

## Announcement

05.02.2020

### BACHELOR THESIS:

# Automatic Query-Based Extractive Summary Creation for German

#### DESCRIPTION

In recent years, the need for multi-document summarization has increased enormously trying to process the ever-growing volume of information on the world wide web. Because of the fact that manual creation of text summaries by professionals is labor intensive, difficult and expensive to repeat, alternative methods like crowd-based summarization or automated summarization has been tackled in the last decade. Especially in case of domain specific summarization, most currently available automated tools suffer from low summary quality. To overcome this challenge, the aim is in this thesis to compare the quality of all the existing automatic summarization tools and to develop a new automatic summarization tool based on existing tools.

The goal of the thesis is to develop an automatic summarization tool for German language. To realize this, following steps are required:

- Literature analysis of current automatic summarization tools.
- Extension of an existing automatic tool: Design and implementation of a automatic summarization tool for query-based extractive summarization for German language
- Testing this new tool with a real crowdsourcing task.

*The thesis may be written in German or in English (preferred).*

#### REQUIREMENTS

- Good statistical knowledge
- Fluent in one of the programming languages Python (NLP knowledge is a plus)
- Experience with RESTful API and JSON
- Interest in crowdsourcing
- Course of study: computer science, computer engineering, industrial engineering, information systems management, media computer science
- English required, German is a plus

#### CONTACT

Neslihan Iskender, [neslihan.iskender@tu-berlin.de](mailto:neslihan.iskender@tu-berlin.de)