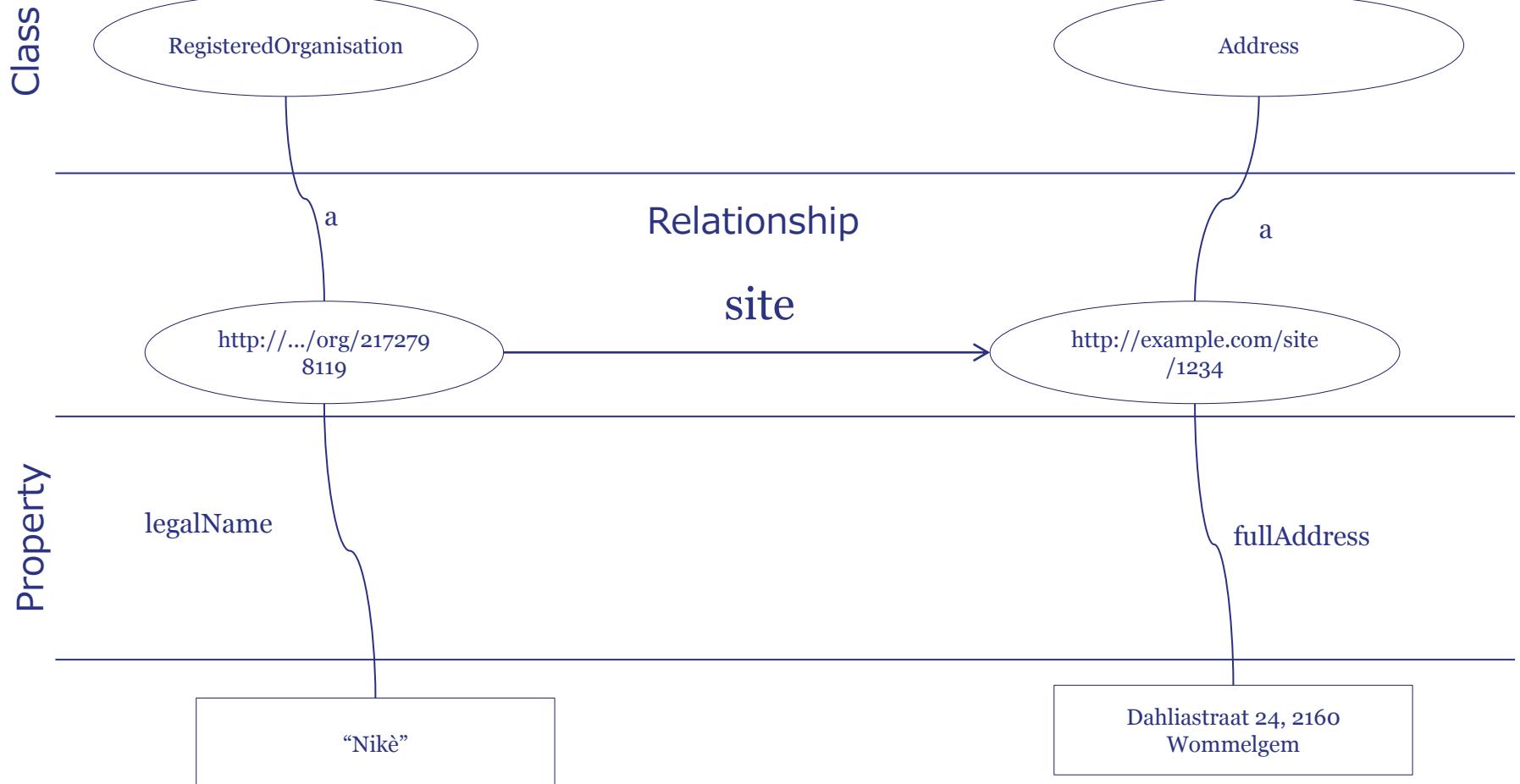
What are classes, relationships and properties?

Class. A construct that represents things in the real and/or information world, e.g. a person, an organisation, a concepts such as “health” or “freedom”.

Relationship. A link between two classes; for the link between a document and the organisation that published it (i.e. organisation *publishes* document), or the link between a map and the geographic region it depicts (i.e. map *depicts* geographic region). In RDF relationships are encoded as object type properties.

Property. A characteristic of a class in a particular dimension such as the legal name of an organisation or the date and time that an observation was made.

Examples of classes, relationships and properties



Model your data

How to reuse from other vocabularies, define your own terms and publish and promote your vocabularies to describe the data.

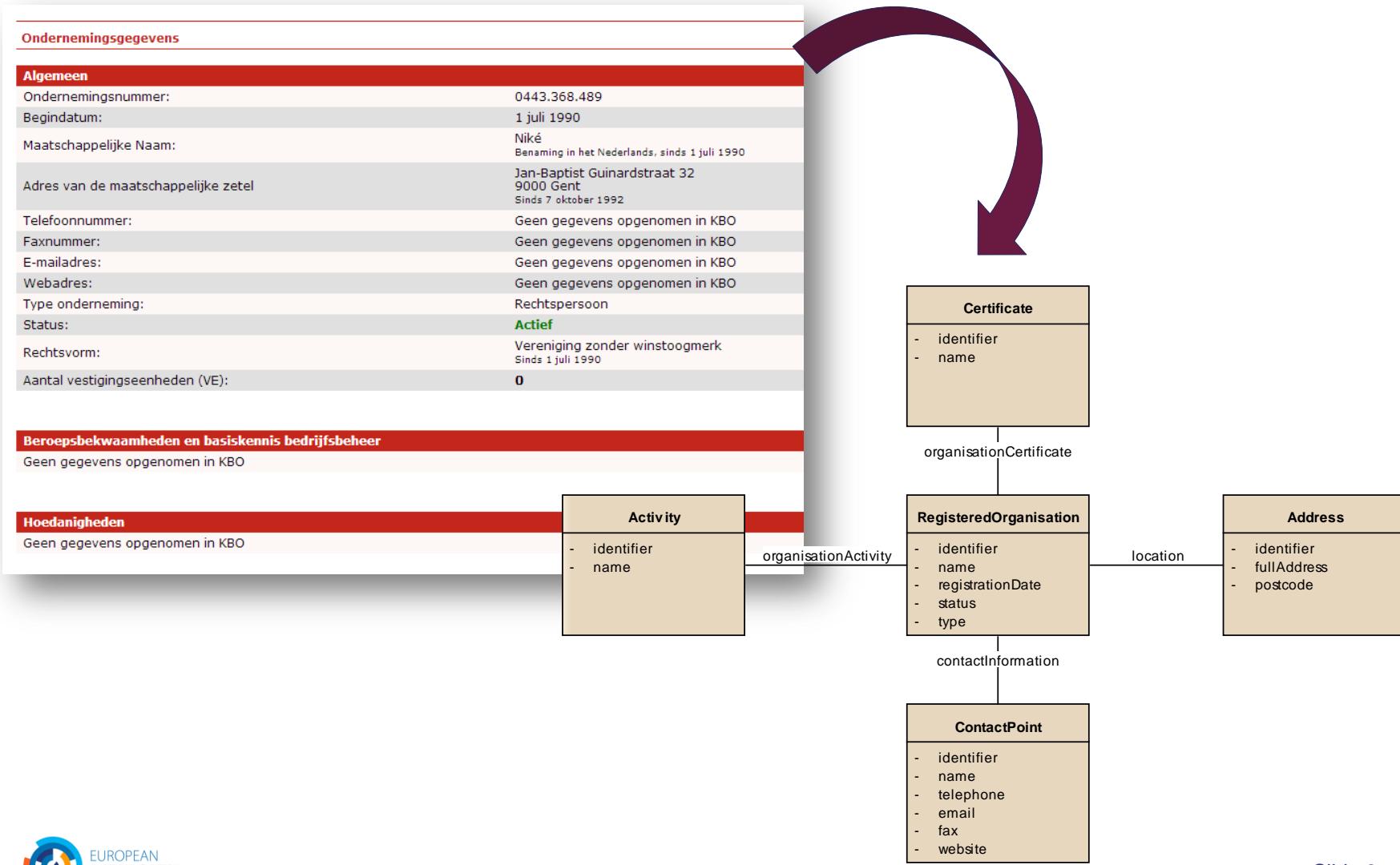
6 steps for modelling your data

- 1** Start with a **robust Domain Model** developed following a structured process and methodology.
- 2** Research **existing terms** and their usage and **maximise reuse** of those terms.
- 3** Where new terms can be seen as specialisations of existing terms, create **sub class** and **sub properties**.
- 4** Where **new terms** are required, create them following **commonly agreed best practice**.
- 5** Publish within a **highly stable environment** designed to be **persistent**.
- 6** Publicise the **RDF schema** by registering it with relevant services.

See also:

[https://joinup.ec.europa.eu/community/semic/document/
cookbook-translating-data-models-rdf-schemas](https://joinup.ec.europa.eu/community/semic/document/cookbook-translating-data-models-rdf-schemas)

1 Start with a robust Domain Model



2 *Reuse existing terms and vocabularies*

- General purpose vocabularies: DCMI, RDFS  Dublin Core Metadata Initiative®
- To name things: rdfs:label, foaf:name, skos:prefLabel 
- To describe people: FOAF, vCard, Core Person Vocabulary
- To describe projects: DOAP, ADMS.SW 

- To describe interoperability assets: ADMS
- To describe registered organisations: Registered Organisation Vocabulary 
- To describe addresses: vCard, Core Location Vocabulary 
- To describe public services: Core Public Service Vocabulary 
- To describe datasets: DCAT, DCAT Application Profile, VoID 

Creating application profiles

- Different domains have different requirements for domain-specific semantics, e.g. classification concepts.
- Generic RDF vocabularies usually provides the **core base classes** needed to allow extensions to add specific sub-class structures or classification schemes as required.
- In such cases, reusers are encouraged to define **application profiles particular to an application domain** by specifying (if required) **sub-classes, sub-properties and controlled vocabularies**.
- For example,
 - [DCAT Application profile for data portals in Europe](#)
 - [Registered Organization vocabulary](#) as an application profile of the [Organization ontology](#).

See also:

joinup.ec.europa.eu/asset/dcat_application_profile/home

Advantages of reuse

- Reuse greatly **aids interoperability** of your data
 - Use of dcterms:created, for example, the value for which should be a data typed date such as 2013-02-21^^xsd:date, is immediately processable by many machines. If your schema encourages data publishers to use a different term and date format, such as ex:date "21 February 2013" – data published using your schema will require further processing to make it the same as everyone else's.
- Reuse **adds credibility** to your schema.
 - It shows it has been published with care and professionalism, again, this promotes its reuse.
- Reuse is **easier and cheaper**.
 - Reusing classes and properties from well defined and properly hosted vocabularies avoids your having to replicate that effort.

You can find reusable RDF vocabularies on...

The screenshot shows the Joinup platform's semantic assets search interface. At the top, there's a navigation bar with links for Contact, Search, Glossary, Help, Partners, Analytics, and a language selector (English (en)). Below the navigation is a search bar with a placeholder 'Search' and a 'Search' button. The main content area has a header 'Semantic Assets' and a sub-header 'Share and reuse semantic interoperability assets'. A sidebar on the left includes links for Semantic Assets, Catalogue of assets, Projects, Federated Repositories, and Advanced Search. The main content area displays a search form with fields for Language (set to 'Any language'), Current version, and Keywords (set to 'organisation RDF'). Below the search form, it says '1 to 20 of 102 results' with a page navigation. There are two asset cards: one for 'Organisation Type List' and another for 'Enterprise Competence Organisation Schema'. Both cards show a small preview image, the title, a brief description, and download links.

<http://joinup.ec.europa.eu/>



The screenshot shows the Linked Open Vocabularies (LOV) search interface. At the top, there's a logo for LOV (Linked Open Vocabularies) and a search bar with the word 'organisation' and a 'Search' button. Below the search bar, it says '147 results in 28 vocabularies'. On the left, there are three filter panels: 'Filter by Domain' (listing City (71), Data & Systems (0), Library (12), Market (4), Media (35), Science (5), Upper & Meta (4), and Where & When (7)); 'Filter by Type' (listing rdfs:Class (53), rdf:Property (139), voaf:Vocabulary (8), and Other (22)); and 'Filter by Vocabulary' (listing org (37)). The main content area is a list of results with columns for the term, its type, and a score. Each result has a 'More' link at the end. The results include various RDF terms like swrc:Organization, rdfs:label Organisation, foaf:Organization, and org:organization, all related to the concept of 'organisation'.

OPEN DATA SUPPORT



Slide 13

3 *Creation of sub-classes and sub-properties*

- RDF schemas and vocabularies often include **terms that are very generic**.
- By creating **sub-class** and **sub-property** relationships, systems that understand the super property or super class may be able to interpret the data even if the more specific terms are unknown.
- **Do not create sub-classes and sub-properties simply to allow you to use your own term** for something that already exists.



Creation of sub-properties - example

The Registered Organization vocabulary defines three sub-properties of org:classification: companyType, companyStatus and companyActivity.

Company Type

Property	Domain	Range
<code>rov:companyType</code> <code>rdfs:subPropertyOf org:classification</code>	org:Organization	skos:Concept

This property records the type of company. Familiar types are SA, PLC, LLC, GmbH etc. At the time of publication, there is no agreed set of company types that crosses borders. The term 'SA' is used in Poland although they mean slightly different things. The UK's LLP and Greece's EPE provide further examples of close, but not exact, matches.

That said, each jurisdiction will have a limited set of recognized company types and these should be expressed in a consistent manner in a SKOS Concept Scheme.

Company Status

Property	Domain	Range
<code>rov:companyStatus</code> <code>rdfs:subPropertyOf org:classification</code>	org:Organization	skos:Concept

Recording the status of a company presents the same issues as its type. The terms 'insolvent', 'bankrupt' and 'in receivership,' for example, are likely to mean slightly different things with different legal implications.

Taking advice from [XBRL Europe](#) as a starting point, however, the term 'Normal Activity' does appear to have cross-border usefulness and this should be used in preference to terms like 'trading' or 'operating.'

Best Practice for recording various other status levels is to use the relevant jurisdiction's terms and to encode these in a SKOS Concept Scheme.

Company Activity

Property	Domain	Range
<code>rov:companyActivity</code> <code>rdfs:subPropertyOf org:classification</code>	org:Organization	skos:Concept

Defining a sub-property in RDF

```
<rdf:Property rdf:about="rov:companyType">  
    <rdfs:label xml:lang="en">company type</rdfs:label>  
    <rdfs:comment xml:lang="en" rdf:parseType="Literal">  
        This property records the type of company. Familiar types are SA, PLC, LLC,  
        GmbH etc. Each jurisdiction will have a limited set of recognised company  
        types and these should be used in a consistent manner using a  
        skos:Concept as described in the <a href="#skos:Concept">Code</a> Class.  
    </rdfs:comment>  
    <rdfs:isDefinedBy rdf:resource="http://www.w3.org/ns/regorg#" />  
    <rdfs:range rdf:resource="skos:Concept"/>  
    <rdfs:subPropertyOf rdf:resource="org:classification" />  
    <dcterms:identifier>legal:companyType</dcterms:identifier>  
</rdf:Property>
```



4 Where new terms are required, create them following commonly agreed best practices

- ✓ Classes begin with a capital letter and are always singular, e.g. skos:Concept.
- ✓ Properties begin with a lower case letter, e.g. rdfs:label.
- ✓ Object properties should be verbs, e.g. org:hasSite.
- ✓ Data type properties should be nouns, e.g. dcterms:description.
- ✓ Use camel case if a term has more than one word, e.g. foaf:isPrimaryTopicOf.

Defining a new class - Organisation

```
<rdf:RDF  
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"  
    xmlns:org="htpp://example.org/organisation-schema">  
<rdf:Class rdf:about="org:Organisation">  
    <rdfs:label xml:lang="en">Organisation</rdfs:label>  
    <rdfs:comment xml:lang:"en">  
        Legal entity that is registered in an official national or regional register.  
    </rdfs:comment>  
</rdf:Class>
```

Defining a new property - registrationNumber

```
<rdf:RDF  
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"  
    xmlns:org="htpp://example.org/organisation-schema">  
  
<rdf:Property rdf:about="org:registrationNumber">  
    <rdfs:label xml:lang="en">registrationNumber</rdfs:label>  
    <rdfs:comment xml:lang:"en">  
        The number that a registered organisation receives upon registration  
        in the official register.  
    </rdfs:comment>  
  
</rdf:Class>
```



Defining domain and range restrictions

```
<rdf:RDF  
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"  
    xmlns:org="htpp://example.org/organisation-schema"  
    xmlns:locn="http://www.w3.org/ns/locn#">  
  
<rdf:Property rdf:about="org:isLocated">  
    <rdfs:label xml:lang="en">isLocated</rdfs:label>  
    <rdfs:comment xml:lang:"en">  
        The official address of the registered organisation's headquarters.  
    </rdfs:comment>  
    <rdfs:domain rdf:resource="org:Organisation"/>  
    <rdfs:range rdf:resource="locn:Address">  
</rdf:Class>  
  
  
http://example.org/org/1234 org:isLocated http://dbpedia.org/page/Brussels
```

5 *Publish within a highly stable environment designed to be persistent*

- Choose a stable namespace for your RDF schema (e.g. W3C, Purl...)
- Use good practices on the publication of persistent Uniform Resource Identifiers (URI) sets, both in terms of format and of their design rules and management.
- Examples:
 - <http://www.w3.org/ns/adms>
 - <http://purl.org/dc/elements/1.1>

See also:

<https://joinup.ec.europa.eu/community/semic/document/cookbook-translating-data-models-rdf-schemas>
<http://www.slideshare.net/OpenDataSupport/design-and-manage-persistent-uris>

6 *Publicise the RDF schema by registering it with relevant services*

Once your RDF schema is published you will want people to know about it. To reach a wider audience register it Joinup and Linked Open Vocabularies.

joinup

Share and reuse interoperability solutions for public administrations

European Commission > ISA > Joinup > Advanced search

Advanced search

Filtering by:

- Semantic Asset Projects
- Semantic Asset Releases
- [Clear filters](#)

Refine your search

Filter by type

Repository of Origin

- OSLO - Open Standards for Local Administrations in Flanders (1)
- ListPoint - the open platform for code list standards (1)
- Linked Open Vocabularies (1)
- Finalisér.dk (2)
- Standards and Technical Reports (3)
- MSI Doc (3)
- XRepository (7)
- EU Semantic Interoperability Catalogue (12)
- ESD Standards (12)
- Dutch Standardisation Forum - "Comply or explain"-standards (50)

Current version

1 

Search results for "organisation"

The advanced search helps you navigate through content available on Joinup by (de)selecting the search filters on the left-hand side.

Keywords: organisation

Retain current filters

1 to 20 of 112 results

Sort by: Relevancy ▾ Popularity ▾ Title ▾ Author ▾

2 

3 

More focused.

More relevant.

<http://lov.okfn.org>

Linked Open Vocabularies (LOV)

Features gives you the possibility to search for an existing element (property, class or vocabulary) in vocabularies Catalogue.

Endpoint and metrics about the use of vocabularies in the Semantic Web are used to bring you some

 MONDECA  Inserm  developed by Pierre-Yves Vandebussche

 Endpoint

organisation

147 results in 28 vocabularies

swc:Organization (owl:Class)	score: 0.904	>>
rdfs:label Organisation @de	score: 0.884	>>
foaf:Organization (owl:Class)		>>
rdfs:label Organisation @fr		>>
swc:organization (owl:ObjectProperty)	score: 0.882	>>
rdfs:label organisation @de	score: 0.882	>>
org:organization (owl:ObjectProperty)	score: 0.882	>>
rdfs:label organisation @fr		>>
org:Organization (owl:Class)	score: 0.882	>>
rdfs:label Organisation @fr		>>
rdfs:comment ...nt que "Agent". Les organisations sont souvent décom.....lisés pour nommer l'Organisation. En particulier sk... @fr		>>
cgov:postin (owl:ObjectProperty)	score: 0.882	>>
rdfs:label Organisation @en		>>
rdfs:comment Indicates the organisation that this is a post... @en		>>
meb:organisation (owl:DatatypeProperty)	score: 0.882	>>
rdfs:label organisation		>>
rdfs:comment ... User is part of an organisation		>>
locatedRP (owl:ObjectProperty)	score: 0.882	>>

<http://lov.okfn.org>

More focused.

More relevant.

<http://joinup.ec.europa.eu>

OPEN DATA SUPPORT Slide 22

Conclusions

Analyse

Model

Publish

Start with a robust Domain Model developed following a structured process and methodology.

Research existing terms and their usage and maximise reuse of those terms.

Where new terms can be seen as specialisations of existing terms, create sub class and sub properties as appropriate.

Where new terms are required, create them following commonly agreed best practice in terms of naming conventions etc

Publish within a highly stable environment designed to be persistent.

Publicise the RDF schema by registering it with relevant services.

Group exercise



<http://www.visualpharm.com>

In groups of 2, create the RDF description of a vocabulary for representing a citizen.



<http://www.visualpharm.com>

According to you, what are the main barriers to the reuse of existing RDF vocabularies?

Take also the online test here!

Thank you! ...and now YOUR questions?



This presentation has been created by Open Data Support

Disclaimers

1. The views expressed in this presentation are purely those of the authors and may not, in any circumstances, be interpreted as stating an official position of the European Commission. The European Commission does not guarantee the accuracy of the information included in this presentation, nor does it accept any responsibility for any use thereof. Reference herein to any specific products, specifications, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favouring by the European Commission. All care has been taken by the author to ensure that s/he has obtained, where necessary, permission to use any parts of manuscripts including illustrations, maps, and graphs, on which intellectual property rights already exist from the titular holder(s) of such rights or from her/his or their legal representative.
2. This presentation has been carefully compiled by PwC, but no representation is made or warranty given (either express or implied) as to the completeness or accuracy of the information it contains. PwC is not liable for the information in this presentation or any decision or consequence based on the use of it. PwC will not be liable for any damages arising from the use of the information contained in this presentation. The information contained in this presentation is of a general nature and is solely for guidance on matters of general interest. This presentation is not a substitute for professional advice on any particular matter. No reader should act on the basis of any matter contained in this publication without considering appropriate professional advice.

Authors:

Nikolaos Loutas, Michiel De Keyzer, and Stijn Goedertier

References

Slides 9:

- Linked Data Cookbook. W3C.
http://www.w3.org/2011/gld/wiki/Linked_Data_Cookbook

Slide 10-23:

- ISA Programme. Cookbook for translating Data Models to RDF Schemas.
<https://joinup.ec.europa.eu/community/semic/document/cookbook-translating-data-models-rdf-schemas>

Slide 16, 18,-21:

- W3C. An organization ontology. <http://www.w3.org/TR/vocab-org/>

Slide 23:

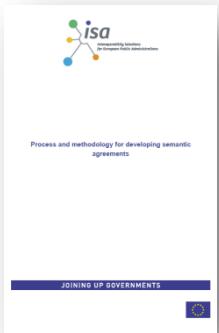
- ADMS Brochure. ISA Programme.
<https://joinup.ec.europa.eu/elibrary/document/adms-brochure>

Further reading



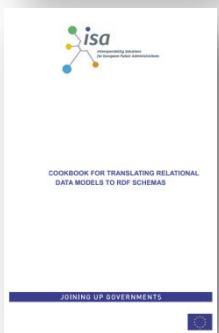
Linked Data Cookbook, W3C Government Linked Data Working Group,

http://www.w3.org/2011/gld/wiki/Linked_Data_Cookbook



EC, ISA Process and methodology for developing semantic agreements,

https://joinup.ec.europa.eu/community/core_vocabularies/document/process-and-methodology-developing-semantic-agreements



EC ISA, Cookbook for translating Data Models to RDF Schemas
<https://joinup.ec.europa.eu/community/semic/document/cookbook-translating-data-models-rdf-schemas>

Related projects and initiatives



Joinup, <http://joinup.ec.europa.eu>



Linked Open Vocabularies (LOV), <http://lov.okfn.org/>



EC ISA, e-Government Core Vocabularies,
https://joinup.ec.europa.eu/community/core_vocabularies/home



W3C Schools – Learn RDF
<http://www.w3schools.com/rdf/default.asp>



EUCLID, <http://euclid-project.eu/>



XML Summer School <http://xmlsummerschool.com/>

Be part of our team...

Find us on



[Open Data Support](#)

<http://www.slideshare.net/OpenDataSupport>



[Open Data Support](#)

<http://goo.gl/y9ZZI>

Follow us



[@OpenDataSupport](#)

Join us on



joinup

<http://www.opendatasupport.eu>

Contact us

contact@opendatasupport.eu