



**OPEN
DATA
SUPPORT**

Module de formation 2.4

Concevoir et
développer des
vocabulaires RDF



PwC firms help organisations and individuals create the value they're looking for. We're a network of firms in 158 countries with close to 180,000 people who are committed to delivering quality in assurance, tax and advisory services. Tell us what matters to you and find out more by visiting us at www.pwc.com. PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details.

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).

© 2013 European Commission

This presentation has been created by PwC

Authors:

Nikolaos Loutas, Michiel De Keyzer and Stijn Goedertier

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.

Les objectifs de cette formation

A la fin de ce module de formation, vous devriez avoir une compréhension de:

- Quelles sont les meilleures pratiques pour la création d'un vocabulaire RDF pour modéliser vos données.
- Où trouver des vocabulaires RDF à réutiliser.
- Comment pouvez-vous créer votre propre vocabulaire RDF.
- Comment faire pour publier votre vocabulaire RDF.
- Le processus et la méthodologie pour l'élaboration d'accords sémantiques développés par le programme ISA de la Commission Européenne

Contenu

Ce module porte sur...

- Les étapes de la modélisation de vos données.
- Comment réutiliser les vocabulaires existants pour modéliser vos données.
- Comment faire pour créer de nouvelles classes et propriétés RDF.
- Comment et où publier votre vocabulaire RDF afin qu'il puisse être réutilisé par d'autres.

Vocabulaire RDF

« Un vocabulaire est un modèle de données comportant des classes, propriétés et relations qui peut être utilisé pour décrire vos données et métadonnées. »

- Les vocabulaires RDF sont des ensembles de termes utilisés pour décrire des choses.
- Un terme est soit une classe, soit une propriété.
 - Propriétés de type d'objet (les relations)
 - Propriétés de type de données (les attributs)

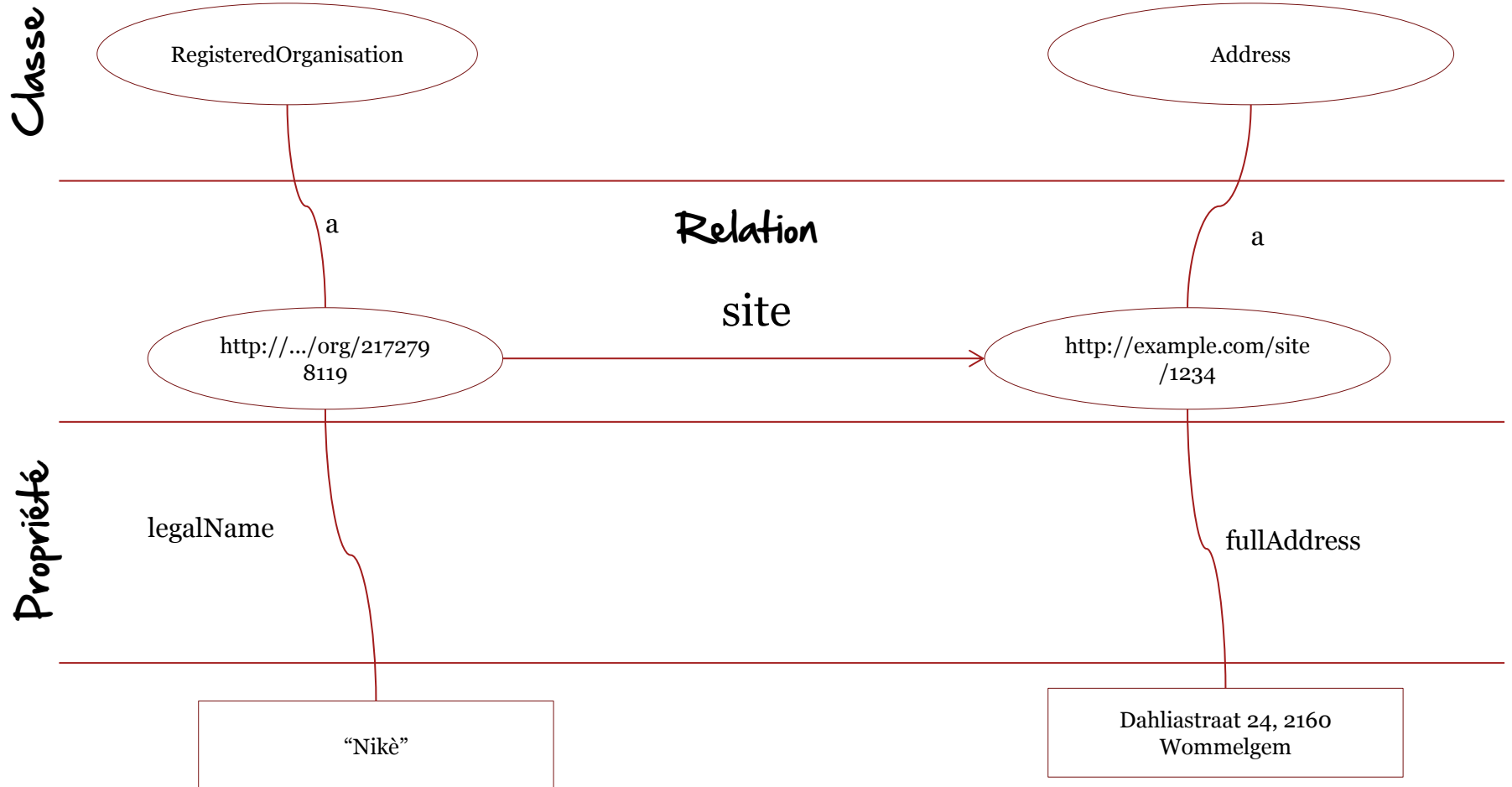
Que sont les classes, les relations et les propriétés?

Classe. *Une construction qui représente des choses dans le monde réel et/ou des informations*, par exemple une personne, une organisation, un concept tel que «santé» ou «liberté».

Relation. *Un lien entre deux classes*; par exemple le lien entre un document et l'organisation qui l'a publié (c-à-d organisation *publie* document), ou le lien entre une carte et la région géographique qu'elle représente (c-à-d carte représente région géographique). En RDF, les relations sont codées comme des propriétés de type d'objet.

Propriété. *Une caractéristique d'une classe* dans une dimension particulière, comme le nom légal d'une organisation ou la date et l'heure où une observation a été faite.

Des exemples de classes, des relations et des propriétés



Modéliser vos données

Comment réutiliser d'autres vocabulaires, définir vos propres termes, publier et promouvoir vos vocabulaires pour décrire les données.

6 étapes pour modéliser vos données

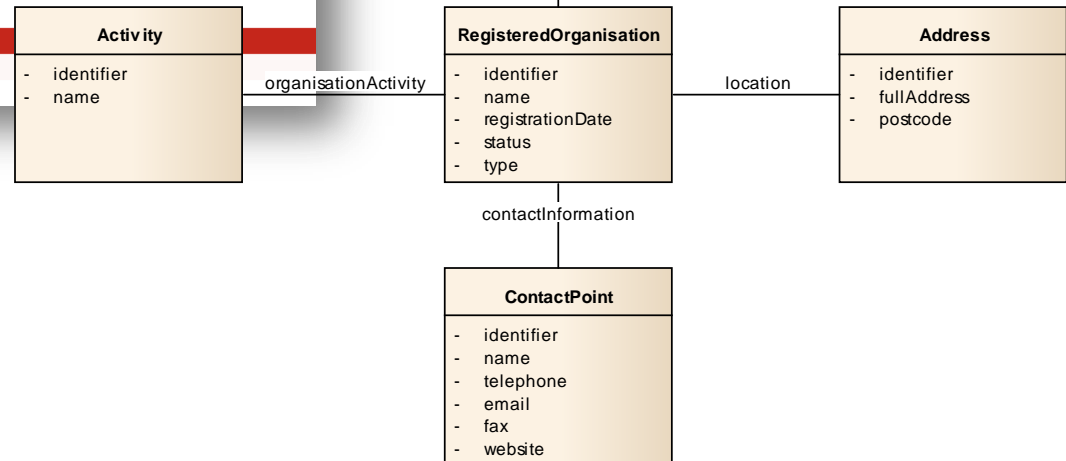
- 1** Commencer avec un **modèle de domaine robuste** élaboré suivant un processus et une méthodologie structurés.
- 2** Rechercher des **termes existants** et leur utilisation pour ensuite **maximiser la réutilisation** de ces termes.
- 3** Lorsque de nouveaux termes peuvent être considérés comme des spécialisations de termes existants, créer des **sous-classes et sous-propriétés**.
- 4** Lorsque de **nouveaux termes** sont nécessaires, créer les, suivant **les meilleures pratiques communément admises**.
- 5** Publier dans un **environnement très stable** conçu pour être persistant.
- 6** **Promouvoir le schéma RDF** en l'enregistrant avec les services pertinents.

Voir aussi:

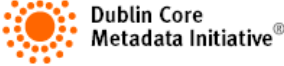








<https://joinup.ec.europa.eu/community/semic/document/cookbook-translating-data-models-rdf-schemas>

1 Commencer avec un modèle de domaine robuste

Ondernemingsgegevens	
Algemeen	
Ondernemingsnummer:	0443.368.489
Begindatum:	1 juli 1990
Maatschappelijke Naam:	Niké Benaming in het Nederlands, sinds 1 juli 1990
Adres van de maatschappelijke zetel	Jan-Baptist Guinardstraat 32 9000 Gent Sinds 7 oktober 1992
Telefoonnummer:	Geen gegevens opgenomen in KBO
Faxnummer:	Geen gegevens opgenomen in KBO
E-mailadres:	Geen gegevens opgenomen in KBO
Webadres:	Geen gegevens opgenomen in KBO
Type onderneming:	Rechtspersoon
Status:	Actief
Rechtsvorm:	Vereniging zonder winst oogmerk Sinds 1 juli 1990
Aantal vestigingseenheden (VE):	0
Beroepsbekwaamheden en basiskennis bedrijfsbeheer	
Geen gegevens opgenomen in KBO	
Hoedanigheden	
Geen gegevens opgenomen in KBO	



2 Réutiliser les termes et vocabulaires existants

- Vocabulaire à usage général : DCMI, RDFS 
- Pour nommer des ressources: `rdfs:label`, `foaf:name`, `skos:prefLabel` 
- Pour décrire des personnes : FOAF, vCard, Core Person Vocabulary 
- Pour décrire des projets : DOAP, ADMS.SW 
- Pour décrire des ressources d'interopérabilité : ADMS 
- Pour décrire des organisations enregistrées: Registered Organisation Vocabulary 
- Pour décrire des adresses: vCard, Core Location Vocabulary 
- Pour décrire des services publics : Core Public Service Vocabulary 
- Pour décrire des ensembles de données: DCAT, DCAT Application Profile, VoID 

Création de profils applicatifs

- Des **domaines différents** ont des **contraintes différentes** pour la sémantique qui est spécifique à ce domaine, par exemple les concepts de classification.
- Les vocabulaires RDF génériques fournissent habituellement les **classes de base nécessaires** pour permettre à des extensions d'ajouter, selon les besoins, des structures spécifiques de sous-classe ou de schémas de classification.
- Dans de tels cas, les réutilisateurs sont encouragés à définir des **profils applicatifs particuliers à un domaine d'application** en spécifiant (si nécessaire) **sous-classes, sous-propriétés et des vocabulaires contrôlés**.
- Par exemple,
 - [Le profil applicatif DCAT pour les portails de données en Europe](#)
 - Le [vocabulaire « Registered Organization »](#) comme profil applicatif de l'ontologie « Organization ».

Voir aussi:

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

Les avantages de la réutilisation

- La réutilisation **facilite l'interopérabilité** de vos données
 - Utilisez `dcterms:created`, par exemple, la valeur de ce qui devrait être une date formatée de la même manière que: `2013-02-21 ^^ xsd:date`. Celle-ci est immédiatement traitable par de nombreuses machines. Si votre schéma encourage les éditeurs de données d'utiliser un terme et un format de date différents, comme par ex: la date «21 Février 2013» - les données publiées en utilisant votre schéma nécessiteront un traitement supplémentaire pour les rendre identiques à toutes les autres.
- La réutilisation de votre schéma leur **ajoute de la crédibilité**.
 - Cela montre qu'il a été publié avec soin et professionnalisme, encore une fois, ceci favorise sa réutilisation.
- La réutilisation est plus facile et moins cher.
 - Réutiliser des classes et des propriétés de vocabulaires bien définis et bien accueillis évite que vous ayez à reproduire cet effort.

Vous pouvez trouver des vocabulaires RDF réutilisables sur ...

The screenshot shows the Joinup website interface. At the top, there is a navigation bar with the European Commission logo and the text "Share and reuse interoperability solutions for public administrations". Below this, there is a search bar with the text "organisation RDF" and a "Search" button. The search results are displayed in a list format, with the first result being "Enterprise Competence Organisation Schema".

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

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

The screenshot shows the Linked Open Vocabularies (LOV) website interface. At the top, there is a navigation bar with the text "Linked Open Vocabularies (LOV)" and logos for MONDECA, Inserm, and others. Below this, there is a search bar with the text "organisation" and a "Search" button. The search results are displayed in a list format, with the first result being "swrc:Organization (owl:Class)".

3 *Création de sous-classes et sous-propriétés*

- Les schémas et les vocabulaires RDF comprennent souvent des **termes qui sont très génériques**.
- En créant des **sous-classes** et des relations de **sous-propriété**, les systèmes qui comprennent la super propriété ou super classe peuvent être capables d'interpréter les données, même si les termes plus spécifiques sont inconnus.
- Ne **pas créer de sous-classes et sous-propriétés simplement pour vous permettre d'utiliser votre propre terme** pour quelque chose qui existe déjà.

Création de sous-propriétés - un exemple

Le vocabulaire “Registered Organization” définit trois sous-propriétés de “org:classification” : “companyType”, “companyStatus” et “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>

Définition d'une sous-propriété en 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 `Code` 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 *Lorsque de nouveaux termes sont nécessaires, créez-les suivant les meilleures pratiques communément admises*

- ✓ Les Classes commencent avec une majuscule et sont toujours au singulier, par exemple, `skos:Concept`.
- ✓ Les propriétés commencent par une lettre en minuscule, par exemple `rdfs:label`.
- ✓ Les propriétés des objets devraient être des verbes, par exemple de `org:hasSite`.
- ✓ Les propriétés de type de données doivent être des noms, par exemple de `dcterms:Description`.
- ✓ Utilisez le « Camel Case » si un terme a plus d'un mot, par exemple `FOAF:isPrimaryTopicOf`.

Définition d'une nouvelle classe - 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>
```

Définition d'une nouvelle propriété - 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>
```

Définir les restrictions de domaine et de portée

<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 Publier dans un environnement très stable conçu pour être persistant

- Choisissez un espace de noms stable pour votre schéma RDF (par exemple W3C, Purl ...)
- Appliquez les bonnes pratiques pour la publication d'ensembles persistants d'identificateurs uniformes de ressources (URI), tant en termes de format que de leurs règles de design et de management.
- Exemples :
 - <http://www.w3.org/ns/adms>
 - <http://purl.org/dc/elements/1.1>

Voir aussi:

<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 Faire connaître le schéma RDF en l'enregistrant auprès des services compétents

Une fois votre schéma RDF publié, vous voulez que les gens le sachent. Pour atteindre un public plus large, enregistrez-le sur Joinup et « Linked Open Vocabularies ».

The image shows two screenshots illustrating the process of registering an RDF schema. The left screenshot is from the Joinup website, showing an advanced search for 'organisation'. A green circle with the number '1' highlights the search input field. A blue circle with the number '2' highlights the 'Filtering by' section on the left, which includes filters for 'Semantic Asset Projects', 'Semantic Asset Releases', and 'Repository of Origin'. A blue annotation 'More targeted.' is next to it, with the text 'Refine the search results via the faceted search filters.' below. A green annotation 'More focused.' is above the search input. A blue circle with the number '3' highlights the 'Organisation Type List' section, which includes 'Enterprise Competence Organisation Schema' and 'EC Fisheries - Regional Fisheries Management Organisation Codes'. An orange annotation 'More relevant.' is next to it. The right screenshot is from the Linked Open Vocabularies (LOV) website, showing search results for 'organisation'. A table lists results with columns for the vocabulary name, URI, and score. The URL 'http://lov.okfn.org' is written below the screenshot.

Joinup: Share and reuse interoperability solutions for public administrations. Search for 'organisation'. Results: 1 to 20 of 112 results. Annotations: 1 (Search input), 2 (Filtering by), 3 (Organisation Type List). Annotations: More targeted., More focused., More relevant.

Linked Open Vocabularies (LOV): Search for 'organisation'. Results: 147 results in 28 vocabularies. Annotations: http://lov.okfn.org

Conclusions

Commencez avec un modèle de domaine robuste élaboré suivant un processus et une méthodologie structurés.

Analyser

Recherchez les termes existants et leur utilisation et maximisez la réutilisation de ces termes.

Lorsque de nouveaux termes peuvent être considérés comme des spécialisations de termes existants, créez des sous-classes et sous-propriétés selon les cas.

Modéliser

Lorsque de nouveaux termes sont nécessaires, créez les suivants les meilleures pratiques communément admises en termes de conventions pour l'attribution de noms, etc.

Publiez dans un environnement très stable conçu pour être persistant.

Publier

Faites connaître le schéma RDF en l'enregistrant avec les services compétents.

Exercice de groupe



<http://www.visualpharm.com>

En groupe de deux, créez la description RDF d'un vocabulaire pour représenter un citoyen.



<http://www.visualpharm.com>

D'après vous, quelles sont les principales barrières pour la réutilisation de vocabulaires RDF existants?

Faites le test en ligne [ici!](#)

Merci!

...et maintenant vos questions?

Références

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>

Lecture supplémentaire



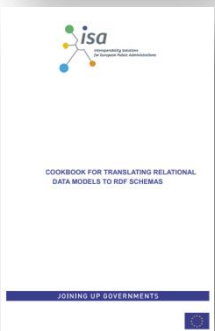
Linked Data Cookbook, W3C Government Linked Data Working Group,

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>

Projets et initiatives apparentés



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)

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



[Open Data Support](http://goo.gl/y9ZZI)

<http://goo.gl/y9ZZI>

Follow us



[@OpenDataSupport](https://twitter.com/OpenDataSupport)

Join us on



joinup

<http://www.opendatasupport.eu>

Contact us

contact@opendatasupport.eu