Distributed Open Data for Research and App Development

An ever growing number of open datasets becomes available online, offering great possibilities for gaining new knowledge and performing comparative studies. Yet how complex is it in practice to match and query the data scattered across multiple open data sources? In this evaluation study you will first come up with interesting queries that can be answered using data available online, e.g. via SPARQL Endpoints, and then analyze the bottlenecks and limitations of existing tools for working with the open data in various scenarios. For those who like to code: develop an app combining data from disparate open datasets. The goal of the study is to generate ideas of tools and services that could make Open Data more accessible for researchers and app developers.

Research problem
Perform empirical study to analyze bottlenecks and issues arising with performing research using Open and Linked Data. Formalize your findings as quality metrics characterizing the accessibility of Open Data.

References


A list of SPARQL implementations (including client software and libraries): http://www.w3.org/wiki/SparqlImplementations

Supervisor: Dr. Vadim Savenkov