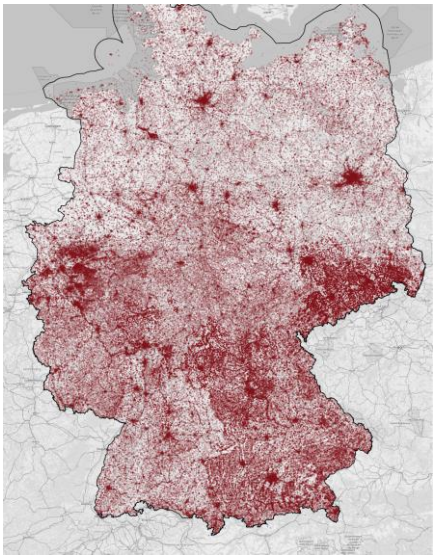
Novel Entity Discovery from Volunteered Geographic Information

Recently, Volunteered Geographic Information (VGI) sources like OpenStreetMap (OSM) (www.openstreetmap.org) became increasingly important for various real-world applications, including navigation and recommendation of Points of Interest. However, the information in OSM lacks a well-defined schema and is only available in specialized data formats. In contrast, Knowledge Graphs such as Wikidata (www.wikidata.org) are built on well-defined ontologies and provide information using standardized formats (SPARQL, RDF), making it easier to utilize this information meaningfully. However, geographic entities captured by popular Knowledge Graphs like Wikidata are far from complete. Therefore, the completion of Knowledge Graphs using VGI is of particular importance.



Geographic Entity Coverage in Wikidata

In this thesis, we tackle the task of novel entity discovery in OSM. This task aims at identifying entities not captured by a reference set of entities in a Knowledge Graph. Related work on novel entity discovery focuses on other types of data, such as tabular data [1]. Related tasks include the alignment of entities from VGI and KGs [2]. In contrast, the discovery of novel entities from VGI addressed in this thesis is currently largely unexplored.

Relevant datasets:
Wikidata, OpenStreetMap

References:

- [1] Shuo Zhang, Edgar Meij, Krisztian Balog, and Ridho Reinanda. 2020. Novel Entity Discovery from Web Tables. In Proceedings of The Web Conference 2020 WWW '20. Association for Computing Machinery, New York, NY, USA, 1298–1308.
- [2] Nicolas Tempelmeier, Elena Demidova: Linking OpenStreetMap with knowledge graphs - Link discovery for schema-agnostic volunteered geographic information. Future Gener. Comput. Syst. 116: 349-364 (2021)

Contact:

Nicolas Tempelmeier tempelmeier@L3S.de