

Techniques for Large Scale Data - Group 23

Julia Szulc - juliasz@student.chalmers.se
Ahmed Groshar - gusgroah@student.gu.se

May 17, 2020

Assignment 4

1.1

The ontology is in `assignemnt4.owl` file.

1.2

Files contents

- `script/queries.py` - strings with Wikidata queries
- `script/data_module.py` - a script obtaining Wikidata records using pre-defined queries
- `script/ontology_creator.py` - a script generating an `.owl` file with a complete ontology

Solution

At first Wikidata queries were constructed using Wikidata Query Service to validate them. Then a Python code that incorporated the queries was written. To communicate with Wikidata server a library `requests` was used. The data obtained this way was then parsed to `.csv` files using `pandas` module.

To create a complete ontology populated with the fetched data, `owlready2` module and an ontology sketch generated before with Protégé were used. The individuals were created by iterating over Wikidata records. Their properties and relationships were set up by defining corresponding triples. The complete ontology was exported to `data/onto.owl` file and loaded in GraphDB.

1.3

The sample SPARQL queries along with comments can be found in `queries` directory.

2.1

We found all Swedish cities that can claim that no more northerly cities have a larger population and same for southerly cities. Below we present the results in tables with three columns (the city name, latitude and population) where the cities are listed according to their latitude. The northerly cities are listed from the north most city and the southerly cities begin from the most southern one.

City	Latitude	Population
Mattila	24.12138889	77
Purra	24.11805556	101
Haparanda Town	24.117033055	3350
Haparanda	24.117032965	6642
Kalix	23.177674141	8395
Luleå	22.191498303	48749
Umeå	20.270566462	89607
Stockholm	18.068611111	972647
Øresund Region	12.816666666	3852993

Table 2.1.1 Northern Cities Table

City	Latitude	Population
Kebal	11.000277777	384
Strömstad	11.186365606	7355
Lysekil	11.451404637	7869
Hönö	11.649711717	9137
Torslanda	11.766666666	22923
Uddevalla	11.914195436	36121
Gothenburg	11.980966694	600473
Øresund Region	12.816666666	3852993

Table 2.1.2 Southern Cities Table