

## Graduiertenkolloquium Angewandte Informatik

### Leveraging Literals for Knowledge Graph Embeddings

M. Sc. Genet Asefa Gesese (FIZ Karlsruhe – Leibniz Institute for Information Infrastructure)

Knowledge Graphs (KGs) are a structured representation of facts pertaining to a specific domain or multiple domains, composed of entities and the relationships between them. They have become essential for managing structured information, with a growing interest in their use in artificial intelligence. KGs can be used to support decision-making and to enhance machine learning tasks such as question-answering, recommender systems, and relation extraction. Some of the most popular publicly available general-purpose KGs include DBpedia, Wikidata, and YAGO. To maximize efficiency, it is beneficial to transform these KGs into a low-dimensional vector space while preserving their underlying semantics. However, KGs are inherently incomplete due to the open-world assumption, meaning there are missing links between entities. This incompleteness negatively impacts their usage for different real-world applications. Hence, there is a need for embedding-based link prediction approaches to perform KG completion tasks. Link prediction involves predicting the likelihood of relationships between entities based on available information in the KG. This task can be executed in two settings: transductive and inductive. In the transductive setting, all entities in the test and validation sets must be included in the training set. Conversely, in the inductive setting, the test and validation sets may contain entities not seen during training.

It is crucial to consider the semantics contained within the KG while performing a link prediction task so as to produce valid facts. Most large-scale KGs, such as Wikidata, comprise a significant number of numerical attributive triples which connect entities to numerical literal nodes. In addition to numerical literals, KGs also contain a substantial amount of both short and long textual literals, such as names, labels, and descriptions of entities and relations. These literals, both numerical and textual, hold crucial information about entities and relations and can be used to generate missing facts. Therefore, developing a link prediction approach that incorporates literals is an important aspect of KG completion.

This talk will focus on discussing link prediction methods, which utilize text and numerical literals, proposed for both transductive and inductive settings. Moreover, benchmark datasets that are created for the evaluation of link prediction methods will also be presented.

**Termin: Freitag, 24.02.2023, 14:00 Uhr**

Ort: Kaiserstr. 89, 76133 Karlsruhe

Kollegiengebäude am Kronenplatz (Geb. 05.20), 1. OG, Raum 1C-04

(Hinweise für Besucher: <http://www.aifb.kit.edu/web/Kontakt>)

Veranstalter: Institut AIFB, Forschungsgruppe Information Service Engineering

Zu diesem Vortrag lädt das Institut für Angewandte Informatik und Formale Beschreibungsverfahren alle Interessierten herzlich ein.

M. Färber, S. Lazarova-Molnar, A. Oberweis, H. Sack (Org.),  
A. Sunyaev, Y. Sure-Vetter, A. Vinel, M. Volkamer, J. M. Zöllner