

Introduction au web de données et au web sémantique



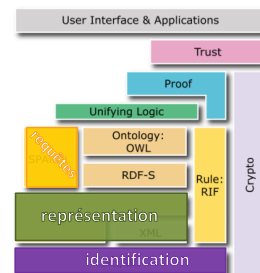
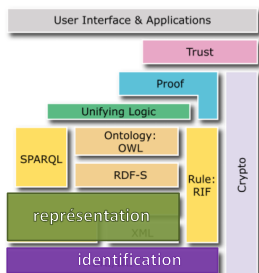
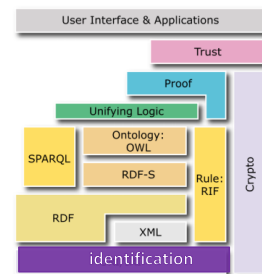
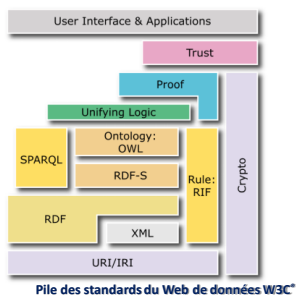
Catherine Faron Zucker, Olivier Corby
INRIA Sophia Antipolis, CNRS, UNS
<http://wimmics.inria.fr/>

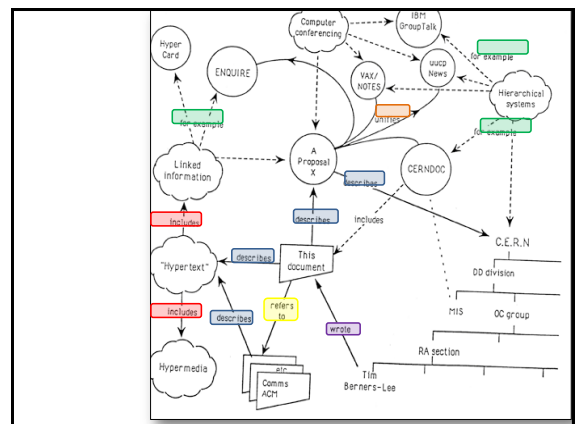
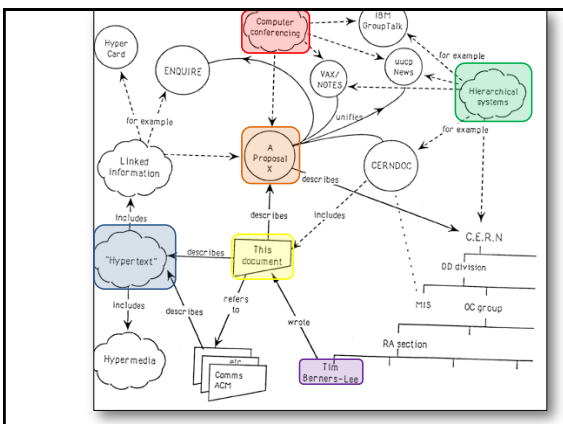
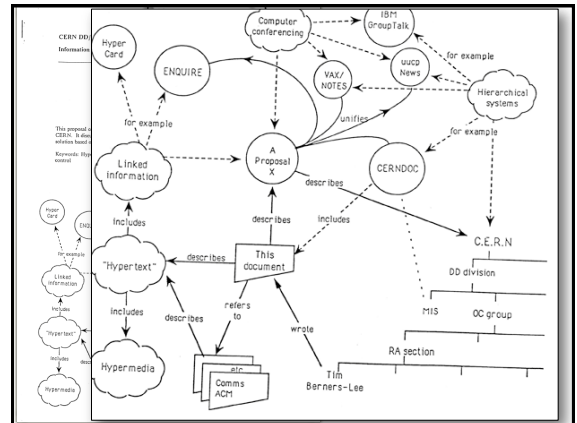
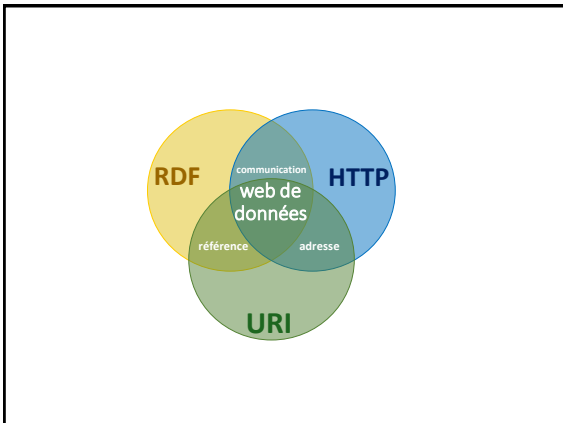
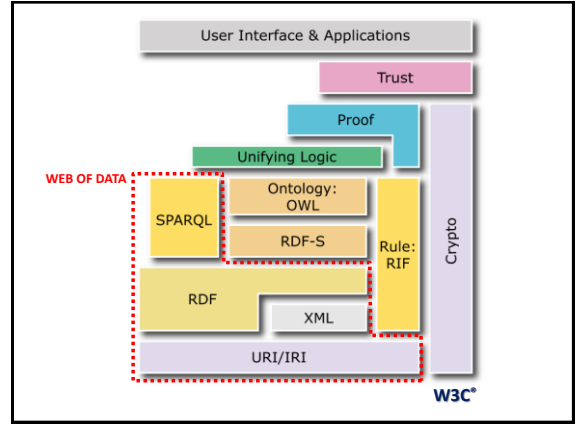
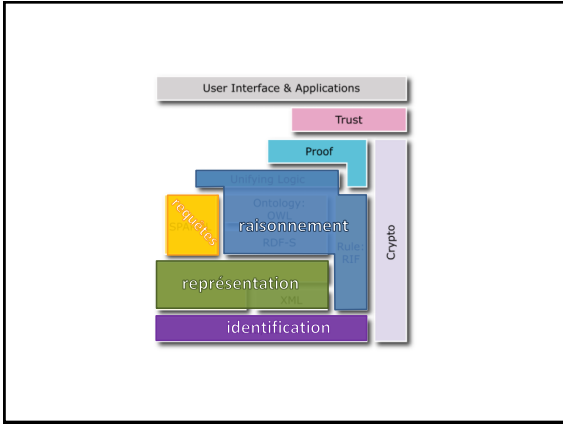
Séminaire INRA Open Data, Saint Martin de Londres, 17/12/2014

Web des données ouvertes

Un mouvement mondial

- utile pour diffuser les données scientifiques produites
- utile pour partager de l'information, des données
- utile pour tirer parti des données publiées par d'autres

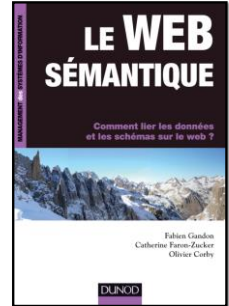




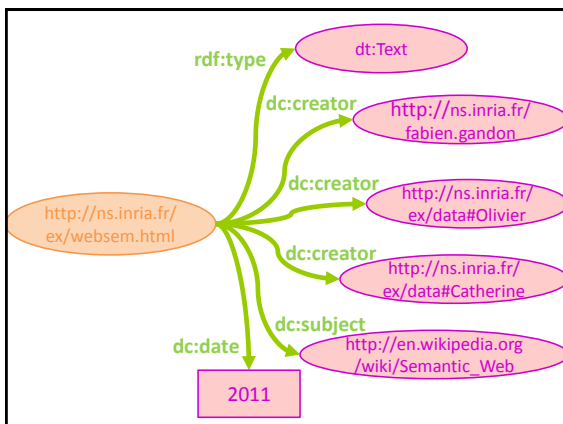
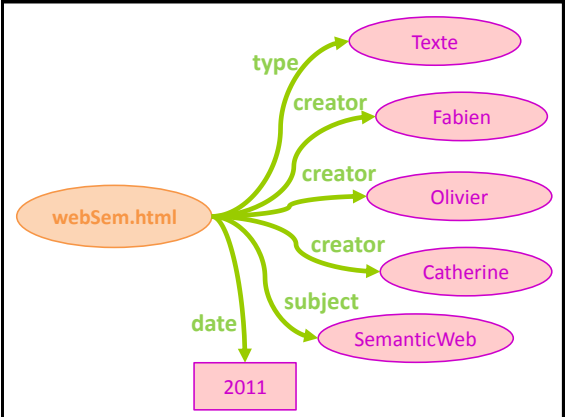
RDF est un modèle de triplets *i.e.*
toute connaissance est décomposée en
(**sujet** , **predicat** , **objet**)



*Le texte websem.html
a pour auteurs Fabien,
Olivier et Catherine,
a pour sujet le web
sémantique et a été
écrit en 2011*



websem.html est un **texte**
websem.html a pour auteur **Fabien**
websem.html a pour auteur **Olivier**
websem.html a pour auteur **Catherine**
websem.html a pour sujet **le web sémantique**
websem.html a été écrit en **2011**



syntaxe XML

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:inria="http://inria.fr/schema#" >
```

```
<rdf:Description
  rdf:about="http://inria.fr/rr/doc.html">
  <inria:author rdf:resource=
    "http://ns.inria.fr/catherine.faron#me"/>
  <inria:theme>Web</inria:theme>
</rdf:Description>
```

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

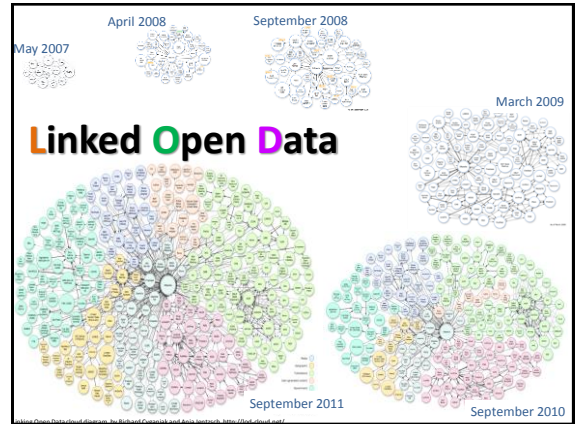
syntaxes N-Triples et Turtle

```
<http://inria.fr/rr/doc.html>
<http://inria.fr/schema#author>
<http://ns.inria.fr/catherine.faron#me> .
```

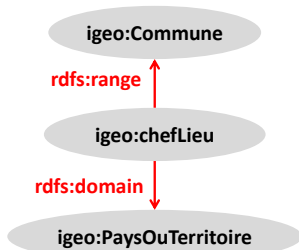
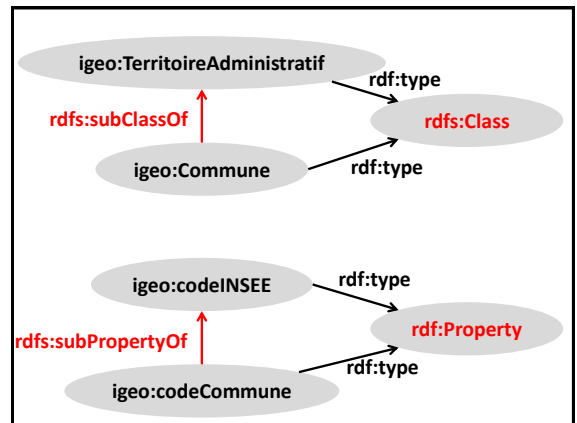
```
<http://inria.fr/rr/doc.html>
<http://inria.fr/schema#theme> "Web" .
```

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
@prefix inria: <http://inria.fr/schema#> .
```

```
<http://inria.fr/rr/doc.html>
inria:author <http://ns.inria.fr/catherine.faron#me> ;
inria:theme "Web" .
```



RDFS schémas pour définir les classes de ressources, leurs propriétés, et organiser leurs hiérarchies



SKOS: Simple Knowledge Organization System

thésaurus, taxinomies, folksonomies, vocabulaires contrôlés...

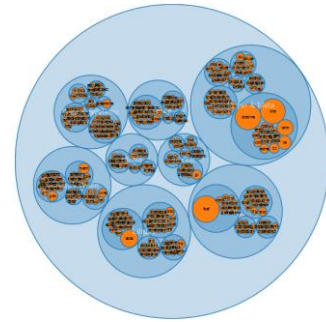
Définition et documentation des systèmes classificatoires

- Concepts et schémas SKOS
 - Les classes skos:Concept et skos:Schema
 - Les propriétés skos:inScheme, skos:topConceptOf
- Labels et codes classificatoires de concepts
 - skos:prefLabel, skos:altLabel, skos:hiddenLabel, skos:notation
- Documentation
 - skos:note, skos:definition, skos:example, skos:changeNote, skos:editorialNote, skos:historyNote et skos:scopeNote.

SKOS: Simple Knowledge Organization System

Définition des relations sémantiques entre concepts SKOS

- Réseaux sémantiques
 - skos:related
- Hiérarchies de concepts
 - skos:broader, skos:narrower
- Alignement de schémas
 - skos:closeMatch, skos:exactMatch
 - skos:relatedMatch, skos:broadMatch, skos:narrowMatch



Linked Open Vocabularies

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

Les 4 règles du Linked Data



1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
4. Include links to other URIs, so that they can discover more things

Alimentation du Web de Données



Alimentation du Web de Données



RDF in HTML attributes

Ajout de données RDF dans des pages HTML

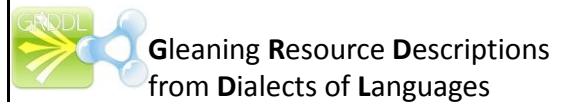
```
<body vocab="http://inra.org/schema/">
<div resource="/inra/program/sw" typeof="Tutorial">
<h2 property="title">Semantic Web</h2>
<p>Date: <span property="date">2014-12-17</span></p>
...
<p>Outstanding introduction to the Semantic Web (of Data)
<a property="location" href="http://www.inra.fr/">INRA</a>
</p>
</div>
</body>
```

Ajout de données RDF dans des pages HTML

```
prefix ex: <http://inra.org/schema/>
<http://example.org/inra/program/sw>
  rdf:type ex:Tutorial ;
  ex:date "2014-12-17" ;
  ex:location <http://www.inra.fr> .
```

```
<body vocab="http://inra.org/schema/">
<div resource="/inra/program/sw" typeof="Tutorial">
<h2 property="title">Semantic Web</h2>
<p>Date: <span property="date">2014-12-17</span></p>
...
<p>Outstanding introduction to the Semantic Web (of Data)
<a property="location" href="http://www.inra.fr/">INRA</a>
</p>
</div>
</body>
```

Alimentation du Web de Données



Profil GRDDL

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
xml:lang="en" lang="en">
<head profile="http://www.w3.org/2003/g/data-view" >
<title>Robin's Schedule</title>
<link rel="transformation"
href="http://www.w3.org/2002/12/cal/glean-hcal" />
</head>
<body> ... </body>
</html>
```

Transformation GRDDL (e.g. XSLT)

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
xml:lang="en" lang="en">
<head profile="http://www.w3.org/2003/g/data-view" >
<title>Robin's Schedule</title>
<link rel="transformation"
href="http://www.w3.org/2002/12/cal/glean-hcal" />
</head>
<body> ... </body>
</html>
```

Alimentation du Web de Données



RDB to RDF Mapping Language

Direct Mapping of Relational Data to RDF

Table: People		
ID	fname	addr (Address)
7	Catherine	18
8	Olivier	22
...

```
<People/ID=7> rdf:type <People> .
<People/ID=7> <People#fname> "Catherine" .
<People/ID=7> <People#ref-addr> <Address/ID=18> .
```

```
<People/ID=8> rdf:type <People> .
<People/ID=8> <People#fname> "Olivier" .
<People/ID=8> <People#ref-addr> <Address/ID=22> .
```

Customized Mapping of Relational Data to RDF

```
<#TriplesMap1>
  rr:logicalTable [ rr:tableName "EMP" ];

  rr:subjectMap [
    rr:template "http://jdev.org/staff/{EMPNO}";
    rr:class ex:Teacher;
  ];

  rr:predicateObjectMap [
    rr:predicate foaf:name;
    rr:objectMap [ rr:column "ENAME" ];
  ].
```

```
<http://jdev.org/staff/7369> rdf:type ex:Teacher.
<http://jdev.org/staff/7369> foaf:name "Catherine".
<http://jdev.org/staff/7370> rdf:type ex:Teacher.
<http://jdev.org/staff/7370> foaf:name "Olivier".
```

interroger avec SPARQL
SPARQL Protocol and RDF
Query Language

LINKED DATA
 On the web, open access
 Machine-readable data
 Non-proprietary format
 RDF standards
 Linked RDF
 YOUR DATA 5

SPARQL Explorer for http://dbpedia.org/sparql

```
SPARQL
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX dbp: <http://www.dbpedia.org/resource/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX foaf: <http://xmlns.foaf.org/2000/01/01/foaf#>
PREFIX ex: <http://example.org/>
PREFIX dbpedia: <http://dbpedia.org/resource/>
PREFIX schema: <http://schema.org/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
```

Results: Browse [Go] [Reset]

Powered by [Cygnum SPARQL](#)

e.g. DBpedia

SPARQL 1.1

Langage de requête (syntaxe Turtle)

- SPARQL 1.1 Query Language - W3C REC 21 Mar. 2013
- SPARQL 1.1 Update - W3C REC 21 Mar. 2013
- 2. Langage de présentation des résultats
- SPARQL Query Results XML Format - W3C REC 21 Mar. 2013

triplets SPARQL

- syntaxe Turtle avec des points d'interrogation pour les variables:

```
?x rdf:type ex:Person
```

- décrire des patrons de graphes à trouver:

```
SELECT ?subject ?property ?value
WHERE { ?subject ?property ?value }
```

- un patron est par défaut une conjonction de triplets:

```
SELECT ?x WHERE
{ ?x rdf:type ex:Person .
  ?x ex:name ?name . }
```

préfixes des espaces de noms

*déclarer des préfixes pour les vocabulaires utilisés dans la requête:

```
PREFIX mit: <http://www.mit.edu#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?student
WHERE {
  ?student mit:registeredAt ?x .
  ?x foaf:homepage <http://www.mit.edu> .
}
```

*déclarer un espace de base par défaut i.e. pour les URI relatifs

```
BASE <...>
```

la langue et le type des littéraux

pour obliger à respecter un attribut : @fr , ^^xsd:integer

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?x ?f WHERE {
  ?x foaf:name "Fabien"@fr ; foaf:knows ?f .
}
```

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?x WHERE {
  ?x foaf:name "Fabien"@fr ;
    foaf:age "21"^^xsd:integer .
}
```

déclarer un motif optionnel

une partie du patron de graphe n'est pas obligatoire.

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?person ?name
WHERE {
  ?person foaf:homepage <http://fabien.info> .
  OPTIONAL { ?person foaf:name ?name . }
}
```

La variable ?name est potentiellement « unbound »

patrons alternatifs

faire l'union des résultats de deux motifs de graphes

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?person ?name
WHERE {
  ?person foaf:name ?name .
  {
    ?person foaf:homepage <http://fabien.info> .
  }
  UNION
  {
    ?person foaf:homepage <http://bafien.org> .
  }
}
```


soustraire un motif des résultats

supprimer des résultats correspondants à un pattern

```
PREFIX ex: <http://www.example.abc#>
SELECT ?x
WHERE {
  ?x a ex:Person
  MINUS { ?x a ex:Man }
}
```

valeurs prédéfinies pour les variables

fournir des résultats où une partie du *binding* est prédéfinie

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?person ?name
WHERE {
  ?person foaf:name ?name .
  VALUES ?name { "Peter" "Pedro" "Pierre" }
}
```

motifs de chemins dans les graphes

des expressions régulières sur des chemins entre ressources

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?friend WHERE {
  ?x foaf:name "Fabien Gandon" ;
  foaf:knows+ ?friend .
}
```

filtres

```
PREFIX ex: <http://inria.fr/schema#>
SELECT ?person ?name
WHERE {
  ?person rdf:type ex:Person ; ex:name ?name ; ex:age ?age .
  FILTER (xsd:integer(?age) >= 18) }
```

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT * where {
  ?x foaf:name ?n .
  FILTER (?n IN ("fabien", "olivier", "catherine") ) }
```

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT * where {
  ?x foaf:name ?name ; foaf:age ?age .
  FILTER( if (langMatches(lang(?name),"FR"), ?age>=18, ?age>=21) ) }
```

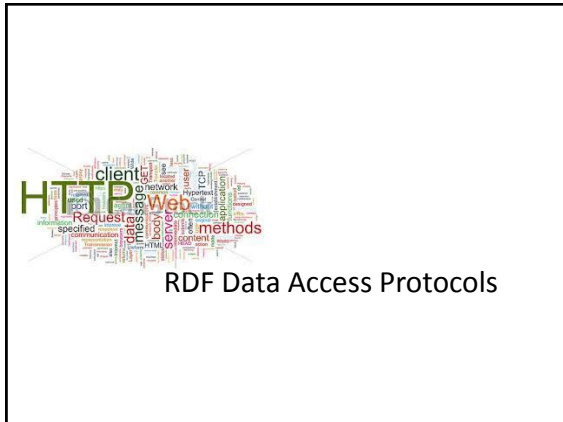
```
SELECT ?name
WHERE {
  ?x foaf:name ?name . FILTER NOT EXISTS { ?x foaf:age -1 }
}
```

from

```
PREFIX mit: <http://www.mit.edu#>
FROM <http://www.mit.edu/data.rdf>
SELECT ?student
WHERE { ?student mit:registeredAt ?x . }
```

XML result

```
<?xml version="1.0"?>
<sparql xmlns="http://www.w3.org/2005/sparql-results#">
  <head> <variable name="student"/> </head>
  <results>
    <result>
      <binding name="student">
        <uri>http://www.mit.edu/data.rdf#ndieng</uri>
      </binding>
    </result>
    <result>
      <binding name="student">
        <uri>http://www.mit.edu/data.rdf#jdoe</uri>
      </binding>
    </result>
  </results>
</sparql>
```



SPARQL 1.1. Protocol

Transport de requête SPARQL vers un service SPARQL
Retour des résultats en HTTP

```
http://fr.dbpedia.org/sparql?query=select distinct * where
{ <http://fr.dbpedia.org/resource/Auguste> ?p ?v }
limit 100
```

```
http://example.org/sparql?update=insert data {<a> <b> <c>}
```

SPARQL 1.1 Graph Store HTTP Protocol

Association des méthodes HTTP à des opérations sur des graphes RDF

- Récupération du contenu d'un graphe RDF
- Destruction, création, remplacement du contenu d'un graphe RDF
- Ajout de triplets RDF à un graphe existant

SPARQL 1.1 Graph Store HTTP Protocol

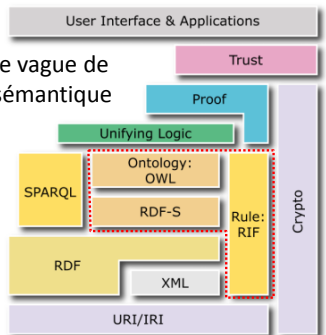
HTTP PUT: DROP + INSERT

HTTP POST: INSERT

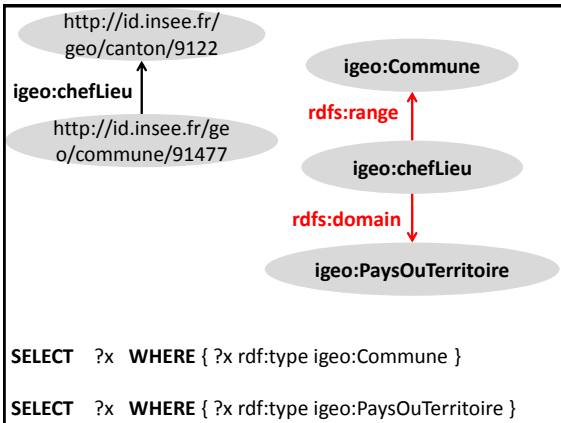
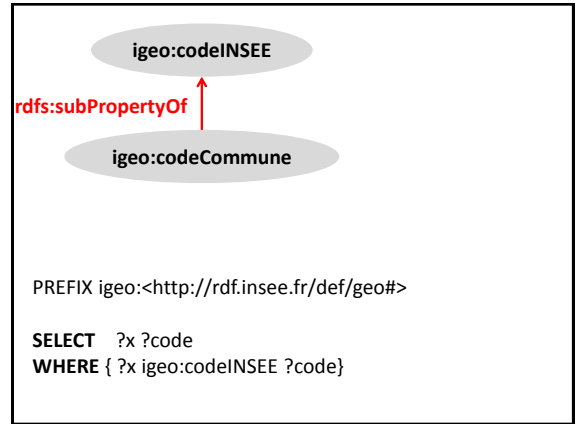
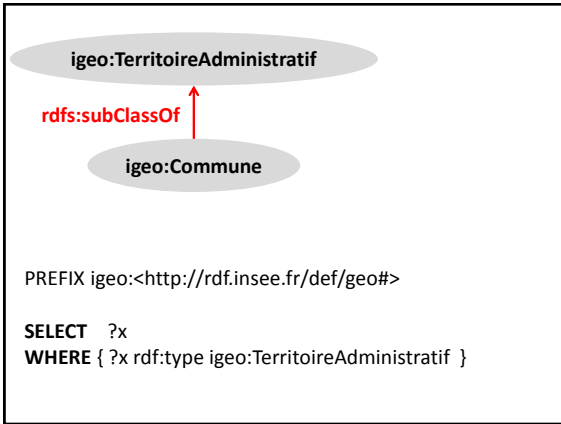
HTTP DELETE

Web de données & Web sémantique

Web donnée: première vague de
déploiement du web sémantique



RDFS Entailment



OWL in one...

Ontology Web Language

- par union
- la disjonction
- par intersection
- par complément
- par restriction
- par cardinalité
- par équivalence
- par énumération
- restrict. valeur
- union disjointe
- clés

prop. algébriques

prop. disjointes

cardinalité qualifiée

négation prop. indiv

prop. chaînées

