Homework 4: SPARQL

Due Date: Friday, 2016/10/07 @11:59pm on Blackboard.

Summary

The goal of this homework is to write SPARQL queries to retrieve data from dbpedia and become familiar with RDF/S data.

Task 1

Write queries to retrieve data on entities of the **Artist** class and its subclasses (http://dbpedia.org/resource/Artist) in dbpedia according to the following specifications (Outputs are underlined, class names are **bolded**). Use rdfs:label for the names of people.

1. Get the names of all **Painter** (10 points)

2. How many **Painter** are in dbpedia? (10 points)

3. What are the subclasses of **Artist**? (10 points)

4. Get the English names of all **Actor** that starred in the movie "A Christmas Tale" (http://dbpedia.org/resource/A_Christmas_Tale) sorted in ascending order by name. (15 points)

    Hint: You can use lang(?yourNameVariable) to get the language.

5. Get the English names and the birth dates of all **Artist** born before the year 1800 with "David" in their name, sorted in descending order of their birth date. (15 points)

6. Get the English names of all **Artist** that plays exactly 10 kinds of instruments, but does not play Guitar (http://dbpedia.org/resource/Guitar). Sort the results in ascending order of names. (20 points)

7. Provide another interesting query related to the **Artist** class and list the returned results. This query should be substantially different from the previous. (20 points)

Task 2

Run the queries you designed from Task 1 with the following command:

curl -X POST -H "Accept: text/csv" --data-urlencode "query@query1.sparql" http://dbpedia.org/sparql

where query1.sparql is the text file containing the query of question 1. Please make sure you can successfully run the command with your query.

You can also explore dbpedia by going to the web interface of its SPARQL endpoint:

http://dbpedia.org/sparql

Hint: the classes in dbpedia have a namespace "dbo" (e.g., dbo:Painter) that the web interface already knows about.

or as webpages:

http://live.dbpedia.org/page/Francisco_Goya

Submission Instructions

You must submit the following files (totally 8 files) in a single .zip archive named `Firstname_Lastname_hw4.zip` and submit it via Blackboard:

- The PDF file `Firstname_Lastname_hw4.pdf` containing the queries and its first 10 results. Also describe the 7th query you designed briefly.
- The 7 query files
  - Query1.sparql
  - Query2.sparql
  - Query3.sparql
  - Query4.sparql
  - Query5.sparql
  - Query6.sparql
  - Query7.sparql

Ground Rules

This homework must be done individually. You can ask others for help with the tools, but the submitted homework needs to be your own work. **We will be running sophisticated scripts on all the submissions to detect potential cheaters.**