

Knowledge Technologies
Fall Semester 2023
Bonus +5% of the course mark
Announced: February 7, 2024
Due: February 29, 2023 before 23:59

Description

Background

Recent advances in satellite technology have led to frequent and high-resolution monitoring of Earth at a global scale, providing an unprecedented amount of Earth observation (EO) data. Through the Copernicus programme, Sentinel satellites reach the scale of more than 10 TB data per day leading to the creation of gigantic repositories of Earth observation images. Naturally, navigating such repositories manually is cumbersome. This has led to the creation of digital assistants and Question Answering engines that target such image archives to make Earth observation data more accessible. This family of tools take as input natural language questions and answer appropriately depending on their design.

Assignment

In this bonus assignment, you are asked to write ten (10) natural language questions that are given as input to a Question Answering engine that targets a Knowledge Graph which contains geo-entities and satellite images of Earth. The engine is then tasked with querying the Knowledge Graph, using SPARQL/GeoSPARQL, and providing an appropriate answer.

In addition to writing new questions, you are also asked to note the expected return type of the question. The expected return type can be any combination of:

$\{Name, Coordinates, Property(), Count(), Image, Boolean, AVG(...), SUM(...)\}$

depending on the question. You should also specify the expected number of tuples returned if it is necessary to correctly answer the question (which would be equivalent to a LIMIT clause in SPARQL).

For example the question “*Where is Athens located?*” has an expected return type of “Coordinates”, since we are asking for the location of a geoentity. While the question “*What is the average lake size of each country in Europe?*” has an expected return type of “Name, AVG(Property(area))”. There is no need to specify a LIMIT clause in either question. However, for the question “*What are the 5 largest rivers in France?*”, the return type would be “Name”, with a limit of 5.

We provide several examples to help you understand the capabilities of the system and to be used as a quality target for your questions. We want you to try and compose varied and complex questions. You can take inspiration from the given examples, but you should strive to write unique and interesting questions. Your questions should request the return of multiple values and employ complex temporal or geospatial filters.

Knowledge Graph

The target knowledge graph (KG) contains information about named, natural geofeatures (mountains, lakes, rivers, hills, plains, forests, etc.) and human-created geofeatures (dams, cities, towns, administrative divisions, countries, archaeological sites, town squares, etc.). There are available geometries for all geofeatures, as well as thematic properties (the height of mountains, the surface area of any geofeature, the population of human settlements, whether a city is a capital or not, etc.). In addition, the KG includes satellite images that comprise of the area that has been captured (as a geometry), the time of capture and information about the contents of the image. The available image information is:

- vegetation coverage
- ice coverage
- snow coverage
- oil spill coverage
- waterbody coverage
- cloud coverage
- medium/high cloud probability
- which satellite produced the image
- when was the image produced
- sun elevation in the satellite image
- sun azimuth in the satellite image

Submission guides

We expect you to submit a .csv file like the one in the example (with 3 columns instead of 5) with the following format:

Question, Return type, Limit

Support

Sergios - Anestis Kefalidis (s.kefalidis[at]di.uoa.gr) will be supervising this assignment. Please submit your questions on Piazza under the corresponding directory (bonus).