Social Media and Political Agenda Setting

Description of Replication Material, Datasets, and R Scripts

Fabrizio Gilardi Lucien Baumgartner Clau Dermont Karsten Donnay
Theresa Gessler Maël Kubli Lucas Leemann Stefan Müller

In this file, we describe all datasets and R scripts required to rerun and reproduce the analyses reported in the paper and Supporting Information. In order to have the data, scripts, and outputs saved in the correct folders, which is required to run the scripts successfully, we recommend to download all files as a single ZIP file (click the "Access Dataset" button in the right-hand corner and select "Original Format ZIP"). If you only need selected files, make sure to download data in the original file format. Please do not hesitate contacting the authors if you have any questions.

Datasets

The repository for replicating the empirical analyses contains the following datasets. The following files (in alphabetic order) are used in the R scripts listed below. All other files in the repository, which are not listed here are intermediate products of the analysis and are there to save time if only some steps of the analysis should be reproduced.

- *auto_dictionaries_lsd.RData*
- *corona_CdT.RDS*
- *corona_facebook.RDS*
- *corona_smd.RDS*
- *corona_smd_cdt.RDS*
- *corona_twitter.RDS*
- *lsde_frenche_germane.RData*

R Scripts

The repository contains the following R scripts. If you want to reproduce the entire analysis, please load and execute the scripts in the following order:

- *01-01-01-smd-data.R:* script downloading data from database with classification of newspaper articles from the SMD. This script cannot be run since the input data cannot be added to the replication file. We are unable to provide the newspaper articles for replication, due to a binding data protection agreement with the Swiss Media Database. To run the script you would need the following data:
    - *auto_dictionaries_lsd.RData*
    - *lsde_frenche_germane.RData*
    - *corona_smd.RDS*
- *01-01-02-add_CdT_to_SMD.R*: script downloading data from database with classification of CdT articles and combination with SMD data. This script cannot be

run since the input data cannot be added to the replication file. We are unable to provide the newspaper articles for replication, due to a binding data protection agreement with the Swiss Media Database. To run the script you would need the following data:
    - *auto_dictionaries_lsd.RData*
    - *lsde_frenche_germane.RData*
    - *corona_smd.RDS*
    - *corona_smd_cdt.RDS*
- *01-02-01-facebook-data.R*: script downloading data from database with classification of Facebook posts. This script cannot be run since the input data cannot be added to the replication file. We are unable to provide the Facebook posts for replication, due to a binding data protection agreement with Crowdtangle (Facebook). To run the script you would need the following data:
    - *auto_dictionaries_lsd.RData*
    - *lsde_frenche_germane.RData*
    - *corona_facebook.RDS*
- *01-03-01-twitter-data.R*: script downloading data from database with classification of Twitter posts. This script cannot be run since the input data cannot be added to the replication file. We are unable to provide the Twitter posts for replication, due to a binding data protection agreement with Twitter. To run the script you would need the following data:
    - *auto_dictionaries_lsd.RData*
    - *lsde_frenche_germane.RData*
    - *corona_twitter.RDS*
- *02-01-01-figures.R*: This script produces Figure 2 (Figure 1 is a flow chart that does not require any code). To run this script, you need the following files:
    - *ggplot_theme_ddl.R*
    - *corona_twitter.RDS*
    - *corona_facebook.RDS*
    - *corona_smd_cdt.RDS*

The beginning of each R script lists the packages that need to be loaded and the package versions that have been used when running the scripts.
The entire analysis has been successfully executed using the following platform and package versions:

```
R version 4.1.0 (2021-05-18)
Platform: x86_64-apple-darwin17.0 (64-bit)
Running under: macOS Big Sur 10.16

Matrix products: default
LAPACK: /Library/Frameworks/R.framework/Versions/4.1/Resources/lib/libRlapack.dylib

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

attached base packages:
[1] parallel  tools     stats     graphics  grDevices utils     datasets  methods
base

other attached packages:
 [1] magrittr_2.0.1    quanteda_3.0.0    pbmcapply_1.5.0   data.table_1.14.0
jsonlite_1.7.2
 [6] rjson_0.2.20      elastic_1.2.0     showtext_0.9-2    showtextdb_3.0
sysfonts_0.8.3
```

```
[11] cowplot_1.1.1     urltools_1.7.3    glue_1.4.2        ggExtra_0.9
ggeffects_1.1.0
[16] jpeg_0.1-8.1      png_0.1-7         forcats_0.5.1     stringr_1.4.0
purrr_0.3.4
[21] readr_1.4.0       tidyr_1.1.3       tibble_3.1.2      ggplot2_3.3.3
tidyverse_1.3.1
[26] dplyr_1.0.6

loaded via a namespace (and not attached):
 [1] httr_1.4.2         modelr_0.1.8       RcppParallel_5.1.4 shiny_1.6.0
 [5] assertthat_0.2.1   triebeard_0.3.0    cellranger_1.1.0   lattice_0.20-44
 [9] pillar_1.6.1       backports_1.2.1    digest_0.6.27      promises_1.2.0.1
[13] rvest_1.0.0        colorspace_2.0-1   Matrix_1.3-3       htmltools_0.5.1.1
[17] httpuv_1.6.1       pkgconfig_2.0.3    httpcode_0.3.0     broom_0.7.6
[21] haven_2.4.1        xtable_1.8-4       scales_1.1.1       later_1.2.0
[25] generics_0.1.0     ellipsis_0.3.2     withr_2.4.2        cli_2.5.0
[29] crayon_1.4.1       readxl_1.3.1       mime_0.10          stopwords_2.2
[33] fs_1.5.0           fansi_0.5.0        xml2_1.3.2         hms_1.1.0
[37] lifecycle_1.0.0    munsell_0.5.0      reprex_2.0.0       compiler_4.1.0
[41] rlang_0.4.11       grid_4.1.0         rstudioapi_0.13    miniUI_0.1.1.1
[45] gtable_0.3.0       DBI_1.1.1          curl_4.3.1         R6_2.5.0
[49] lubridate_1.7.10   fastmap_1.1.0      utf8_1.2.1         fastmatch_1.1-0
[53] cld2_1.2.1         stringi_1.6.2      crul_1.1.0         Rcpp_1.0.6
[57] vctrs_0.3.8        dbplyr_2.1.1       tidyselect_1.1.1
```