

Schema Alignment of OpenStreetMap and Knowledge Graphs on Property Level

OpenStreetMap (OSM)¹ is one of the most important sources of Volunteered graph information (VGI). OSM contains geographical data from millions of volunteers for 188 countries. Although OSM has over 6 billion geographic entities, the data is not directly accessible to semantic applications. Knowledge graphs (KGs), in contrast, provide semantic information for real-world entities and can be easily accessed via SPARQL endpoints, but have very limited coverage of geographic entities. Aligning these resources on schema level can boost data fusion and entity alignment that in turn will help real-world applications such as POI recommendation, question answering and real-time navigation systems.

In our previous work, we proposed a neural schema alignment [Dsouza et al. \[2021a\]](#) wherein we mapped OSM tags to KG classes. For example, in Figure 2, we mapped class *mountain* from Wikidata to class `wkgs:Peak`. For faster access to OSM data, we built WorldKG² [Dsouza et al. \[2021b\]](#) knowledge graph that has data about OSM nodes.

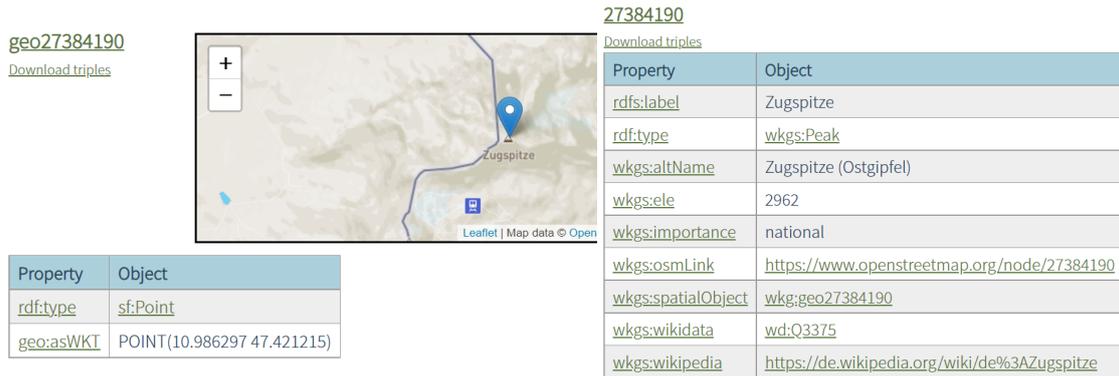


Figure 1: Entity description and map view of *Zugspitze* mountain in WorldKG.

Wikidata			WorldKG		
Subject	Predicate	Object	Subject	Predicate	Object
Q3375	<i>name</i>	<i>Zugspitze</i>	27384190	<code>rdfs:label</code>	<i>Zugspitze</i>
Q3375	<i>coordinate</i>	47.25, 10.59	27384190	<code>geo:asWKT</code>	10.98629747.421215
Q3375	<i>elevation above sea level</i>	2962	27384190	<code>wkgs:ele</code>	2962
Q3375	<i>instance of</i>	<i>mountain</i>	27384190	<code>rdf:type</code>	<code>wkgs:Peak</code>

Figure 2: Snippets of *Zugspitze* entity description in Wikidata and WorldKG.

To build a more complete WorldKG to empower various applications, we not only need to align KGs at the class level, but also at property level. The goal of this thesis is to develop machine learning methods property alignment [Iyer et al. \[2020\]](#), [Xue et al. \[2021\]](#) between the WorldKG knowledge graph and other general-purpose knowledge graphs such as Wikidata and DBpedia. An example alignment for the given example in Figure 2 would look like: `wkgs:ele` property from WorldKG mapped to *elevation above sea level* property from Wikidata. In this thesis, you will:

- Research on state-of-the-art methods for property alignment and design approach for the task in this thesis accordingly.

¹www.openstreetmap.org

²<http://www.worldkg.org/>

- Develop the method to align properties within the WorldKG and Wikidata/DBpedia.
- Evaluate the developed method on large-scale dataset, and compare its performance with state-of-the-art baselines, e.g. from OAEI initiative³.

Prerequisites:

- Knowledge of Semantic Web technologies and Knowledge Graphs, e.g. obtained through successful participation in Advanced Methods of Information Retrieval course, or other relevant courses.
- Theoretical and practical knowledge of machine learning
- Good programming skills

Contact: Alishiba Dsouza

Please email dsouza@cs.uni-bonn.de along with your transcript and a link to GitHub (if available).

References

- Alishiba Dsouza, Nicolas Tempelmeier, and Elena Demidova. Towards neural schema alignment for openstreetmap and knowledge graphs. In *Proc. of the ISWC 2021*, Lecture Notes in Computer Science. Springer, 2021a. URL <http://arxiv.org/abs/2107.13257>.
- Alishiba Dsouza, Nicolas Tempelmeier, Ran Yu, Simon Gottschalk, and Elena Demidova. Worldkg: A world-scale geographic knowledge graph. abs/2109.10036, 2021b. URL <http://arxiv.org/abs/2109.10036>.
- Vivek Iyer, Arvind Agarwal, and Harshit Kumar. Veealign: a supervised deep learning approach to ontology alignment. In *OM@ ISWC*, pages 216–224, 2020. URL http://ceur-ws.org/Vol-2788/oei20_paper13.pdf.
- Xingsi Xue, Chaofan Yang, Chao Jiang, Pei-Wei Tsai, Guojun Mao, and Hai Zhu. Optimizing ontology alignment through linkage learning on entity correspondences. *Complexity*, 2021, 2021. URL <https://doi.org/10.1155/2021/5574732>.

³<http://oei.ontologymatching.org/>