



HAL
open science

Attitudinally-positioned European sample dataset

Pedro Ramaciotti Morales, Jimena Royo-Letelier, Jean-Philippe Cointet, Armin Pournaki

► **To cite this version:**

Pedro Ramaciotti Morales, Jimena Royo-Letelier, Jean-Philippe Cointet, Armin Pournaki. Attitudinally-positioned European sample dataset. sciences po. 2024, 18 p. <hal-04807916>

HAL Id: hal-04807916

<https://hal.science/hal-04807916v1>

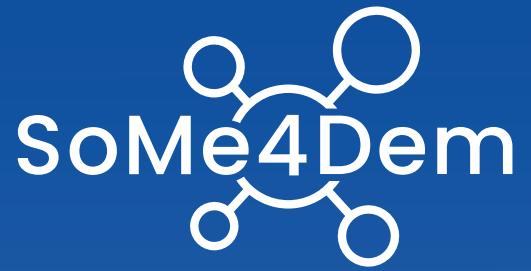
Submitted on 27 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



Attitudinally-positioned European sample dataset

Pedro Ramaciotti^{1,2}, Jimena Royo-Letelier²,
Jean-Philippe Cointet², Armin Pournaki^{2,3}

¹CNRS Complex Systems Institute of Paris

²Sciences Po médialab

³Max Planck Institute for Mathematics in Science

August 12, 2024



Funded by the European Union
under the Horizon Europe framework
(HORIZON-CL2-2022-DEMOCRACY-01-07).

Disclaimer

Funded by the European Union in the HORIZON EUROPE framework under grant agreement No 101094752. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Horizon Europe. Neither the European Union nor the granting authority can be held responsible for them.

Deliverable

D2.1 Attitudinally-positioned European sample dataset

Dissemination level

Public

Type

Report

Executive Summary

This deliverable provides the first version of a dataset of attitudinally positioned populations together with the conceptual and methodological framework required to leverage attitudinal inference for large populations of social media users in Europe. Building on recent advancements on large-scale multi-dimensional political attitude inference in social networks and text, we show how to create a European sample of attitudinally-positioned users along a Left-Right and a Anti-elite dimension measuring attitudes towards elites and trust in institutions. These two dimensions are shown to be relevant to conduct both traditional political analysis on social media and analyses accounting for new forms of polarization related to democratic backsliding. This dataset of users will serve as a frame of reference for the development of case studies exploring different links between activity in online platforms, evidence and impacts in politics in other tasks of the project.

Contents

- 1 Introduction 5
- 2 Background 6
- 3 Summary 6
- 4 Methods 7
 - 4.1 Seeding social networks connected with political debate 7
 - 4.2 Collecting the data 7
 - 4.3 Filtering the data 8
 - 4.4 Labeling users 8
 - 4.5 Positioning users in latent homophily spaces 9
 - 4.6 Matching homophily spaces with dimensions of political expert surveys and positioning users 10
- 5 Data Records 10
- 6 Technical Validation 14
- 7 Code and Data Availability 14
- 8 Ethics, Security and Compliance 16

1 Introduction

Social media play an ever-increasing role in the public sphere and are now a significant element of the media ecosystem and politics (Benkler et al., 2018; Stewart et al., 2020).

Initially, social media was thought to usher in utopic forms of direct democracy and immediate participation (Shirky, 2009), and were heralded as “liberation technology” owing to their role in the “Arab spring” and other social movements related to online (Tucker et al., 2017). But almost three decades later, more dystopian visions are prevalent in discourse regarding social media (Morozov, 2016; Schradie, 2019), with narratives focusing on loss of social cohesion, segregation, and polarization (Lewandowsky et al., 2022). Social media has also been blamed for exposing the public to misinformation and disinformation (Guess et al., 2019). Furthermore, these developments have unfolded in parallel with a trajectory of democratic decline permeated by conflating economic, migratory, political, and security crises in European countries (Zeitlin et al., 2019; Algan et al., 2017).

A widespread causal account of the role of social media in democratic decline involves selective exposure (through personalization, in how users set up their digital ego-ecosystems but also how algorithms curate what they are shown), which would lead to further online segregation and increased polarization. Different works associate polarization and segregation with weakened capacity for collective deliberation and compromise (Huckfeldt et al., 2004) and as one of the main drivers of the spread of misinformation (Osmundsen et al., 2021). A large number of works explore this narrative linking polarization and segregation with negative societal outcomes, popularly conceptualized as “echo chambers” or “filter bubbles”. However, recent works have painted a more nuanced picture. Contrary to popular conceptions, algorithms on social platforms may increase diversity of exposure, or do not always significantly decrease diversity, and in a very wide array of settings (Aiello et al., 2017; Bakshy et al., 2015; Haim et al., 2018; Puschmann, 2019). Beyond the question of whether social media changes selective exposure at relevant scales in society (importantly in comparison with what baseline) and under which circumstances, the effects of this exposure is not completely understood. Recent works also suggest that proposed solutions based on diversifying exposure (traditionally held as a central goal for a large community in research) might backfire. Bail et al., 2018, for instance, showed that forcing cross-exposure between individuals with opposed political views might exacerbate polarization. Other proposed solutions, such as deplatforming radical individuals might backfire too, forcing already polarized individuals to new platforms with less regulation and moderation systems in place (Horta Ribeiro et al., 2021).

To advance the science needed to better understand the effects of different types of selective exposure, the extent to which they exist, and circumstances under which they manifest in different platforms, improvements must be made in the political characterization of platform users. Any measurement, experiment or reported result on political diversity, segregation, or polarization must necessarily hinge on political attributes of users on which to compute metrics and report findings. One traditional approach to fulfilling this requirement is to conduct surveys among online populations. This approach poses the challenge of matching survey respondents with online user identifiers, which is difficult for several reasons: it is intrusive and increases operational risks (e.g., with regards to Art. 9 of GDPR about political profiling), it requires that researchers reach relevant and representative parts of online populations with survey requests, it relies on cooperation on the part of users (i.e., they must volunteer their user identifiers to be matched with platform data), and it needs that respondents are truthful and coherent in how they report their own political stances. An alternative approach is to infer estimates on political stances indirectly from platform behavioral data traces that have statistical properties making aggregate computations useful in studies.

Studies pioneering the use of behavioral data traces from platforms in political stances estimation were first developed in the US; e.g., using follower networks on Twitter/X (Barberá, 2015) – hereinafter Twitter – and likes on Facebook (Bond et al., 2015). These works have focused on, however, on single dimensional representations of political stances, on Liberal-Conservative or Left-Right dimensions on which to position users for large scale social media studies. This focus on a single dimension has since permeated many research design strategies in

social media studies, in part because of the influence of these research works conducted in the US setting where a single dimension is sufficiently explicative of observed data, but also because redistributive Left-Right cleavage dimensions are also relevant across several national settings. This single-dimensional approach, however, has shown to be inadequate in the context of several European countries, which are known to display political systems organized along several cleavage and strategic issue and ideology dimensions (Bakker et al., 2012). Recent results have shown that some online political environments in Europe require several dimensions to explain observed online behavior data (Ramaciotti et al., 2021). More recently, progress in multidimensional political stance estimation has shown that it is possible to infer the positions of large online population samples on Left-Right dimensions but also others that are relevant to understanding European politics and phenomena such as democratic backsliding in social media studies (Ramaciotti Morales, Cointet, et al., 2022).

In this document we detail the construction of a dataset containing estimated political positions for large online populations in multiple European countries. In the construction of our dataset we leverage both platform and political survey data with political positions of parties. This latter source of data allows us to position these large populations along the dimensions of the survey instrument and that come endowed with reference points (e.g., a specific spatial reference point for centrist users on the Left-Right dimension). Additionally, we use political survey data to position users along different political dimensions. For the sake of comparability across countries we will position users in a Left-Right dimension provided by the survey (see details in the following sections). In order to appraise changing polarization and political competition linked to democratic decline, we also position users in a dimension intended to measure the degree to which they subscribe populist rhetorics, in particular in the form of anti-establishment and anti-elite rhetoric.

2 Background

This document describes the results under Deliverable D2.1 of project SoMe4Dem: “Attitudinally-positioned European sample dataset”, in which we position large online populations across the EU on comparable ideology and issue dimensions using social media behavioral trace data. These populations with multidimensional political positions are then to be used in other tasks in the project (e.g., in deliverables D2.2 and D5.2).

For the construction of these datasets we begin analyzing Twitter data on the 8 most populous countries in the EU: Belgium, France, Germany, Italy, the Netherlands, Poland, Romania, Spain. To this list we included Slovenia in order to foster collaborations with new partners in the SoMe4Dem project joining via the Hop-on facility funding program: the IRRIS Institute for Research, Development and Strategies of Society, Culture and the Environment (*Inštitut Irris za raziskave, razvoj in strategije družbe, kulture in okolja*).

This document describes the protocols and procedure leading to the construction of these datasets, as well as a description of the data records, its location for open access, and means of validation.

3 Summary

The rest of this document is structured in six sections:

- **Methods** (Section 4): providing an explanation of the data delimitation, collection, and treatment leading to the computation of political positions for large Twitter populations in the 9 selected countries.
- **Data Records** (Section 5): describing the format of the data rendered publicly available.

- **Technical Validation** (Section 6): describing procedures and results that assess the quality of the computed political positions.
- **Code and Data Availability** (Section 7): detailing the permanent location and referencing of the code and the data rendered publicly available.
- **Ethics, Security and Compliance** (Section 8): describing the compliance with regulations in place, as well as a legal deposit, and ethical considerations and protocols.

4 Methods

4.1 Seeding social networks connected with political debate

To seed our online populations we begin by identifying accounts of political figures in our selected set of nine countries. These accounts are then used to search for additional user accounts connected to these political figures. This logic, owing to several previous works (Barberá, 2015; Bond et al., 2015; Ramaciotti et al., 2021), is intended to yield sets of users that are sufficiently connected to the political public sphere of each country. This approach does not address the question of the multiple forms of identification of individuals with countries (e.g., citizenship, residency, or different degrees of participation in the economy). Instead, this approach treats the digital public sphere on a platform as a population of interests in itself, in which individuals can potentially join in freely. This also means that users can belong to datasets of different countries simultaneously.

To fix political figures with comparable functional roles across countries we restricted political figures to Members of Parliament (MPs). Additionally, we avoid distinctions of roles between different chambers of parliament by considering MPs from a country as a single set. We conducted a manual annotation of MPs in selected countries in the first several months of 2023, manually looking and annotating the corresponding Twitter accounts. Table 4.1 reports the number of accounts identified.

4.2 Collecting the data

We then proceeded to collect the followers of our MPs on Twitter. This owes again to previous work (Barberá, 2015; Bond et al., 2015; Ramaciotti et al., 2021), and is in particular a requisite of the positioning method we will use, the so-called Language-Independent Network Attitudinal Embedding (LINATE; Ramaciotti Morales, Cointet, et al., 2022). The rationale behind this choice is multifaceted. On the one hand, we want to rely on the visibility of political figures in countries to serve as one delimitation criterion for our populations. Alternative criteria include, for instance, followers of accounts of news media outlets (Barberá, 2015). We rejected additional criteria for the sake of comparability across countries. On the other hand, the choice to follow (or not) any given political figure that has sufficient visibility allows us to cast the problem of political positioning as an ideal point estimation problem based on Item Response Theory frameworks (Clinton et al., 2004).

We collected followers of these accounts of MPs using Twitter’s API in March 2023. The numbers of MPs and followers per country in Table 4.1 reflect the potential size of our populations, allowing for users to belong to spheres of several countries simultaneously. We additionally collected the profile bios of the selected accounts for the purposes of proposing validation metrics for political positions based on text written by users (see Section 6).

Table 4.1: Number of Members of Parliament (MPs) Twitter accounts used to seed the collection of Twitter accounts (acc.) in each country, followers, and followers that are included depending on the filtering criteria established in Section 4.3.

	Belgium	France	Germany	Italy	Netherlands	Poland	Romania	Slovenia	Spain
MP acc.	196	881	676	469	197	468	81	85	468
Collected acc.	5 44 373	5 093 576	3 371 918	6 809 729	1 852 339	1 965 669	48 389	137 379	5 403 193
Selected acc.	64 726	494 065	443 852	777 069	180 063	224 962	1 822	13 894	846 230
Bios	37 508	299 850	235 632	330 637	109 270	105 375	1 023	5 841	474 214

4.3 Filtering the data

Next, we apply several criteria to filter these sets of potential populations. The purpose of these filters is two-fold: 1) to improve the quality of the samples minimizing users that are not active on the platform and decreasing the probability of including bots, and 2) to assure that each user in our final sample follows a sufficiently large number of MPs. This second criteria is needed to use follower networks for inference, taking links (i.e., edges or follow relations) as signals, minimizing the probability that a user follows an MPs for reasons other than ideological homophily. This condition also aims at improving the “political sophistication” (Luskin, 1990) on the part of users and the degree to which spatial political models are suited to analyze political behavior among our populations. Formally, we apply the following criteria to our populations, deleting from our populations users that 1) are not followed by at least 25 users; 2) follow less than 3 MPs; and 3) that are themselves MPs. We consider separately MPs and their followers, deleting from the list of followers users that are themselves MPs. In our Item Reponse Theory framework, MPs are *items* that are chosen by their followers. MPs will thus have a position as items of choice, while users will have a position by virtue of their choices. After having filtering users, one MP in Poland and 2 MPs in Romania stand with no followers and are also removed from our dataset. Table 4.1 reports the number of MPs after having removed these three MPs.

These criteria are commonly applied to ideal point estimation methods in social media in other works (Barberá, 2015; Bond et al., 2015; Ramaciotti et al., 2021). The resulting populations in the countries analyzed are reported in Table 4.1.

4.4 Labeling users

In Section 6 we will propose a validation of positions based on annotations produced for each user. To compute annotations we will use a language model to translate profile bios into English language. For this, we used the nllb-200-distilled-600M language model¹ (Costa-Jussà et al., 2022). Profile bios are text strings of 160 characters with which users can present themselves publicly. Many users volunteer in their profiles information about their political preferences, writing, for instance, text that might be similar to “I am proud to be right-leaning” or “Always have been a leftist”².

Because of the size of our populations, we cannot manually annotate each profile bios. We choose to produce annotations with another language model. We use the zephyr-7b-beta language model³ to submit each profile bios text to prompts leading to outputs labeling each individual according to two criteria: 1) whether they are Left- or Right-leaning, and 2) whether they subscribe anti-elite rhetorics or if they are members of political, economical or social elites. Our aim is to show that our two selected CHES dimensions and the positions we infer for our large populations enable accurate classification of these groups. The details of how these annotations are used in validation metrics will be presented in Section 6.

We submit each English-translated profile bio to our language model four times, producing binary annotations for four labels, using the following prompts:

¹Available online at <https://huggingface.co/facebook/nllb-200-distilled-600M>.

²These examples are fabricated. We do not transcribe real profile bios text that would allow identification of users.

³Available online at <https://huggingface.co/HuggingFaceH4/zephyr-7b-beta>.

Table 4.2: Number of users annotated as begging *Left-leaning*, *Right-leaning*, subscribing populist rhetorics (*Populist*) or as perceived as being part of elite groups (*Elite*) according to a language model inspecting their Twitter profile bios.

	Belgium	France	Germany	Italy	Netherlands	Poland	Romania	Slovenia	Spain
Unparsed (Left)	17.2%	17.5%	19.2%	21.5%	19.9%	31.1%	20.3%	26.5%	21.7%
Unparsed (Right)	17.8%	17.8%	20.4%	21.0%	20.3%	29.4%	19.1%	25.5%	21.7%
Unparsed (Populist)	10.3%	11.8%	12.7%	15.5%	12.8%	22.8%	13.2%	19.4%	14.7%
Unparsed (Elite)	13.4%	14.2%	16.0%	17.5%	16.7%	23.0%	15.0%	20.8%	16.4%

- *labeled_left*:
You are an expert in European politics. Please classify the following Twitter profile bio as “Left-leaning” or “Not-Left” according to whether the author of the text (who is from [COUNTRY]) is politically Left-leaning or not. The response should be in the form of a single term with the name of the category: “Left-leaning” or “Not-Left”: [TEXT OF THE BIO].
- *labeled_right*:
You are an expert in European politics. Please classify the following Twitter profile bio as “Right-leaning” or “Not-Right” according to whether the author of the text (who is from [COUNTRY]) is politically Right-leaning or not. The response should be in the form of a single term with the name of the category: “Right-leaning” or “Not-Right”: [TEXT OF THE BIO].
- *labeled_populist*:
You are an expert in European politics. Please classify the following Twitter profile bio as “Populist” or “Not-Populist” according to whether the author of the text (who is from [COUNTRY]) holds populist views or not. Populist views include, among others, believing that society is split between the people and elites, or that political elites are corrupt. The response should be in the form of a single term with the name of the category: “Populist” or “Not-Populist”: [TEXT OF THE BIO].
- *labeled_elite*:
You are an expert in European politics. Please classify the following Twitter profile bio as “Elite” or “Not-Elite” according to whether the author of the text (who is from [COUNTRY]) belongs to an elite group, including political or economic elites. The response should be in the form of a single term with the name of the category: “Elite” or “Not-Elite”: [TEXT OF THE BIO].

In each one of the previous prompts we replace [COUNTRY] with the country under consideration for the users, and [TEXT OF THE BIO] with the string of the profile bio. Our prompts produce a very large proportion of outputs in the intended requested form. We discarded outputs that do not correspond to one of the two requested allowed categories specified in the prompt. A large majority of outputs from our language model queries yield results in the form requested in the prompts. Table 4.2 reports the percentage of bios per country and label for which the result could not be parsed assuring a label, with the worst case being Poland when queries whether bios are Left-leaning or not: around 31% of outputs could not be parsed into our binary labels.

Then we examine pairs of related annotations to further discard annotations that are contradictory. Concretely, if a user is identified as both *labeled_left* = *Left-leaning* and *labeled_right* = *Right-leaning* we considered that it is neither. Similarly, if a user is identified as both *labeled_populist* = *Populist* and *labeled_elite* = *Elite* we consider that it is neither. The results of this annotation strategy are reported in Table 4.3.

4.5 Positioning users in latent homophily spaces

Following the LINATE method, we next proceed to compute homophily multidimensional embeddings for each country. This procedure takes as input the selected user accounts and their following relations linking them

Table 4.3: Number of users annotated as begging *Left-leaning*, *Right-leaning*, subscribing populist rhetorics (*Populist*) or as perceived as being part of elite groups (*Elite*) according to a language model inspecting their Twitter profile bios.

	Belgium	France	Germany	Italy	Netherlands	Poland	Romania	Slovenia	Spain
<i>labeled_left</i>	1 442	16 932	12 352	7 351	2 793	2 243	18	99	2 5261
<i>labeled_right</i>	536	7 904	3 019	3 315	1 819	2 904	28	120	8 835
<i>labeled_populist</i>	300	4 055	2 006	2 774	1 362	1 171	6	52	7 164
<i>labeled_elite</i>	6 455	45 763	31 111	34 426	17 207	10 485	261	731	45 622

to MPs, producing a multidimensional spatial model where distances are related to the similarity of neighbors: users positioned in close proximity in space follow similar sets of MPs, and MPs in close proximity in space are followed by similar sets of users. We employ a generative Item Response Theory homophily model and compute the values of the parameters using the observed observation:

$$P(i \rightarrow j) = \text{logistic}(\alpha - \beta \|\phi_i - \phi_j\|^2)$$

where $P(i \rightarrow j)$ is the probability of observing user i following MP j , modeled as dependent on shape parameters α and β , and on the pairwise distances between their unobservable positions in an homophily space: ϕ_i and ϕ_j .

Figure 4.1 shows the distribution of the positions of the users in our selected populations along the first two dimensions (δ_1 and δ_2) of the latent homophily spaces. Because for annotated accounts of MPs their political parties are also known (shown in Figure 4.1 with colored crosses +), we can compute proxy positions for a parties in latent homophily space as the centroid of the positions of MPs from that party (shown in Figure 4.1 with colored circles \circ).

4.6 Matching homophily spaces with dimensions of political expert surveys and positioning users

We select the Chapel Hill Expert Survey (CHES 2019 edition; Jolly et al., 2022) as our survey instrument of spatial reference. We focus on two dimensions from this dataset, with the following descriptions and reference points:

- Left-Right ideology (*lrgen*): “position of the party in 2019 in terms of its overall ideological stance”, with 0 being the left-most position, 10 being the right-most position, and 5 being the political center in the Left-Right scale.
- Anti-elite salience (*antielite_salience*): “salience of anti-establishment and anti-elite rhetoric”, with 0 standing for “no importance at all” and 10 for “extremely important” (referring to the importance granted to the establishment and elites).

Because party positions exist in our latent homophily space and in the reference CHES space, we compute for each CHES dimension a map going from the former to the latter matching party positions. For each country, let P be the number of parties that can be identified with parties listed in the CHES instrument. We consider the first $P - 1$ dimensions of this latent space as possibly containing information leverageable in a linear map, and fit the parameters of this map as a Linear Least Squares Ridge regression.

Using the linear map computed for each country, we map the position of all identified MPs and selected users per country onto the two CHES dimensions. Figure 4.2 shows the density of the positions in populations in shades of blue along the two selected CHES dimensions: Left-Right and Anti-elite salience. Positions of MPs are shown in crosses (+) colored by political parties. Party proxy positions are computed as the centroid of MPs and shown in circles (\circ).

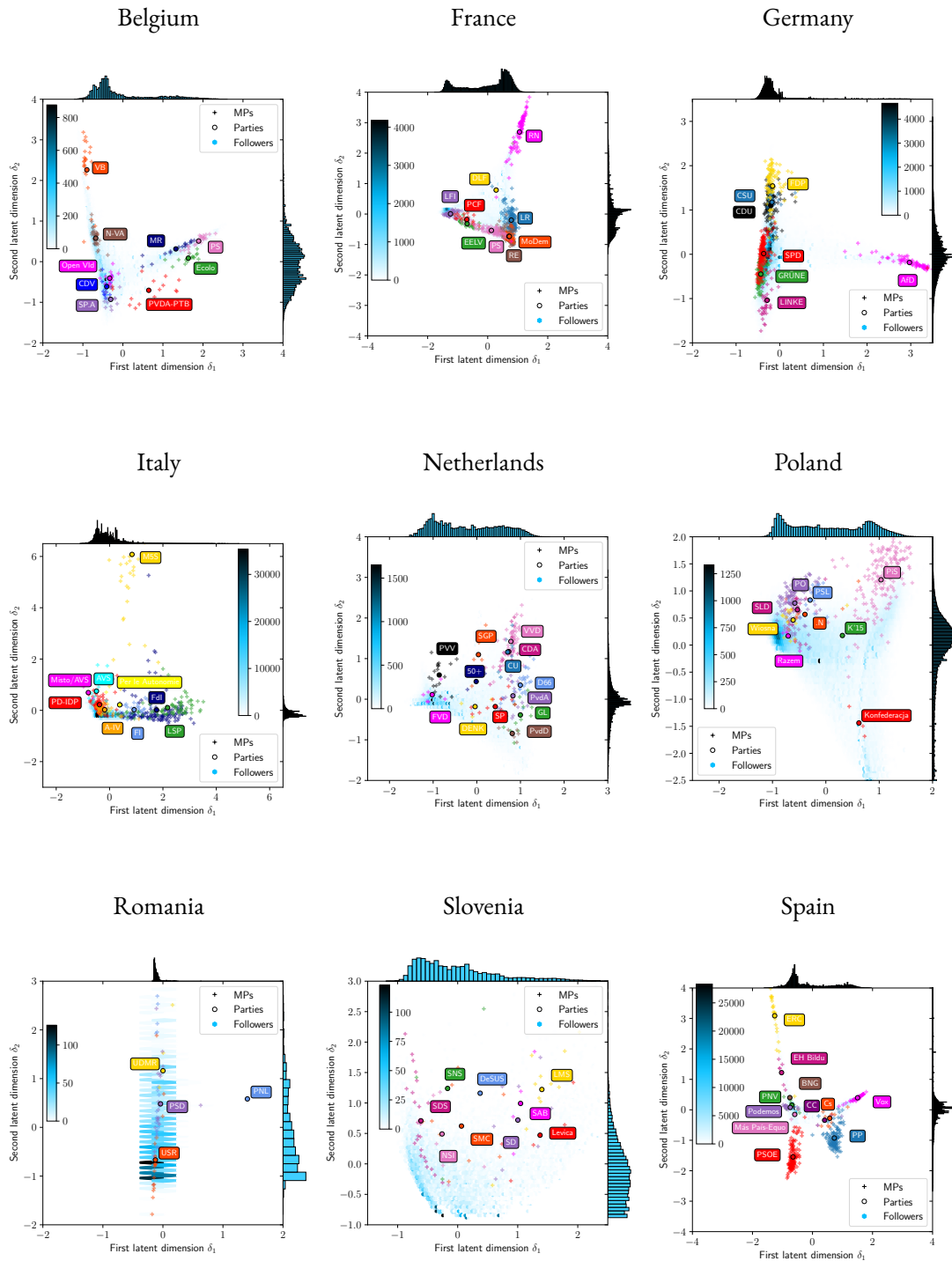


Figure 4.1: First two dimensions of the homophily spaces (δ_1 and δ_2) for our Twitter populations in the 9 countries considered for the project. The density of the positions of the populations is shown in shades of blue. Positions of MPs are shown in crosses (+) colored by political parties. Party proxy positions are computed as the centroid of MPs and shown in circles (o).

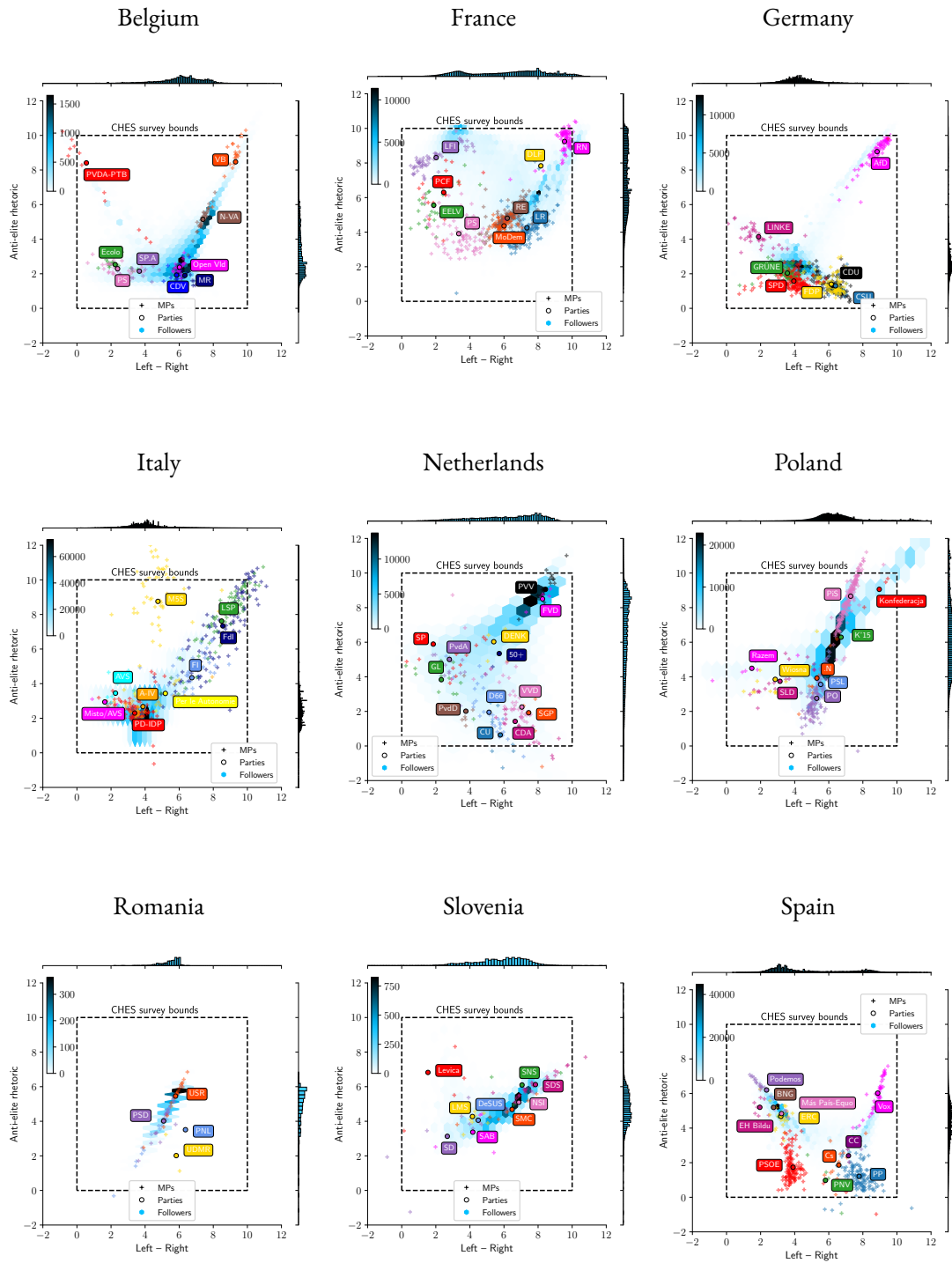


Figure 4.2: Distributions of online populations on the two selected CHES dimensions, Left-Right and Anti-elite salience, for the 9 countries considered for the project. The density of the positions of the populations is shown in shades of blue. Positions of MPs are shown in crosses (+) colored by political parties. Party proxy positions are computed as the centroid of MPs and shown in circles (o).

5 Data Records

These datasets are stored in separate files according to countries. Additionally, for each country, the positions of MPs and users are stored in separate csv tables. For Belgium, for example, two files are available: *belgium_MPs.csv* and *belgium_users.csv*. We exclude Romania from the data, because our labeling strategy in Section 4.4 does not allow us to compute quantitative assessments of the accuracy of positions on the CHES dimensions.

The files containing information about MPs are structured in the following way:

- Each row is an MPs for which we identified a Twitter account.
- Columns are:
 - *delta_1* (specifying the position along the first dimension of the latent homophily space specified in Section 4.5);
 - *delta_2* (specifying the position along the second dimension of the latent homophily space specified in Section 4.5);
 - *left_right* (specifying the Left-Right coordinate position of the MP as described in Section 4.6);
 - *antielite* (specifying the coordinate position of the MP on a scale measuring the salience of anti-establishment and anti-elite rhetoric, as described in Section 4.6);
 - *party* (specifying the political party to which the MPs is affiliated)

The files containing information about users are structured in the following way:

- Each row is a user.
- Columns are:
 - *delta_1* (specifying the position along the first dimension of the latent homophily space specified in Section 4.5);
 - *delta_2* (specifying the position along the second dimension of the latent homophily space specified in Section 4.5);
 - *left_right* (specifying the Left-Right coordinate position of the user as described in Section 4.6);
 - *antielite* (specifying the coordinate position of the user on a scale measuring the salience of anti-establishment and anti-elite rhetoric, as described in Section 4.6);
 - *labeled_left* (a binary variable specifying whether the user is labeled as being Left-leaning by the language model inspecting its text profile bio, as described in Section 4.4);
 - *labeled_right*: (a binary variable specifying whether the user is labeled as being Right-leaning by the language model inspecting its text profile bio, as described in Section 4.4);
 - *labeled_populist*: (a binary variable specifying whether the user is labeled as subscribing populist rhetorics by the language model inspecting its text profile bio, as described in Section 4.4);
 - *labeled_elite*: (a binary variable specifying whether the user is labeled as belonging to an elite group by the language model inspecting its text profile bio, as described in Section 4.4);

6 Technical Validation

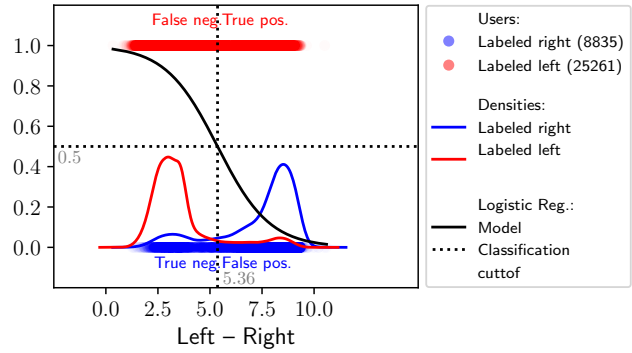
While the correctness of the logic behind the embedding of MPs and users in the selected CHES was argued on the basis of an Item Response Theory framework, it does not offer *ex post* validations. In this section, we present metrics aiming to show that the positions of users in the CHES dimensions are coherent with other estimations of political positions that can be made for users on the basis of the text that they choose to write in their profile bios. The method proposed in Section 4 results in positions along a continuous scale, while the methods based on language models considered in this document yield categorical classifications (see Section 4.4). The aim of this section is to show that, for the fraction of users for which categorical political variables or labels can be estimated from text, there is high coherence between political estimation based on text and on the CHES embedding or positioning.

Our proposed validation builds on the assumption that the positions of the users that have labels produced in Section 4.4 must be coherently distinguished using one of the selected CHES dimensions. We expect that the positions of users labeled as Left- and Right-leaning must be distinguishable in the Left-Right dimension, and that the positions of users labeled as subscribing populist rhetoric and belonging to elite groups must be distinguishable in the Anti-elite salience dimension.

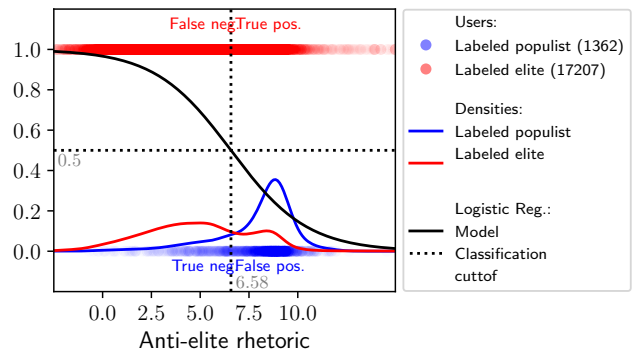
We test this hypothesis by assessing the degree to which a linear classifier trained on the corresponding dimension can distinguish the two relevant groups of labeled users (Ramaciotti Morales and Muñoz Zolotoochin, 2022). We perform this assessment, for each CHES dimension, by considering only users that have a relevant label (e.g., Left- and Right-leaning for the Left-Right dimension). We take one label (e.g., Right-leaning) as having values equal to 1 and the complementary label (Left-leaning, in this example) as having values equal to 0, and we fit a logistic regression on the positions of these users on the CHES dimensions. For the obtained logistic model, we then compute the error in classifying (at threshold probability equal to 0.5) users with labels. We report the accuracy of this classification as the F-score, namely F1 (thus accounting for both precision and recall).

Because pairs of complementary labeled users per dimension can be highly unbalanced we systematically subsample the majority group to match the size of the minority group. We draw 100 such subsampling operations for each dimension and for each country, and report the average F1-score and its standard deviation (up to two decimal positions). We exclude from our analysis the dataset from Romania, because the small number of labeled users is not deemed sufficient to establish a trend along the respective dimension (*labeled_left*=18, *labeled_right*=28, *labeled_populist*=6, *labeled_elite*=261; see Table 4.3). We conclude that our adopted strategy fails to provide assurances regarding the quality of the positions of the users in the CHES dimensions and we exclude this country from the datasets that we render public. Table 6.1 summarizes these averages and standard deviations. Figure 6.1 illustrates this procedure by reporting logistic regression classifications covering the spectrum of F1-scores obtained, highlighting the varying strength of evidence supporting the accuracy of positioning along the chosen CHES dimensions. We show an example of high (Spain on the Left-Right dimension), intermediate (Netherlands on the Anti-elite dimension) and low F1-score (Spain on the Anti-elite dimension, the lowest F1-Score in our datasets). These qualifications into high, intermediate, and low are relative within our datasets and not an assessment on the usability of the data for different purposes. The accuracy of all datasets, except for that of Romania, are deemed adequate for the rest of the tasks in the project.

Spain: precision=0.939, recall=0.882, F1=0.910



Netherlands: precision=0.979, recall=0.721, F1=0.830



Spain: precision=0.909, recall=0.551, F1=0.687

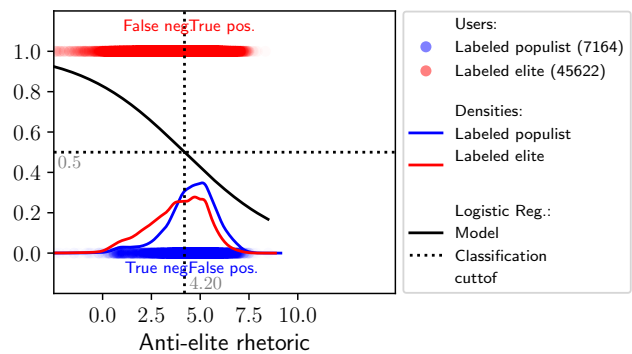


Figure 6.1: Illustration of the text-based validation of the estimated political positions high, intermediate, and low F1-score (relative within the datasets). F1-score measures the accuracy of a logistic regression model in distinguishing complementary groups of users (labeled as Left- and Right-leaning for the Left-Right dimension) on three countries: Spain on the Left-Right dimension (highest F1-score), Netherlands on the Anti-elite dimension (intermediate F1-score), and Spain on the Anti-elite dimension (the lowest F1-score). F1-score is computed subsampling the majority group (matching the size of the minority group) 100 times and reporting average and standard deviation of the F1-score.

Table 6.1: F1-score measuring the accuracy of a logistic regression model in distinguishing complementary groups of users (labeled as Left- and Right-leaning for the Left-Right dimension, and labeled as populist and elite for the anti-elite dimension) on the nine countries considered in the project. F1-score is computed subsampling the majority group matching the size of the minority group, subsampling 100 times and reporting average and standard deviation of the F1-score.

	Belgium	France	Germany	Italy	Netherlands	Poland	Slovenia	Spain
Left – Right	0.84 ± 0.00	0.86 ± 0.00	0.92 ± 0.00	0.85 ± 0.00	0.85 ± 0.00	0.76 ± 0.00	0.74 ± 0.01	0.91 ± 0.00
Anti-elite rhetoric	0.86 ± 0.00	0.84 ± 0.00	0.91 ± 0.00	0.82 ± 0.00	0.83 ± 0.00	0.73 ± 0.00	0.72 ± 0.00	0.69 ± 0.00

7 Code and Data Availability

The anonymized data are available in a Figshare repository at:

<https://doi.org/10.6084/m9.figshare.25288210.v1>

The reproducibility code is available in a GitHub repository at:

https://github.com/pedroramaciotti/some4dem_d21.

Because of the low number of labeled bios obtained for Romania, we do not include this country in the dataset.

8 Ethics, Security and Compliance

This work has been funded by the SoMe4Dem Horizon Europe project, Grant No. 101094752. This work also relies on data collected by the “European Polarisation Observatory” (EPO) of CIVICA Research (co-)funded by EU’s Horizon 2020 programme under Grant No 101017201.

Further information about the project can be found on its website <https://some4dem.eu/>. The data used in the project are included in legal deposits made with the French data protection authority CNIL (*Commission nationale de l’informatique et des libertés*) on 23 August 2023. The methodologies of project SoMe4Dem conducted at Sciences Po are sanctioned by Opinion n°2023-038, dated 18 October 2023, of the Research Ethics Committee CDR (*Comité de Déontologie de la Recherche*) of the Paris Institute of Political Sciences FNSP (*Fondational Nationale de Sciences Politiques*), Sciences Po.

According to the project’s Data Management Plan, the identified form of these data is stored with the data controlled at Sciences Po and publicly available only in anonymized form.

References

- Aiello, Luca Maria et al. (2017). “*Evolution of ego-networks in social media with link recommendations*”. In: *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pp. 111–120.
- Algan, Yann et al. (2017). “*The European trust crisis and the rise of populism*”. In: *Brookings papers on economic activity* 2017.2, pp. 309–400.
- Bail, Christopher A et al. (2018). “*Exposure to opposing views on social media can increase political polarization*”. In: *Proceedings of the National Academy of Sciences* 115.37, pp. 9216–9221.
- Bakker, Ryan et al. (2012). “*Complexity in the European party space: Exploring dimensionality with experts*”. In: *European Union Politics* 13.2, pp. 219–245.
- Bakshy, Eytan et al. (2015). “*Exposure to ideologically diverse news and opinion on Facebook*”. In: *Science* 348.6239, pp. 1130–1132.
- Barberá, Pablo (2015). “*Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data*”. In: *Political analysis* 23.1, pp. 76–91.
- Benkler, Yochai et al. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.
- Bond, Robert et al. (2015). “*Quantifying social media’s political space: Estimating ideology from publicly revealed preferences on Facebook*”. In: *American Political Science Review* 109.1, pp. 62–78.
- Clinton, Joshua et al. (2004). “*The statistical analysis of roll call data*”. In: *American Political Science Review* 98.2, pp. 355–370.
- Costa-Jussà, Marta R. et al. (2022). “*No Language Left Behind: Scaling Human-Centered Machine Translation*”. arXiv: 2207.04672 [cs.CL].
- Guess, Andrew et al. (2019). “*Less than you think: Prevalence and predictors of fake news dissemination on Facebook*”. In: *Science advances* 5.1, eaau4586.
- Haim, Mario et al. (2018). “*Burst of the filter bubble? Effects of personalization on the diversity of Google News*”. In: *Digital journalism* 6.3, pp. 330–343.
- Horta Ribeiro, Manoel et al. (2021). “*Do platform migrations compromise content moderation? evidence from r/the_donald and r/incels*”. In: *Proceedings of the ACM on Human-Computer Interaction* 5.CSCW2, pp. 1–24.
- Huckfeldt, Robert et al. (2004). *Political disagreement: The survival of diverse opinions within communication networks*. Cambridge University Press.
- Jolly, Seth et al. (2022). “*Chapel Hill expert survey trend file, 1999–2019*”. In: *Electoral studies* 75, p. 102420.
- Lewandowsky, Stephan et al. (2022). “*Technology and democracy: A paradox wrapped in a contradiction inside an irony*”. In: *Memory, mind & media* 1, e5.

- Luskin, Robert C (1990). “*Explaining political sophistication*”. In: *Political behavior* 12.4, pp. 331–361.
- Morozov, Evgeny (2016). “*The net delusion: How not to liberate the world*”. In: *Democracy: A Reader*. Columbia University Press, pp. 436–440.
- Osmundsen, Mathias et al. (2021). “*Partisan polarization is the primary psychological motivation behind political fake news sharing on Twitter*”. In: *American Political Science Review* 115.3, pp. 999–1015.
- Puschmann, Cornelius (2019). “*Beyond the bubble: Assessing the diversity of political search results*”. In: *Digital Journalism* 7.6, pp. 824–843.
- Ramaciotti, Pedro et al. (2021). “*Unfolding the dimensionality structure of social networks in ideological embeddings*”. In: *Proceedings of the 2021 IEEE/ACM international conference on advances in social networks analysis and mining*, pp. 333–338.
- Ramaciotti Morales, Pedro, Jean-Philippe Cointet, et al. (2022). “*Inferring attitudinal spaces in social networks*”. In: *Social Network Analysis and Mining* 13.1, p. 14.
- Ramaciotti Morales, Pedro and Gabriel Muñoz Zolotoochin (2022). “*Measuring the accuracy of social network ideological embeddings using language models*”. In: *International Conference on Information Technology & Systems*. Springer, pp. 267–276.
- Schradie, Jen (2019). *The revolution that wasn't: How digital activism favors conservatives*. Harvard University Press.
- Shirky, Clay (2009). *Here comes everybody: How change happens when people come together*. Penguin UK.
- Stewart, Evan et al. (2020). “*The new structural transformation of the public sphere*”. In: *Sociological Theory* 38.2, pp. 170–191.
- Tucker, Joshua A et al. (2017). “*From liberation to turmoil: Social media and democracy*”. In: *J. Democracy* 28, p. 46.
- Zeitlin, Jonathan et al. (2019). “*Introduction: The European Union beyond the polycrisis? Integration and politicization in an age of shifting cleavages*”.