

Technologies du Web sémantique

Project 1 - Publish integrated data for a tourist area

The general goal of this project is to produce semantically rich description of a tourist area (e.g. a ski resort, a historical city, a tourist trail). For this you must integrate/interconnect data from

- Open Street Map (OSM): a set of geographic entities covering the whole earth
- DBpedia: an RDF graph derived from Wikipedia and Wikidata
- at least one additional source

To do

- Select a tourist area (resort, city, trail, ...)
- Create an initial RDF schema for the data you are interested in
- Find at least one data source (in addition to OSM and DBpedia) that is specific to the selected area
- Create an RDF graph by transforming data from the sources to RDF triples and importing them into a triple store (GraphDB). Your graph must represent a significant number of objects of interest (hotels, skilifts, hiking paths, museums, ...)
- Write SPARQL queries that demonstrate that your graph integrates data from the selected sources
- Write a short report with the following structure
 - introduction: selected area and information you want to collect
 - data sources (≤ 1 page)
 - RDF schema (in a graphical form)
 - source transformation methods (≤ 1 page/source)
 - queries and their results on the obtained graph

Technical information

Open Street Map

Open Street Map data can be loaded with the following URL scheme:

<https://api.openstreetmap.org/api/0.6/map?bbox=longitude-min,latitude-min,longitude-max,latitude-max>

for example

<https://api.openstreetmap.org/api/0.6/map?bbox=11.54,48.14,11.543,48.145>

More information at https://wiki.openstreetmap.org/wiki/Downloading_data .

An OSM file describes geographic entities either as nodes

```
<node id="369314798" visible="true" version="2" changeset="29685197"
  timestamp="2015-03-23T18:38:56Z" user="tartiflex" uid="507417"
  lat="46.2523028" lon="6.5775254">
  <tag k="ele" v="1650"/>
  <tag k="name" v="Pointe des Riandets"/>
  <tag k="natural" v="peak"/>
  <tag k="source" v="cadastre-dgi-fr"/>
</node>
```

or as “ways”

```
<way id="93493533" visible="true" version="4" changeset="71444385"
timestamp="2019-06-20T13:21:01Z" user="Yugo74" uid="1887927">
  <nd ref="1083904565"/>
  <nd ref="1083904609"/>
  <nd ref="5777821791"/>
  <nd ref="1083904616"/>
  <nd ref="1083904577"/>
  <nd ref="1083904584"/>
  <nd ref="1083904561"/>
  <nd ref="1083904565"/>
  <tag k="addr:city" v="Boège"/>
  <tag k="addr:housenumber" v="227"/>
  <tag k="addr:postcode" v="74420"/>
  <tag k="addr:street" v="Rue de la Vallée Verte"/>
  <tag k="brand" v="Spar"/>
  <tag k="brand:wikidata" v="Q610492"/>
  <tag k="brand:wikipedia" v="en:Spar (retailer)"/>
  <tag k="building" v="supermarket"/>
  <tag k="building:levels" v="1"/>
  <tag k="name" v="Spar"/>
  <tag k="official_name" v="Spar Boège"/>
  <tag k="operator" v="Spar"/>
  <tag k="phone" v="+33 4 50 39 10 49"/>
  <tag k="shop" v="supermarket"/>
  <tag k="source" v="cadastre-dgi-fr "/>
  <tag k="website" v="https://magasins.spar.fr/fr/s/spar-boege"/>
</way>
```

The complete description is in https://wiki.openstreetmap.org/wiki/OSM_XML
and <https://wiki.openstreetmap.org/wiki/Elements> .

DBpedia

The dbpedia graph use the vocabulary defined in the DBpedia ontology. This ontology can be browsed from <http://dbpedia.org/ontology/Place> (Place is probably the most important class for our application)

To access a SPARQL endpoint, in this case dbpedia, from a python program use the *sparqlwrapper* library. Everything is explained in <https://github.com/RDFLib/sparqlwrapper> , with an example.

To interactively query dbpedia go to <http://dbpedia.org/sparql>