

RDF, SPARQL et Web sémantique

- Sémantique ?
- Vers un Web sémantique
- Dublin Core
- Microformats
- RDF
- Pile technologique
- Triplets RDF
- RDF/XML
- RDFS
- SPARQL
- Ontologies
- Linking Open Data
- Outils

- Qui a trait au sens

- Disciplines :

- IA, réseaux de neurones (boîte noire)
- Bases de connaissance
- Web sémantique

- Représentation des connaissances
- Relations entre les informations
- Capitalisation du savoir
- => Inférences

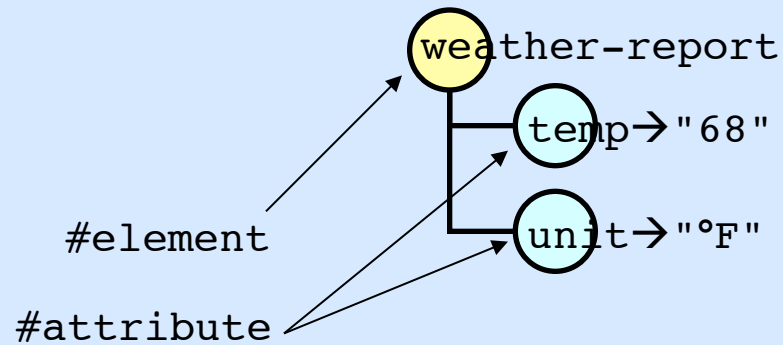
Model 0: vue binaire : 3C 3F 78 6D 6C 20 76 65 72 73 69 6F 6E 3D 22 31
2E 30 22 3F 3E 3C 77 65 61 74 68 65 72 2D 72 65 70 6F 72 74 20
74 65 6D 70 3D 22 36 38 22 20 75 6E 69 74 3D 22 B0 46 22 3E

Model 1: vue chaîne de caractère :

```
<?xml version="1.0"?><weather-report temp="68" unit="°F">
```

Model 0: vue binaire : 3C 3F 78 6D 6C 20 76 65 72 73 69 6F 6E 3D 22 31
2E 30 22 3F 3E 3C 77 65 61 74 68 65 72 2D 72 65 70 6F 72 74 20
74 65 6D 70 3D 22 36 38 22 20 75 6E 69 74 3D 22 B0 46 22 3E

Model 2: vue XML :
(vue abstraite)



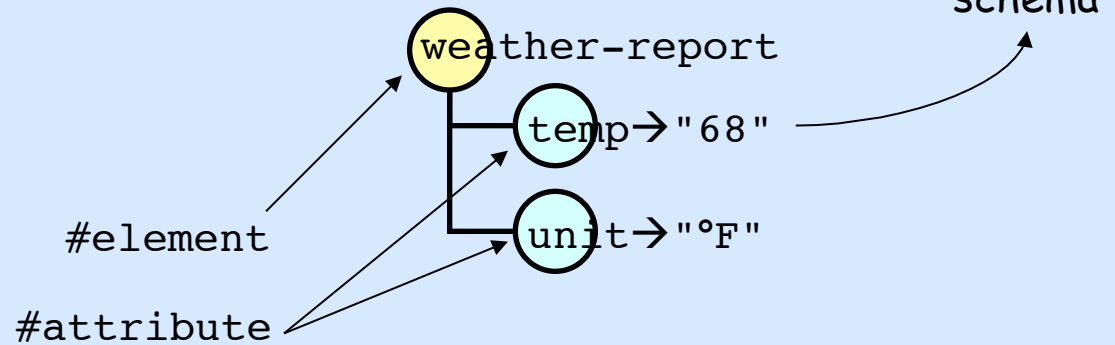
Model 1: vue chaîne de caractère :

```
<?xml version="1.0"?><weather-report temp="68" unit="°F">
```

Model 0: vue binaire : 3C 3F 78 6D 6C 20 76 65 72 73 69 6F 6E 3D 22 31
2E 30 22 3F 3E 3C 77 65 61 74 68 65 72 2D 72 65 70 6F 72 74 20
74 65 6D 70 3D 22 36 38 22 20 75 6E 69 74 3D 22 B0 46 22 3E

Model 3: vue de données typées : `#xs:decimal temp=68`
`#unit:temp unit=°C|°F`

Model 2: vue XML :
 (vue abstraite)

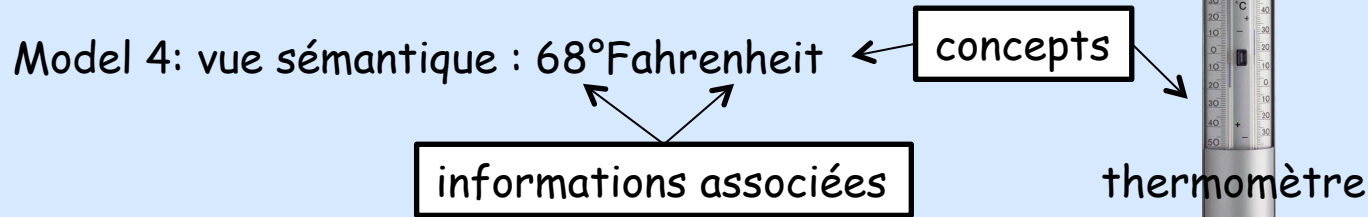


Model 1: vue chaîne de caractère :

`<?xml version="1.0"?><weather-report temp="68" unit="°F">`

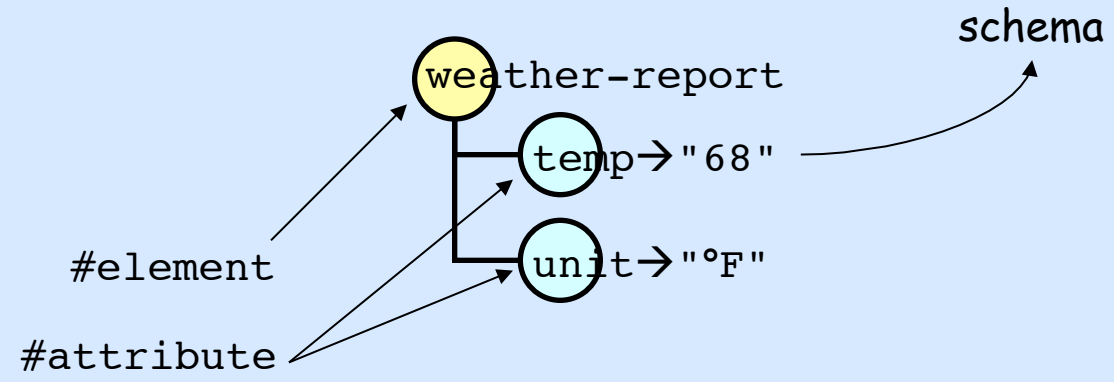
Model 0: vue binaire : 3C 3F 78 6D 6C 20 76 65 72 73 69 6F 6E 3D 22 31
 2E 30 22 3F 3E 3C 77 65 61 74 68 65 72 2D 72 65 70 6F 72 74 20
 74 65 6D 70 3D 22 36 38 22 20 75 6E 69 74 3D 22 B0 46 22 3E

Ontologie



Model 3: vue de données typées : #xs:decimal temp=68
#unit:temp unit=°C|°F

Model 2: vue XML :
(vue abstraite)



Model 1: vue chaîne de caractère :

```
<?xml version="1.0"?><weather-report temp="68" unit="°F">
```

Model 0: vue binaire : 3C 3F 78 6D 6C 20 76 65 72 73 69 6F 6E 3D 22 31
2E 30 22 3F 3E 3C 77 65 61 74 68 65 72 2D 72 65 70 6F 72 74 20
74 65 6D 70 3D 22 36 38 22 20 75 6E 69 74 3D 22 B0 46 22 3E

- Augmentation du niveau d'abstraction
- Augmentation du montant d'informations (schéma)
- Mise en relation des données
- Ajout de concepts associés

Métadonnées
Descriptions de données

Documentalistes



« J'ai fait un rêve pour le Web [dans lequel les ordinateurs] deviennent capables d'analyser toutes les données sur le Web — le contenu, liens, et les transactions entre les personnes et les ordinateurs. Un « Web Sémantique », qui devrait rendre cela possible, n'a pas encore émergé, mais quand ce sera fait, les mécanismes plan-plan d'échange, de bureaucratie et de nos vies quotidiennes seront traités par des machines dialoguant avec d'autres machines. Les « agents intelligents » qu'on nous promet depuis longtemps vont enfin se concrétiser. »

— Tim Berners-Lee, Weaving the Web

Web sémantique
Web 3.0

- « le Web sémantique fournit un framework qui permet aux données d'être partagées et réutilisées entres plusieurs applications, entreprises et groupes d'utilisateurs »
- Objectif : rendre les données du Web exploitable par des machines
 - Traiter la connaissance au lieu de traiter du texte
 - Inférer de nouvelles informations
- Repose sur le Web des données (structuré)
 - *Linked Data, Web of Data*
 - URI pour identifier les ressources, HTTP pour y accéder
- Technologies principales :
 - RDF
 - XML
 - OWL
- Autres initiatives :
 - Dublin Core (historique)
 - Microformats (voie détournée)

TITLE	Nom de la ressource
CREATOR	Personne ou organisation créatrice de la ressource
SUBJECT	Sujet de la ressource
DESCRIPTION	Description de la ressource (résumé)
PUBLISHER	L'éditeur de la ressource
CONTRIBUTOR	Personne ayant collaboré à la création de la ressource
DATE	Date de publication de la ressource
TYPE	Type de la ressource (page Web, poème, photographie...)
FORMAT	Format de la ressource (PDF, HTML, JPEG...)
IDENTIFIER	Identifiant unique de la ressource (ISBN, URI...)
SOURCE	Information indiquant que la ressource est dérivée d'une autre source
LANGUAGE	Code langue ISO 639 dans lequel la ressource est rédigée
RIGHTS	Note de droits de copie et de propriété intellectuelle

- Objectif : apport sémantique à HTML (et XHTML)
- Moyen : par les attributs standards HTML class, rel, rev

<http://microformats.org/>

```
Les oiseaux se sont perchés à
<span class="geo">
  <span class="latitude">52.48</span>,
  <span class="longitude">-1.89</span>
</span>
```

```
<div>
  <div>Philippe Poulard</div>
  <div>Inria</div>
  <div>+33 (0)4 92 38 77 77 </div>
  <a href="http://reflex.gforge.inria.fr">http://reflex.gforge.inria.fr</a>
</div>
```

hCard = HTML vCard (*Visit Card*)



```
<div class="vcard">
  <div class="fn">Philippe Poulard</div>
  <div class="org">Inria</div>
  <div class="tel">+33 (0)4 92 38 77 77 </div>
  <a class="url" href="http://reflex.gforge.inria.fr">http://reflex.gforge.inria.fr</a>
</div>
```

hAtom, hCalendar, hResume, xFolk, XHTML Friends Network (XFN), XOXO...

Web Sémantique

Resource Description Framework



Permet de décrire des dépendances entre des ressources

C' est un *framework* : peut être décliné de différentes façons selon les domaines d' application

Méta-données : ce ne sont pas les données qui relient les différentes ressources, mais les informations qui décrivent ces données

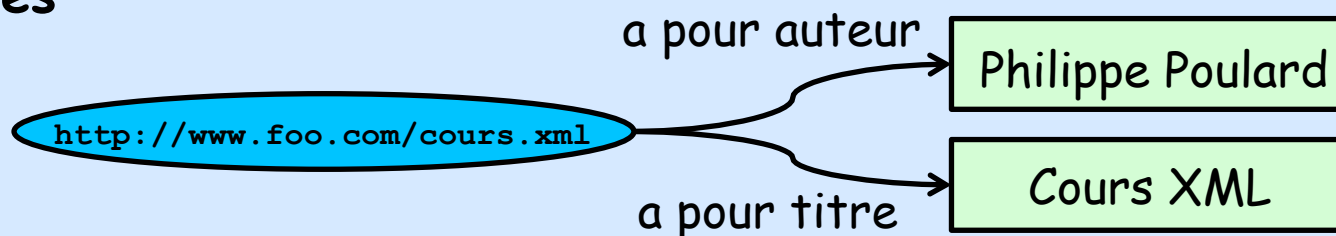
Dublin-core : vocabulaire RDF qui définit les méta-données sur le contenu des documents du Web

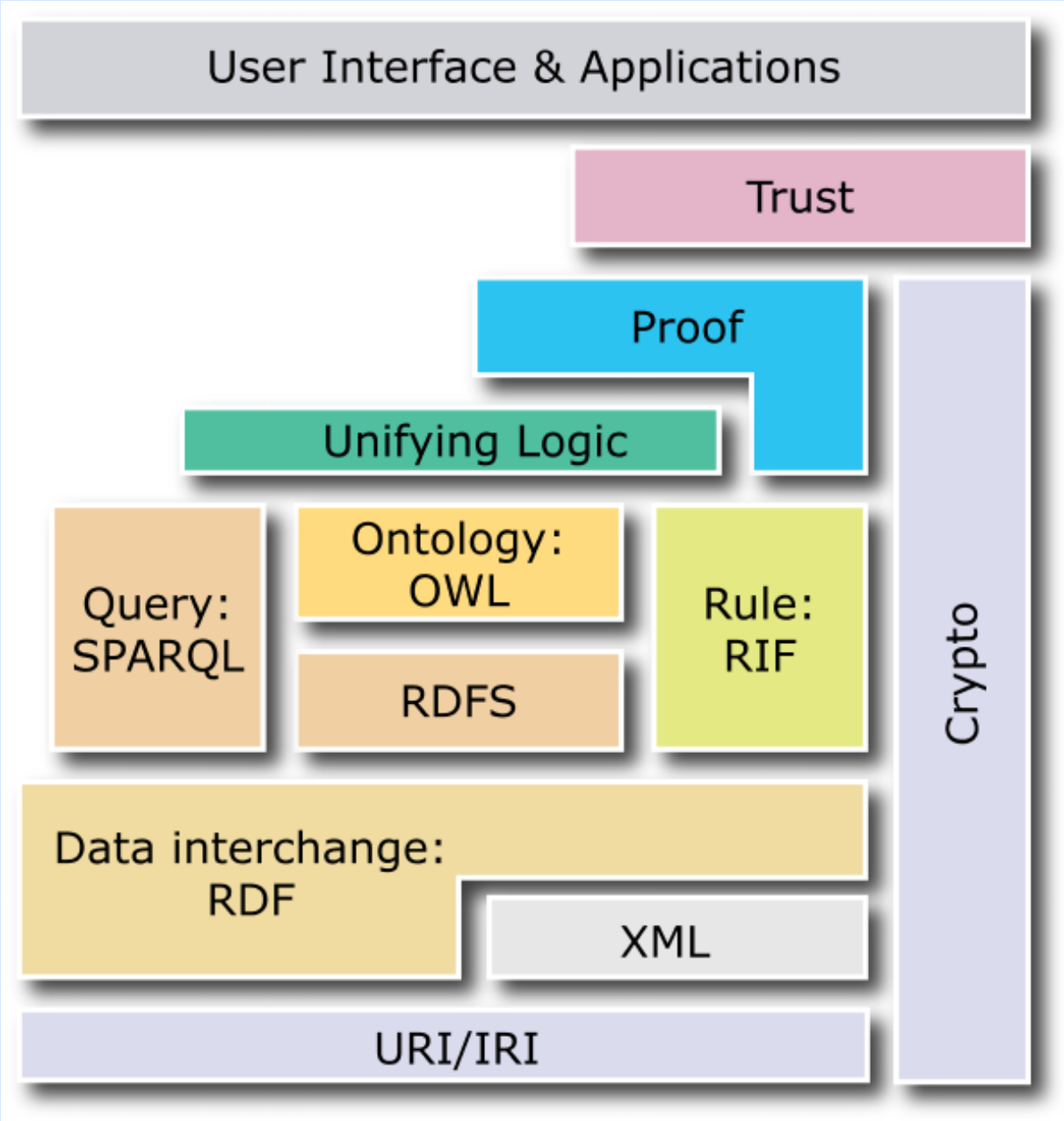
Un vocabulaire RDF définit un ensemble d' éléments et le contenu pour les méta-données de chaque domaine.

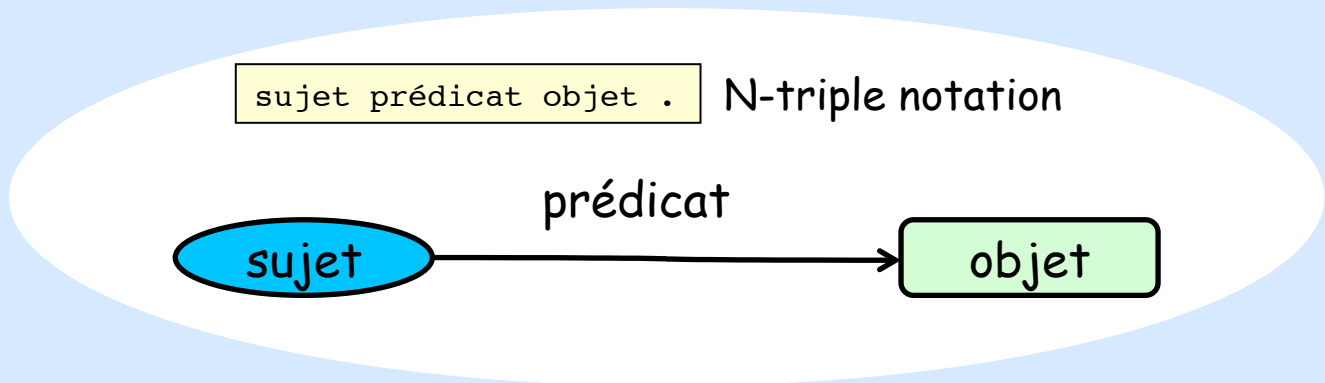


Ontologies

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/DC/">
  <rdf:Description about="http://www.foo.com/cours.xml">
    <dc:CREATOR>Philippe Poulard</dc:CREATOR>
    <dc:TITLE>Cours XML</dc:TITLE>
  </rdf:Description>
</rdf:RDF>
```



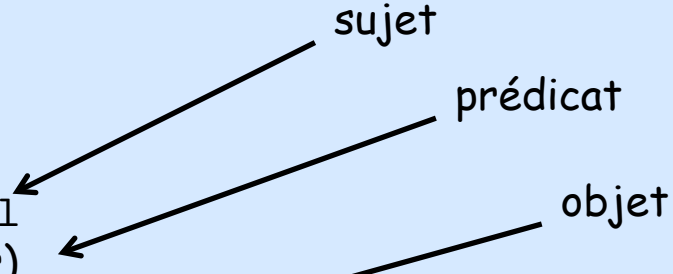




sujet prédictat objet . N-triple notation

Déclaration sur une ressource :

`http://example.org/index.html`
 a un créateur (creator)
 dont la valeur est "Philippe Poulard"



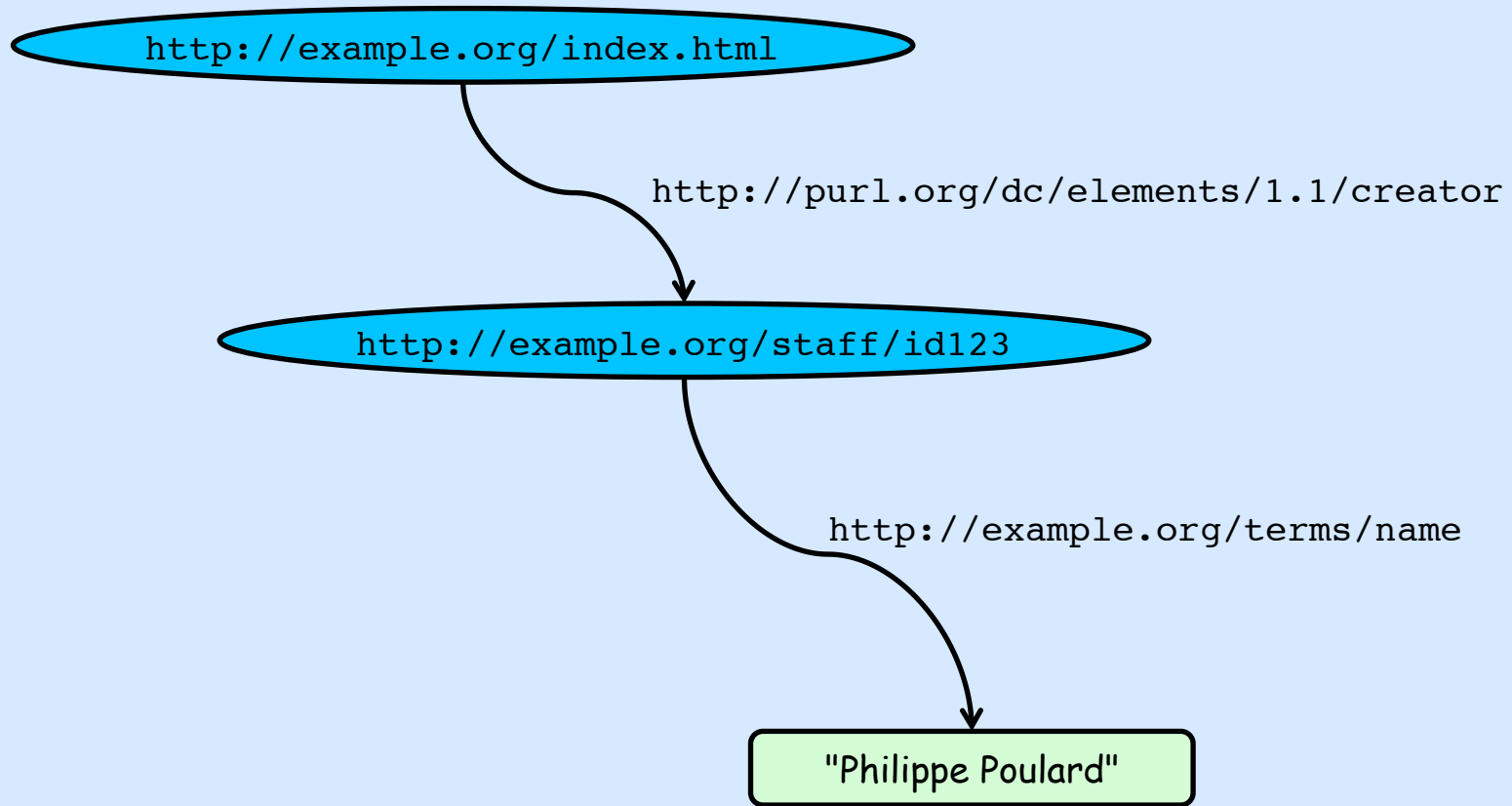
`<http://example.org/index.html> <http://purl.org/dc/elements/1.1/creator> "Philippe Poulard" .`

(N-triple notation)

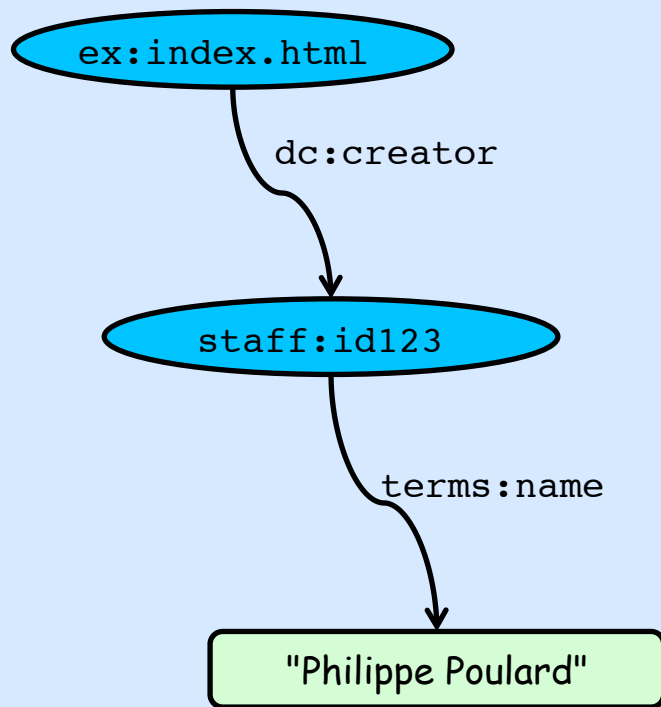
`http://example.org/index.html`

`http://purl.org/dc/elements/1.1/creator`

`"Philippe Poulard"`



```
<http://example.org/index.html>  
  <http://purl.org/dc/elements/1.1/creator> <http://example.org/staff/id123> .  
<http://example.org/staff/id123> <http://example.org/terms/name> "Philippe Poulard" .
```



Vocabulary mapping :

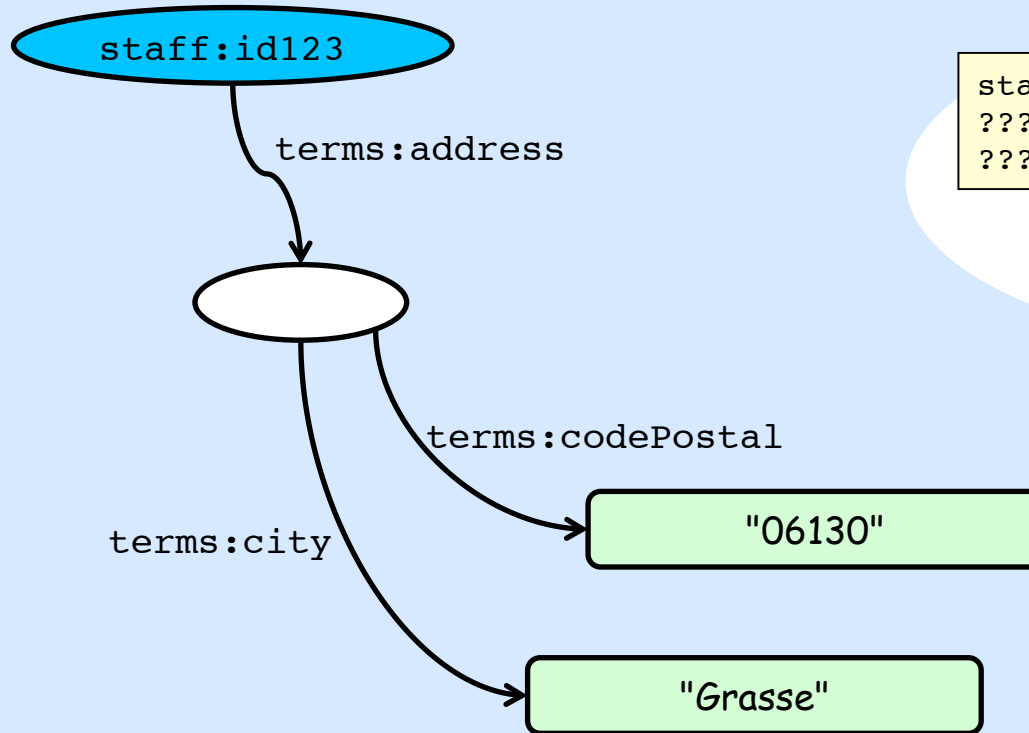
ex → http://example.org/
 dc → http://purl.org/dc/elements/1.1/
 staff → http://example.org/staff/
 terms → http://example.org/terms/

Syntaxe :

```
@prefix ex:      <http://example.org/> .
@prefix dc:     <http://purl.org/dc/elements/1.1/> .
@prefix staff:  <http://example.org/staff/> .
@prefix terms:  <http://example.org/terms/> .
```

- Les IDs des ressources se construisent par la concaténation de l'URL de base et du nom local.
- L'espace de nommage sert d'URL de base
- Comportement différent des espaces de nommage XML

```
ex:index.html    dc:creator    staff:id123 .
staff:id123     terms:name    "Philippe Poulard" .
```



```

staff:id123 terms:address   ??? .
???        terms:codePostal "06130" .
???        terms:city      "Grasse" .
  
```

Comment distinguer les nœuds blancs différents ?

Blank node identifier

```

staff:id123 terms:address   _:ppaddress .
_:ppaddress terms:codePostal "06130" .
_:ppaddress terms:city      "Grasse" .
  
```

```
staff:id123      terms:address      _:ppaddress .  
_:ppaddress     terms:codePostal "06130"^^xs:integer .  
_:ppaddress     terms:city      "Grasse" .  
ex:index.html   dc:date           "2012-12-25"^^xs:date .
```

W3C XML Schema datatypes



Une (des) syntaxe(s) XML pour RDF

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:ex="http://example.org/"
  xmlns:staff="http://example.org/staff/"
  xmlns:terms="http://example.org/terms/"

  <rdf:Description rdf:about="http://example.org/index.html">
    <dc:date rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2012-12-25</dc:date>
    <dc:language>fr</dc:language>
    <dc:creator rdf:resource="http://example.org/staff/id123"/>
  </rdf:Description>

</rdf:RDF>
```

```
<rdf:Description rdf:about="http://example.org/index.html">
  <dc:date rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2012-12-25</dc:date>
</rdf:Description>
<rdf:Description rdf:about="http://example.org/index.html">
  <dc:language>fr</dc:language>
</rdf:Description>
<rdf:Description rdf:about="http://example.org/index.html">
  <dc:creator rdf:resource="http://example.org/staff/id123"/>
</rdf:Description>
```

Astuces syntaxiques

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [<!ENTITY xsd "http://www.w3.org/2001/XMLSchema#">]>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:terms="http://example.com/terms/"
  xml:base="http://example.com/products">

  <rdf:Description rdf:ID="item1234">
    <terms:model rdf:datatype="&xsd:string">Overnighter</terms:model>
    <terms:sleeps rdf:datatype="&xsd:integer">2</terms:sleeps>
    <terms:weight rdf:datatype="&xsd:decimal">2.4</terms:weight>
    <terms:packedSize rdf:datatype="&xsd:integer">784</terms:packedSize>
  </rdf:Description>

</rdf:RDF>
```

xml:base permet de résoudre les ID RDF `rdf:ID="http://example.com/products#item1234"`
 L'expansion des entités permet de factoriser les noms des types RDF

Nœuds blancs

```
<dc:editor rdf:nodeID="abc" />
...
<rdf:Description rdf:nodeID="abc">
...
```

- Définition de classes, spécialisation, etc

```

ex:MotorVehicle    rdf:type      rdfs:Class .
ex:PassengerVehicle  rdf:type      rdfs:Class .
ex:Van              rdf:type      rdfs:Class .
ex:Truck            rdf:type      rdfs:Class .
ex:MiniVan          rdf:type      rdfs:Class .

ex:PassengerVehicle  rdfs:subClassOf  ex:MotorVehicle .
ex:Van               rdfs:subClassOf  ex:MotorVehicle .
ex:Truck              rdfs:subClassOf  ex:MotorVehicle .

ex:MiniVan           rdfs:subClassOf  ex:Van .
ex:MiniVan           rdfs:subClassOf  ex:PassengerVehicle .

ex:age               rdf:type      rdf:Property .
ex:age               rdfs:range     xsd:integer .

```

- `rdfs:domain` pour exprimer que la propriété `author` s'applique aux instances de la classe `Book` :

```

ex:Book    rdf:type      rdfs:Class .
ex:author  rdf:type      rdf:Property .
ex:author  rdfs:domain   ex:Book .

```

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [<!ENTITY xsd "http://www.w3.org/2001/XMLSchema#">]>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="http://example.org/schemas/vehicles">

  <rdfs:Class rdf:ID="MotorVehicle"/>

  <rdfs:Class rdf:ID="PassengerVehicle">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Truck">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Van">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="MiniVan">
    <rdfs:subClassOf rdf:resource="#Van"/>
    <rdfs:subClassOf rdf:resource="#PassengerVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Person"/>

  <rdfs:Datatype rdf:about="&xsd;integer"/>

  <rdf:Property rdf:ID="registeredTo">
    <rdfs:domain rdf:resource="#MotorVehicle"/>
    <rdfs:range rdf:resource="#Person"/>
  </rdf:Property>
```



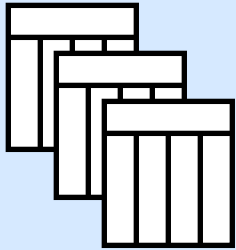
```
<rdf:Property rdf:ID="rearSeatLegRoom">
  <rdfs:domain rdf:resource="#PassengerVehicle"/>
  <rdfs:range rdf:resource="&xsd;integer"/>
</rdf:Property>

<rdf:Property rdf:ID="driver">
  <rdfs:domain rdf:resource="#MotorVehicle"/>
</rdf:Property>

<rdf:Property rdf:ID="primaryDriver">
  <rdfs:subPropertyOf rdf:resource="#driver"/>
</rdf:Property>

</rdf:RDF>
```

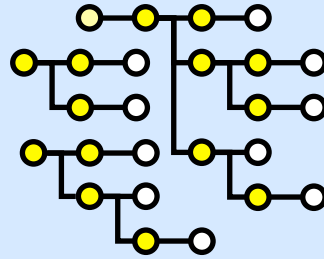
SGBDR



tables

SQL

BD XML

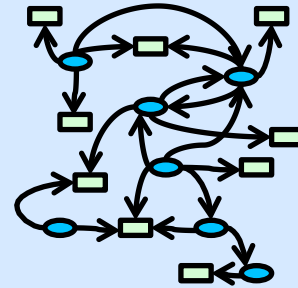


collections

XQuery

(triplestore)

BD de triplets



graphes

SPARQL

```

@prefix vCard:    <http://www.w3.org/2001/vcard-rdf/3.0#> .
@prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

<http://somewhere/MattJones/> vCard:FN    "Matt Jones" .
<http://somewhere/MattJones/> vCard:N      _:b0 .
_:b0 vCard:Family "Jones" .
_:b0 vCard:Given  "Matthew" .

<http://somewhere/RebeccaSmith/> vCard:FN    "Becky Smith" .
<http://somewhere/RebeccaSmith/> vCard:N      _:b1 .
_:b1 vCard:Family "Smith" .
_:b1 vCard:Given  "Rebecca" .

<http://somewhere/JohnSmith/>    vCard:FN    "John Smith" .
<http://somewhere/JohnSmith/>    vCard:N      _:b2 .
_:b2 vCard:Family "Smith" .
_:b2 vCard:Given  "John" .

<http://somewhere/SarahJones/>   vCard:FN    "Sarah Jones" .
<http://somewhere/SarahJones/>   vCard:N      _:b3 .
_:b3 vCard:Family "Jones" .
_:b3 vCard:Given  "Sarah" .

```

```

SELECT ?x
WHERE { ?x <http://www.w3.org/2001/vcard-rdf/3.0#FN> "John Smith" }

```



x	

<http://somewhere/JohnSmith/>	

Solutions :

- paire {nom, valeur}
- jeu de résultat (peut aussi être un graphe)
- La solution est la manière de faire correspondre le modèle

```
SELECT ?x ?fname
WHERE {?x <http://www.w3.org/2001/vcard-rdf/3.0#FN> ?fname}
```

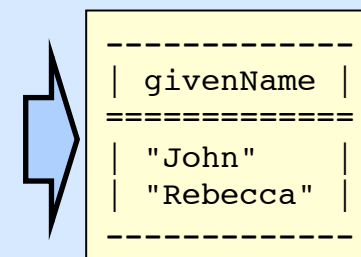


x	name
<http://somewhere/RebeccaSmith/>	"Becky Smith"
<http://somewhere/SarahJones/>	"Sarah Jones"
<http://somewhere/JohnSmith/>	"John Smith"
<http://somewhere/MattJones/>	"Matt Jones"

```
SELECT ?givenName
WHERE
  { ?y <http://www.w3.org/2001/vcard-rdf/3.0#Family> "Smith" .
    ?y <http://www.w3.org/2001/vcard-rdf/3.0#Given> ?givenName .
  }
```

```
PREFIX vcard:      <http://www.w3.org/2001/vcard-rdf/3.0#>

SELECT ?givenName
WHERE
  { ?y vcard:Family "Smith" .
    ?y vcard:Given  ?givenName .
  }
```



```
PREFIX vcard:      <http://www.w3.org/2001/vcard-rdf/3.0#>

SELECT ?y ?givenName
WHERE
{ ?y vcard:Family "Smith" .
  ?y vcard:Given  ?givenName .
}
```



y	givenName
_:b0	"John"
_:b1	"Rebecca"

```
<http://somewhere/RebeccaSmith/>
  info:age "23"^^xsd:integer ;
  vCard:FN "Becky Smith" ;
  vCard:N [ vCard:Family "Smith" ;
            vCard:Given  "Rebecca" ] .
```

```
PREFIX info: <http://somewhere/peopleInfo#>

SELECT ?resource
WHERE
{
  ?resource info:age ?age .
  FILTER (?age >= 24)
}
```



```
-----
| resource                               |
-----
| <http://somewhere/JohnSmith/>         |
-----
```

Types de données

```
1 = 1.0  
"1"^^xsd:integer = "1.0"^^xsd:decimal  
?x + ?y < 0.1 * ?z  
xsd:integer(?age) >= 18
```

Fonctions

```
isURI(?x)  
isLiteral(?y)  
isBlank(?z)  
bound(?t)
```

Langage

```
lang(?x) = "en"  
"engineer"@en -> true  
"ingénieur"@fr -> false  
datatype(?y) = xsd:string  
! (?x || ?y && ?z)
```



```
FILTER regex(?x, "pattern" [, "flags"])
```

- Flag: "i" → insensible à la casse
- Voir <http://www.w3.org/TR/xpath-functions/#regex-syntax>

```
PREFIX vcard: <http://www.w3.org/2001/vcard-rdf/3.0#>

SELECT ?g
WHERE
{ ?y vcard:Given ?g .
  FILTER regex(?g, "r", "i") }
```



g
"Rebecca"
"Sarah"

```
PREFIX info: <http://somewhere/peopleInfo#>
PREFIX vcard: <http://www.w3.org/2001/vcard-rdf/3.0#>
```

```
SELECT ?name ?age
```

```
WHERE
```

```
{
  ?person vcard:FN ?name .
  ?person info:age ?age .
}
```



name	age
"Becky Smith"	23
"John Smith"	25

```
PREFIX info: <http://somewhere/peopleInfo#>
PREFIX vcard: <http://www.w3.org/2001/vcard-rdf/3.0#>
```

```
SELECT ?name ?age
```

```
WHERE
```

```
{
  ?person vcard:FN ?name .
  OPTIONAL { ?person info:age ?age }
}
```



name	age
"Becky Smith"	23
"Sarah Jones"	
"John Smith"	25
"Matt Jones"	

```
PREFIX info: <http://somewhere/peopleInfo#>
PREFIX vcard: <http://www.w3.org/2001/vcard-rdf/3.0#>
```

```
SELECT ?name ?age
```

```
WHERE
```

```
{
  ?person vcard:FN ?name .
  OPTIONAL { ?person info:age ?age . FILTER ( ?age > 24 ) }
}
```



name	age
"Becky Smith"	
"Sarah Jones"	
"John Smith"	25
"Matt Jones"	

```
PREFIX info:      <http://somewhere/peopleInfo#>
PREFIX vcard:    <http://www.w3.org/2001/vcard-rdf/3.0#>

SELECT ?name ?age
WHERE
{
  ?person vcard:FN ?name .
  OPTIONAL { ?person info:age ?age . }
  FILTER ( !bound(?age) || ?age > 24 )
}
```



name	age
"Sarah Jones"	
"John Smith"	25
"Matt Jones"	

```

@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix vcard: <http://www.w3.org/2001/vcard-rdf/3.0#> .

_:a foaf:name "Matt Jones" .
_:b foaf:name "Sarah Jones" .
_:c vcard:FN "Becky Smith" .
_:d vcard:FN "John Smith" .

```

```

PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX vCard: <http://www.w3.org/2001/vcard-rdf/3.0#>

SELECT ?name
WHERE
{
  { [] foaf:name ?name } UNION { [] vCard:FN ?name }
}

```

```

PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX vCard: <http://www.w3.org/2001/vcard-rdf/3.0#>

SELECT ?name
WHERE
{
  [] ?p ?name
  FILTER ( ?p = foaf:name || ?p = vCard:FN )
}

```



name
"Matt Jones"
"Sarah Jones"
"Becky Smith"
"John Smith"

Nœud blanc : []

Forme des résultats :

- SELECT : retourne un tableau de résultats
- CONSTRUCT : retourne un graphe RDF, basé sur un modèle dans la requête
- DESCRIBE : retourne un graphe RDF, basé sur ce que le processeur de requêtes est configuré pour renvoyer
- ASK : pose une requête booléenne.

Modificateurs de résultat :

- OFFSET / LIMIT : couper le nombre de solutions
- ORDER BY : trier les résultats
- DISTINCT : ne rendre qu'une seule ligne pour une combinaison de variable et de valeur.

```
ORDER BY ?x ?y
```

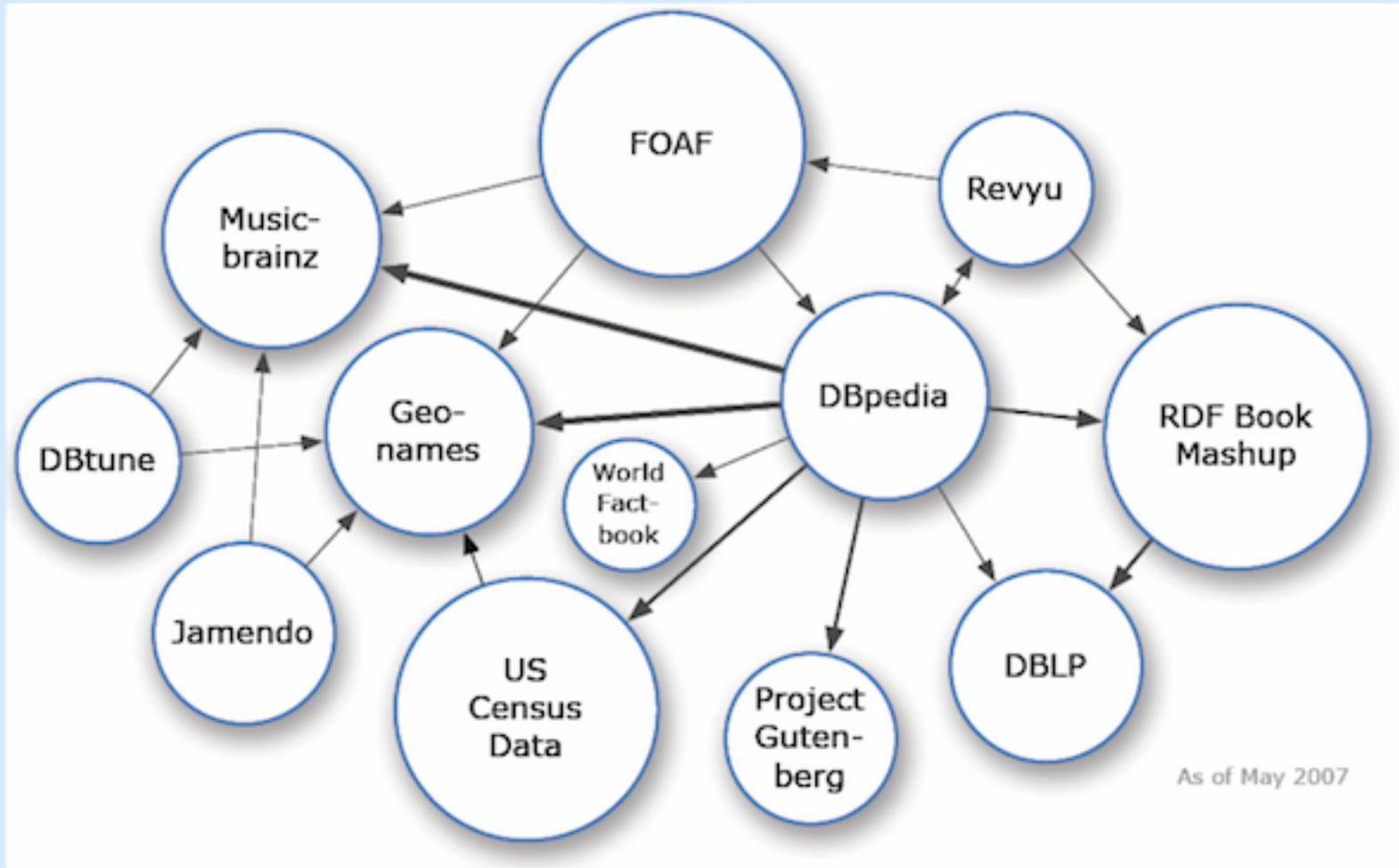
```
ORDER BY DESC(?x)
```

```
ORDER BY x:func(?x) # condition de tri personnalisé
```

- Faculté de déduire des résultats non directement annotés
- Le moteur exploite les ontologies dont il dispose (RDFS, OWL)

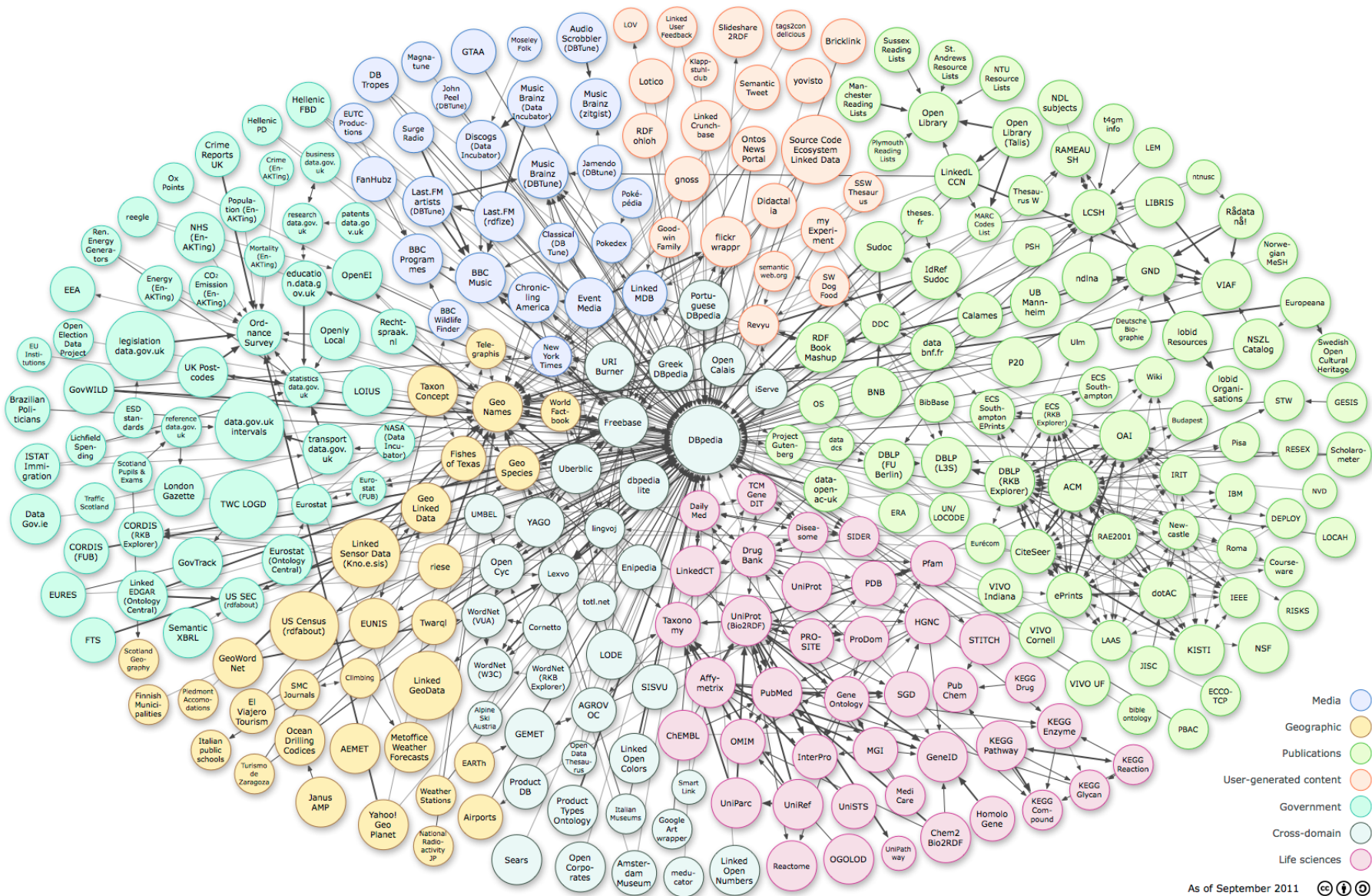
Ontologies

- Désigne un ensemble structuré de savoirs dans un domaine de connaissance particulier
- Vocabulaire, concepts associés → dans un domaine



Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch.

<http://lod-cloud.net/>



Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch.
<http://lod-cloud.net/>



