

## Computational Text Analysis

Instructor: Theresa Gessler

10-11-12 May 2022

Over the past years, the digitalization of legal, political, journalistic corpora as well as the growth of online sources has allowed researchers to address new questions across political and social science disciplines. This has led to a growth of new methods to analyze text as data. The aim of this course is to introduce students to the quantitative analysis of textual data. We will discuss the theoretical underpinnings of text-as-data approaches, the implementation of these methods through hands-on exercises using the R statistical programming language and discuss examples of recent empirical research that uses these techniques. The course covers the collection of text data with webscraping techniques, text preprocessing, dictionaries and descriptive analysis of texts, as well as supervised and unsupervised learning methods to classify the content of text corpora.