

Web sémantique et représentation de connaissance

Jérôme Euzenat, Jean-François Baget



`Jerome.Euzenat@inrialpes.fr`

`Jean-Francois.Baget@inrialpes.fr`

But du cours

Présenter les moyens de représenter la connaissance
et leur sémantique

en considérant les développements autour du web
sémantique

et les difficultés théoriques qu'ils soulèvent

Plan du cours

- 1 Introduction au web sémantique
- 2 Sémantique et théorie des modèles
- 3&4 Un langage assertionnel à base de graphe (RDF et graphes conceptuels)
- 5&6 Un langage définitionnel (OWL et logiques de descriptions)
- 7 Règles et requêtes
- 8 Transformations et préservations de propriétés

3

Ressources

Page web:

<http://www.inrialpes.fr/exmo/teaching/src2>

Quelques références.

Copie des transparents (non pas en avance...).

L'ancien polycopié, peut être utile:

4

But de l'introduction

Comment utiliser
les données et les ressources du web sémantique

(au travers d'un exemple)

5

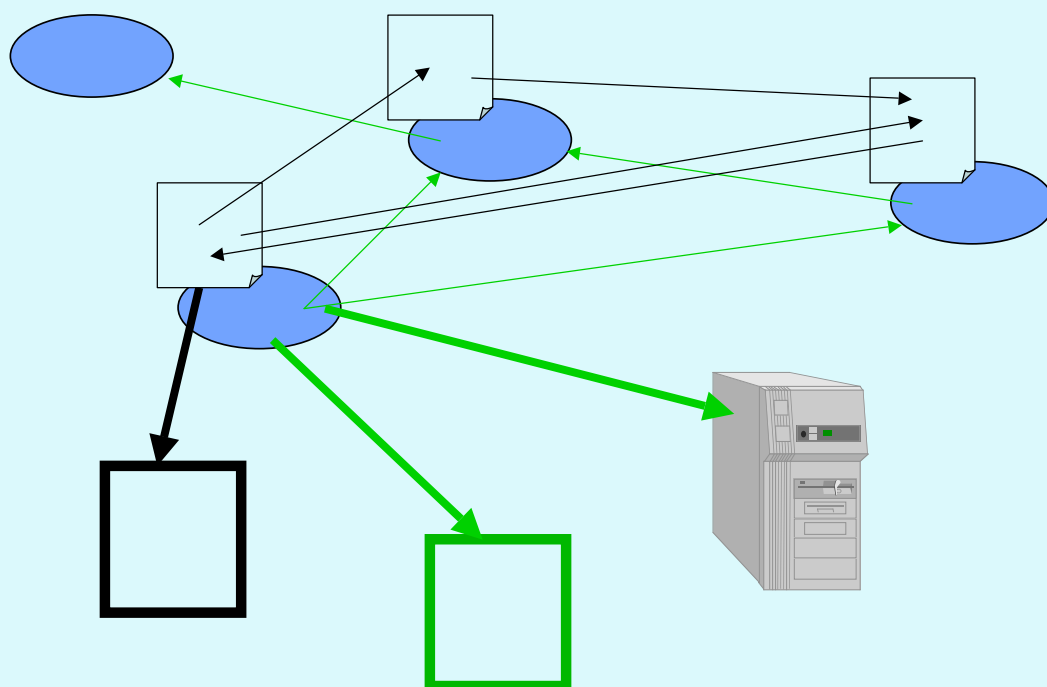
Plan de l'introduction

- Le web sémantique (introduction)
- Exemples (RSS, Dublin Core, Creative Commons, etc.) (qu'est-ce qui existe)
- Modéliser son domaine (de quoi ai-je besoin)
- Produire des descriptions en RDF (et hop, c'est parti!)
- Le cas multimédia

6

Le web sémantique

Semantic web



Apports

- Plein texte (ou mots-clé):
“livre” “Bertrand Russell”
- Requête structurée:
“livres sur Bertrand Russell”
- Requête renseignées:
“autobiographies de Bertrand Russell”
- Requêtes composites:
“quelle bibliothèque la plus proche peut me prêter l’autobiographie de Bertrand Russell”

9

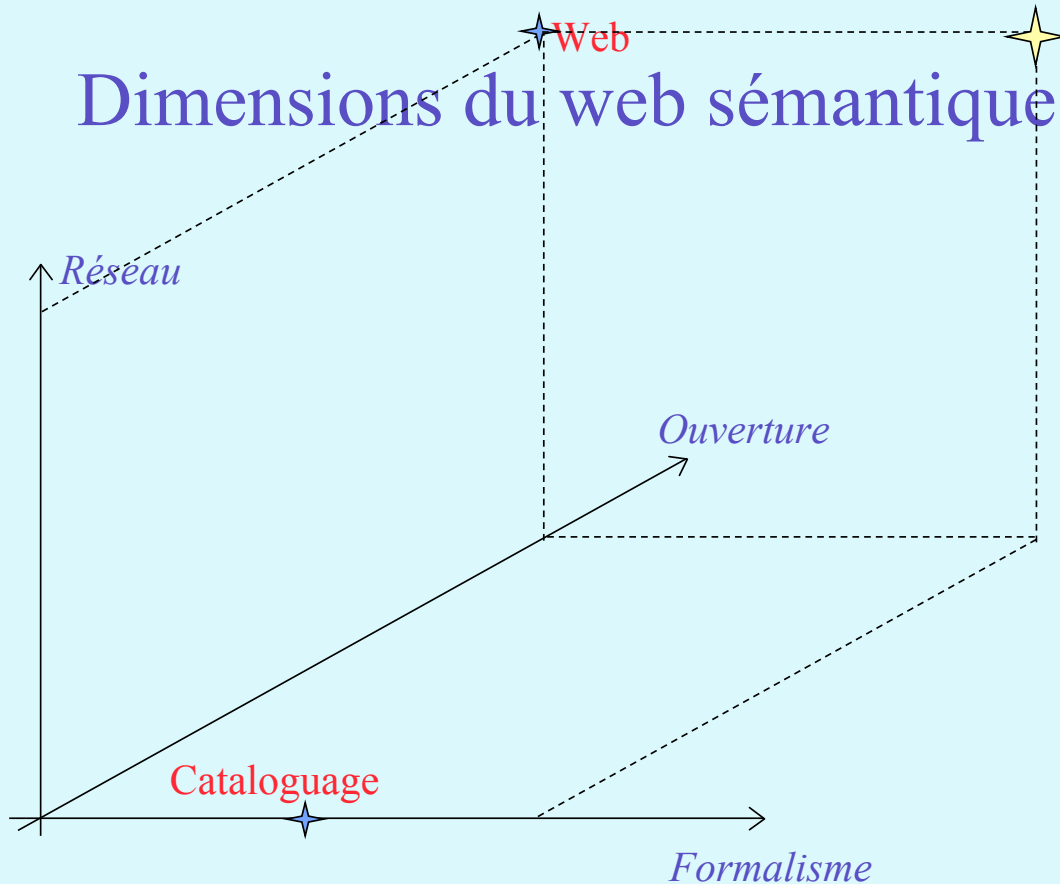
Définitions

10000 pieds: « un web compréhensible par des machines ».

50 mm: « une infrastructure composée de langages pour exprimer la connaissance, pour décrire les relations entre connaissance, pour décrire les conditions d’utilisation, pour décrire les garanties et les modes de paiement et de dispositifs permettant de trouver les ressources. »

10

Dimensions du web sémantique



11

Quels sont les besoins ?

Une *identification* et un *accès* aux ressources du Web de façon simple, robuste, et efficace (URI).

Des langages pour décrire le *contenu* des documents/ressources et des moteurs d'inférences pour les manipuler...

Des *ressources* de plusieurs types : ontologies (bases de connaissances), bases de données, éléments de code, etc.

12

Caractéristiques

Un web pour les machines...

- ouvert
 - décentralisé
 - interopérable
 - formel
- } web
- sémantique

+ personnel

15

Notre exemple

Élaborer un schéma permettant d'indexer une collection d'ouvrages, tout en tirant parti du web sémantique:

- lier les ouvrages à leur auteur;
- trouver les ouvrages accessibles sur le web;
- comprendre les droits associés;
- diffusion d'information via RSS.

16

Quelques ressources

17

Dublin Core

Le Dublin Core définit un certain nombre de champs de métadonnées utilisables pour décrire les ressources du Web

15 *champs* ou *éléments* standardisés

- titre, créateur, description, sujet, date, source, droits...

Les champs peuvent être précisés à l'aide de *raffinements*

- description : résumé, table des matières

Les valeurs de ces champs reprennent généralement d'autres standards

18

Dublin Core

```
<html>
  <head><title>Cours INRIA</title>
  <link rel="schema.DC"
    href="http://purl.org/dc/elements/1.1/" />
  <meta name="DC.Title" lang="fr" content="Cours INRIA
    en HTML" />
  <meta name="DC.Date.created" scheme="W3CDTF"
    content="2004-09-27" />
  <meta name="DC.Date.modified" scheme="W3CDTF"
    content="2004-09-28" />
  <meta name="DC.Subject" lang="fr" content="HTML,
    document, Dublin Core" />
  <meta name="DC.Language" scheme="RFC3066" content="fr-
    FR" />
</head>
<body></body>
</html>
```

19

Dublin Core

Traduction en triplets RDF:

```
<rdf:RDF
  xmlns:rdf='http://www.w3.org/1999/02/22-rdf-syntax-
    ns#'
  xmlns:dc='http://purl.org/dc/elements/1.1/'>
  <rdf:Description about="Cours_INRIA.html">
    <dc:title>Cours INRIA en HTML</dc:title>
    <dc:created>2004-09-27</dc:created>
    <dc:modified>2004-09-28</dc:modified>
    <dc:language>fr</dc:language>
  </rdf:Description>
</rdf:RDF>
```

20

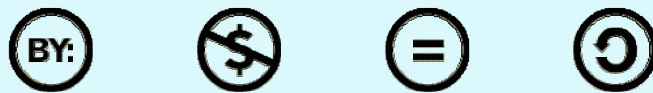
Creative Commons

Objectifs :

- adapter les droits des auteurs au nouveau médium qu'est Internet,
- fournir un cadre juridique au partage sur le web d'œuvres de l'esprit comme les images, les sons ou les textes

Principe : "un jeu de logos"

- la combinaison de ces 4 briques fournit 6 types de licence

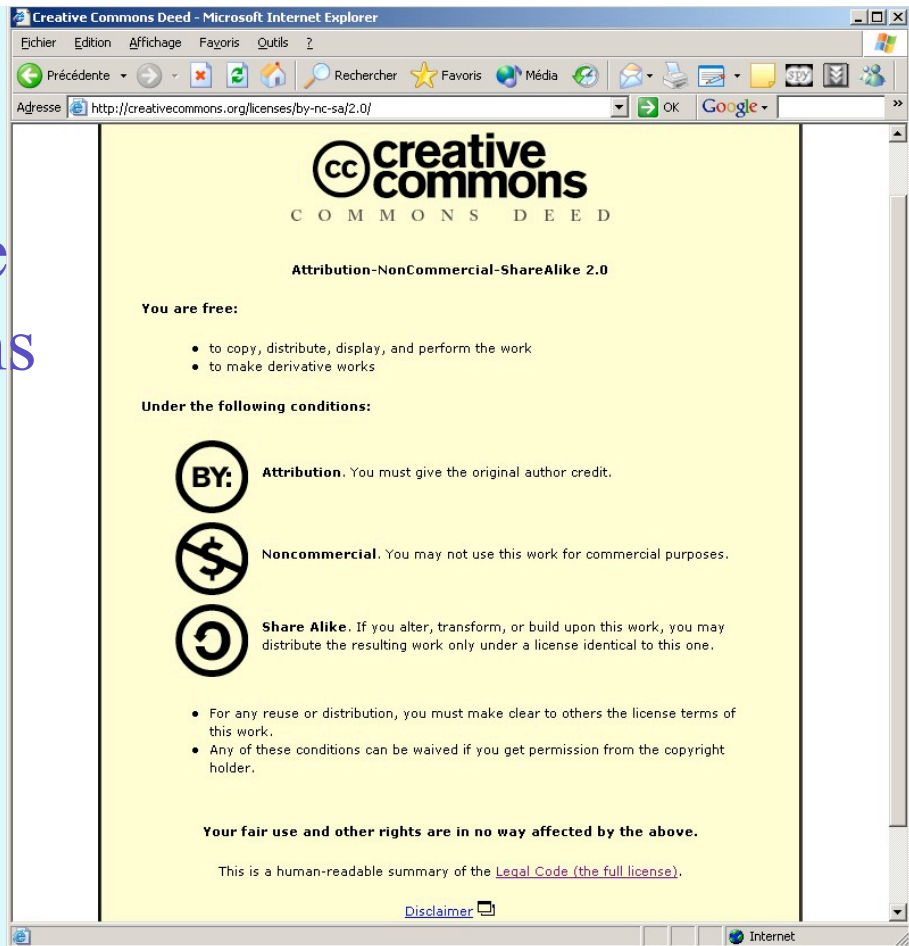


21

Creative Commons

A screenshot of the Creative Commons website's 'Choose a License' page, viewed in Microsoft Internet Explorer. The browser's address bar shows 'http://creativecommons.org/license/'. The page features a navigation menu with links for 'home', 'learn more', 'get content', 'discuss', 'technology', and 'choose license'. On the left side, there is a vertical sidebar with four steps: 'step 1 choose license', 'step 2 review choice', 'step 3 mark content', and 'step 4 publicize'. Below these steps, there is a section titled 'Or Choose:' with several license icons: 'pd Public Domain', 'cc Recombo', 'fc Founder's Copyright', 'cc-gnu-gpl CC-GNU GPL', 'cc-gnu-lgpl CC-GNU LGPL', and 'share music Share Music'. The main content area is titled 'Choose License' and contains a form with the following text: 'With a Creative Commons license, you keep your copyright but allow people can copy and distribute your work, provided they give you credit -- and only on the conditions you specify here. If you want to offer your work with no conditions, choose the public domain. Do you want to:'. Below this text are two radio button options: 'Allow commercial uses of your work?' with 'Yes' and 'No' options, and 'Allow modifications of your work?' with 'Yes', 'Yes, as long as others share alike', and 'No' options. There are also dropdown menus for 'Jurisdiction of your license' (set to 'Generic') and 'Tell us the format of your work' (set to 'Other'). A 'Select a License' button is at the bottom of the form. The footer of the page includes links for 'home | learn more | get content | discuss | technology | choose license'.

Creative Commons



Creative Commons

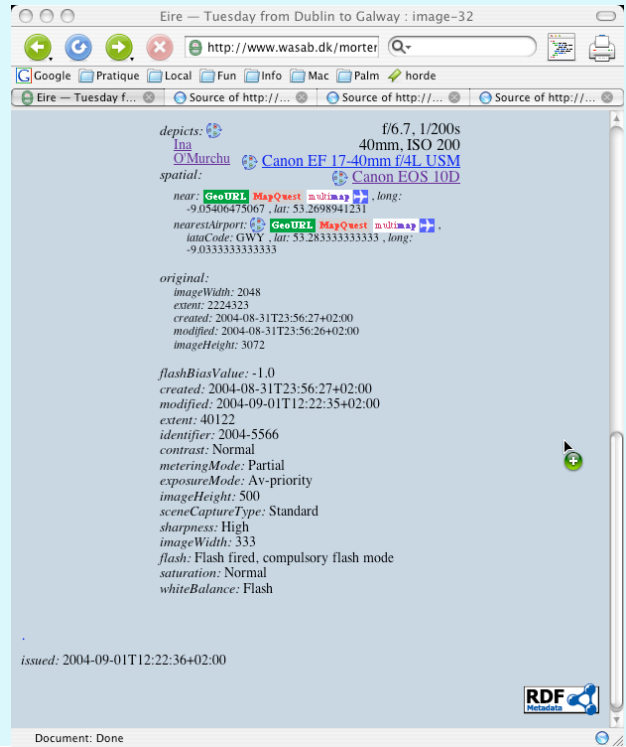
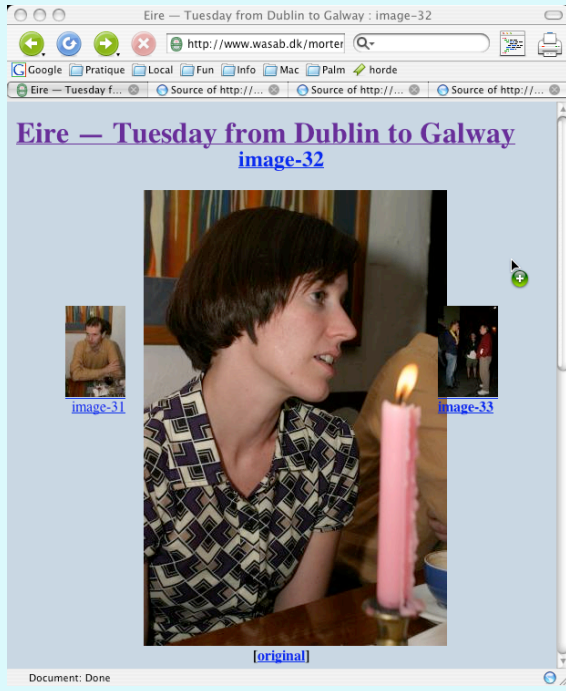
```
<rdf:RDF xmlns='http://web.resource.org/cc/'  
  xmlns:dc='http://purl.org/dc/elements/1.1/'  
  xmlns:rdf='http://www.w3.org/1999/02/22-rdf-syntax-ns#' >
```

```
<Work rdf:about="">  
  <license rdf:resource="http://creativecommons.org/licenses/by-nc-sa/2.0/">  
</Work>
```

```
</rdf:RDF>
```

```
<License rdf:about="http://creativecommons.org/licenses/by-nc-sa/2.0/">  
  <permits rdf:resource="http://web.resource.org/cc/Reproduction"/>  
  <permits rdf:resource="http://web.resource.org/cc/Distribution"/>  
  <requires rdf:resource="http://web.resource.org/cc/Notice"/>  
  <requires rdf:resource="http://web.resource.org/cc/Attribution"/>  
  <prohibits rdf:resource="http://web.resource.org/cc/CommercialUse"/>  
  <permits rdf:resource="http://web.resource.org/cc/DerivativeWorks"/>  
  <requires rdf:resource="http://web.resource.org/cc/ShareAlike"/>  
</License>
```

FOAF



Ina O' Murchu

[foafnaut web view plink]

FOAF



Knows:

Kerry Santo
[foafnaut web view plink sha1sum of a personal mailbox URI name:
5e942989a5d4e85fa6d41b123ffadf77aca2fd8f]

workplace homepage:

[DERI](#)

weblog:

[Ecademy Weblog](#)

homepage:

[Ecademy Profile Page](#)

Address:

Country:

Ireland

Region/State:

Locality/Suburb:

Galway

Organization:

Referenced Vocabularies

- <http://www.w3.org/1999/02/22-rdf-syntax>
The RDF Vocabulary (RDF)
- <http://xmlns.com/foaf/0.1/>
- <http://www.w3.org/2000/01/rdf-schema>

FOAF

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:bio="http://purl.org/vocab/bio/0.1/"
  xmlns:vCard="http://www.w3.org/2001/vcard-rdf/3.0#"
  xmlns:geo="http://www.w3.org/2003/01/geo/wgs84_pos#"
  xmlns:admin="http://webns.net/mvcb/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/">

  <foaf:PersonalProfileDocument rdf:about="">
    <foaf:primaryTopic rdf:nodeID="_40800"/>
    <foaf:maker rdf:nodeID="_40800"/>
    <dc:title>FOAF for Ina O' Murchu at Ecademy</dc:title>
    <dc:description>Ecademy Friend-of-a-Friend description for Ina O' Murchu</dc:description>
    <admin:generatorAgent rdf:resource="http://www.ecademy.com"/>
    <admin:errorReportsTo rdf:resource="mailto:webmaster@ecademy.com"/>
  </foaf:PersonalProfileDocument>
```

```
<foaf:Person rdf:nodeID="_40800">
  <foaf:name>Ina O' Murchu</foaf:name>
  <foaf:title>Miss</foaf:title>
  <foaf:firstName>Ina</foaf:firstName>
  <foaf:surname>O' Murchu</foaf:surname>
  <foaf:nick>inaomurchu</foaf:nick>
  <foaf:mbox_sha1sum>d21dd8e89548cf79fa8bb462561711fc09540856</foaf:mbox_sha1sum>
  <foaf:Organization>DERI</foaf:Organization>
  <vCard:ADR rdf:parseType="Resource">
    <vCard:Locality>Galway</vCard:Locality>
    <vCard:Country>Ireland</vCard:Country>
  </vCard:ADR>
  <foaf:workplaceHomepage dc:title="DERI" rdf:resource="http://www.deri.ie"/>
  <foaf:img rdf:resource="http://www.ecademy.com/images/photo/40800.jpg"/>
```

27

RSS

Permet de publier une succession d'éléments d'information (fil).

Dés qu'il est présenté en RDF, il permet la manipulation et la sélection automatique.

28

RSS+DC+CC

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF xmlns:ct="http://xmlfr.org/dtd/content" ...>
  <channel rdf:about="http://xmlfr.org/documentations/articles/channel.rss10">
    <title>Articles</title>
    <link>http://xmlfr.org/documentations/articles/</link>
    <description>Articles de fond sur XML</description>
    <dc:source rdf:resource="http://xmlfr.org/documentations/articles/" />
    <dc:language>fr</dc:language>
    <dc:publisher>Eric van der Vlist (mailto:vdv@dyomedeia.com)</dc:publisher>
    <dc:creator>Eric van der Vlist (mailto:vdv@dyomedeia.com)</dc:creator>
    <cc:license rdf:resource="http://creativecommons.org/licenses/by-nc-sa/2.0/" />
    <image rdf:resource="http://xmlfr.org/bandeaux/xmlfr_88x31.gif" />
    <items>
      <rdf:Seq>
        <rdf:li
rdf:resource="http://xmlfr.org/documentations/articles/040331-0001" />
        <rdf:li
rdf:resource="http://xmlfr.org/documentations/articles/040130-0001" />
        <rdf:li
rdf:resource="http://xmlfr.org/documentations/articles/031126-0001" />
      </rdf:Seq>
    </items>
  </channel>
```

29

RSS+TAXO

```
<item rdf:about="http://xmlfr.org/documentations/articles/040331-0001">
  <title>2004, l'annee RSS ?</title>
  <link>http://xmlfr.org/documentations/articles/040331-0001</link>
  <dc:description>Les deux dernieres annees furent certainement determinantes
en matiere de XML. Son adoption dans de nombreux contextes en a fait une
technologie pivot pour l'echange de donnees. Parallelement, le phenomene
weblog a connu un franc succes en 2003, et a permis de mettre en avant une
technologie XML existante depuis 1999 : RSS.</dc:description>
  <taxo:topics>
    <rdf:Bag>
      <rdf:li resource="http://xmlfr.org/index/object.title/rss/" />
      <rdf:li resource="http://xmlfr.org/index/object.title/xml/" />
      <rdf:li resource="http://xmlfr.org/index/object.title/weblog/" />
      <rdf:li resource="http://xmlfr.org/index/person/frederic+laurent/" />
      <rdf:li resource="http://xmlfr.org/index/object.title/blogger/" />
      <rdf:li resource="http://xmlfr.org/index/object.title/syndication/" />
      <rdf:li resource="http://xmlfr.org/index/object.title/rdf/" />
    </rdf:Bag>
  </taxo:topics>
  <dc:publisher>XMLfr</dc:publisher>
  <dc:type>text</dc:type>
  <dc:language>fr</dc:language>
</item>
```

30

Synthèse

- Il existe de nombreuses ressources en RDF;
- Elles sont utilisées conjointement.

Ces ressources constituent un embryon de web sémantique...

Modéliser le domaine

Utiliser d'autres ressources

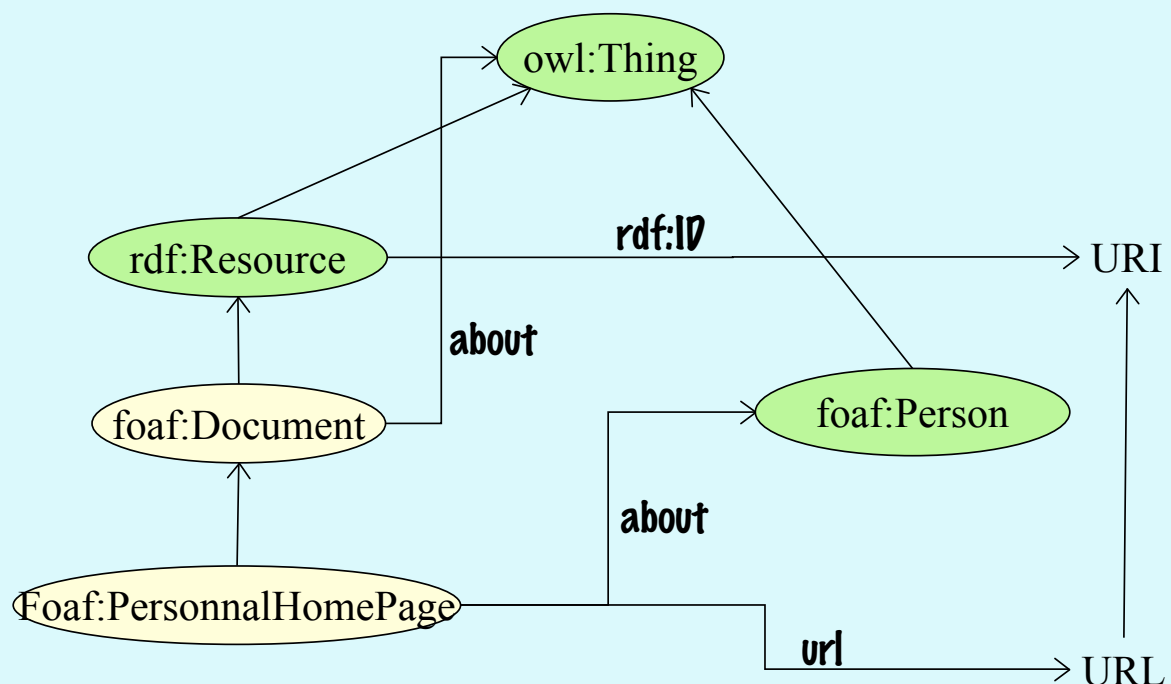
Quels besoins?

- Liens vers les documents en ligne;
- Décrire les méta-données (DublinCore);
- Connaissance des droits (Creative Commons);
- Liens vers les personnes et les organisations (FOAF);
- Référence à des évènements (iCalendar);
- Engendrer des fils (RSS).

Bénéfice: acquérir les informations disponibles ailleurs.

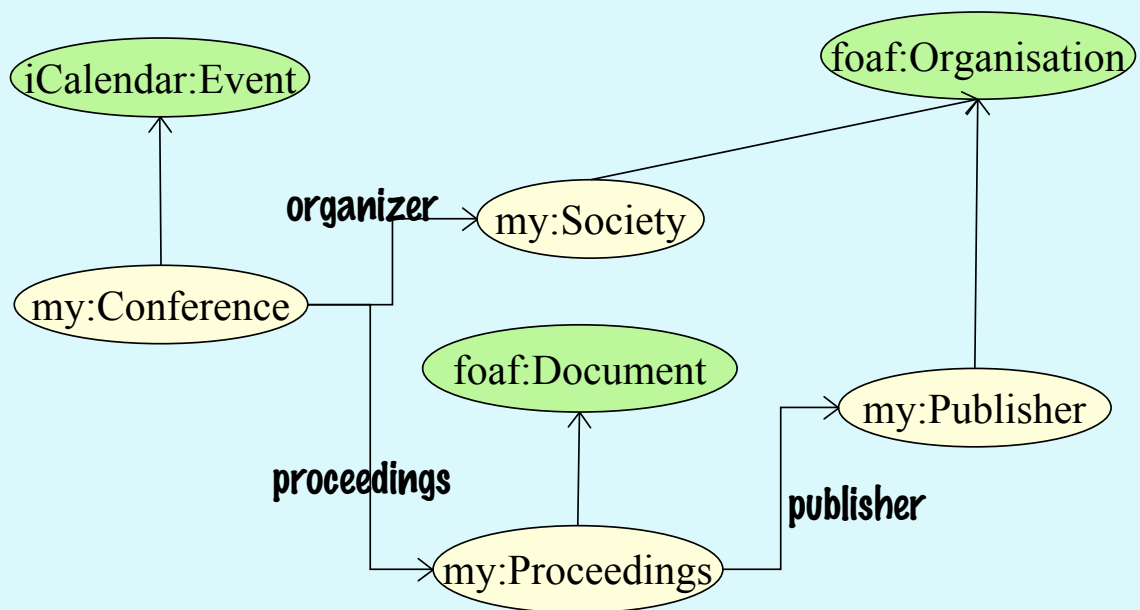
33

Utiliser d'autres ressources (étendre)



34

Utiliser d'autres ressources (étendre)

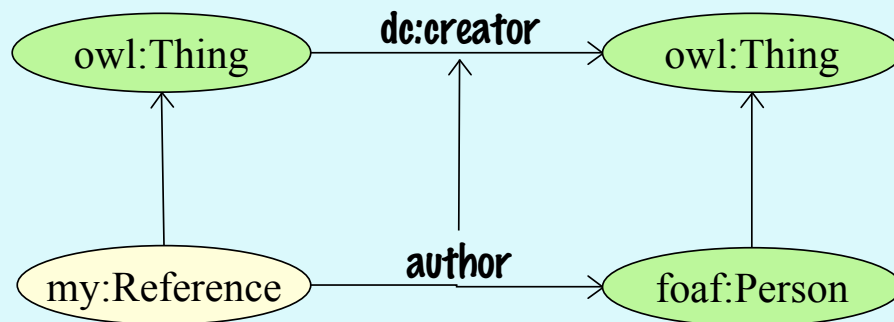


35

Je dois ajouter une extension par le haut

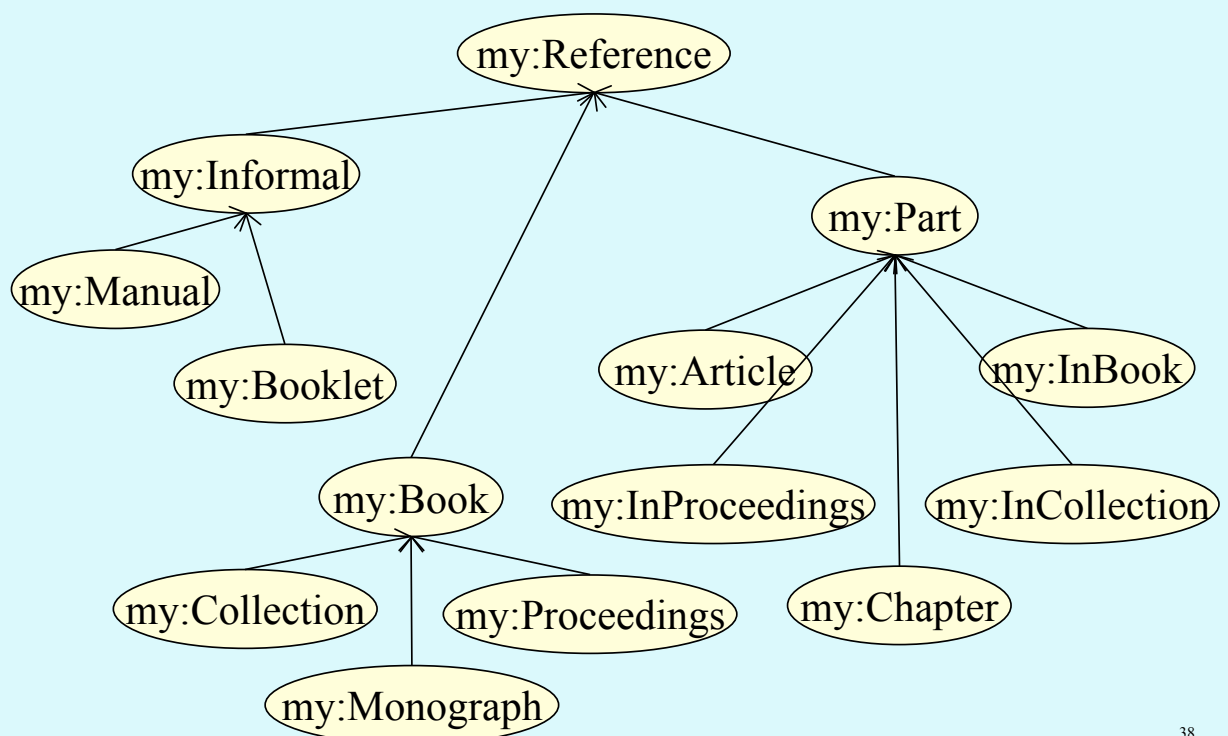
36

Utiliser d'autres ressources (étendre)



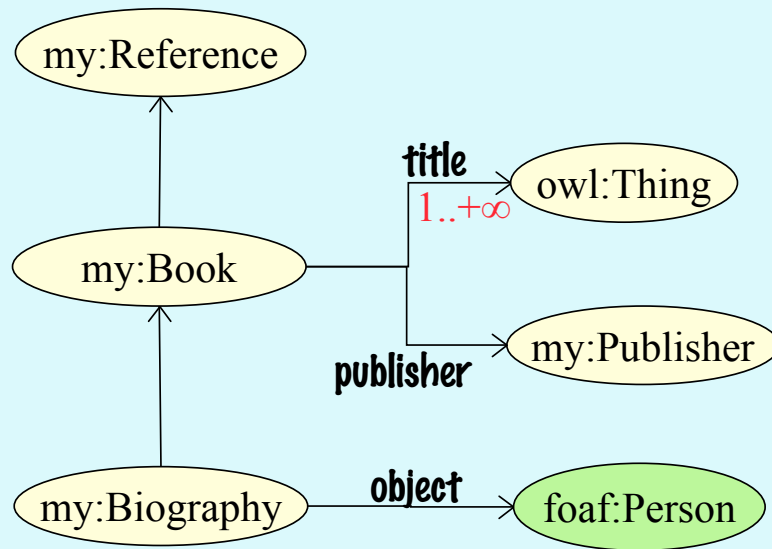
37

Créer ses propres ressources

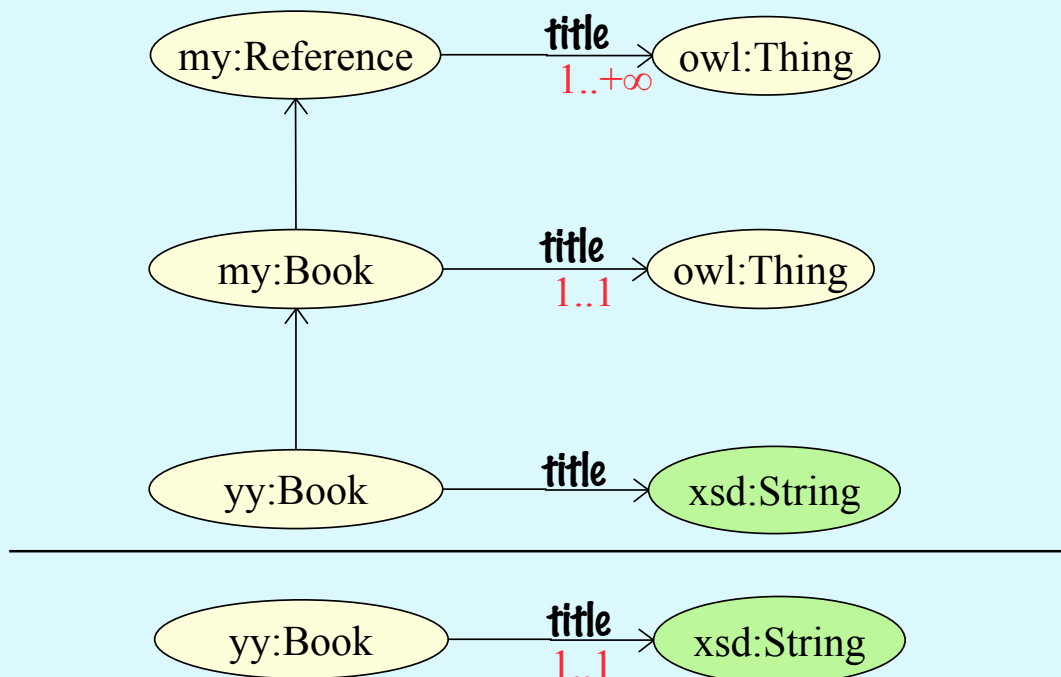


38

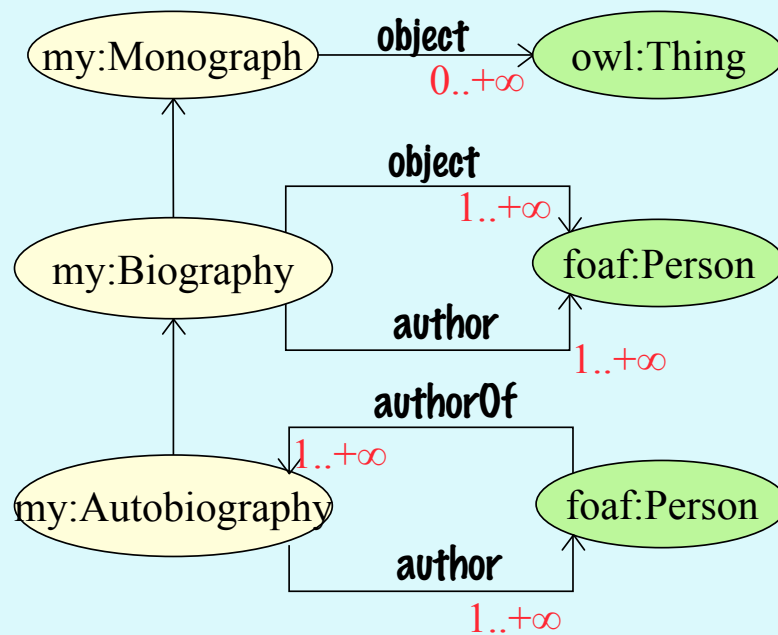
Ajouter des contraintes



Ajouter des contraintes (2)



Ajouter des contraintes (limitation)



41

RDF Schéma



- Introduction de classes et de propriétés;
- Spécialisation de classes et de propriétés;
- Restriction du domaine et codomaine des propriétés.

+

- Méta-modèle réflexif (tout est ressource, les classes et propriétés spécialisent les ressources, ressources et propriétés sont des classes...).

42

Vocabulaire réservé

`rdfs:Resource`, `rdfs:subClassOf`, `rdfs:Literal`, `rdfs:Datatype`,
`rdfs:Class`, `rdfs:subPropertyOf`

`rdfs:domain`, `rdfs:range`

`rdfs:member`, `rdfs:Container`, `rdfs:ContainerMembershipProperty`

`rdfs:comment`, `rdfs:seeAlso`, `rdfs:isDefinedBy`, `rdfs:label`

43

OWL: constructeurs



Hiérarchie de classes et de propriétés

typant nœuds et arcs RDF;

Construction de classes par:

constructeurs ensemblistes (union, intersection);

contraintes sur les attributs (cardinalité, type);

Contraintes globales (exclusivité).

44

OWL (vocabulaire réservé)

owl:Class, owl:DatatypeProperty, owl:ObjectProperty
owl:complementOf, owl:unionOf, owl:intersectionOf
owl:oneOf

owl:Restriction, owl:allValuesFrom, owl:someValueFrom,
owl:hasValue, owl:value, owl:maxCardinality,
owl:minCardinality

owl:disjointWith, rdfs:subClassOf, rdfs:subPropertyOf,
owl:equivalentProperty, owl:sameAs, owl:differentFrom
owl:TransitiveProperty, owl:FunctionalProperty...

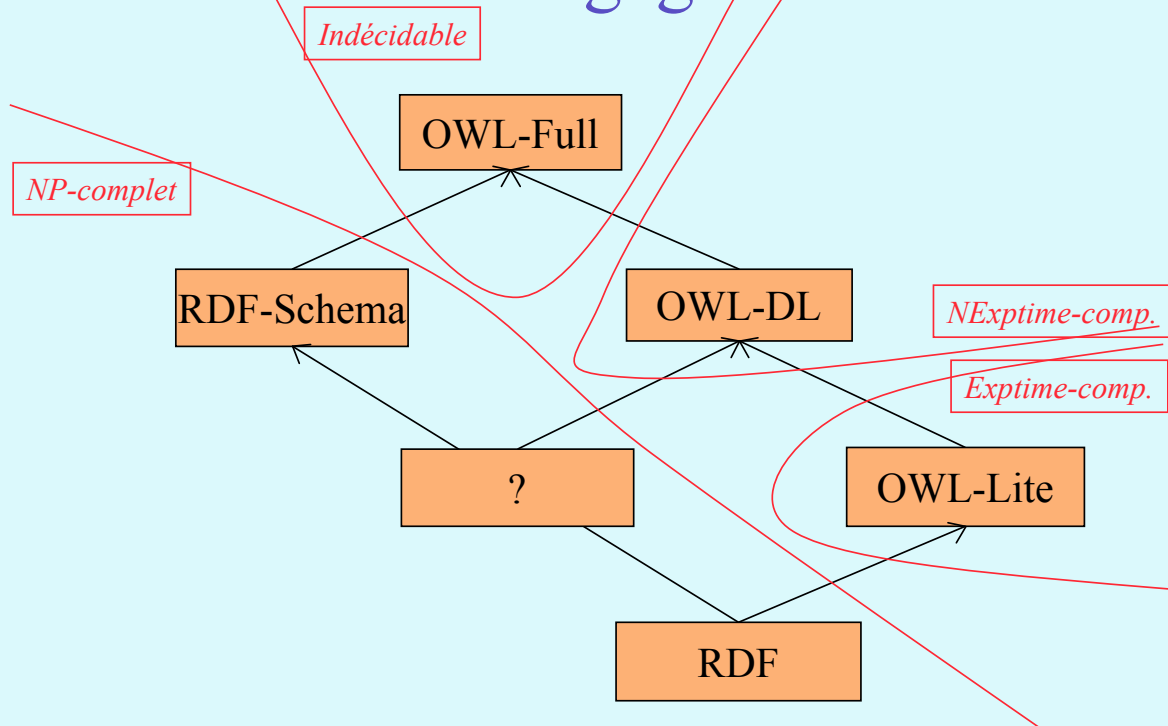
45

OWL (syntaxe)

```
<owl:Class rdf:ID="Book">  
  <owl:intersectionOf>  
    <owl:Class rdf:resource="#Reference" />  
    <owl:Restriction>  
      <owl:onProperty rdf:resource="#title" />  
      <owl:minCardinality>1</owl:minCardinality>  
    </owl:Restriction>  
    <owl:Restriction>  
      <owl:onProperty rdf:resource="#publisher" />  
      <owl:allValuesFrom rdf:resource="#Publisher" />  
    </owl:Restriction>  
  </owl:intersectionOf>  
</owl:Class>
```

46

Des langages...



47

Synthèse

Décrire ses données c'est:

- D'abord tirer parti des autres sources;
- Les étendre;
- Les restreindre;
- Les rendre disponibles sur le web!

48

Exprimer ses données

49

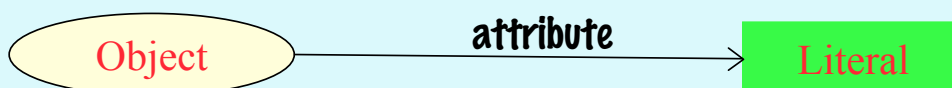
RDF (resource description framework)



Conçu pour annoter les documents XML.

Fondé sur un modèle de multi-graphe étiqueté :

Objet -- attribut -> Objet|Littéral

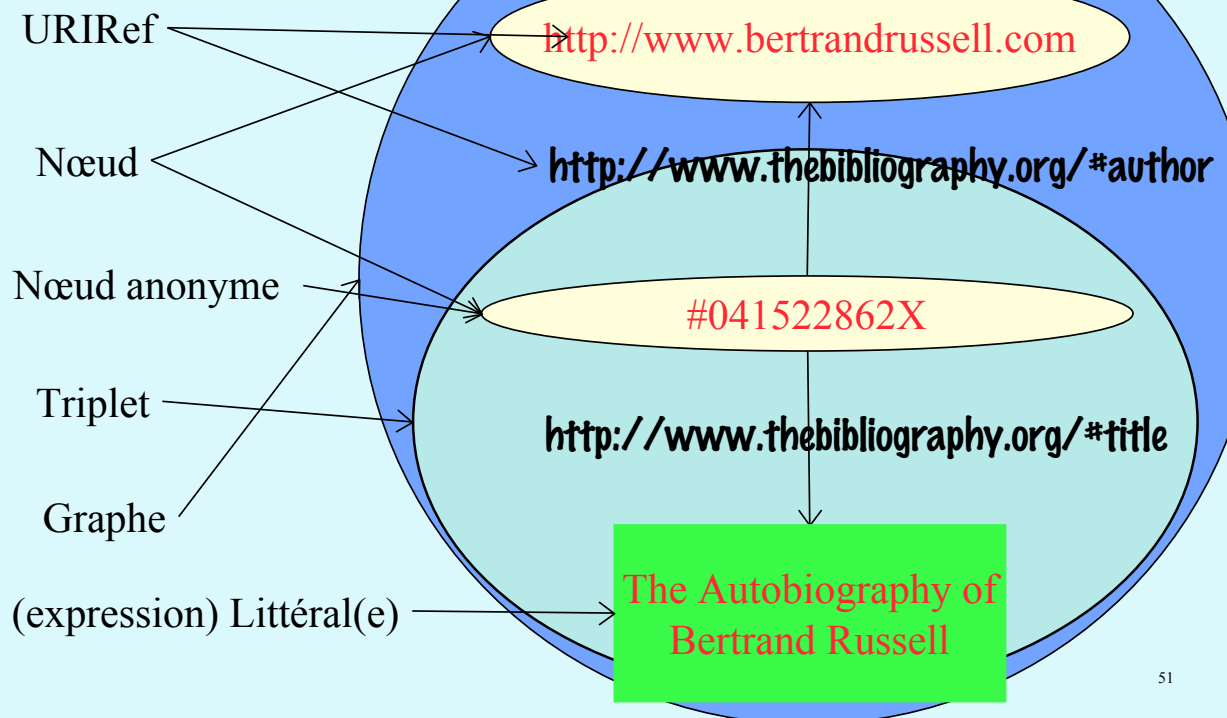


++ collections, “réification”...

A commencé “quick & dirty”. A récemment été doté d’une sémantique en théorie des modèles.

50

RDF: éléments

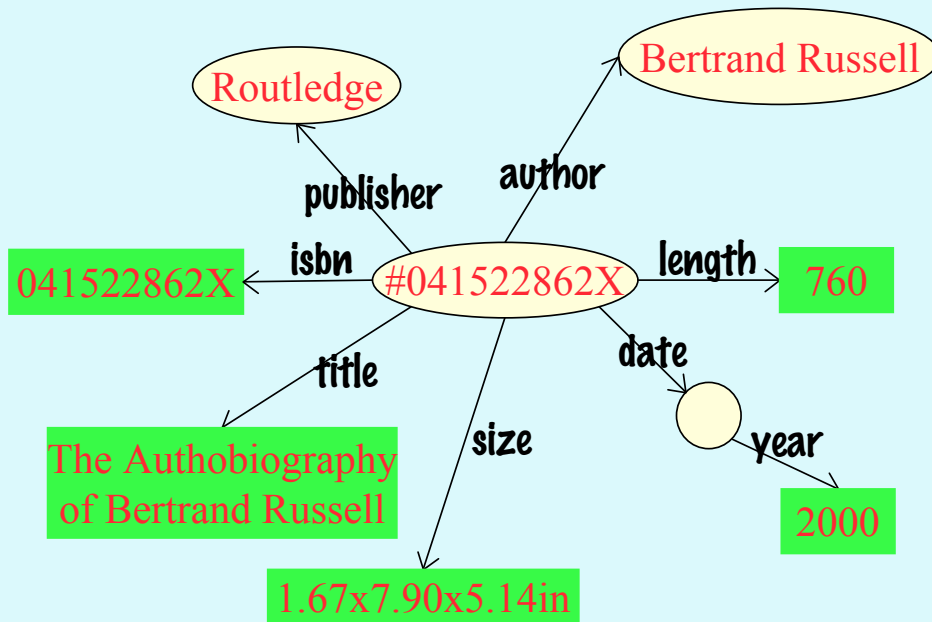


RDF (figure 1)

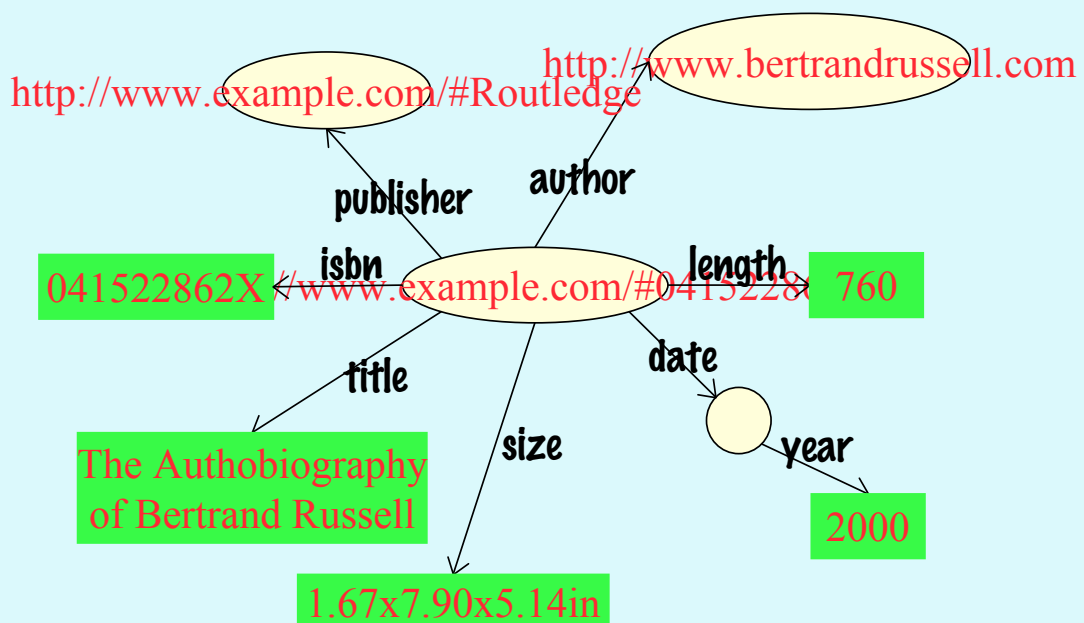
```
<Description rdf:ID="041522862X">
  <rdf:type rdf:resource="&bib;Autobiography"/>
  <bib:author>
    <foaf:Person
      rdf:about="http://www.bertrandrussell.com"/>
  </bib:author>
  <bib:title>The      Autobiography      of      Bertrand
  Russell</bib:title>
  <bib:isbn>041522862X</bib:isbn>
  <bib:length>760</bib:length>
  <bib:size>1.67x7.90x5.14in</bib:size>
  <bib:publisher>
    <Publisher rdf:resource="#Routledge"/>
  </bib:publisher>
  <bib:date>
    <Date year="2000"/>
  </bib:date>
</Description>
```

52

RDF (figure 1)



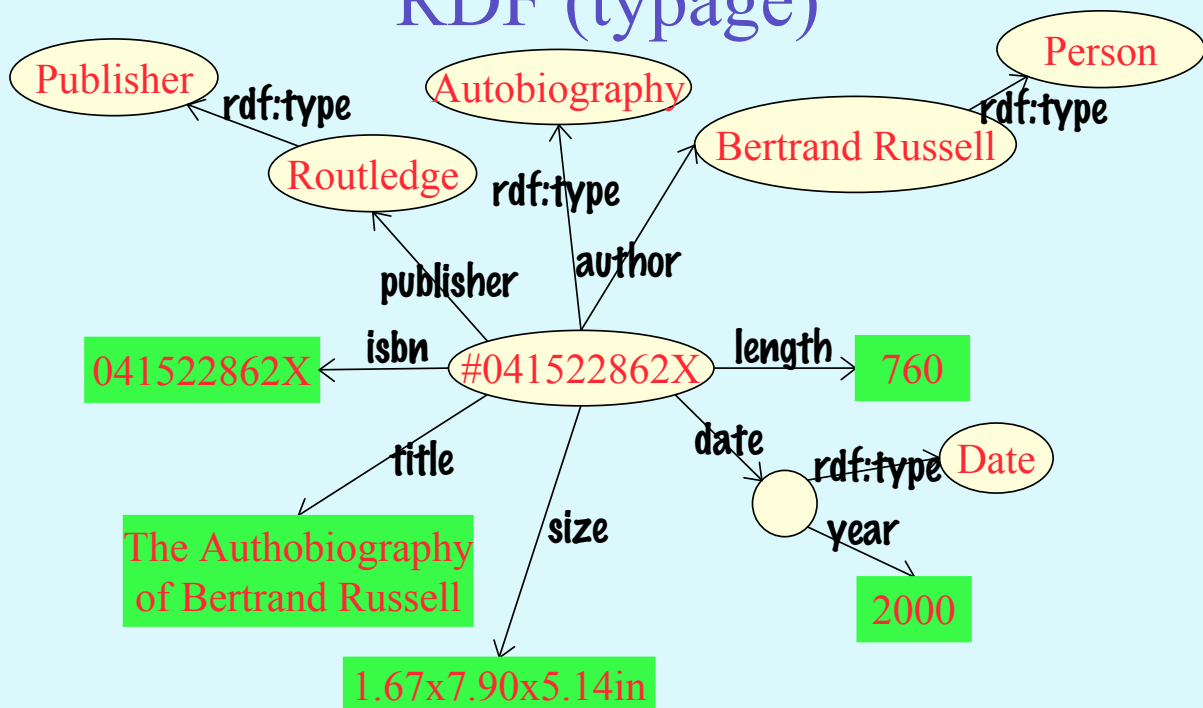
RDF (moins simple)



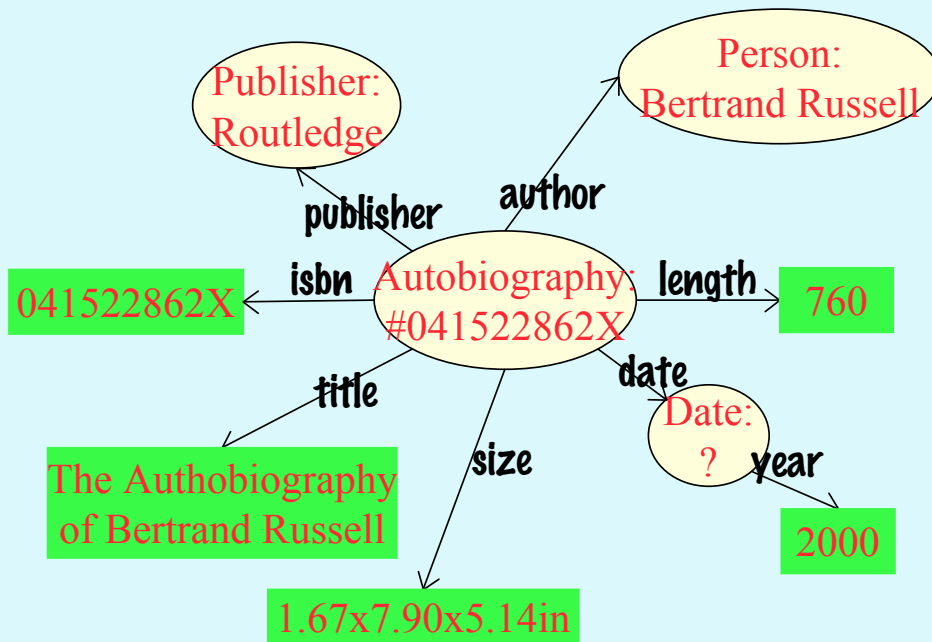
RDF (plus complexe)



RDF (typage)

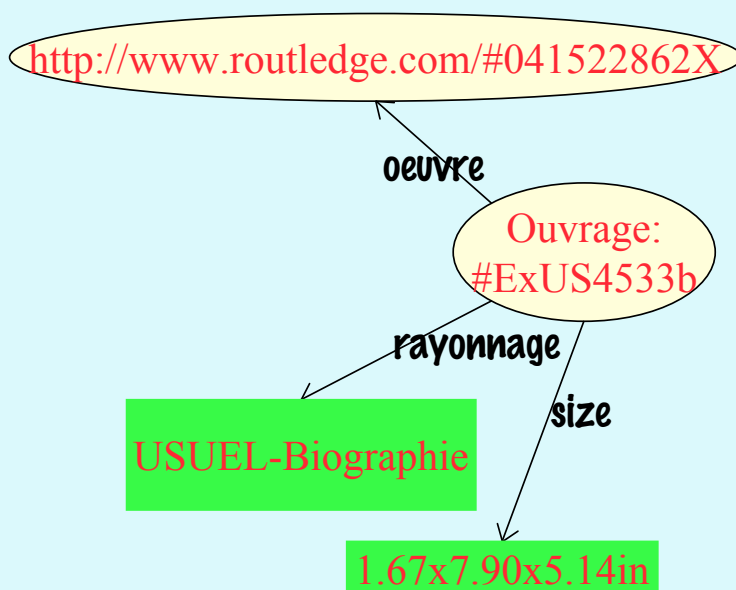


RDF (simplifiée)



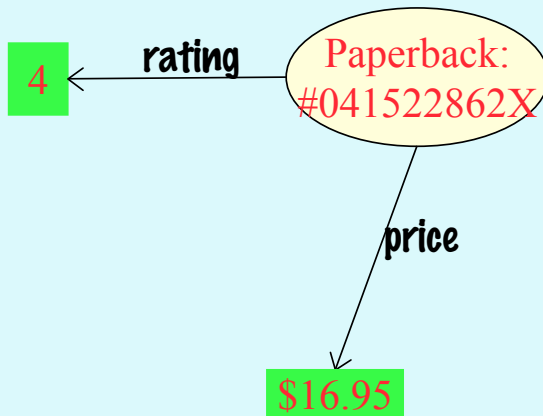
57

RDF (figure 2)



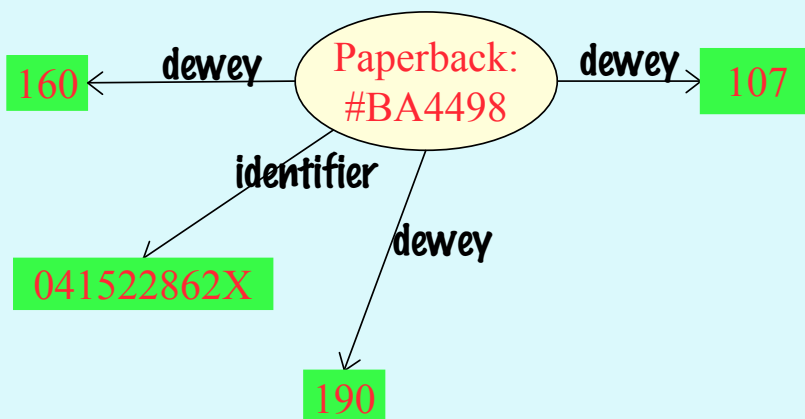
58

RDF (figure 3)



59

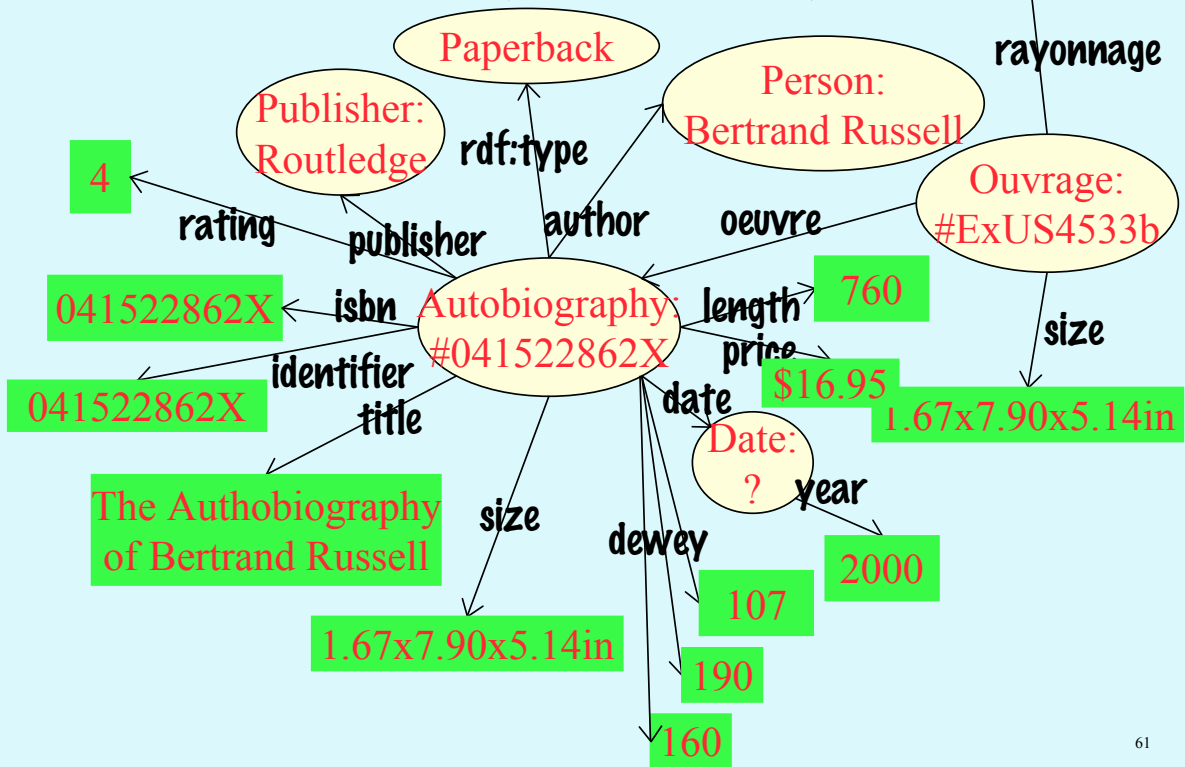
RDF (figure 4)



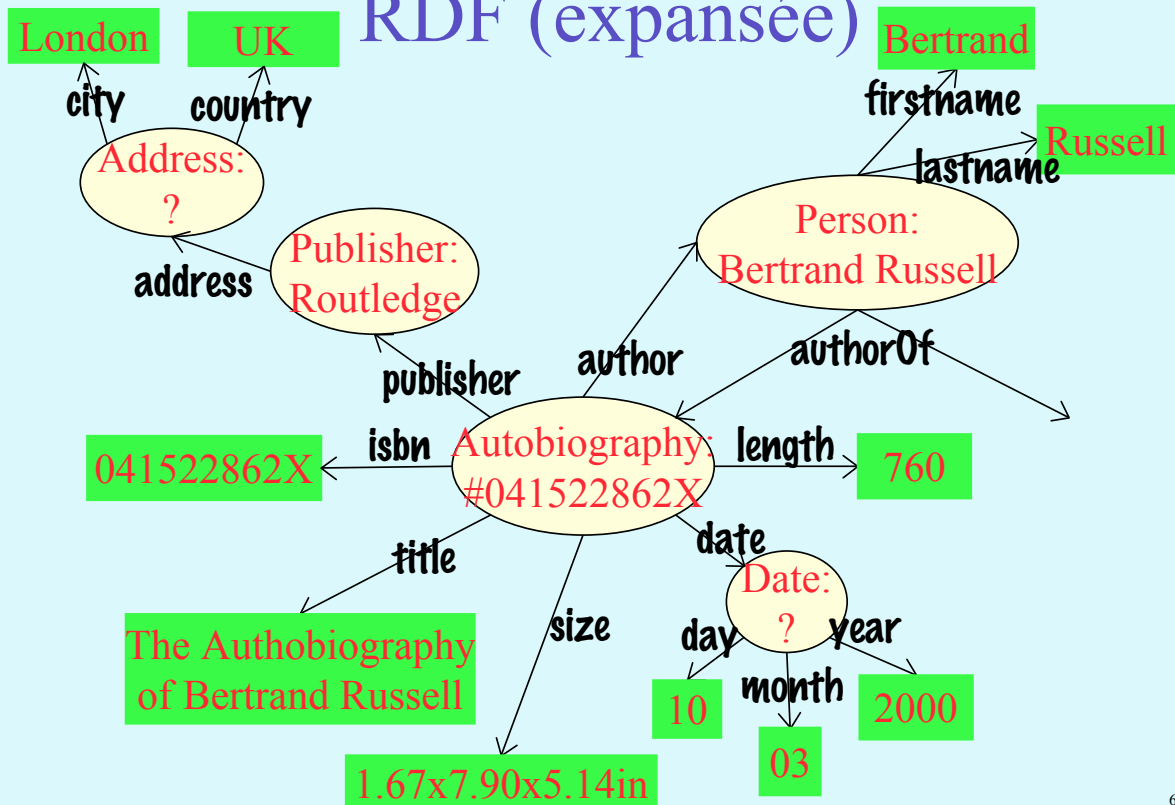
60

RDF (fusionnée)

USUEL-Biographie



RDF (expansée)



RDF (figure 2, revue)

```
<rss:item rdf:ID="041522862X">
  <rdf:type rdf:resource="&bib;Autobiography" />
  <rdf:type rdf:resource="&cc;Work" />
  <bib:author>
    <foaf:Person rdf:about="http://www.bertrandrussell.com" />
  </bib:author>
  <rss:title>The Autobiography of Bertrand Russell</rss:title>
  <dc:title>The Autobiography of Bertrand Russell</dc:title>
  <rss:link>http://www.bertrandrussell.com/041522862X</rss:link>
  <cc:license
    rdf:resource="http://creativecommons.org/licenses/by-nc-nd/2.0/" />
  <rss:description>The autobiography of Bertrand Russell is now
  available for download under a Creative Commons
  licence!</rss:description>
  <bib:publisher>
    <Publisher rdf:resource="#Routledge" />
  </bib:publisher>
</rss:item>
```

63

Synthèse

Les descriptions sont distribuées:

- Elles sont identifiées grace aux URI;
- Elles peuvent toujours être étendues;
- Elles peuvent être controlées par une ontologie.

64

Une application du web sémantique

- Consomme du RDF

Ex. les informations de l'éditeur, celles sur l'auteur, celles sur les bibliothèques, sur les thématiques, les droits associés au document...

- Produit du RDF

Ex. le classement du document, sa disponibilité...

65

Différences avec XML

Sémantique formelle

Ouverture:

- on peut *toujours* ajouter des attributs aux objets, on peut *toujours* ajouter des sous-classes.

Ex. l'ontologie the-bibliography utilise celle de iCalendar et foaf.

Ex. 041522862X caractérisé ici comme une autobiographie, peut l'être ailleurs comme un "paperback".

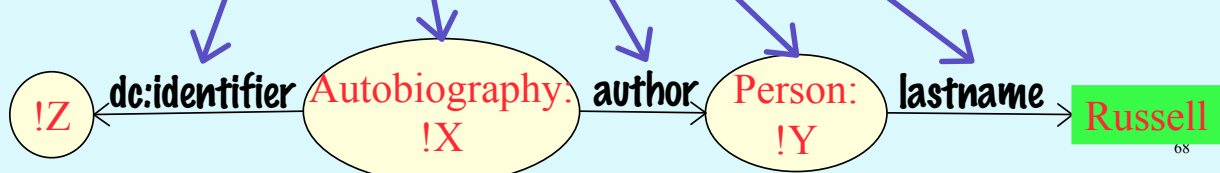
66

Utiliser les descriptions

67

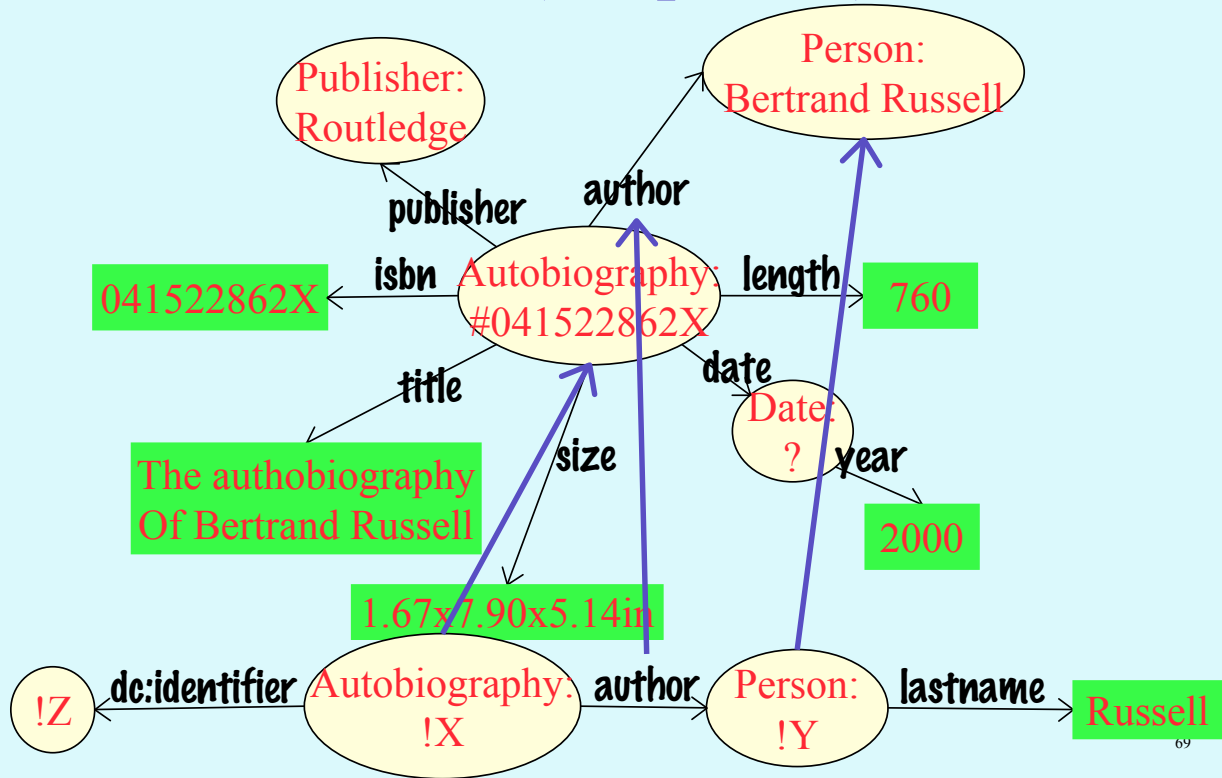
Requête (RDQL)

```
SELECT ?x, ?y, ?z
WHERE (?x, <rdf:type>, <bib:Autobiography>),
      (?x, <bib:author>, ?y),
      (?y, <rdf:type>, <foaf:Person>),
      (?y, <foaf:lastname>, Russell),
      (?x, <dc:identifier>, ?z)
USING rdf FOR <http://www.w3.org/1999/02/22-rdf-syntax-ns#>,
      bib FOR <http://www.thebibliography.org/>,
      foaf FOR <http://xmlns.com/foaf/0.1/>,
      dc FOR <http://dublincore.org/2001/08/14/dces#>
```

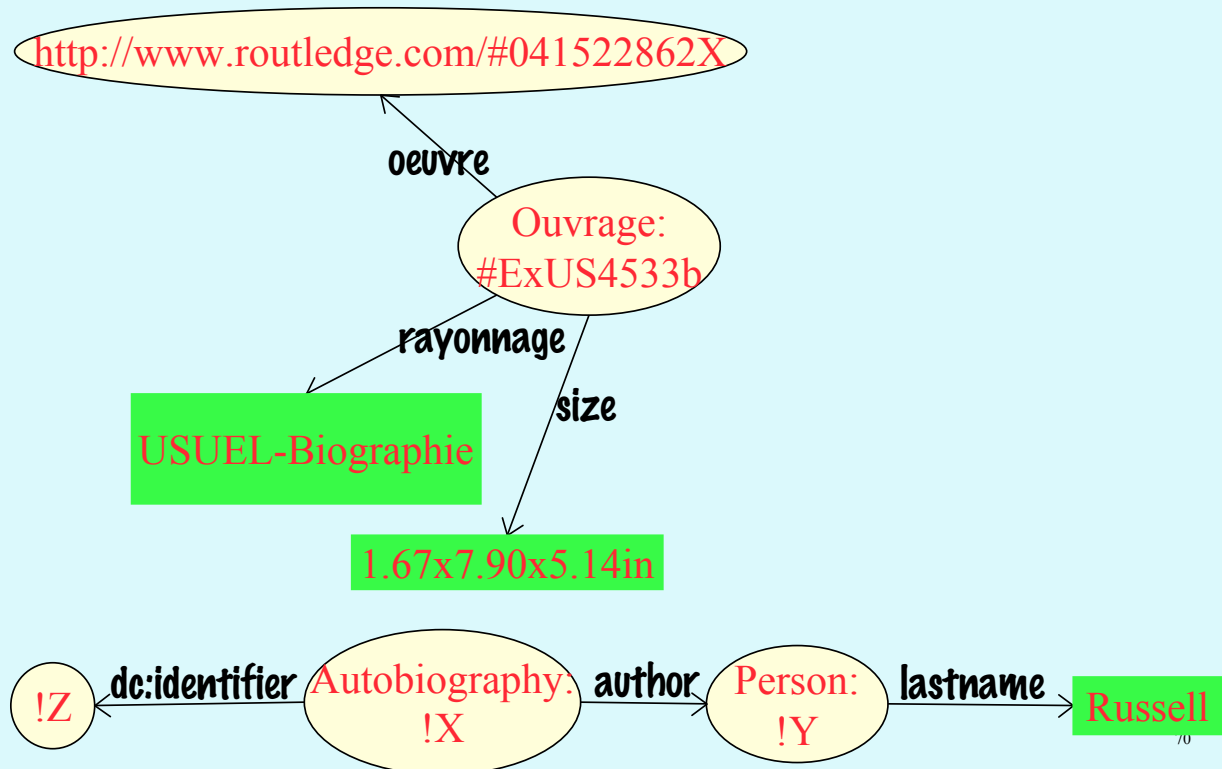


68

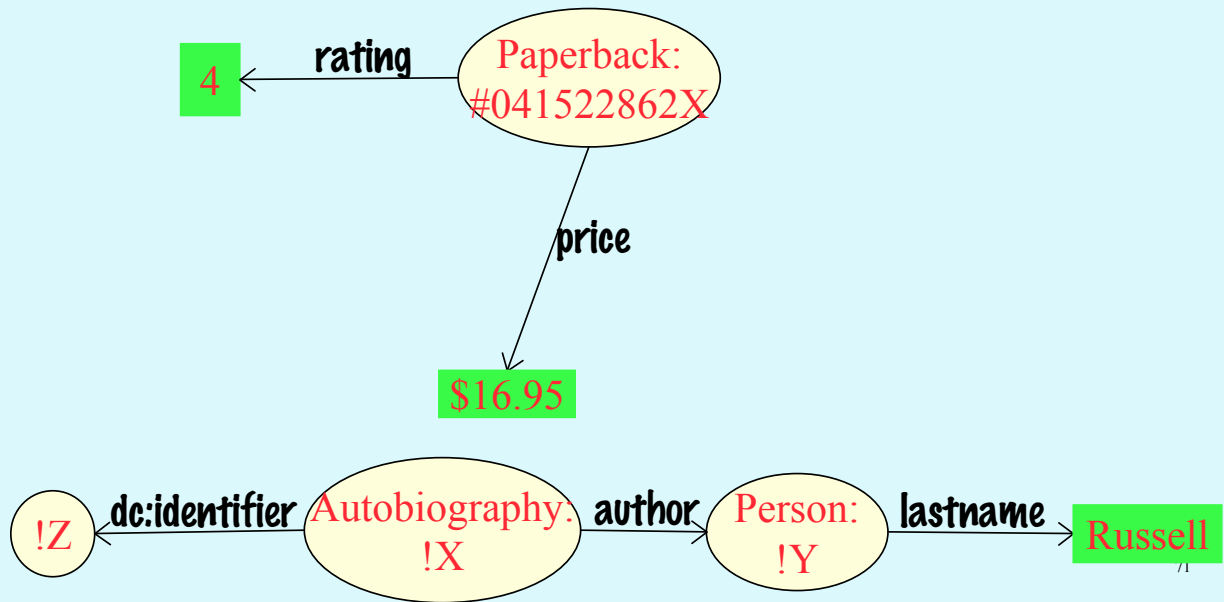
RDF (simplifiée)



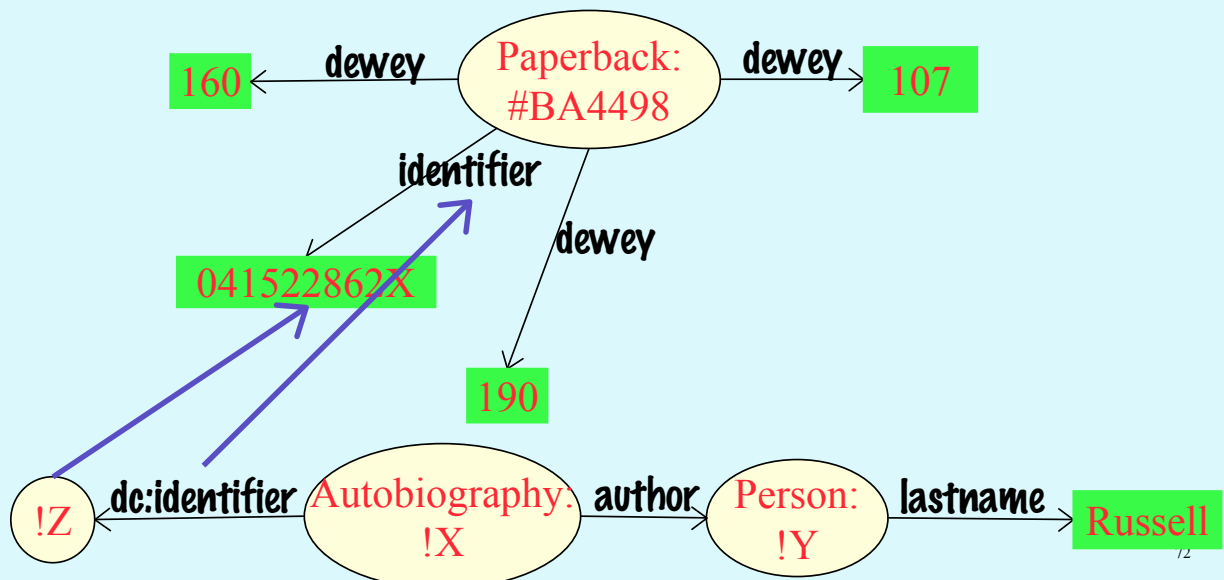
RDF (figure 3)



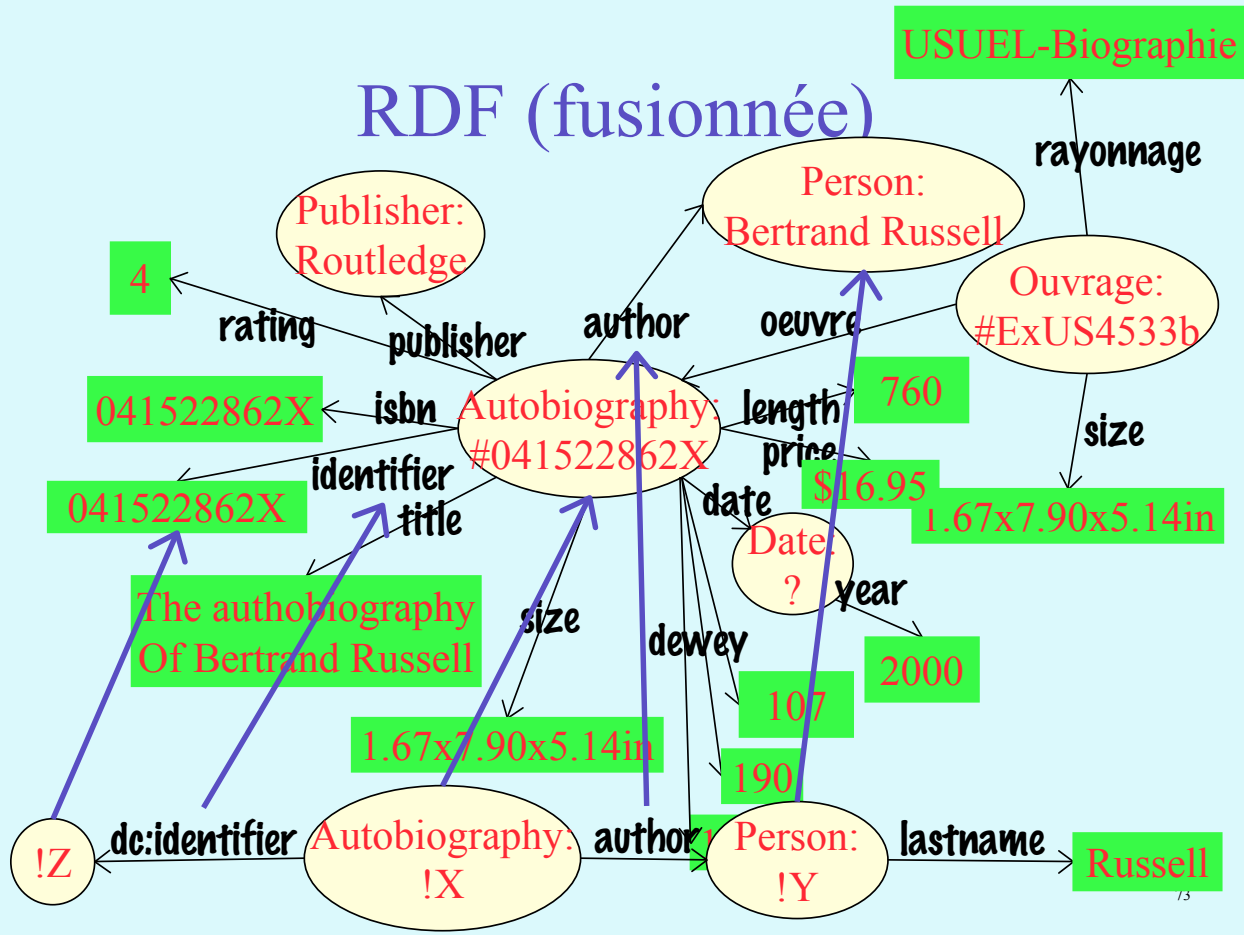
RDF (figure3)



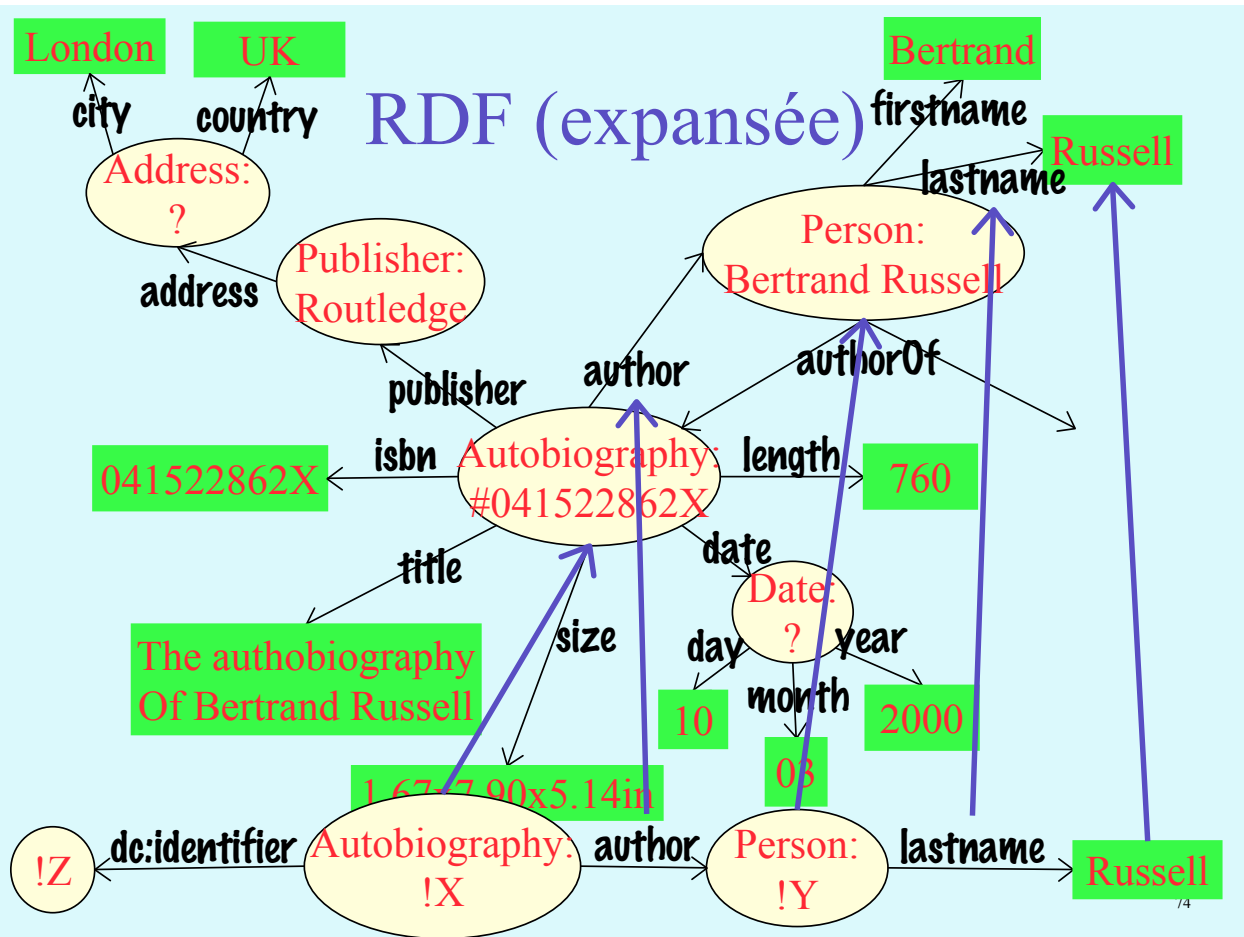
RDF (figure3)

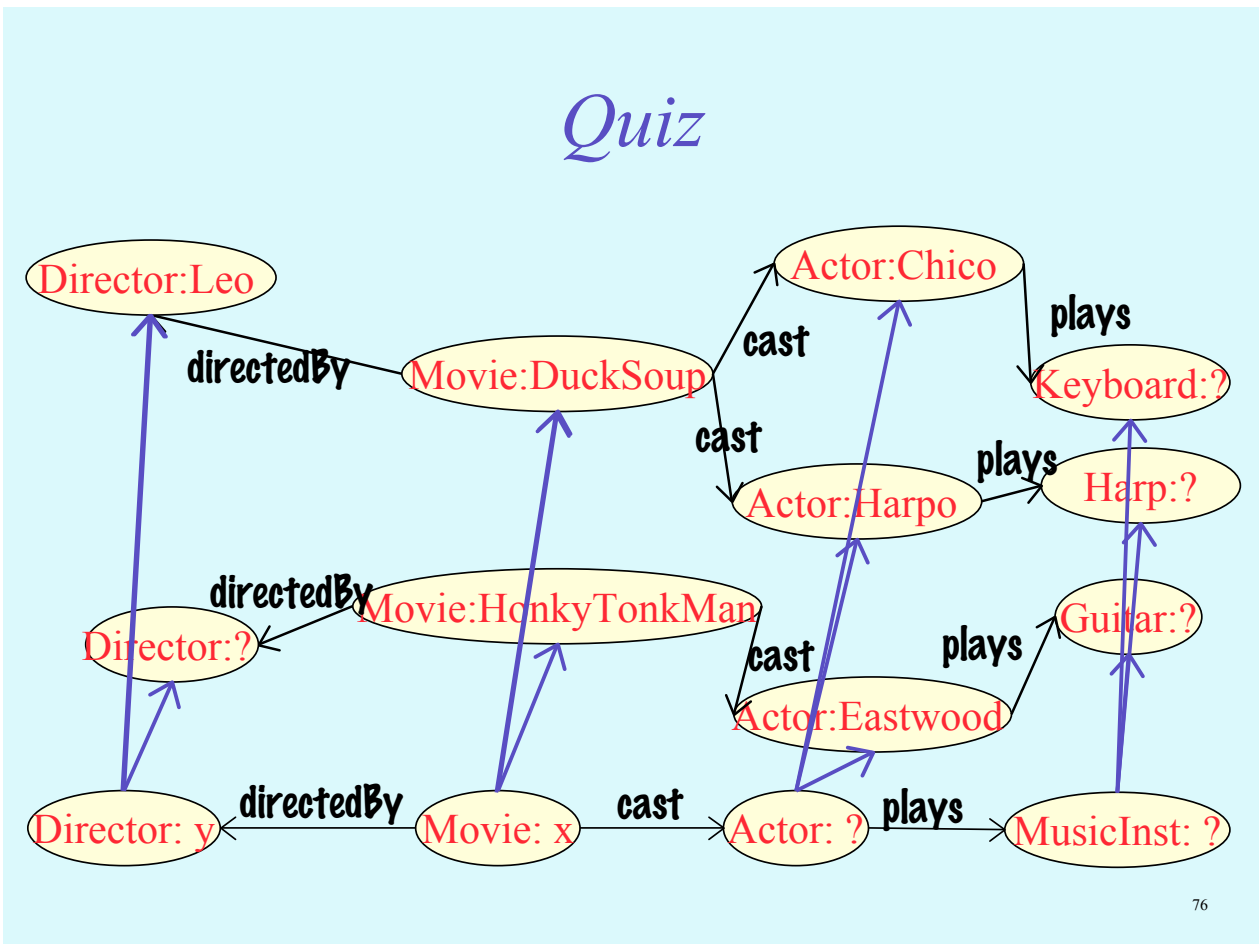
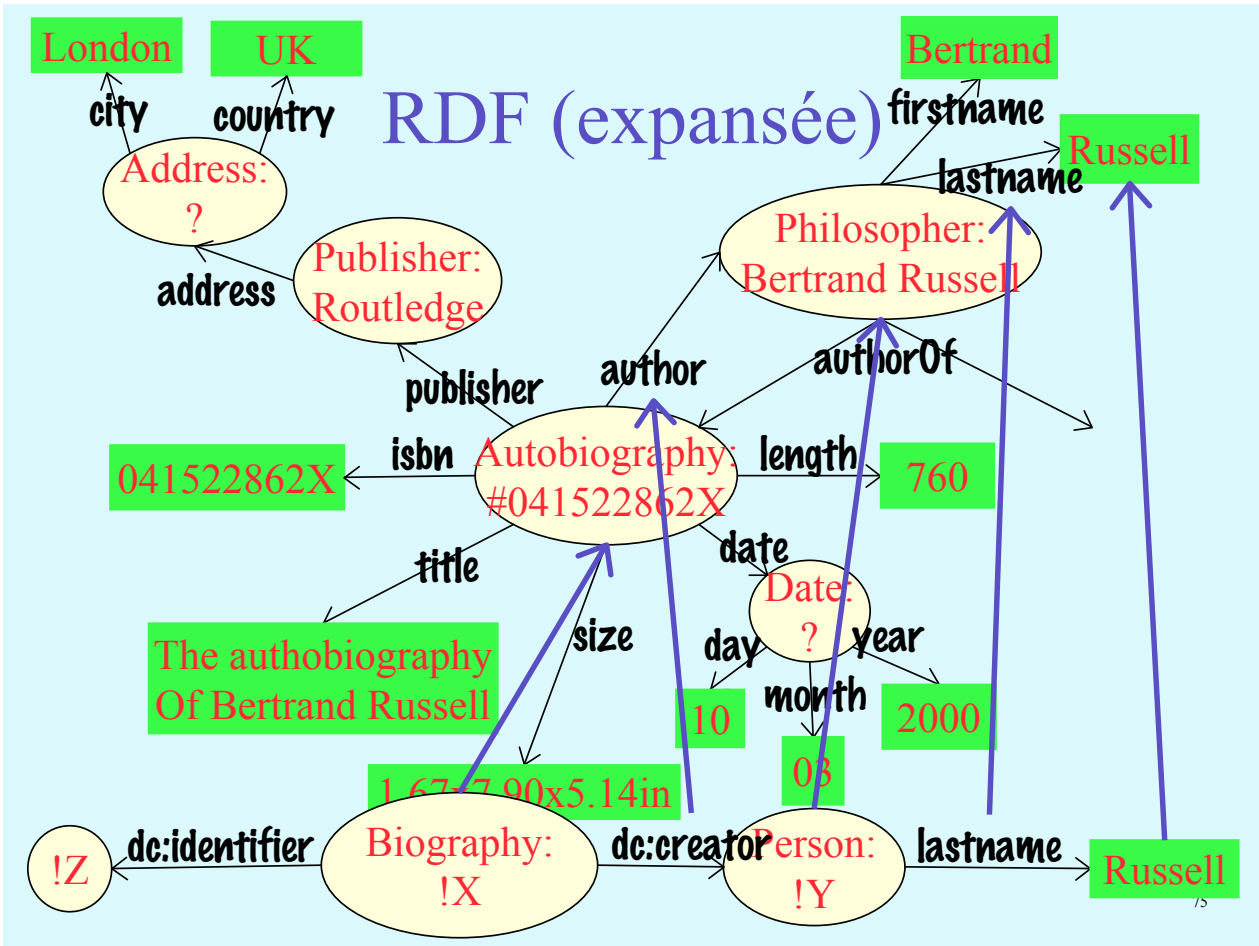


RDF (fusionnée)



RDF (expansée)





Synthèse

Les requêtes peuvent être vues comme des graphes!

- Répondre à une requête peut prendre en compte toutes les descriptions disponibles;
- Répondre à une requête utilise les ontologies pour compléter la requête.

Adresses pratiques et conclusions

Conclusion

À l'instar du web, le web sémantique sera réalisé par ses utilisateurs.

Il reste beaucoup de développements à réaliser.

Mais les premières pierres sont disponibles et seront utilisables facilement dans quelques temps.

79

Adresses pratiques

W3C:

<http://www.w3.org/2001/sw/>

Trouver des documents RDF:

<http://pear.cs.umbc.edu/swoogle>

The semantic web, an introduction:

<http://infomesh.net/2001/swintro/>

RDF resource guide

<http://www.ilrt.bris.ac.uk/discovery/rdf/resources/>

80

Swoogle

search and metadata for the semantic web

Navigation

- [Home](#)
- [Documentation](#)
- [Ontology Dictionary](#)
- [Swoogle Statistics](#)
- [Ontology Submission](#)
- [Topics](#)
- [Visitor Statistics](#)
- [Feedbacks](#)
- [Surveys](#)
- [Top 10 of this site](#)
- [Your Account](#)

Document Search:

[advanced](#)

1 - 1 of total 1 results for **euzenat** in 0.9691 seconds

<http://www.inrialpes.fr/exmo/people/euzenat/euzenat.rdf>

Suffix: rdf Encoding: RDF/XML Last modified: 2004-02-20 08:49:39

Classes defined: 0 Properties defined: 0 Instances defined: 10


Triples: 57 Namespaces used: 13 Ontology Ratio: 0

Cached: [Original File](#) [N-Triples](#)

Swoogle view: [Document Properties](#) [Term Properties](#)

[Swoogle Search](#) ©

Swoogle Today

2004-09-28 

total files	560252 (100.00%)
-------------	---------------------

81

?

<http://www.inrialpes.fr/exmo/>

Jerome.Euzenat@inrialpes.fr

Raphael.Troncy@ina.fr