Topic: Interlinking the Swiss Feed Database with Linked Open Data Cloud

Recent technological advancements have substantially reduced the cost of generating and storing data. Web 2.0 is now full of a lot of data (either crowd sourced like Tumblr and Wikipedia or contributed by official organizations like Eurostat). In order to make this data better exploitable technologies of Semantic Web and Linked Data have recently started being developed. According to the W3C, "The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries", while Linked Data is a term describing best methods of publishing structured data so that it can be interlinked.
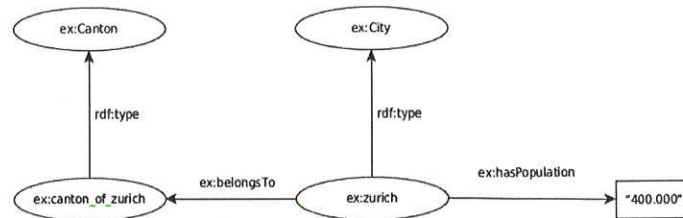

Figure 1: An RDF graph about Zurich

In the context of Semantic Web data are represented using the graph model RDF[1]. The Resource Description Framework (RDF) is a language for representing information about resources in the World Wide Web. An example of the following group of statements "City of Zurich belongs to the canton of Zurich and has population 400,000 people." could be represented as the RDF graph in Figure 1.

In the context of RDF, data properties are described by ontologies (using OWL[2] and RDFS[3] knowledge description languages) which can easily evolve, if needed, without requiring all the data consumers to be changed. A simple ontology describing data similar to the RDF graph of Figure 1 is shown in Figure 2.
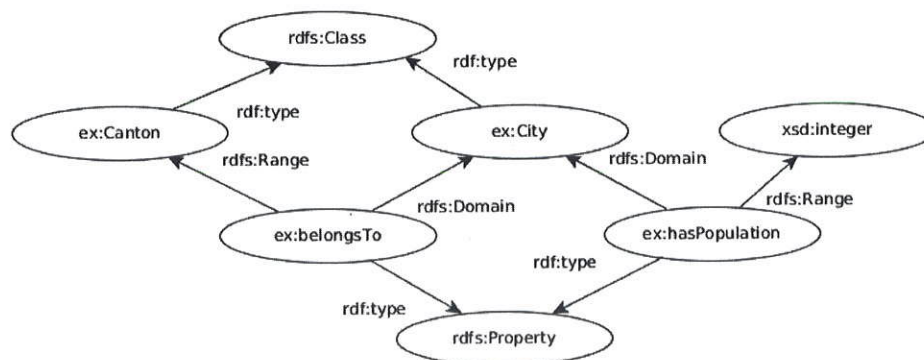

Figure 2: An RDFS ontology describing properties of cities and cantons

Goal of this project is to get familiar with Semantic Web and Linked Data technologies and relevant tools. For this purpose student should accomplish the following tasks:

---

1   http://www.w3.org/2007/02/turtle/primer/
2   http://www.w3.org/TR/2009/REC-owl2-overview-20091027/
3   http://www.w3.org/TR/rdf-schema/

1. Construct an ontology in OWL (using the tool Protégé[4]) that describes properties of the Swiss Feed Database[5].
2. Transform the dataset of the Swiss Feed Database to RDF using standard tools (e.g., GeoTriples[6], D2R[7]).
3. Optional: Interlink the dataset with one of the datasets provided by the Linked Open Data Cloud (e.g., LinkedGeoData, GeoNames), using relevant tools (e.g., Silk[8]).
4. Write a report about performing previous tasks and demonstrate the new RDF dataset and the interlinking with an auxiliary dataset, if done, by posing various queries (using SPARQL, GeoSPARQL) using RDF stores (e.g., Strabon[9], Virtuoso[10]).

4   http://protege.stanford.edu/
5   http://www.ifi.uzh.ch/dbtg/research/sfdb.html
6   http://sourceforge.net/projects/geotriples/
7   http://d2rq.org/
8   http://silk-framework.com/
9   http://www.strabon.di.uoa.gr/
10 http://virtuoso.openlinksw.com/