

Concepts liés au cours de SdC

PICARD DRUET David

30 novembre 2015

Par ce que je suis un noob et je le vis mal.
Et puis le fait que ça puisse mener à tirer dans des cochons avec un pistolet
à clou est très motivant.

La grande majorité (si ce n'est la totalité) des informations sont
retrouvables avec n'importe quel moteur de recherche, en moins de temps
qu'il n'en faut pour dire : "RTFM, sale noob!".

Le seul intérêt de ce doc en fait, c'est de les rassembler.

1 URI

Courte chaîne de caractères identifiant une ressource sur un réseau. Doit permettre d'identifier une ressource de manière permanente, même si la ressource est déplacée ou supprimée. ex : Dans le monde réel, y'a l'ISBN pour les livres. Ben l'URI c'est pareil, mais sur le web. Les adresses IRI incluent les adresses URI et les adresses URL.

2 Web sémantique

Le Web sémantique est un mouvement collaboratif mené par le World Wide Web Consortium (W3C) qui favorise des méthodes communes pour échanger des données.

Le Web sémantique est une vision de l'information qui permet d'être lisible par les humains et par les machines. Cela permettra d'effectuer les travaux fastidieux et répétitifs dans le domaine de la recherche d'information par des machines tout en améliorant et consolidant l'information sur le Web pour ses utilisateurs.

Le Web sémantique, comme prévu initialement, est un système qui permet aux machines de "comprendre" et de répondre aux demandes complexes de l'homme en fonction du sens de ces demandes. Une telle "compréhension" exige que les sources d'information pertinentes aient été sémantiquement structurées au préalable.

Le Web sémantique propose des langages spécialement conçus pour les données : RDF (Resource Description Framework), OWL (Ontology Web

Language), et XML (eXtensible Markup Language). Ces technologies sont combinées afin de fournir des descriptions qui complètent ou remplacent le contenu des documents Web.

3 Ontologies

3.1 C'est quoi ?

- Représentation formelle d'une conceptualisation partagée.
 - Représentation des connaissances en décrivant les concepts par leur signification et leurs relations entre eux.
 - Généralement, c'est ce qui est implicite aux documents.
- ⇒ taxonomie avec des relations supplémentaires, des contraintes et des définitions formelles.

3.2 Principes

- Individus → sont des instances de classes
- Propriétés → relations binaires entre individus
- Classes → ensemble d'individus

4 RDF : Resource Description Framework

C'est un modèle de graph destiné à la description (de façon formelle) des ressources Web et leurs métadonnées (= données liées). Une des syntaxe de RDF est le RDF/XML.

4.1 Principes fondamentaux

Document RDF = ensemble de triplets

triplet RDF = (sujet, prédicat, objet)

Avec : sujet -> ressource à décrire; prédicat -> propriété applicable à la ressource; objet -> valeur de la propriété.

Le prédicat est ****obligatoirement**** identifié par une URI.

ex :

Un document RDF ainsi formé correspond à un multigraphe orienté étiqueté. Chaque triplet correspond alors à un arc orienté dont le label est le prédicat, le nœud source est le sujet et le nœud cible est l'objet.

La structure de RDF est extrêmement générique et sert de base à un certain nombre de schémas ou vocabulaires dédiés à des applications spécifiques comme les langages d'ontologie RDFS et OWL.

4.2 Languages de requête

Il y'en a un certain nombre. Le langage SPARQL est à la fois celui évoqué dans le cours, et destiné à devenir le standard dans ce domaine.

ex :

```
PREFIX foaf:    <http://xmlns.com/foaf/0.1/>
```

```
SELECT ?name ?mbox
WHERE
{ ?x foaf:name ?name.
  ?x foaf:mbox ?mbox }
```

5 SPARQL : SPARQL Protocol and RDF Query Language

5.1 Caractéristiques

Technologie clé du web sémantique, adapté à la structure spécifique des graphes RDF. Assez proche de SQL, permet d'exprimer des requêtes interrogatives (SELECT) ou constructives (CONSTRUCT).

- Requête SELECT : extrait du graphe RDF un sous-graphe vérifiant les conditions définies dans une clause WHERE.
- Requête CONSTRUCT : engendre un nouveau graphe, que complète le graphe interrogé.
- Valeurs optionnelles : OPTIONAL

5.2 Exemples

Les exemples 1 et 2 vont porter sur un graphe RDF fait en TP, exemple_graph_rdf.txt, normalement fourni avec le document.

Ex 1 On veut obtenir toutes les reactions qui consomment du glucose

```
#### On prend chaque reaction du graphe , sans redondance
SELECT DISTINCT ?reaction
WHERE{
    :glucose :estConsommePar ?reaction .
}
```

Ex 2 On veut obtenir toutes les reactions qui produisent du E.

```
SELECT DISTINCT ?reaction
WHERE{
    ?reaction :produit :E.
```

```
}
```

Ex 3 Cet exemple porte sur ce graphe :

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
_:a foaf:name "Johnny Lee Outlaw" .
_:a foaf:mbox <mailto:jlow@example.com> .
_:b foaf:name "Peter Goodguy" .
_:b foaf:mbox <mailto:peter@example.org> .
_:c foaf:mbox <mailto:carol@example.org> .
```

Requête :

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

```
#### On prend tout name et mbox dans le graph
```

```
SELECT ?name ?mbox
```

```
#### Ou name et mbox sont associes au meme sujet
```

```
WHERE
```

```
{ ?x foaf:name ?name .
  ?x foaf:mbox ?mbox }
```

	name	mbox
Résultats	"Johnny Lee Outlaw"	<mailto:jlow@example.com>
	"Peter Goodguy"	<mailto:peter@example.org>

Ex 4 A partir d'ici, ce seront des exemples de requêtes sur CHEBI, une base de données sur les petites molécules.

```
#### IRI
```

```
PREFIX chebi: <http://purl.obolibrary.org/obo/CHEBI_>
```

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
```

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
```

```
#### Pour obtenir tout les descendants du glucose (id chebi: 17234)
```

```
#### On recupere les descendants, avec label si il existe.
```

```
SELECT DISTINCT ?desc (str(?descLabel) as ?label)
```

```
WHERE{
```

```
####Qui sont sous classe du glucose
```

```
?desc rdfs:subClassOf+ chebi:17234.
```

```
#### si il existe un label du descendant, il est retourne
```

```
OPTIONAL {?desc rdfs:label ?descLabel.}
```

```
}
```

```
#### Pour obtenir tout les ancetres du glucose:
SELECT DISTINCT ?anc (str(?ancLabel) as ?label)
WHERE {
    chebi:17234 rdfs:subClassOf+ ?anc .
    ?anc rdf:type owl:Class .
    OPTIONAL { ?anc rdfs:label ?ancLabel . }
}
```

```
#### Ancetres communs du glucose et myricoside
SELECT DISTINCT ?anc (str(?ancLabel) as ?label)
WHERE {
    chebi:17234 rdfs:subClassOf+ ?anc .
    chebi:7055 rdfs:subClassOf+ ?anc .
    ?anc rdf:type owl:Class .
    OPTIONAL { ?anc rdfs:label ?ancLabel . }
}
```