## Knowledge Graph Completion Leveraging Neural Language Models

Open Knowledge Graphs (KGs) such as DBpedia, Freebase, etc. have been recognized as the foundations for diverse machine learning based applications. Most of these KGs are created either via automated information extraction from Wikipedia snapshots, information accumulation provided by the users or by using heuristics. However, these KGs are often incomplete, i.e., there are missing links between the entities. The goal of Knowledge Graph Completion (KGC) is to automatically predict these missing links. Numerous cutting-edge KGC approaches, including those for data mining, machine learning, and natural language processing, have been published at renowned conferences. However, the impact of Neural Language Models (NLMs) on KGC is still an open research.

Neural Language Models (NLMs) leverage Neural Networks to learn the distributed continuous representations of words in a low dimensional vector space. Different non-contextual NLMs such as Word2vec, GloVe, etc. as well as contextual NLMs such as BERT, etc. have been proposed in the literature which provide the backbone on Natural Language Processing based applications.

The **main focus of this thesis** is to explore if these **NLMs can be leveraged to improve KGC**. The model will be evaluated on benchmark datasets extracted from the open KGs Freebase, WordNet, and YAGO.

Knowledge Graph Completion
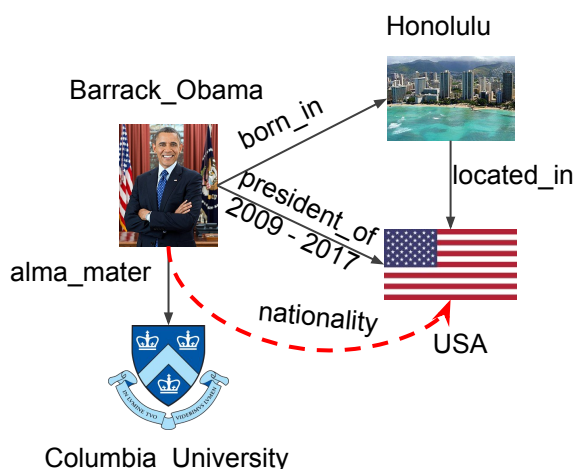
This thesis will be supervised by **Russa Biswas, Dr. Mehwish Alam, and Prof. Dr. Harald Sack, Information Service Engineering at Institute AIFB, KIT, in collaboration with FIZ Karlsruhe**.

Prerequisites
- Good programming skills in Python
- Interest in Deep Learning technologies
- Interest in Machine Learning approaches
- Interest in Semantic Web technologies

Contact persons:

Russa Biswas
russa.biswas@kit.edu

Dr. Mehwish Alam
mehwish.alam@kit.edu

aifb

**Institute of
Applied Informatics and
Formal Description Methods**
http://www.aifb.kit.edu/

KIT
Karlsruhe Institute of Technology