

Examen du cours Recherche d'information / web sémantique

ENSEEIHT 3A Info – année 2017-2018

Les documents sont autorisés pour les 2 parties. Merci d'utiliser des feuilles différentes pour chaque partie afin de faciliter la correction.

Partie « recherche d'information »

(durée indicative : 1h – barème 10 pts)

Partie « Web sémantique »

Documents autorisés - durée indicative : 50 mn – barème 10 pts

1. Questions de cours (2 points). Réponses COURTES attendues

1.1 Quel est l'objectif général du web sémantique ? donner trois objectifs plus spécifiques

Rendre accessible les contenus du web (pages html et données) aux programmes informatiques, ce qui passe par

- . la standardisation des formats de représentation des connaissances
- . la constitution d'ontologies, de vocabulaires formels et de bases de connaissances partagées pour décrire les contenus du web
- . l'ajout d'annotations sémantiques aux pages web pour représenter les métadonnées et les contenus sous la forme de graphes de connaissances

1.2 En quoi le web des données est-il une évolution du web sémantique ?

Le web des données applique les mêmes objectifs aux données ouvertes du web. Ces données sont décrites par des vocabulaires plus souvent que par des ontologies. La priorité est mise sur la réutilisabilité et la simplicité des vocabulaires, et ensuite sur le partage des sources ainsi représentées sous forme de données ouvertes et liées. Cette approche favorise

2. Exercices : (8 points)

exercice 2.1 (4 points)



DBpedia

Browse using Formats

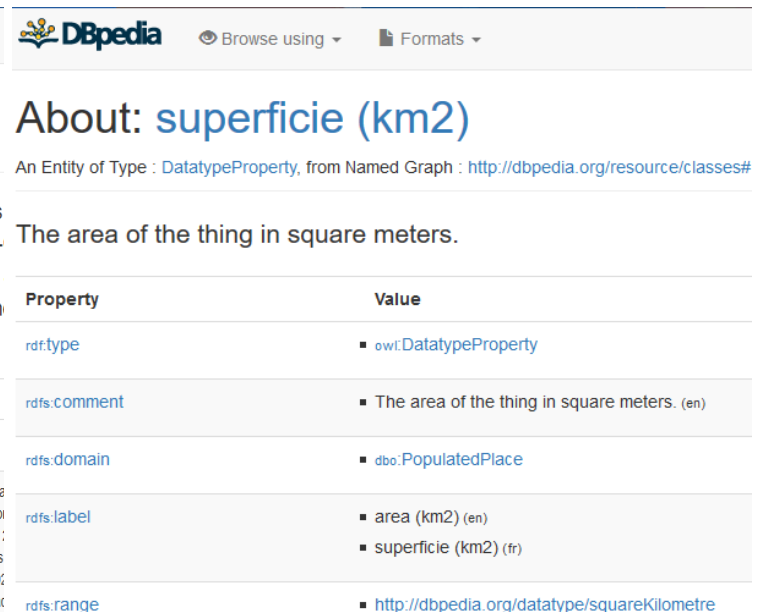
About: Toulouse

An Entity of Type : lieu, from Named Graph : <http://dbpedia.org>, within Data Space : dbpedia.org

Toulouse est une commune du Sud ou Midi de la France. Capitale pendant près de 100 ans capitale historique du Languedoc, elle est aujourd'hui chef-lieu du département de la Haute-Garonne, Languedoc-Roussillon-Midi-Pyrénées et le siège de Toulouse Métropole. Reliant Toulouse à classé au Patrimoine mondial de l'Unesco depuis 1996. La basilique Saint-Sernin, plus grande est également inscrite depuis 1998 au titre des chemins de Saint-Jacques de Compostelle.

Property	Value
dbo:PopulatedPlace/area	118.3
dbo:abstract	Toulouse est une commune du Sud ou Midi de la France. Capitale pendant près de 100 ans capitale historique du Languedoc, elle est aujourd'hui chef-lieu du département de la Haute-Garonne, Languedoc-Roussillon-Midi-Pyrénées et le siège de Toulouse Métropole. Avec 458 298 habitants au 1er janvier, la plus peuplée de France, après Paris, Marseille et Lyon. Ses habitants sont appelés les toulousains. Toulouse regroupe 1 291 517 habitants en 2013, ce qui en fait aussi la quatrième du pays. Avec 92 est la sixième derrière Lille et Nice et devant celle de Bordeaux. Ville à l'architecture caractéristique, Toulouse est surnommée la « ville rose » en raison de la couleur du matériau de construction.

Fig. 1 <http://dbpedia.org/page/Toulouse>



DBpedia

Browse using Formats

About: superficie (km2)

An Entity of Type : DatatypeProperty, from Named Graph : <http://dbpedia.org/resource/classes#>

The area of the thing in square meters.

Property	Value
rdfs:type	owl:DatatypeProperty
rdfs:comment	The area of the thing in square meters. (en)
rdfs:domain	dbo:PopulatedPlace
rdfs:label	area (km2) (en) superficie (km2) (fr)
rdfs:range	http://dbpedia.org/datatype/squareKilometre

Fig.2 <http://dbpedia.org/ontology/PopulatedPlace/area>

<http://dbpedia.org/ontology/area> About: [superficie \(m2\)](#)

An Entity of Type : [DatatypeProperty](#), from Named Graph : <http://dbpedia.org/resource/classes#>, within Data Space : dbpedia.org

Property	Value
rdf:type	<ul style="list-style-type: none"> rdf:Property owl:DatatypeProperty
rdfs:comment	<ul style="list-style-type: none"> The area of the thing in square meters. (en)
rdfs:isDefinedBy	<ul style="list-style-type: none"> http://dbpedia.org/ontology/
rdfs:label	<ul style="list-style-type: none"> area (m2) (en) superficie (m2) (fr)
rdfs:range	<ul style="list-style-type: none"> xsd:double
is owl:equivalentProperty of	<ul style="list-style-type: none"> dbo:landArea ; dbo:waterArea

Figure 3 : <http://dbpedia.org/ontology/area>

- Quels sont tous les Types de chacune de ces entités ? Ecrire en Turtle cette information.
<http://dbpedia.org/page/Toulouse> [rdf:type](#) <http://dbpedia.org/ontology/lieu>
<http://dbpedia.org/ontology/PopulatedPlace/area> [rdf:type](#) [owl:DatatypeProperty](#)
<http://dbpedia.org/ontology/area> [rdf:type](#) [rdf:Property](#)
<http://dbpedia.org/ontology/area> [rdf:type](#) [owl:DatatypeProperty](#)

en turtle
[dbp:Toulouse](#) [a](#) [dbo:lieu](#) .
[dbo:PopulatedPlace/area](#) [a](#) [owl:DatatypeProperty](#) .
<http://dbpedia.org/ontology/area> [a](#) [rdf:Property](#), [owl:DatatypeProperty](#) .
- Etant données les valeurs des propriétés [rdfs:domain](#) et [rdfs:range](#) sur la figure 2, et des informations de la figure 1 : si on lance un raisonneur, quel nouveau type serait associé à <http://dbpedia.org/page/Toulouse> ? et que représente « 118,3 » ? expliquer la sémantique de [rdfs](#) que vous utilisez pour affirmer cela.
<http://dbpedia.org/ontology/PopulatedPlace/area> en plus du type actuel
la surface en km²
- Si la superficie de <http://dbpedia.org/page/Toulouse> sur la figure 1 avait été donnée avec la propriété <http://dbpedia.org/ontology/area>, quelle en aurait été la valeur ? Si on lançait alors un raisonneur, quel type serait associé à <http://dbpedia.org/page/Toulouse> ?
118,3 x 10⁶ m²
le même que pour le moment
- Que diriez-vous du [rdfs:comment](#) de la figure 2 ?
Il est faux

Exercice 2.2 (4 points)

Soit un extrait de la représentation de Toulouse dont le début est présenté figure 1.

```
@prefix dbr: <http://dbpedia.org/resource/> .
dbr:Toulouse rdf:type      dbo:Location , dbo:Place , yago:YagoPermanentlyLocatedEntity .
dbr:Toulouse dbp:population 461190 ;
                dbo:major    dbr:Jean-Luc_Moudenc ;
```

```
dbp:website <http://www.toulouse.fr/> ;
dbo:country dbr:France .
rdf:label    Toulouse
```

- a. Que retourne la requête SPARQL suivante :

```
SELECT ?truc ?label_truc
WHERE { ?truc dbo:country dbr:France .
        ?truc rdfs:label ?label_truc . }
ORDER BY ?label_truc
```

Elle retourne toutes les entités localisées en France (pour lesquelles on a la relation `dbo:country dbr:France`), et affiche l'étiquette (`rdfs:label`) de ces entités. Les résultats sont classés par ordre alphabétique de leur étiquette.

- b. Ecrire une requête SPARQL qui retourne les 100 premiers endroits (`dbo:Location`) situés en France, leur nom et leur population. Cette requête retourne-t-elle `dbr:Toulouse` ? pourquoi ?

```
SELECT ?lieu ?label_lieu ?pop
WHERE { ?lieu rdf:type dbo:Location .
        ?lieu dbo:country dbr:France .
        ?lieu rdfs:label ?label_lieu .
        ?lieu dbp:population ?pop .}
LIMIT 100
```

Non car `dbr:Toulouse` n'est pas de type `dbo:Location`

- c. Ecrire une requête SPARQL qui retourne tous les endroits situés en France, leur population et dont la population est supérieure à 10 000 (habitants), classés par leur taille de population. Cette requête retourne-t-elle `dbr:Toulouse` ? pourquoi ?

```
SELECT ?lieu ?pop
WHERE { ?lieu rdf:type dbo:Location .
        ?lieu dbo:country dbr:France .
        ?lieu dbp:population ?pop .}
FILTER (?pop >= 10000)
ORDER BY ?pop
LIMIT 100
```

Non, toujours pour la même raison

- d. Ecrire une requête qui cherche tous les lieux d'habitation (`dbo:Settlement`) vérifiant ces mêmes critères. Cette requête permet-elle de retourner `dbr:Toulouse` ? Pourquoi ? Si vous répondez non, pourquoi et que faut-il faire pour que la réponse soit Oui ?

```
SELECT ?lieu ?pop
WHERE { ?lieuHab rdf:type dbo:Settlement .
        ?lieuHab dbo:country dbr:France .
        ?lieuHab dbp:population ?pop .}
FILTER (?pop >= 10000)
ORDER BY ?pop
LIMIT 100
```

Non toujours pas. Il faudrait rajouter `dbr:Toulouse rdf:type dbo:Settlement`