Linking of Named Events in Text with the EventKG Knowledge Graph

The extraction of machine-readable information from text, i.e., Information Extraction, is a crucial prerequisite for many tasks, including knowledge graph population and question answering. One typical task of Information Extraction is Named Entity Linking (NEL), where mentions of entities representing real-world objects (persons, places, companies, …) are linked to the nodes in a target knowledge graph. While this task can be rather straightforward in the case of many entities (e.g., *Barack Obama* was speaking in *Washington, D.C.*"), this task can become particularly challenging in the case of event mentions, as events typically have a more elaborate description that can include dates, locations and related entities, (e.g., "Barack Obama was inaugurated as President of the United States on Tuesday, January 20, 2009" → First inauguration of Barack Obama).

To approach the challenge of Named Event Linking, in this thesis, you will develop an approach that links event mentions in text to the corresponding event representations in the EventKG knowledge graph http://eventkg.l3s.uni-hannover.de/ .

In the first step, you will create a dataset, which focuses on named event annotations in text. Then, you will develop a tool tailored to Named Event Linking that uses parts of this dataset for training and evaluation. In detail, the following tasks will be covered in this thesis:

- Creation of an annotated text corpus that contains named events mentions in text and the links to the event representation in a target knowledge graph (EventKG).
- Analysis of the performance of the state-of-the-art methods for Named Entity Linking (NEL) on the resulting event dataset: Can state-of-the-art NEL tools also deal with named events? Such existing tools will be used as a baseline for the evaluation of the new event linking method to be developed.
- Creation of a Named Event Linking tool: Can NEL tools be improved and deal with events more accurately?

References:
- Gottschalk, Simon, and Elena Demidova. "EventKG: A Multilingual Event-centric Temporal Knowledge Graph." *European Semantic Web Conference*. Springer, Cham, 2018.
- Röder, Michael, Ricardo Usbeck, and Axel-Cyrille Ngonga Ngomo. "Gerbil– Benchmarking Named Entity Recognition and Linking Consistently." *Semantic Web* 9.5 (2018): 605-625.

Datasets:
- EventKG, NEL datasets (AIDA CoNLL-YAGO Dataset, AQUAINT, …)

Contact:
Simon Gottschalk
gottschalk@L3S.de