

# Transformer-based coreference resolution for German

Coreference resolution (CR) is the task of finding all expressions that refer to the same entity in a text. Transformer-based neural models such as BERT have improved many NLP tasks drastically including coreference resolution. However, at this moment there is no version of a neural network-based approach for the German language. The BA thesis is requested to conduct the following work:

- Research on transformer-based model for CR to identify best fitting approach for German data
- Preprocess German CR data for model training
- Train and evaluate transformer-based CR model
- Wrap model in docker-based microservice architecture
- Publish model and docker image together documentation

Suggested literature / links:

- <https://github.com/mandarjoshi90/coref>
- <https://github.com/huggingface/neuralcoref>