

# Turning Annotated Recipe Data Into a Knowledge Graph

Semantic Computing Group  
Jan-Philipp Töberg  
jtoeberg@techfak.uni-bielefeld.de

Intelligent agents are challenged by unknown situations in open worlds. They cannot perform everyday tasks like cutting food or setting a table without encountering unknown motions, objects or environments. To mitigate this problem, providing these robots with commonsense knowledge is a possible way to increase their world understanding and support their planning capabilities [1]. However, this household-specific commonsense knowledge is, despite its potential, not often benchmarked or compared between different approaches.

In previous work [2], we collected recipe data and annotated it with aspects relevant for robotic *Table Setting* scenarios (e.g. What type of plate to use? What cutlery to use?). Due to the size of the annotated data, future work investigates how generalised rules can be extracted based on the available data. To increase the rule mining capabilities, this annotated data should be turned into a single knowledge graph, linked to available background knowledge.

In this thesis you will create this knowledge graph and connect it to other, large-scale knowledge graphs providing background knowledge like DBpedia [3]. In this thesis, you need to design the underlying ontology, turn the data into concrete knowledge triples and connect them to background knowledge in other knowledge graphs through entity linking and classification techniques.

Prior knowledge regarding knowledge graphs and ontologies is advised, but not required. You can use the programming language of your choice, but Python is recommended. The thesis can be taken in English or German.

## Related literature

- [1] J.-P. Töberg, A.-C. N. Ngomo, M. Beetz, and P. Cimiano, 'Commonsense knowledge in cognitive robotics: a systematic literature review', *Front. Robot. AI*, vol. 11, 2024, doi: 10.3389/frobt.2024.1328934.
- [2] J.-P. Töberg, S. Kenneweg, and P. Cimiano, 'RoboCSKBench: Benchmarking Embodied Commonsense Capabilities of Large Language Models', in *UR2025*, College Station, Texas, USA, Submitted.
- [3] C. Bizer et al., 'DBpedia - A crystallization point for the Web of Data', *Journal of Web Semantics*, vol. 7, no. 3, pp. 154–165, Sep. 2009, doi: 10.1016/j.websem.2009.07.002.

The Semantic Computing Group researches and develops methods that enable machines to acquire relevant knowledge as well as linguistic capabilities. Using methods from *natural language understanding* and *machine learning*, we are aiming at machines that are capable of knowledge acquisition by reading unstructured textual data. In particular, the group focuses on methods for information extraction, semantic parsing, ontology learning, sentiment analysis, entity linking, as well as question answering.

More information is available at: <http://www.sc.cit-ec.uni-bielefeld.de/>

Interested? @mail to [jtoeberg@techfak.uni-bielefeld.de](mailto:jtoeberg@techfak.uni-bielefeld.de)