

LES LANGAGES DU WEB SÉMANTIQUE

M2 IFI 2013-2014

Catherine Faron Zucker faron@polytech.unice.fr

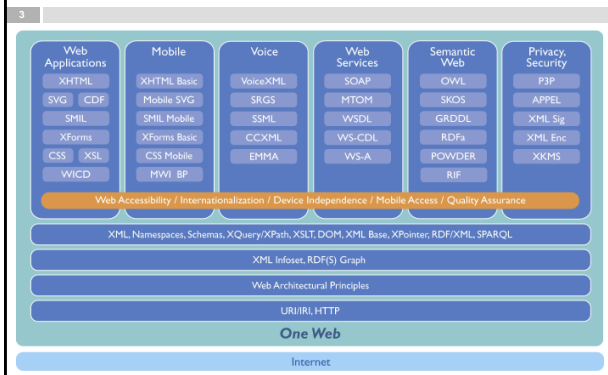
The Semantic Web in a nutshell

World Wide Web Consortium : www.w3.org

« The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation »

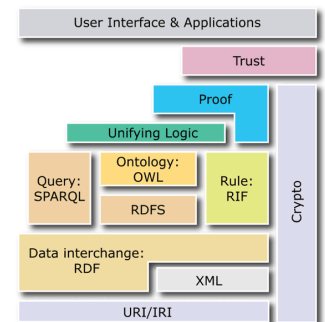
www.scientificamerican.com/2001/0501issue/0501berners-lee.html

An extension of the current web



The W3C Semantic Web Stack

- Nommage: URI
- Syntaxe: XML
- Sémantique:
 - RDF
 - RDFS
 - OWL
 - SPARQL
 - RIF



XML et les langages du WS

- XML : Extensible Markup Language
 - Standard du W3C
 - Langage de description et d'échange de :
 - documents structurés
 - données structurées
- XML permet d'exprimer la structure d'un document comme un arbre de syntaxe abstraite
- XML ne permet pas d'exprimer la signification des balises, leur sémantique

XML et les langages du WS

```
<book>
  <author>
    <Person>
      <name> David Lodge </name>
    </Person>
  </author>
  <title> Small World </title>
</book>
```

```
<aaa>
  <bbb>
    <ddd>
      <eee> David Lodge </eee>
    </ddd>
  </bbb>
  <ccc> Small World </ccc>
</aaa>
```

XML et les langages du WS

7

- A Book is a kind of Document
- Documents have an author, which is a Person
- Documents have a title which is a Literal

- Décrire le sens selon un modèle: RDF,
 - par exemple dans sa syntaxe XML
- En utilisant des vocabulaires standardisés, ontologies: RDFS, OWL
- Pour répondre à des requêtes : SPARQL

Resource Description Framework RDF

<http://www.w3.org/RDF>

Resource Description Framework

9

- Langage de représentation de connaissances
 - sur les ressources du Web
 - sur des entités identifiées sur le Web (par un uri)
- Standard dédié
 - à l'échange des informations sur le web
 - au traitement des informations sur le web

- > Modèle
 - + Syntaxes RDF/XML, N-Triples, N3 Turtle, ...

RDF Design Goal

10

- Modèle simple
- Sémantique formelle
- Vocabulaire extensible basé sur les URI
- Syntaxe XML
- XML schema datatypes
- Autorise quiconque à faire des déclarations sur n'importe quelle ressource: modularité

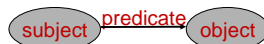
- <http://www.w3.org/TR/rdf-concepts/>

Modèle simple

11

- Décrire les entités avec
 - des propriétés
 - des valeurs de propriétés
- Modèle de triplets
 - ressource propriété valeur

- Modèle de graphe
 - Jointure (conjonction) de triplets



Modèle simple

12

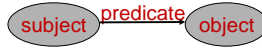
- Web-based design : Propriétés vs Classes

 - Le web est un réseau où tout le monde peut participer
-
- A diagram showing multiple overlapping triplets. Each triplet consists of 'subject', 'predicate', and 'object' ovals. The triplets are arranged in a way that they overlap, illustrating a network structure.
- Une classe n'est pas définie entièrement localement : elle peut être réutilisée, complétée ailleurs par de nouvelles propriétés

Modèle simple

13

- Modèle de triplets
 - ressource propriété valeur



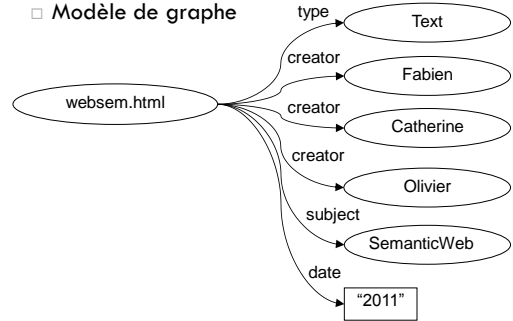
Le texte websem.html a pour auteurs Fabien, Catherine et Olivier, a pour sujet le web sémantique et a été écrit en 2011

(websem.html , type , Text)
 (websem.html , creator , Fabien)
 (websem.html , creator , Catherine)
 (websem.html , creator , Olivier)
 (websem.html , subject , SemanticWeb)
 (websem.html , date , "2011")

Modèle simple

14

- Modèle de graphe



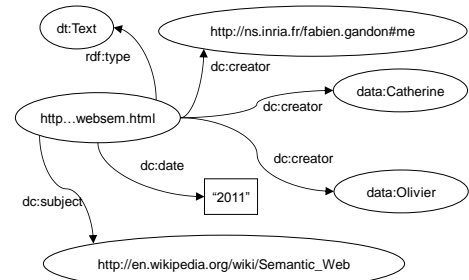
Vocabulaire basé sur les URI

15

(<http://ns.inria.fr/ex/websem.html> , rdf:type , dt:Text)
 (<http://ns.inria.fr/ex/websem.html> , dc:creator ,
 <http://ns.inria.fr/fabien.gandon#me>)
 (<http://ns.inria.fr/ex/websem.html> , dc:creator , #Catherine)
 (<http://ns.inria.fr/ex/websem.html> , dc:creator , #Olivier)
 (<http://ns.inria.fr/ex/websem.html> , dc:subject ,
 http://en.wikipedia.org/wiki/Semantic_Web)
 (<http://ns.inria.fr/ex/websem.html> , dc:date , "2011")

Vocabulaire basé sur les URI

16



Vocabulaire basé sur les URI

17

- Un nœud d'un graphe RDF est soit une ressource identifiée (URI) ou anonyme, soit un littéral
 - Le sujet d'une propriété est une ressource
 - La valeur d'une propriété est une ressource ou un littéral
 - Les nœuds internes sont donc tous des ressources et les littéraux n'apparaissent qu'aux feuilles
- Un arc d'un graphe RDF (une propriété) est un URI

Syntaxe N-Triples

18

```
<http://ns.inria.fr/ex/websem.html>
  <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
  <http://purl.org/dc/dcmitype/Text> .
<http://ns.inria.fr/ex/websem.html> <http://purl.org/dc/elements/1.1/title>
  "Web Sémantique" .
<http://ns.inria.fr/ex/websem.html> <http://purl.org/dc/elements/1.1/subject>
  <http://en.wikipedia.org/wiki/Semantic\_Web> .
<http://ns.inria.fr/ex/websem.html> <http://purl.org/dc/elements/1.1/creator>
  <http://ns.inria.fr/fabien.gandon#me> .
<http://ns.inria.fr/ex/websem.html> <http://purl.org/dc/elements/1.1/creator>
  <http://ns.inria.fr/ex/data#Catherine> .
```

Syntaxe Notation3 ou N3

```

@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix dt: <http://purl.org/dc/dcmitype/>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix dc: <http://purl.org/dc/elements/1.1/>.
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
<http://ns.inria.fr/ex/websem.html>
  a dt:Text, <http://xmlns.com/foaf/0.1/Document> ;
  dc:title "Web Sémantique" ;
  dc:creator <http://ns.inria.fr/ex/data#Catherine>,
    <http://ns.inria.fr/ex/data#Olivier>,
    <http://ns.inria.fr/fabien.gandon#me> ;
  dc:subject <http://en.wikipedia.org/wiki/Semantic_Web>.
  
```

Syntaxe XML

□ <http://www.w3.org/TR/rdf-syntax-grammar/>

Namespaces, Qualified Names

- <http://www.w3.org/TR/REC-xml-names/>
- Un espace de noms est désigné par un URI
- On reconnaît les termes issus d'un même espace de noms à ce qu'ils commencent par l'URI de cet espace de noms.
- Un préfixe désigne l'URI d'un espace de noms
 - En XML: `xmlns:dc=http://purl.org/dc/elements/1.1`
 - En N3: `@prefix dc: <http://purl.org/dc/elements/1.1/>`
- Nom qualifié = préfixe + ":", " + nom local
 - `dc:title` signifie `<http://purl.org/dc/elements/1.1/title>`

Namespaces, Qualified Names

Namespaces nommés

```

<?xml version="1.0"?>
<rdf:RDF xmlns:rdf=http://www.w3.org/1999/02/22-rdf-syntax-ns#
  xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dt="http://purl.org/dc/dcmitype/"
  xml:base="http://ns.inria.fr/ex/data">
  <rdf:Description rdf:about="http://ns.inria.fr/ex/websem.html">
    <type><rdf:Description rdf:about="http://xmlns.com/foaf/0.1/Document"/>
  </type>
</rdf:Description >
...
</rdf:RDF>
  
```

Namespace par défaut

Tout est ressource :
`dt:Text` désigne `http://purl.org/dc/dcmitype/Text`
`type` désigne `http://www.w3.org/1999/02/22-rdf-syntax-ns#type`

Base

```

□ http://www.w3.org/TR/xmlbase/ base
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf=http://www.w3.org/1999/02/22-rdf-syntax-ns#
  xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dt="http://purl.org/dc/dcmitype/"
  xml:base="http://ns.inria.fr/ex/data">
  <rdf:Description rdf:about="http://ns.inria.fr/ex/websem.html">
    <dc:creator><rdf:Description rdf:about="#Catherine"/>
  </dc:creator>
</rdf:Description >
...
</rdf:RDF>
  
```

Tout est ressource :
`#Catherine` désigne `http://ns.inria.fr/ex/data#Catherine`

Entités

Déclarations d'entités

```

<!DOCTYPE rdf:RDF [
  <ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#">
  <ENTITY xsd "http://www.w3.org/2001/XMLSchema#">
]>
<rdf:RDF xmlns:rdf=&rdf; xmlns:rdfs=&rdfs; xmlns:xsd=&xsd; >
...
</rdf:RDF>
  
```

Appels d'entités

Syntaxe XML

25

```
<rdf:RDF>
  <rdf:Description rdf:about="http://ns.inria.fr/ex/websem.html">
    <dc:creator>
      <rdf:Description rdf:about="#Catherine"/>
    </dc:creator>
    <dc:title>Web sémantique</dc:title>
  </rdf:Description>
</rdf:RDF>
```

Syntaxe(s) XML

26

```
<rdf:RDF>
  <rdf:Description rdf:about="http://ns.inria.fr/ex/websem.html">
    <dc:creator rdf:resource="#Catherine"/>
    <dc:title>Web sémantique</dc:title>
  </rdf:Description>
</rdf:RDF>
```

Syntaxe(s) XML

27

```
<rdf:RDF>
  <rdf:Description rdf:about="http://ns.inria.fr/ex/websem.html">
    <dc:creator rdf:resource="#Catherine"/>
    <dc:title>Web sémantique</dc:title>
  </rdf:Description>
</rdf:RDF>
```

Syntaxe(s) XML

28

```
<rdf:RDF>
  <rdf:Description rdf:about="http://ns.inria.fr/ex/websem.html"
    dc:title="Web sémantique"/>
    <dc:creator rdf:resource="#Catherine"/>
  </rdf:Description>
</rdf:RDF>
```

Exercice

29

Typage des ressources

30

```
<rdf:RDF xmlns:rdf=&rdf; xmlns:rdfs=&rdfs; ... >
  <rdf:Description rdf:about="http://ns.inria.fr/ex/websem.html">
    <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Document"/>
    <rdf:type rdf:resource="http://purl.org/dc/elements/1.1/Text"/>
  </rdf:Description>
</rdf:RDF>
```

http://ns.inria.fr/ex/websem.html **rdf:type** dt:Text.
http://ns.inria.fr/ex/websem.html **a** foaf:Document.

- Une ressource peut avoir plusieurs types

Typage des ressources

31

```
<rdf:RDF xmlns:rdf=&rdf; xmlns:rdfs=&rdfs; ... >
  <rdf:Description rdf:about="http://ns.inria.fr/ex/websem.html">
    <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Document"/>
    <rdf:type rdf:resource="http://purl.org/dc/elements/1.1/Text"/>
  </rdf:Description>
</rdf:RDF>

<rdf:RDF xmlns:rdf=&rdf; xmlns:rdfs=&rdfs; ... >
  <dt:Text rdf:about="http://ns.inria.fr/ex/websem.html">
    <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Document"/>
  </dt:Text>
</rdf:RDF>
```

Ressources anonymes

32

- Ressource non identifiée (sans URI), blank node, noeud vide
- Sémantique existentielle : il existe une ressource

```
<rdf:RDF xmlns:rdf=&rdf; ... >
  <dt:Text rdf:about="http://ns.inria.fr/ex/websem.html">
    <dc:creator>
      <rdf:Description>
        <foaf:firstName>Catherine</foaf:firstName>
      </rdf:Description>
    </dc:creator>
  </dt:Text>
</rdf:RDF>
```

Ressources anonymes

33

```
<rdf:RDF xmlns:rdf=&rdf; ... >
  <dt:Text rdf:about="http://ns.inria.fr/ex/websem.html">
    <dc:creator rdf:parseType="Resource">
      <foaf:firstName>Catherine</foaf:firstName>
    </dc:creator>
  </dt:Text>
</rdf:RDF>
```

Ressources anonymes

34

- Identification des blank nodes
 - Il faut pouvoir différencier les différents blank nodes d'un graphe dans une représentation par triplets ou dans certaine syntaxe XML
 - Syntaxe XML: `rdf:nodeID = "abc"`
 - N3: `_:abc`

Ressources anonymes

35

```
<rdf:RDF xmlns:rdf=&rdf; ... >
  <dt:Text rdf:about="http://ns.inria.fr/ex/websem.html">
    <dc:creator rdf:nodeID="abc"/>
  </dt:Text>
  ...
  <rdf:Description rdf:nodeID="abc">
    <foaf:firstName>Catherine</foaf:firstName>
  </rdf:Description>
</rdf:RDF>
```

Ressources anonymes

36

```
<http://ns.inria.fr/ex/websem.html> a dt:Text; dc:creator _:abc.
_:abc foaf:firstName "Catherine".
```

Réification par Statement

37

- Un triplet est réifié par un Statement
 - Considérer un triplet comme une ressource
 - Décrire cette ressource

```
< http://ns.inria.fr/fabien.gandon#me> dc:creator _:s.  
_:s rdf:type rdf:Statement.  
_:s rdf:subject <http://ns.inria.fr/ex/websem.html>.  
_:s rdf:predicate dc:title.  
_:s rdf:object "Web sémantique".
```

XML schema datatypes

38

- RDF repose sur XML Schema datatypes pour typer les littéraux

- <http://www.w3.org/2001/XMLSchema#>

```
xsd:integer  
xsd:float  
xsd:string  
xsd:dateTime  
xsd:boolean  
...  
rdf:XMLLiteral  
rdfs:Literal
```

XML schema datatypes

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.  
@prefix dc: <http://purl.org/dc/elements/1.1/>.  
<http://ns.inria.fr/ex/websem.html> dc:date "2011"^^xsd:Year.
```

- Le type par défaut est xsd:string
"Catherine" ⇔ "Catherine"^^xsd:string

XML schema datatypes

40

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"  
  <dt:Text rdf:about="http://ns.inria.fr/ex/websem.html">  
    <dc:date rdf:datatype="xsd:Year">2011</dc:date>  
  </dt:Text>  
  ...  
</rdf:RDF>
```

Conteneurs

41

- rdfs:Container
- `_:x` rdf:type rdf:Bag
Groupe de ressources ou littéraux, sans ordre
- `_:y` rdf:type rdf:Seq
Séquence **ordonnée** de ressources ou littéraux
- `_:z` rdf:type rdf:Alt
Groupe de ressources ou littéraux **alternatifs**

Conteneurs

42

- Syntaxe(s) XML

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"  
  <rdf:Seq rdf:about="http://example.org/favourite-fruit">  
    <rdf:_1 rdf:resource="http://example.org/banana"/>  
    <rdf:_2 rdf:resource="http://example.org/apple"/>  
    <rdf:_3 rdf:resource="http://example.org/pear"/>  
  </rdf:Seq>  
  
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"  
  <rdf:Seq rdf:about="http://example.org/favourite-fruit">  
    <rdf:li rdf:resource="http://example.org/banana"/>  
    <rdf:li rdf:resource="http://example.org/apple"/>  
    <rdf:li rdf:resource="http://example.org/pear"/>  
  </rdf:Seq>
```

Collections

43

- `_w: rdf:type rdf:List`
 - Collection ne contenant **que** les éléments spécifiés
 - à la différence des conteneurs qui restent "ouverts"

Collections

44

```
<rdf:Description rdf:about="http://example.org/courses/6.001">
  <s:hasFruit rdf:nodeID="f1"/>
</rdf:Description>
<rdf:Description rdf:nodeID="f1">
  <rdf:first rdf:resource="http://example.org/banana"/>
  <rdf:rest rdf:nodeID="f2"/>
</rdf:Description>
<rdf:Description rdf:nodeID="f2">
  <rdf:first rdf:resource="http://example.org/apple"/>
  <rdf:rest rdf:nodeID="f3"/>
</rdf:Description>
<rdf:Description rdf:nodeID="f3">
  <rdf:first rdf:resource="http://example.org/pear"/>
  <rdf:rest rdf:resource=rdf:nil/>
</rdf:Description>
```

Collections

45

- Syntaxe XML abrégée

```
<rdf:Description rdf:about="http://example.org/basket">
  <ex:hasFruit rdf:parseType="Collection">
    <rdf:Description rdf:about="http://example.org/banana"/>
    <rdf:Description rdf:about="http://example.org/apple"/>
    <rdf:Description rdf:about="http://example.org/pear"/>
  </ex:hasFruit>
</rdf:Description>
```

Sémantique

46

- FOL : RDF en est un sous-ensemble
 - Prédicats binaires
 - quantification existentielle
 - Conjonction
 - Pas de prédicats n-aires
 - Pas de disjonction
 - Pas de négation
 - Pas de quantification universelle
- Tout énoncé RDF est considéré comme vrai
- RDF est monotone
 - Ce qui est vrai et ce que l'on peut déduire restent vrai au fur et à mesure que l'on ajoute de nouveaux énoncés RDF

Sémantique formelle

47

- <http://www.w3.org/TR/rdf-mt/>
- Théorie des modèles