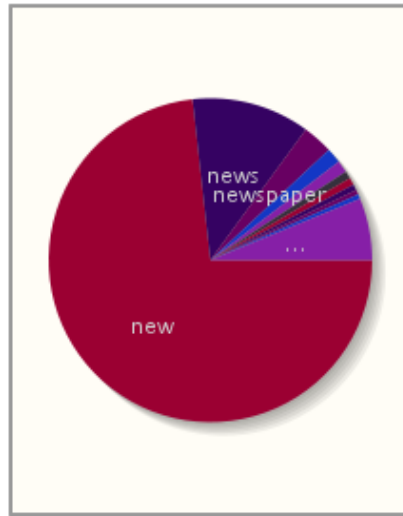
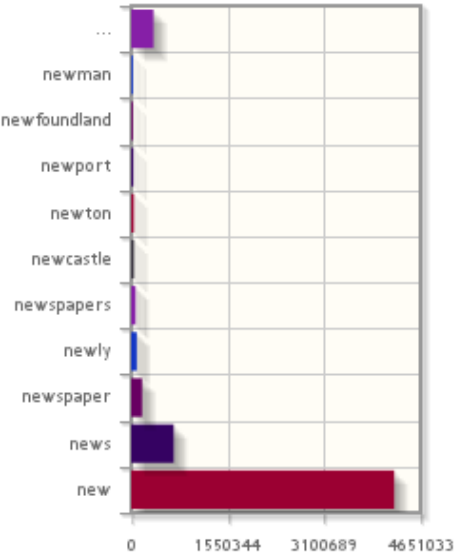


Extraction and analysis of Wikipedia language models: Corpora Explorer (Corpex)



BENEFITS

- Write your thesis
- in a context of the research project RENDER
- in touch with AIFB, Wikimedia and other partners
- with potential further collaboration

**Bachelor
or
Master Thesis**

GOALS

- Extract language models from the different language versions of Wikipedia
- Analyze and visualize the extracted language models
- Extend the Corpora Explorer (Corpex)
- Automate the Corpex extraction and analysis pipeline

REQ.

- Solid understanding of (semantic) web technologies
- Appropriate programming skills
- Good command in English (thesis in English or German)

CANDIDATES

- Students in
- Information Engineering (Informationswirtschaft)
 - Business Engineering (Wirtschaftsingenieurwesen)
 - Economics Engineering (Technische VWL)
 - Computer Science (Informatik)
 - Other subjects areas with background and interest in computer science

PDF/MORE



CONTACT

- Elena Simperl
- elena.simperl@kit.edu
- +49 721 608 45778

