

EIS / SemWeb Lab 2016

Topic: Semantic Lifting of Budget Data

Mentor: Fathoni Musyaffa, musyaffa@iai.uni-bonn.de

Team:

- Florian Weile, weile@cs.uni-bonn.de (team leader)
- Tatiana Novikova, s6tanovi@uni-bonn.de
- Aberham Gebreyohannes, s6abgebr@uni-bonn.de
- Samuel Y. Ayele, ayele@informatik.uni-bonn.de

Topic description

Transforming available real government budget data from raw data (tabular/XML) into a predefined OpenBudgets.eu RDF data model. The datasets mostly contain "code lists", a specific code that is reusable across the datasets, for example EUR code encodes Euro currency in the currency code list. Code lists in the datasets needs to be extracted and interlinked with Linked Open Data sets like DBpedia and other already existing code lists. Open-source tools to perform the extraction, transformation, loading, and quality assurance of the dataset can be used as a backend (e.g. LinkedPipes ETL, UnifiedViews). A validation and the publication of the transformation result following Linked Data principles is also part of this task. The overall goal of this task is to publish budget data following Linked Data principles to foster transparency in the government and improve participatory budgeting by the citizen.

Preparation Tasks

1. Understand the datasets
 - a. See the examples at <https://github.com/openbudgets/datasets>
2. Play with the linkedpipes demo at: <http://etl.linkedpipes.com>
 - a. Read the documentation at: <http://etl.linkedpipes.com/documentation/>

Goal of the Lab

1. Implement transformation of government data published in various formats like xml and csv to into the openbudgets.eu RDF data model.
2. Implement a user interface for non technical users for the above task.
 - a. Using the REST service of <http://etl.linkedpipes.com/>

Problems

- The data published by the governments are in various formats (xml, csv)
- The governments publish the data in the official language of the respective countries.

TODO

- Check out www.openbudgets.eu project.
 - Check out <http://openbudgets.eu/about/deliverables/>
 - Deliverables interesting for the lab topic are 1.2, 1.3 and 1.4:
 - <http://openbudgets.eu/assets/deliverables/D1.2.pdf>
 - <http://openbudgets.eu/assets/deliverables/D1.3.pdf>
 - <http://openbudgets.eu/assets/deliverables/D1.4.pdf>
 - Deliverable 1.1 is interesting as well, it describes how the data format for openbudget.eu was chosen.
- Check out <http://etl.linkedpipes.com/> project.
- Prepare a Presentation for next weeks meeting on thursday.