

2022-02-03

## Call for Bachelor/Master Thesis: „Creating a Large Knowledge Graph about Scientific Publications for Innovation Forecast“

### What is the goal?

The goal is to create an RDF knowledge graph about scientific publications in a similar way as described in [1][2]. This includes the following steps:

1. Transforming the data of OpenAlex.org into RDF (considering existing vocabulary in the linked open data cloud).
2. Improving the knowledge graph (e.g., machine learning-based author name disambiguation).
3. Automatically interlinking the knowledge graph with existing knowledge graphs in the linked open data cloud (e.g., Wikidata).



Source: <https://openalex.org>

The contents of the Bachelor's/Master's thesis are already well defined.

The student will get the chance to publish the work as a research paper and to provide the knowledge graph on the web in collaboration with industry partners. In this way, the student can gain insights into the scientific communities and practices of IT companies.

### What are the prerequisites?

You should have interest in knowledge graphs and data mining, as well as solid programming skills (e.g., in Python) to deal with large data files.

[1] <http://www.semantic-web-journal.net/system/files/swj2779.pdf>

[2] [http://dbis.informatik.uni-freiburg.de/content/team/faerber/papers/MAKG\\_ISWC2019.pdf](http://dbis.informatik.uni-freiburg.de/content/team/faerber/papers/MAKG_ISWC2019.pdf)

Please send your requests with a transcript  
of records and a short CV to:  
**Dr. Michael Färber**  
**michael.farber@kit.edu**