FACULTY OF ELECTRICAL ENGINEERING

Combining cycling with public transport in Prague (Checkpoint 2)

OSW - Ontologies and Semantic Web

Michal Cvach
January 6, 2019

1 TASK ONTOLOGY

For the final checkpoint, I am submitting a **task ontology**, which is based on the supplied **OSW Ontology**. Since my data sets are rather simple, I did not need to add too much to the ontology. I basically added a few clases (public transport stops, cycling routes...) and then a few object properties to link the data.

The final ontology can be found in the file `cvachmic_OSW2018_ontology.ttl`.

2 LINKING AND QUERIES

The data are linked based on their spatial coordinates (latitude and longitude) in **Krovak** (102067) spatial system. Now we can however extract additional knowledge, which I tried to demonstrate with my three queries.

- The first query (`query1.sparql`) shows basically the simplest use case, where we want to find all public transport stations, which allow bicycle transport (which means all means of transport excluding the bus based on information from PID [1]) that are close to some cycling route.
- The second query (`query2.sparql`) is a little bit more specific in a sense that we only care about tram stations, and we only care about those, that work during the night time.

- The third query (`query3.sparql`) is about finding all public transport stations that allow bicycle transport, that are in close proximity to some chosen coordinates (in this scenario close to −1045164.02 latitude and −735544.04 longitude in the Krovak spatial system).

3 Discussion and conclusion

I currently treat the spatial information like normal data, which brings a lot of problems. Mostly the performance is not really great, because I am treating the latitude and longitude information like normal doubles, so I have to do some mathematics all the time. It would have been a better idea to use the support of spatial objects provided by the **GeoSPARQL** plugin. With the use of **contains** and **buffer**, I would probably achieve better results. Unfortunately, I was not able to find a simple and fast way to transform my data sets (which were in **GeoJSON**) into RDF with geometries in a good way using **OntoRefine** or other tools I know, so I decided to stick with my representation from previous checkpoints, even though it would have to change for real use.

I have also made some changes to the data for this checkpoint in OntoRefine, so I am also submitting new OntoRefine projects and insert queries. This mostly meant I have added some additional columns and I have changed some values in some columns from meaningless numbers to string literals or boolean values.

To conclude this all up, I think I have obtained a sensible idea of why we might want to link our data and why it might be a good idea to then use **Ontologies** in combination with our data to obtain new information. I have also realized that my topic choice and some decisions in previous checkpoints maybe were not the greatest. But all in all it was a good experience. The outcome of my semestral work unfortunately is not of the quality I have hoped for, but I have definitely learned some valuable lessons throughout the process.

References

1. Available also from: `http://www.dpp.cz/skolem/`.