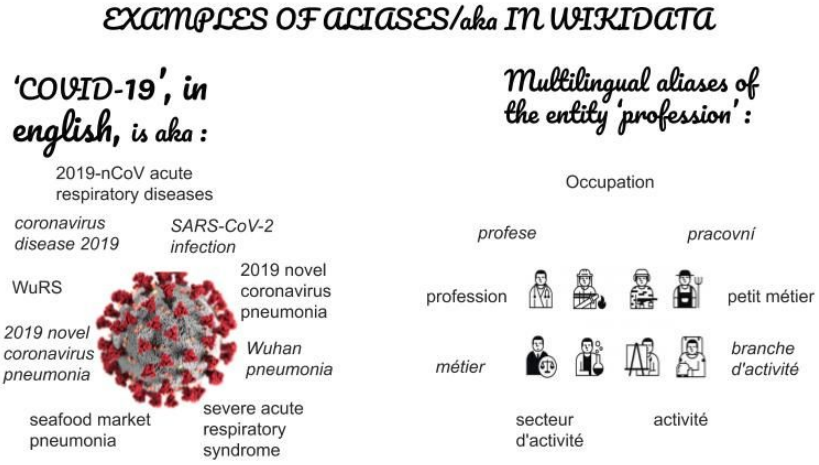
Call for Master Thesis

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

How Well Do Aliases Represent an Entity?

In most Knowledge Graphs (KGs) such as Wikidata [1], it is common to have aliases (also known as) for entity labels or names. For instance in Wikidata, the entity 'COVID-19' has multiple aliases which provide valuable information about the entity (refer to the figure in the right hand side). Moreover, aliases of entities can also be provided in more than one language which may contain complementary information.



For instance, as shown in the figure above, the entity 'profession' has aliases in multiple languages which provide complimentary/additional semantics.

Different KG embedding techniques such as DKRL [2], which map KGs to a low dimensional vector space, have been proposed. Such learned embeddings are usually applied in various downstream tasks such as machine translation and question answering. However, the multilingual aliases of entities have not been leveraged by KG embedding techniques to enhance entity representations.

Therefore, in this thesis, the advantages of leveraging the additional semantics which are present in such aliases for the purpose of KG representation will be investigated.

This thesis will be supervised by **Prof. Dr. Harald Sack and Genet Asefa Gesese, Information Service Engineering at Institute AIFB, KIT, in collaboration with FIZ Karlsruhe**.

[1] https://www.wikidata.org/wiki/Wikidata:Main_Page
[2] https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12216/12004

WIKIPEDIA
The Free Encyclopedia

Which prerequisites should you have?

- Very Good programming skills in Python
- Interest in Machine/Deep Learning technologies

Contact person:
Dr. Genet Asefa Gesese
genet-asefa.gesese@partner.kit.edu
genet-asefa.gesese@fiz-karlsruhe.de

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

KIT Karlsruhe Institute of Technology