

First Semantic Web Experiments

Petr Křemen

October 4, 2018

1 Introduction

The goal of this seminar is to become familiar with existing open linked data resources and the GraphDB platform.

2 Linked Data

2.1 Exploring DBPedia

DBPedia is a machine-readable dataset extracted from Wikipedia, published as Linked Data.

1. Explore the resource http://dbpedia.org/page/The_Shining_%28film%29. Notice that properties and their values have also their linked data identifiers.
2. Click on the `dbo:director` link and explore the property.
3. Using the linked data exploration, find two other movies directed by the same person as this one.
4. Try to lookup other DBPedia resources of your interest at <http://dbpedia.org/fct/> and explore them.

2.2 Exploring LinkedMDB

LinkedMDB is a machine-readable dataset about movies extracted from IMDB, Freebase and other sources.

1. Explore the resource <http://data.linkedmdb.org/resource/film/2014>.
2. Check `owl:sameAs` properties – they link the resource to other resources (from other dataset).
3. Try to find out how many writers does the movie have and what are their names.

3 Experiments with Open Data

Take a look at the CSSZ open data catalog (<https://data.cssz.cz/web/otevrena-data/katalog-otevrenych-dat>).

Ex. 1 — Get the CSVs of the datasets Počet nově přiznaných invalidních důchodů v České republice dle skupin diagnóz <https://data.cssz.cz/web/otevrena-data/-/invalidita> and Posudky provedené lékařskou posudkovou službou <https://data.cssz.cz/web/otevrena-data/-/posudky-provedene-lps>. Try to find out what is the ratio of newly approved disability pensions comparing to the total number of conducted medical reports related to disabilities in individual years. You can use any tool you are fluent in (Google Drive, Excel, R, Matlab, etc.)

Answer (Ex. 1) — For example, a simple solution in Google Spreadsheet is available at <https://docs.google.com/spreadsheets/d/10oX-noLtx7JwIRbDPZMvK3aSqIp5M1WnJJb/edit?usp=sharing>

Ex. 2 — Take a look at the resource https://data.cssz.cz/resource/observation/invalidita/2011/pk_id/t/60-64. After dereferencing this resource, you get an RDF document, representing one row in the dataset <https://data.cssz.cz/web/otevrena-data/-/invalidita>. Try to find out the original row in the CSV file.

Answer (Ex. 2) — Row No. 3125

4 Relevant References

- Národní katalog otevřených dat (NKOD) <https://data.gov.cz/datov-sady>.