

Präsentation Metadaten

Open Data Support wird von der Europäischen Kommission finanziert, gemäß 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' (Vertrag No. 30-CE-0530965/00-17).

© 2013 European Commission



**OPEN
DATA
SUPPORT**

Trainingsmodul 2.4

Vokabulare in RDF entwerfen und entwickeln

Lernziele

Am Ende dieses Trainingsmoduls sollten Sie verstehen:

- was die besten Praktiken für die Erstellung eines RDF Vokabulars sind, um Daten zu modellieren.
- wo Sie RDF Vokabulare zur Wiederverwendung finden können.
- wie Sie Ihr eigenes RDF-Vokabular erstellen können.
- wie Sie Ihre RDF-Vokabulare veröffentlichen können.
- den Prozess und die Methodik für die Entwicklung semantischer Vereinbarungen, die vom ISA-Programm der Europäischen Kommission entwickelt werden.

Inhalt

In diesem Modul geht es...

- um die Schritte für die Modellierung Ihrer Daten.
- darum, wie Sie vorhandenen Vokabulare wiederverwenden können, um Ihre Daten zu modellieren.
- darum, wie Sie neue Klassen und Eigenschaften in RDF erstellen können.
- darum, wie und wo Sie Ihr RDF-Vokabular veröffentlichen können, so dass es von anderen wiederverwendet werden kann.

RDF-Vokabular

“Ein Vokabular ist ein Datenmodell, das Klassen, Eigenschaften und Beziehungen umfasst, die für die Beschreibung Ihrer Daten und Metadaten verwendet werden können.”

- RDF-Vokabulare sind **Sätze von Begriffen**, die verwendet werden, um Dinge zu beschreiben.
- Ein Begriff ist entweder eine **Klasse oder eine Eigenschaft**.
 - Objekttyp Eigenschaften (Beziehungen)
 - Datentyp Eigenschaften (Attribute)

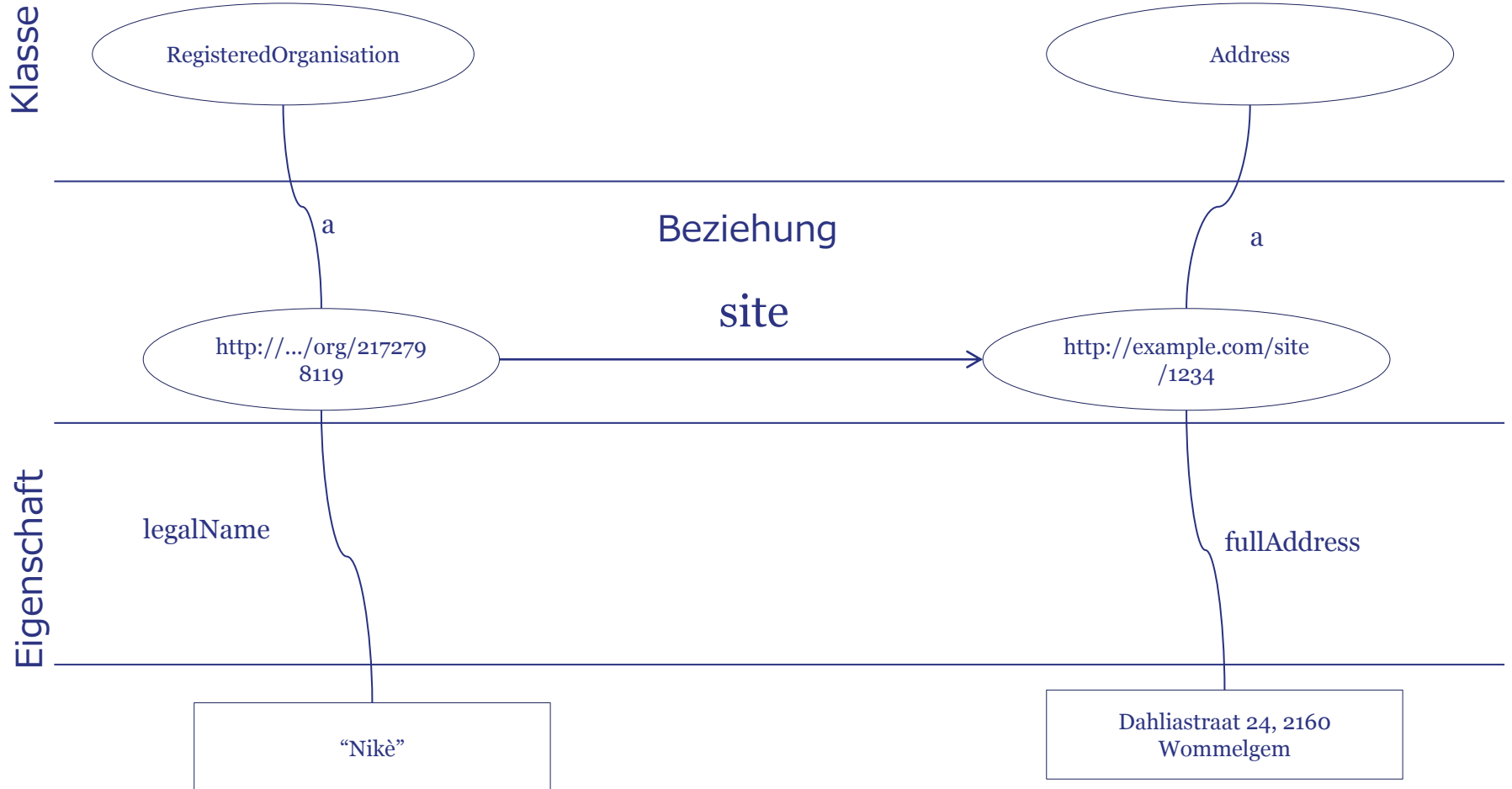
Was sind Klassen, Beziehungen und Eigenschaften?

Klasse. *Ein Konstrukt, das Dinge in der realen und/oder virtuellen Welt darstellt*, z.B. eine Person, eine Organisation, ein Konzept wie “Gesundheit” oder “Freiheit”.

Beziehung. *Eine Verbindung zwischen zwei Klassen*; für die Verbindung zwischen einem Dokument und der Organisation, die es veröffentlicht (d.h. Organisation *veröffentlicht* Dokument), oder zwischen einer Landkarte und der geografischen Region, die es zeigt (d.h. Landkarte *zeigt* geografische Region). In RDF werden Beziehungen als Objekttyp-Eigenschaften kodiert.

Eigenschaft. *Ein Merkmal einer Klasse in einer bestimmten Dimension*, wie dem gesetzlichen Namen einer Organisation oder dem Datum und der Zeit, wann eine Beobachtung gemacht wurde.

Beispiele von Klassen, Beziehungen und Eigenschaften



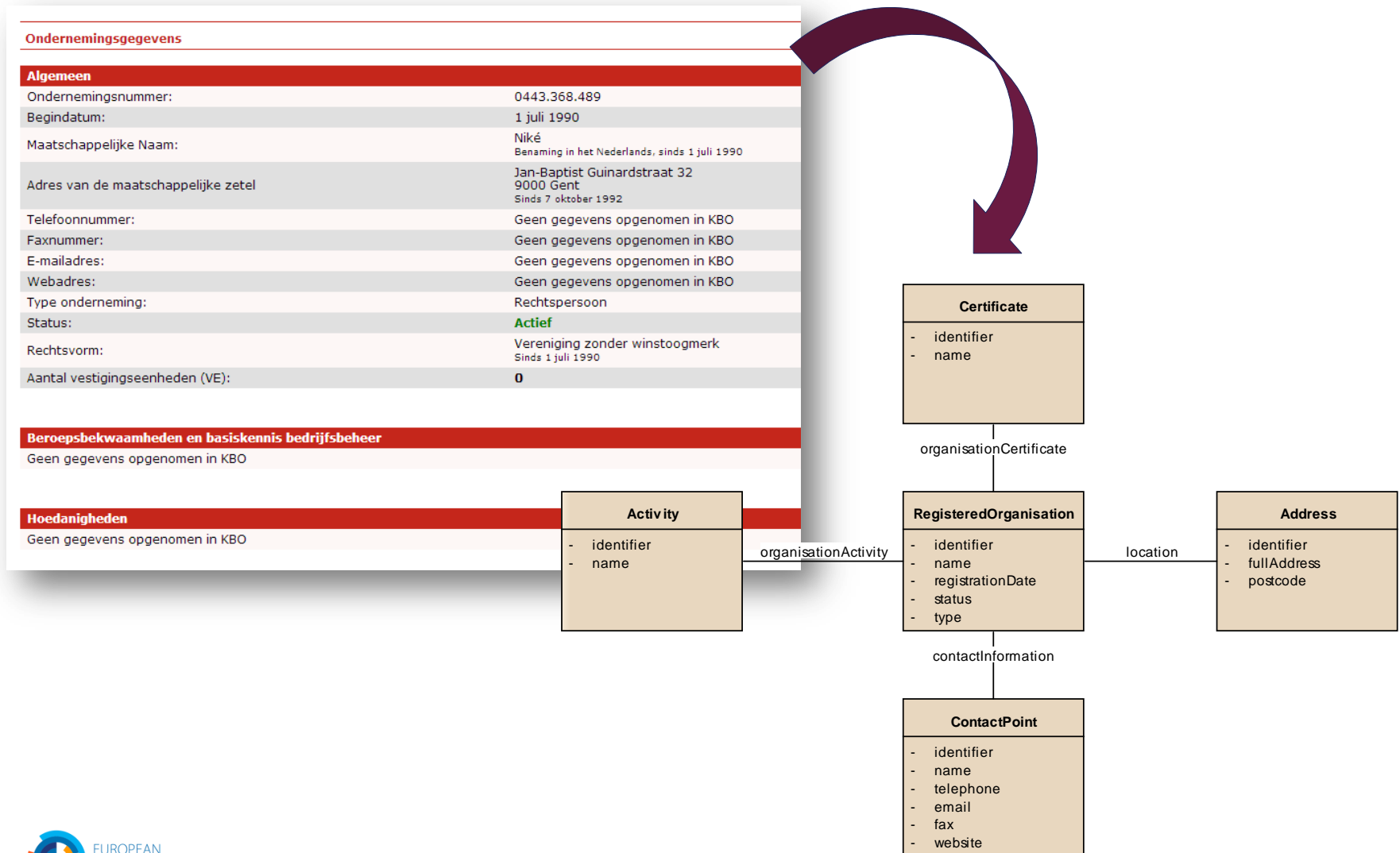
Modellieren Sie Ihre Daten

Wie Sie andere Vokabulare wiederverwenden, Ihre eigenen Begriffe definieren und Ihre Vokabulare veröffentlichen und bewerben können, um die Daten zu beschreiben.

6 Schritte zur Modellierung Ihrer Daten

- 1** Beginnen Sie mit **einem festen Domainmodell**, das nach einem strukturierten Prozess und Methodik entwickelt wurde.
- 2** Erforschen Sie **vorhandene Begriffe** und ihre Nutzung, und **maximieren Sie die Weiterverwendung** dieser Begriffe.
- 3** Dort, wo neue Begriffe als Spezialisierungen von vorhandenen Begriffen betrachtet werden können, erstellen Sie **Unterklassen** und **Untereigenschaften**.
- 4** Wenn **neue Begriffe** erforderlich sind, erstellen Sie sie nach **gemeinsam vereinbarten Best-Practice-Beispielen**.
- 5** Veröffentlichen Sie innerhalb einer **sehr stabilen Umgebung**, die, auf **Langlebigkeit** hin entworfen wurde.
- 6** **Publizieren Sie das RDF-Schema**, indem sie es mit relevanten Dienstleistungen registrieren.

1 *Beginnen Sie mit einem festen Domainmodell*



2 *Verwenden Sie vorhandene Begriffe und Vokabulare wieder*

- Universalvokabular: DCMI, RDFS 
- Um Dinge zu benennen: rdfs:label, foaf:name, skos:prefLabel 
- Um Leute zu beschreiben: FOAF, vCard, Core Person Vocabulary 
- Um Projekte zu beschreiben: DOAP, ADMS.SW 
- Um eine semantische Interoperabilität zu beschreiben: ADMS 
- Um registrierte Organisationen zu beschreiben : Registered Organisation Vocabulary 
- Um Adressen zu beschreiben : vCard, Core Location Vocabulary 
- Um öffentliche Dienstleistungen zu beschreiben : Core Public Service Vocabulary 
- Um Datensätze zu beschreiben : DCAT, DCAT Application Profile, VoID 

Anwendungsprofile erstellen

- **Verschiedene Domains** haben **verschiedene Anforderungen** an eine domain-spezifische Semantik, z.B. Klassifikationskonzepte.
- Generische RDF-Vokabulare liefern normalerweise die **Basisklassen**, die gebraucht werden, um Erweiterungen zu erlauben, mit denen, wenn erforderlich, spezifische Unterklassenstrukturen oder Klassifikationsschemen hinzugefügt werden können.
- In solchen Fällen werden Wiederverwender ermutigt, **Anwendungsprofile** einer **speziellen Anwendungsdomain** zu definieren, indem sie (falls erforderlich) **Unterklassen, Untereigenschaften und kontrollierte Vokabulare** definieren.
- Zum Beispiel,
 - [DCAT Anwendungsprofil für Datenportale in Europa](#)
 - [Registrierte Organisationsvokabulare](#) als Anwendungsprofil der [Organisations-Ontologie](#).

Siehe auch:

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

Vorteile der Weiterverwendung

- Weiterverwendung hilft in hohem Maße bei der **Interoperabilität** Ihrer Daten.

Zum Beispiel macht die Verwendung von <dcterms:created>, dessen Wert ein Datei vom Typ 2013-02-21[^]xsd:date sein sollte, dass die Daten sofort von viele Maschinen verarbeitbar sind. Wenn Ihr Schema Datenherausgeber dazu ermuntert, ein unterschiedliches Format für Begriff Datum, wie ex:date "21 February 2013“, zu benutzen, dann wird bei veröffentlichten Daten, bei denen Ihr Schema genutzt wird, eine weitere Verarbeitung erforderlich sein, um es das Gleiche wie alle anderen zu machen.

- Weiterverwendung **erhöht die Glaubwürdigkeit** Ihres Schemas.

Es zeigt, dass es mit Sorge und Professionalität veröffentlicht worden ist, und dieses fördert einmal mehr seine Weiterverwendung.

- Weiterverwendung ist **einfacher** und **billiger**.

Durch die Weiterverwendung von Klassen und Eigenschaften von gut definierten und richtig gehosteten Vokabeln vermeiden Sie, dass Sie diese Bemühung wiederholen müssen.

Hier können Sie vorhandene Vokabulare finden...

<http://lov.okfn.org/>

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

3 *Erstellen Sie Unterklassen und Untereigenschaften*

- RDF-Schemas und Vokabulare enthalten oft **Begriffe, die sehr generisch sind.**
- Durch die Erstellung von **Unterklassen-** und **Untereigenschafts-** Beziehungen können Systeme, die die Supereigenschaft oder Superklasse verstehen, in der Lage sein, die Daten zu interpretieren, selbst wenn die spezifischeren Begriffe unbekannt sind.
- **Erstellen Sie nicht Unterklassen und Untereigenschaften, nur um Ihnen zu erlauben, Ihren eigenen Begriff für etwas, das schon existiert, zu verwenden.**

Erstellung von Untereigenschaften - Beispiel

Das Vokabular der Registrierten Organisationen definiert drei Untereigenschaften von `org:classification`: `companyType`, `companyStatus` und `companyActivity`

Company Type

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

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 example 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>	<code>org:Organization</code>	<code>skos:Concept</code>

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>	<code>org:Organization</code>	<code>skos:Concept</code>

Eine Untereigenschaften in RDF definieren

```
<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 [Code](#skos:Concept) 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 *Wenn neue Begriffe erforderlich sind, erstellen Sie diese nach gemeinsam vereinbarten Best-Practices.*

- ✓ Klassen beginnen mit einem Großbuchstaben und sind immer Singular, z.B. skos:Concept.
- ✓ Eigenschaften beginnen mit einem Kleinbuchstaben, z.B. rdfs:label.
- ✓ Objekt-Eigenschaften sollte Verben sein, z.B. org:hasSite.
- ✓ Datentyp-Eigenschaften sollte Substantive sein, z.B. dcterms:description.
- ✓ Verwenden Sie Binnenversalien, wenn ein Begriff mehr als ein Wort hat, z.B. foaf:isPrimaryTopicOf.

Eine neue Klasse definieren - Organisation

```
<rdf:RDF
```

```
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
```

```
  xmlns:org="http://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>
```

Eine neue Eigenschaft definieren - registrationNumber

```
<rdf:RDF
```

```
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:org="http://example.org/organisation-schema">
```

```
<rdf:Property rdf:about="org:registrationNumber">
```

```
  <rdfs:label xml:lang="en">registrationNumber</rdfs:label>
```

```
  <rdfs:comment xml:lang="en">
```

in the The number that a registered organisation receives upon registration
 official register.

```
  </rdfs:comment>
```

```
</rdf:Class>
```

Domain- und Bereichseinschränkungen definieren

<rdf:RDF

xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"

xmlns:org="http://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 Veröffentlichen Sie innerhalb einer sehr stabilen Umgebung, die auf Langlebigkeit hin entworfen wurde.

- Wählen Sie einen stabilen Namensraum für Ihr RDF-Schema (z.B. W3C, Purl...)
- Verwenden Sie Good Practices bei der Veröffentlichung von langlebigen Uniform Resource Identifiers (URI)-Sätzen, sowohl in Bezug auf das Format als auch auf ihre Gestaltungsregeln und Verwaltung.
- Beispiele:
 - <http://www.w3.org/ns/adms>
 - <http://purl.org/dc/elements/1.1>

Siehe auch:

<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 Publizieren Sie das RDF-Schema, indem Sie es mit relevanten Diensten registrieren.

Sobald Ihr RDF-Schema veröffentlicht wird, werden Sie wollen, dass Leute darüber Bescheid wissen. Um ein breiteres Publikum zu erreichen, registrieren Sie es auf Joinup und Linked Open Vocabularies.

The image shows two overlapping screenshots. The background screenshot is the Joinup website, displaying an advanced search for 'organisation'. A green circle with the number '1' is placed over the search input field. A blue circle with the number '2' is placed over the 'Refine your search' filters on the left. A red circle with the number '3' is placed over the search results. Handwritten annotations include 'More focused.' in green above the search field, 'More targeted.' in blue to the left of the filters, and 'More relevant.' in orange above the results. The foreground screenshot is the Linked Open Vocabularies (LOV) website, showing search results for 'organisation' with a list of URIs and scores.

More focused.

More targeted.

More relevant.

1

2

3

<http://lov.okfn.org>

Refine the search results via the faceted search filters.

Schlussfolgerungen

Analysieren

Beginnen Sie mit einem Domainmodell, das fest ist und nach einem strukturierten Prozess und Methodik entwickelt wurde.

Erforschen Sie vorhandene Begriffe und ihre Nutzung, und maximieren Sie die Weiterverwendung dieser Begriffe.

Modellieren

Wenn neue Begriffe als Spezialisierungen von vorhandenen Begriffen betrachtet werden können, erstellen Sie Unterklassen und Untereigenschaften.

Wenn neue Begriffe erforderlich sind, erstellen Sie diese nach gemeinsam vereinbarten Best-Practices.

Veröffentlichen Sie innerhalb einer sehr stabilen Umgebung, die auf Langlebigkeit hin entworfen wurde.

Veröffentlichen

Publizieren Sie das RDF-Schema durch Registrierung mit relevanten Dienstleistungen.

Gruppenübung



<http://www.visualpharm.com>

Erstellen Sie in Zweiergruppen die RDF Beschreibung eines Vokabulars, um einen Bürger darzustellen.



<http://www.visualpharm.com>

Was sind Ihrer Meinung nach die größten Hindernisse für die Weiterverwendung von vorhandenen RDF-Vokabularen?

Nehmen Sie auch die Online-Test!

Vielen Dank!

...und jetzt IHRE Fragen?

Diese Präsentation wurde von Open Data Support erstellt

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.

Autoren:

Nikolaos Loutas, Michiel De Keyzer, and Stijn Goedertier

Referenzen

Folie 9:

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

Folie 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>

Folie 16, 18-21:

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

Folie 23:

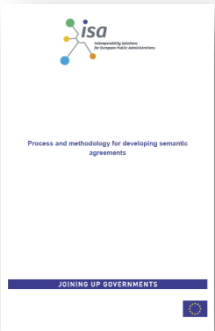
- ADMS Brochure. ISA Programme.
<https://joinup.ec.europa.eu/elibrary/document/adms-brochure>

Weiter lesen



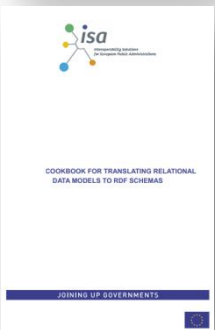
Linked Data Cookbook, W3C Government Linked Data Working Group,

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>

Verwandte Projekte und Initiativen



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/>

Werden Sie Teil unseres Teams...

*Finden Sie
uns auf*



Open Data Support

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



Open Data Support

<http://goo.gl/y9ZZI>

*Folgen Sie
uns auf*



@OpenDataSupport

*Begleiten Sie
uns auf*



joinup

<http://www.opendatasupport.eu>

*Kontaktieren
Sie uns unter*

contact@opendatasupport.eu