Where do parties talk about what? Party issue salience across communication channels

Description of replication material

Christoph Ivanusch

Contents

1 Introduction **2**

2 Main scripts **2**

3 Additional scripts **2**

4 Data sets **3**

5 Manual "gold standard" **4**

6 Additional data sets **4**

1 Introduction

This file describes the code and data sets required to rerun and reproduce the analysis reported in the paper "Where do parties talk about what? Party issue salience across communication channels". To run the code successfully, the replication data set needs to be download and stored as a single file (ZIP-file). Overall, the replication material contains the following types of scripts and data sets:

2 Main scripts

The following scripts are necessary for replicating the empirical analyses. The scripts (R, RMarkdown and Python) are named numerically (01-04) and listed in the correct order to perform the analysis.

- *01_Setup.R*: This script loads the correct package versions to make sure that all R code runs successfully.
- *02_Classification.py*: This script performs the cross-domain topic classification with the transformer-based language model BERT.
  Note: I recommend to run this script via Google Colab!
  Attention: To run the script users need to provide an API key for downloading the manifestos!
- *03_Analysis_Preparation.R*: This script creates the main data set used for the statistical analysis.
  Attention: To run the script users need to provide an API key for downloading the manifestos!
- *04_Tables_and_Figures.Rmd*: This script performs the statistical analysis and creates the figures and tables reported in the paper.

3 Additional scripts

.

The following scripts (R) are called from within the main scripts.

- *Validation_ParlSpeeches.R*: This script calculates validation statistics for the topic classification of the parliamentary speeches.
- *Validation_Pressreleases.R*: This script calculates validation statistics for the topic classification of the press releases.
- *Validation_Tweets.R*: This script calculates validation statistics for the topic classification of the tweets.

4 Data sets

The following files contain the individual text corpora analysed in the paper and further data sets. Unfortunately, I am not allowed to publish the texts themselves in the replication material due to copyright reasons. Hence, I can often only provide meta data (e.g. party name, date) on the individual texts in the replication material. More detailed information on each data set and its content is listed below.

- *Corpus_ParlSpeeches.csv*: This file contains all parliamentary speeches held by MPs in the the Austrian, Swiss and German parliaments between January 1, 2019 and September 26, 2021. The corpus is based on an updated version of ParlSpeech V2 data set provided by Rauh and Schwalbach (2020) (for Austria and Germany) and data downloaded from the webservice of the Swiss parliament through the R package *swissparl* (Zumbach 2020).
  Note: Unfortunately, I can only provide meta data in the replication material and not the full texts because of copyright reasons.

- *Corpus_Pressreleases.csv*: This file contains all press releases published by Austrian, Swiss and German parties and their parliamentary party groups (PPGs) between January 1, 2019 and September 26, 2021. The data for Austria is based on webscraping a webservice of the Austrian Press Agency (*APA-OTS*), the data for Germany is based on webscraping of party and PPG websites and the data for Switzerland was provided by the DigDemLab at the University of Zurich (Gilardi et al. 2022).
  Note: Unfortunately, I can only provide meta data in the replication material and not the full texts because of copyright reasons.

- *Corpus_TweetsParty.csv*: This file contains all tweets posted by Austrian, Swiss and German parties and their parliamentary party groups (PPGs) between January 1, 2019 and September 26, 2021. The data was collected through the Twitter Researcher API.
  Note: Unfortunately, I can only provide some meta data (e.g. party name, id) in the replication material and not the full texts because of copyright reasons.

- *Corpus_TweetsPerson.csv*: This file contains tweets posted by party leaders, lead candidates, general secretaries and MPs of Austrian, Swiss and German parties between January 1, 2019 and September 26, 2021. The data was collected through the Twitter Researcher API.
  Note: Unfortunately, I can only provide some meta data (e.g. party name, id) in the replication material and not the full texts because of copyright reasons.

- *Results_ParlSpeeches.csv*: This file contains the classification results for the parliamentary speeches (*Corpus_ParlSpeeches.csv*).

Note: Unfortunately, I can only provide meta data in the replication material and not the full texts because of copyright reasons.

- *Results_Pressreleases.csv*: This file contains the classification results for the press releases (*Corpus_Pressreleases.csv*).
  Note: Unfortunately, I can only provide meta data in the replication material and not the full texts because of copyright reasons.

- *Results_TweetsParty.csv*: This file contains the classification results for the tweets from party accounts (*Corpus_TweetsParty.csv*).
  Note: Unfortunately, I can only provide some meta data (e.g. party name, id) in the replication material and not the full texts because of copyright reasons.

- *Results_TweetsPerson.csv*: This file contains the classification results for the tweets from individual party member accounts (*Corpus_TweetsPerson.csv*).
  Note: Unfortunately, I can only provide some meta data (e.g. party name, id) in the replication material and not the full texts because of copyright reasons.

5 Manual "gold standard"

The folder "ManualGoldStandard" contains samples of manually coded parliamentary speeches, press releases and tweets. These files constitute the manual "gold standard" and are called from the validation files (*Validation_ParlSpeeches.R*, *Validation_Pressreleases.R* and *Validation_Tweets.R*).

- *ManualGoldStandard/ParlSpeeches/Sample_ParlSpeeches.RDS*:This file contains a sample of manually coded parliamentary speeches from Austria, Germany and Switzerland.

- *ManualGoldStandard/Pressreleases/Sample_Pressreleases.RDS*: This file contains a sample of manually coded press releases from parties and parliamentary party groups (PPGs) in Austria, Germany and Switzerland.

- *ManualGoldStandard/Pressreleases/Sample_Tweets.RDS*: This file contains a sample of manually coded tweets from parties and parliamentary party groups (PPGs) in Austria, Germany and Switzerland.

6 Additional data sets

The following data sets contain additional data required during the analysis (e.g. intermediate results, codebook).

- *Main_Data.RDS*: This file contains the main data set used for the analysis. It is produced by the script *03_Analysis_Preparation.R* and called from *04_Tables_and_Figures.R*.

- *Codebook.xlsx*: This file contains information on our codebook and relates the issue categories from the coding scheme to the Manifesto Project (MARPOR) coding scheme (Werner et al. 2021).
- *Twitter_Person_Accounts_List.xlsx*: This file contains information on the accounts from individual party members included in the text corpus.

References

APA-OTS. https://www.ots.at/.

Gilardi, Fabrizio, Lucien Baumgartner, Clau Dermont, Karsten Donnay, Theresa Gessler, Maël Kubli, Lucas Leemann, and Stefan Müller. 2022. "Building Research Infrastructures to Study Digital Technology and Politics: Lessons from Switzerland." *PS: Political Science & Politics* 55 (2): 354–359. https://doi.org/10.1017/S1049096521000895.

Rauh, Christian, and Jan Schwalbach. 2020. *The Parlspeech V2 Data Set: Full-Text Corpora of 6.3 Million Parliamentary Speeches in the Key Legislative Chambers of Nine Representative Democracies*. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/L4OAKN.

Werner, Anika, Onawa Lacewell, Andrea Volkens, Theres Matthieß, Lisa Zehnter, and Leila van Rinsum. 2021. *Manifesto Coding Instructions: 5th Fully Revised Edition.*

Zumbach, David. 2020. *Swissparl: Interface to the Webservices of the Swiss Parliament. R Package Version 0.2.1.* https://CRAN.R-project.org/package=swissparl.