



---

*Statistical Consulting Project in Collaboration with the LMU Institute of Political Science*

## Twitter in the Parliament: A Text-Based Analysis of German Political Entities

---

The rise in popularity of social media throughout the last two decades has contributed decisively to the era of big data. In fact, the majority of what is called big data is made up of large-scale unstructured data, oftentimes text. Advancements in the area of Natural Language Processing (NLP) have provided new ways to analyze such data to researchers and practitioners alike, for instance in the field of topic modeling. Topic modeling is an instance of unsupervised learning which seeks to identify latent thematic clusters within large-scale unstructured texts.

We construct a dataset containing tweets from German Members of the Bundestag during the 19th election period (up until April 2020), as well as a variety of personal, socioeconomic, and election data and apply the Structural Topic Model (STM) to it. This model allows for the incorporation of document-level metadata. Accordingly, we discuss the technical and practical aspects of estimating relationships between latent topic proportions and observable covariates. Finally, we address the issue of how to make causal inference within a topic modeling context.

---

<b>Date and Time</b>	Thursday, July 16, 2020, 11:00 AM
<b>Location</b>	Zoom <sup>1</sup> (Meeting ID: 913 4778 7905, Password: 569844)
<b>Project Partner</b>	Prof. Paul W. Thurner, LMU Institute of Political Science
<b>Project Supervisors</b>	Prof. Christian Heumann, Matthias Aßenmacher
<b>Speakers</b>	Patrick Schulze, Simon Wiegerebe

---

<sup>1</sup> <https://lmu-munich.zoom.us/j/91347787905?pwd=TDRUVzlwcmVrdG5RVVJ3YnBja2hvUT09>