HAL open science

Measuring the accuracy of social network ideological embeddings using language models

Pedro Ramaciotti Morales, Gabriel Muñoz Zolotoochin

To cite this version:

Pedro Ramaciotti Morales, Gabriel Muñoz Zolotoochin. Measuring the accuracy of social network ideological embeddings using language models. 2022 International Conference on Information Technology & Systems (ICITS22), Feb 2022, San José, Costa Rica. hal-03385061

HAL Id: hal-03385061

https://hal.science/hal-03385061v1

Submitted on 19 Oct 2021

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.

2022 International Conference on Information Technology & Systems (ICITS22)

Measuring the accuracy of social network ideological embeddings using language models

Pedro Ramaciotti Morales[1] and Gabriel Muñoz Zolotoochin[2]

[1] médialab, Sciences Po, Paris, France,
pedro.ramaciottimorales@sciencespo.fr
[2] Facultad de Matemáticas, P. Universidad Catótica de Chile

Abstract. The opinions of people on different issues are traditionally studied through polls. Recently, ideological embedding methods have proposed to position social network users in spaces where positions are informative of their opinions. These methods have been shown to be effective in the US and in European settings. However, validating their results is challenging, as required data must not rely on the social network structure used in the embedding (panel data, for example). In this article, we propose a validation method based on language models for classification of users in ideological spaces. We illustrate our methodology using political manifestos, political surveys on party positions, and text utterances produced by Twitter users in Chile and in France, two countries with different party structure. Using text classification related to issues relevant to the ideological spatialization, we show that positions can be shown to be highly accurate, thus allowing for the robust inference of opinions of users at large scales.

Keywords: Network embedding, multidimensional network scaling, text-classification, party systems, ideology, ideology scaling, polarization.

1 Introduction

Polls and surveys are traditional instruments for estimating opinions in populations. More recently, network ideological embedding –or scaling– methods have shown that digital traces in media platforms, such as friendship networks, can be used to infer opinions at massive scales. These methods embed users in spaces where dimensions stand for indicators of opinions, and where positions along dimensions indicate opinions ranging from most opposed to most favorable to different issues. Validating the opinion positions obtained with ideological embedding methods is a challenging task, as they require costly additional data such as vote registries or panel data. In this article we propose a method, based on the automated analysis of text written by users, to validate their ideological positions computed using the network structure via ideological embedding

(and thus not dependent on language). We illustrate our method using Twitter data from two countries with different party structures and languages, Chile and France, showing that ideological embedding has demonstrable accuracy.

Ideological embedding methods have become an active research stream as they hold the promise to provide opinion models for studying many interesting phenomena including polarization and extremism [1–3], to produce contributions bridging opinion dynamics and social network analysis [4, 5], and finally to deliver valuable applications such as passive, inexpensive, and massive polling. These methods work by collecting social network data structures (*i.e.*, the social graph), postulating latent *ideological* variables that may be structuring the network through some social mechanism such as homophily [6] (people with similar opinions, thus with similar variable values, tend to become friends), and use the network data to infer these ideological parameters. Through this procedure, users are embedded in the parameter space, called *ideological space*. When validating the results of these methods, different additional data have been leveraged, such as voting records and self-declared political affiliations [1] (which may be difficult or impossible to obtain), and even panel data obtained through surveys [7] (which may be expensive to obtain). Furthermore, newer network ideological embedding methods [8] propose the possibility of embedding users in multitudes of dimensions serving as indicators to a diversity of issues, which makes it difficult to rely on surveys or voting records for validation. While validations produced using text written by users themselves have been proposed (*e.g.*, in the cited example from [1]), these do not address quantitative measures of fit between identified issues related to dimensions of the ideological space, and the accuracy with which they capture opinions toward those issues.

In this article we propose the first validation method for the accuracy of positions computed using ideological network embedding, and that can be systematically applied to a wide range of issues. We begin by addressing the different previous works relating to our results. We then present the experimental Twitter data from Chile and France, and external opinion position data for parties that is generally used to compute the interpretation of the dimensions of the embedding space. We then present a family of ideological embedding methods and a computation of the interpretation of first dimensions of the embedding space in terms of issues of the political debate for which they act as indicators. Using the position of users and the text they have written in their online profiles to describe themselves, we propose our method for assessing the accuracy of their positions. Finally, we discuss our results and their consequences.

2 Related Work

Measuring opinions in populations often involves *a priori* distributional assumptions (*e.g.*, normality) and establishing the parameters of polls. Alternatively, behavioral traces can be analyzed to infer opinions. The foremost example is the NOMINATE method [9] for the estimation of opinion positions (*e.g.*, left-right scales [10]) of members of parliament (MPs) analyzing how they vote on pro-

posed bills. Ideological network embedding extend the principle of these methods to social networks. Instead of inferring the ideological position of MPs that vote for bills, they infer the positions of users that follow politicians online [3, 1]. These methods have been used to infer the liberal or conservative attitudes of millions of US Twitter and Facebook users from the observed friendship –who follows whom– networks (replacing *voting* by *following*). To perform the inference of ideological parameters, some methods use the Correspondance Analysis (CA) method, shown to be equivalent to a log-spatialization for homophily models [11]. Importantly, these methods do not rely on textual data, making them language-independent. Very recently, these methods have been adapted to work on European settings, and inferring ideological dimensions other than Left-Right divides [12, 8]. In terms of modeling, choice data (be it MPs choosing bills or users choosing friends online) is often leveraged as bipartite network structures. The reader is referred to [13] for survey on ideology estimation methods from bipartite network structures.

Empirical validation using external data (such as self-declared political affiliation or voting registries) has shown qualitatively that it is possible to recover ideological positions in the US and European countries. However, quantitative measurements of the accuracy of these methods have proved elusive, as it is difficult to obtain alternative data on the position of users to contrast the result of these methods. To obtain ideological positions without using social network graphs, researchers have relied on media outlet citation [14, 15] and positions of parliamentarians as computed with the NOMINATE method based on voting on bills [16, 17]. Automated text analysis has been used to infer ideological positions, but in classification tasks. Researchers successfully developed methods to classify political text, *e.g.*, by analysis whether different textual content can be classified coming from Democratic- or Republican-leaning sources in the US [18]. Other methods propose to extract positions of texts, and thus of users who produce them, discovering the objects towards which positions are taken [19]. Yet, other families of methods aim at inferring the sentiment of an uttered text [20, 21], but without identifying the issues to which text is referring nor latent ideological positions of the person uttering it. None of these methods has yet been combined with ideological embedding procedures to assess their accuracy in a way that tackles the challenge of required flexibility in dealing with multiple issues towards which users take positions.

3 Social Network Data

To test our method we use Twitter data. We select two countries with different attitudinal reference data for parties, Chile and France, and we manually identify the Twitter accounts of MPs. We build on the identification criteria and collections of the "Twitter Parliamentarian Database" [22, Section 3] to identify 187 Chilean and 515 French parliamentarians on Twitter. We then collect

the followers of these accounts using Twitter's API[3]. Following [1], we filter out followers that have posted less than 100 tweets or have less than 25 followers to improve the ratio of true users to bots, and retain users that follow at least 3 MPs, and we call them *politicized users*. This collection was conducted on October 2020 and resulted in 418.979 politicized users in Chile and 288.496 in France (see Table 1), for which we also collected their Twitter *bios* (a 160 character profile description written by users). The bipartite network of the MPs and their followers will be used for the ideological embedding procedure, and the text *bios* to validate the ideological positions.

For each country, we consider attitudinal references for parties, which allow to position them in multi-dimensional opinion spaces where positions are indicators of attitudes towards separated, well-identified issues of public debate, and will allow for the interpretation of the content of dimensions in the ideological embedding space. For Chile we use the 2020 Manifesto Project (MARPOR) data [23], which is derived from human annotations of political manifesto texts produced by political parties; annotators associate phrases of manifestos as addressing any of a number of predefined issues, and the number of phrases per issue is used to compute positions. For France, we use the 2019 Chapel Hill Expert Survey (CHES) data [24], which is derived from answers provided by a survey answered by 421 political scientists specializing in European politics. The MARPOR data allows us to position Chilean political parties in 147 attitudinal scales, while the CHES data allows us to position French political parties in 48 scales. Scales mark with a numerical value a favorable or opposed position towards issues such as, *e.g.*, attitudes towards the political left or right, European integration, immigration, or welfare state. The MARPOR dataset includes ideological positions for parties in more than 50 countries, while the CHES dataset includes positions for 34 European countries. We chose to use CHES for France because it is generally considered as more reliable [25], and to show that our approach works when the embedding procedure is used with either of the two referential sources: surveys for France and annotated manifestos for Chile.

Table 1. Twitter datasets used in ideological embedding and validation.

	MPs Twitter Acc. (out of)	Followers	Politized Fol.	Attitudinal References
Chile	187 (198)	3.507.374	418.979	MARPOR
France	515 (925)	4.894.333	288.496	CHES

4 Social Network Ideological Embedding

Using the bipartite networks collected for Chile and France, embed individuals in ideological spaces, which we will validate using text *bios* in the next section.

[3] In compliance with GDPR, the data for this study was declared on July the 15th, 2021, to Sciences Po's data processing registry.

For each country we represent this bipartite sub-graph of the Twitter network as an adjacency matrix $A \in \{0,1\}^{F \times P}$, where P is the number of *politicians* or MPs, and F is the number their politicized followers. We then produce a reduced-dimensionality representation of the MPs and their followers using the coordinates of accounts in the principal components of the CA of matrix A. We denote by δ_i the i-th dimension of this reduced-dimensionality representation. Figure 1 shows the positions of politicized followers on the first two dimensions (δ_1 & δ_2) along with the position of the MPs and their parties (computed as the mean position of MPs from each party), and a point cloud scatterplot in the first three dimensions (δ_1, δ_2 & δ_3). We manually inverted the direction of δ_2 in France for convenience of visualization. This has no effect in the rest of our analysis.

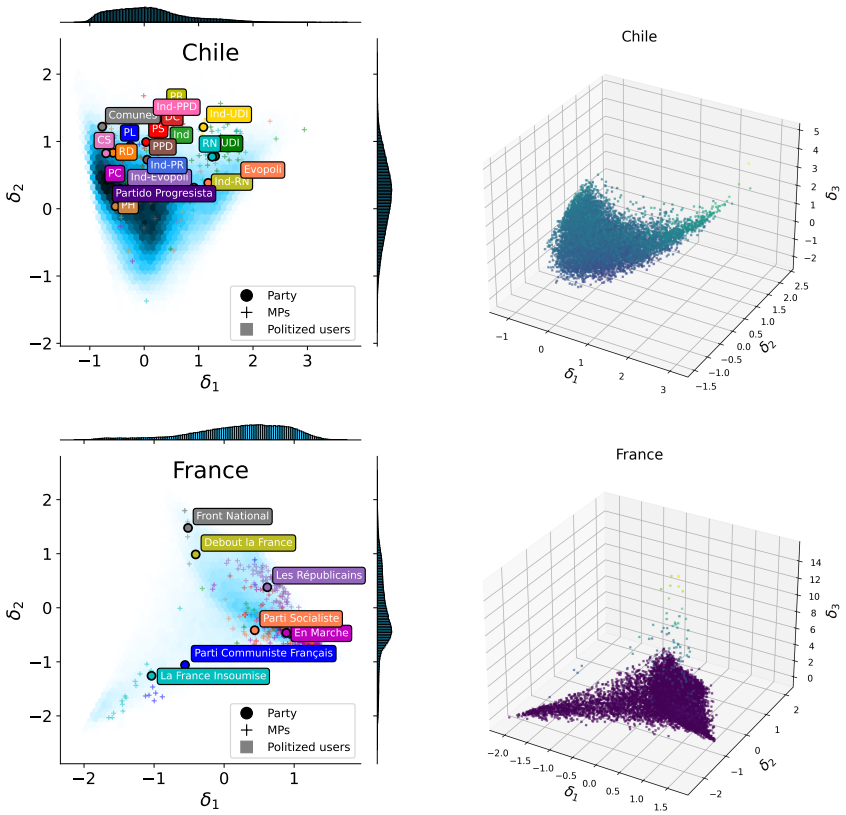

Fig. 1. Ideological embedding of Twitter networks collected for Chile and France. Parliamentarians, party positions computed as mean of parliamentarians per party, and density of their followers (left), and three-dimensional cloud points (right).

To associate an interpretation to dimensions of this ideological space (an issue for which dimensions act as indicators of opinions) we use external attitudinal reference from CHES and MARPOR datasets. For the purposes of validating these spatial positions, we will focus on the single, broadly-investigated issue of Left-Right cleavages [2]. To identify which dimension is related to Left-Right cleavages in each country we compare the position of parties according to the dimensions of ideological space, and the position of parties as provided by chosen external sources (MARPOR in Chile and CHES in France). Figure 2 show the correlation of positions using ideological space and external sources for the issue of Left-Right cleavages, present in both CHES and MARPOR data.

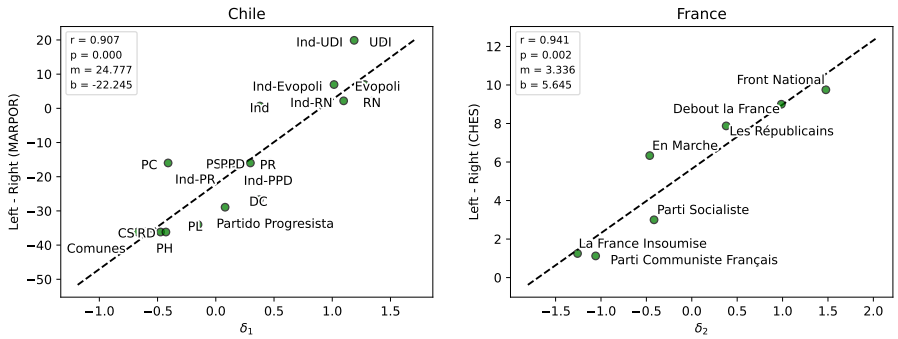

Fig. 2. Party positions in ideological embedding (δ_1 in Chile and δ_2 in France) and in attitudinal sources for Left-Right cleavages (MARPOR for Chile and CHES for France), with reported Pearson r, p-value, slope m, and intersect b.

Readers familiarized with Chilean and French politics, will have identified what correlations confirm: first Chilean dimension δ_1 and second French dimension δ_2 are latent variables related to users' positions on some Left-Right scale. For further details on the dimensionality structure of the French political space (*e.g.*, showing how δ_1 relates to attitudes towards the European Union), the reader is referred to [8].

5 Validating User Positioning Via Text Classification and Language Models

Thus far, analyses have focused on validating ideological embeddings using the position of MPs, as illustrated by Figure 2. In this section we propose a method to validate the position of their followers by assessing the accuracy of positioning in the Left-Right scale. To validate the ideological positions on the Left-Right scale we propose a model based on text classification showing the accuracy achievable by using latent dimensions (δ_1 in Chile and δ_2 in France) as indicators for Left

and Right stances of users. In each country, we classify users as being Left- and Right-leaning based exclusively on the text of their Twitter *bios* using automated text analysis. First, we filter users by the language in which they write their *bio* (Spanish in Chile and French in France) using CLD3 model for language identification [26]. This allowed us to identify 230.834 users in Chile (out of 418.979) and 148.145 users in France (out of 288.496) that write their *bios* in the language of each country. For each *bio*, we infer the sentiment of the text using the BERT multilingual base model for uncased words [27], which classify sentiments in 5 levels, from most negative to most positive. We classify *bios* as having positive sentiment if the sentiment predicted using the BERT language model is equivalent to 4 or 5 (on a scale from 1 to 5), and as having negative sentiment if the BERT-predicted sentiment is equivalent to 1 or 2. Additionally, we label users as belonging to two groups: 1) *Left* if they use the keywords "izquierda" or "izquierdista" in Spanish, and "gauche" or "gauchiste" in French ("left" or "leftist") with their variations in gender and number, and the *bio* has an identified positive sentiment, and 2) *Right* if they use the keywords "derecha" or "derechista" in Spanish, and "droite" in French ("right") with their variations, and the *bio* has an identified positive sentiment. This resulted in 1.886 users labeled *Left* and 1.550 users labeled *Right* in Chile (for a total of 3.436 users using "left" or "right" with positive sentiment in their profile *bios*), and in 891 users labeled *Left* and 739 users labeled *Right* in Chile (for a total of 1.630 users using "left" or "right" with positive sentiment in their profile *bios*). We consider these labeled users as the subset of users for which left or right stances can be determined using automated text analysis. We aim to show that our ideological dimensions serve as a good classifier for them too. Figure 3 shows the empirical densities of these groups of users per country along the pertinent ideological space dimension selected in Section 4.

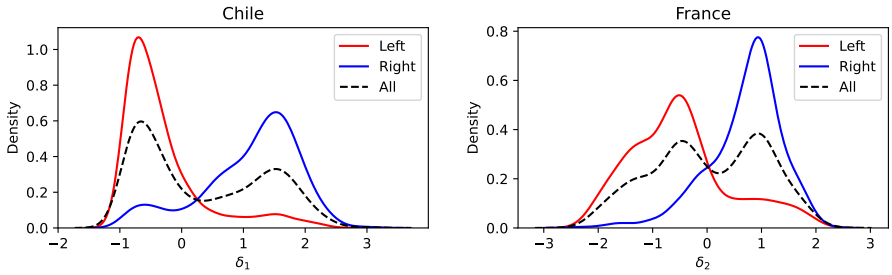

Fig. 3. Empirical density of distribution of the groups of users labeled *Left* or *Right* using automated text analysis along the ideological dimensions previously identified as containing Left-Right cleavages.

In order to show that the selected ideological dimension containing Left-Right cleavages (δ_1 in Chile and δ_2 in France), we fit a logistic model for the

classification of users based the positions of the labeled users identified in the *Left* and *Right* groups. To fit a logistic model for users in the *Right* based on ideological positions from Section 4, we consider *Right* users as *successes* of the model, and users on the *Left* as *failures* of the model. Because groups are complementary (individuals in our model construction are either in the *Left* or in the *Right*), we use the same logistic model to classify them (instead of using a separate model with inverted direction of the ideological dimension and in which the *Left* would be the *success* group and the *Right* the *failure* group). Figure 4 shows, for each country, the positions of users in both groups (represented by colored dots), the cumulative distribution function, and the fitted logistic model for classification. For each country, and for each group, we compute the goodness of fit of our logistic models as the precision and the recall if they were used to classify our labeled users.

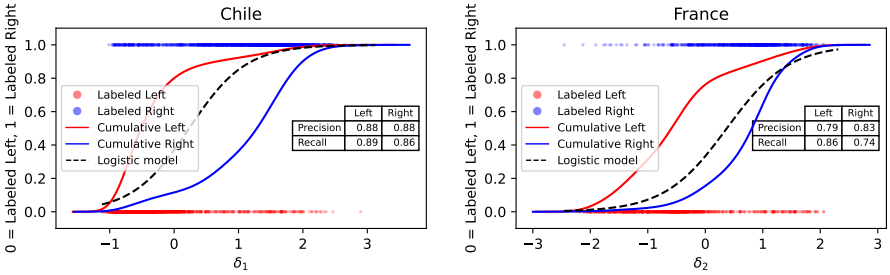

Fig. 4. Users in groups • *Left* and • *Right*, their cumulative distribution function, and the fitted logistic model for classification. Precision and recall for logistic classification (0.5 threshold) is reported in the embedded table.

After fitting a logistic model, we classify users with it by establishing a 0.5 threshold: users below the 0.5 threshold are classify as *Left*, while users over the threshold are classified as *Right*. We compute the performance as the comparison between their groups as classified using text analysis, and as resulting from the logistic model. The values from Figure 4 are read as follows. The Left-Right scale axis can classify users in Left and Right groups such that, in Chile 88% of users classified Left or Right are rightly classified, and 89% of true Left and 86% of true Right users are identified as such. In France, 79% of users classified Left and 83% classified Right are rightly classified, and 86% of true Left and 74% of true Right users are identified as such.

6 Conclusions and Discussion

We presented ideological embedding methods for Chile and France, two countries with different socio-political systems using the Correspondence Analysis

method. Using external attitudinal data we provided meaning to dimensions of this space identifying for both countries dimensions that capture Left-Right cleavages. We proposed a method based on language models to classify users (independently form ideological embedding procedures) as being Left or Right sympathizers, we showed that dimensions of the ideological space can provide accurate classification of users, and we provided measurements of accuracy of the positions. This amounts to the first validation of the positions achieved by ideological embedding in way that is applicable to a wide range of issues that may be captured by other dimensions by changing associated keywords (*e.g.*, "migrant", "immigration", *etc.*, to test the accuracy of dimensions capturing attitudes towards immigration). This method allows for the validation of the position of a large quantity of users (as opposed to validating the positions or MPs), paving the way for the use of ideological embedding methods as a massive, passive, frequent, and inexpensive polling tool. Besides polling, validated ideological positions of social network users in multidimensional ideological spaces in which the substance and meaning of dimensions is well-identified, opens an interesting field of research into many online phenomena at the intersection of Social Network Analysis, Opinion Dynamics, sociology, and political sciences. The positions of users can be used to measure polarization in different online arenas and the effect of algorithmic recommendation on them [4]. The positions of MPs and party sympathizers opens an interesting path for conducting studies of party competitions and ideological party diversity [28, 29].

Acknowledgments. This work has been funded by the French National Agency for Research under grant ANR-19-CE38-0006 "Geometry of Public Issues" (GOPI) and by CIVICA's "European Polarisation Observatory" (EPO).

References

1. Barberá P. Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. Political analysis. 2015;23(1):76–91.
2. Barberá P, Jost JT, Nagler J, Tucker JA, Bonneau R. Tweeting from left to right: Is online political communication more than an echo chamber? Psychological science. 2015;26(10):1531–1542.
3. Bond R, Messing S. Quantifying social media's political space: Estimating ideology from publicly revealed preferences on Facebook. American Political Science Review. 2015;109(1):62–78.
4. Ramaciotti Morales P, Cointet JP. Auditing the Effect of Social Network Recommendations on Polarization in Geometrical Ideological Spaces. In: Fifteenth ACM Conference on Recommender Systems; 2021. p. 627–632.
5. Ramaciotti Morales P, Cointet JP, Laborde J. Your most telling friends: Propagating latent ideological features on Twitter using neighborhood coherence. In: 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM); 2020. p. 217–221.
6. McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: Homophily in social networks. Annual review of sociology. 2001;27(1):415–444.
7. Temporão M, Kerckhove CV, van der Linden C, Dufresne Y, Hendrickx JM. Ideological scaling of social media users: a dynamic lexicon approach. Political Analysis. 2018;26(4):457–473.

8. Ramaciotti Morales P, Cointet JP, Zolotoochin GM. Unfolding the dimensionality structure of social networks in ideological embeddings. In: 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM); 2021. .
9. Poole KT, Rosenthal H. A spatial model for legislative roll call analysis. American Journal of Political Science. 1985:357–384.
10. Clinton J, Jackman S, Rivers D. The statistical analysis of roll call data. American Political Science Review. 2004;98(2):355–370.
11. Lowe W. Understanding wordscores. Political Analysis. 2008;16(4):356–371.
12. Barberá P, Rivero G. Understanding the political representativeness of Twitter users. Social Science Computer Review. 2015;33(6):712–729.
13. Imai K, Lo J, Olmsted J, et al. Fast estimation of ideal points with massive data. American Political Science Review. 2016;110(4):631–656.
14. Bakshy E, Messing S, Adamic LA. Exposure to ideologically diverse news and opinion on Facebook. Science. 2015;348(6239):1130–1132.
15. Baumann F, Lorenz-Spreen P, Sokolov IM, Starnini M. Modeling echo chambers and polarization dynamics in social networks. Physical Review Letters. 2020;124(4):048301.
16. Ecker A. Estimating policy positions using social network data: cross-validating position estimates of political parties and individual legislators in the Polish parliament. Social Science Computer Review. 2017;35(1):53–67.
17. Cointet JP, Ramaciotti Morales P, Cardon D, Froio C, Mogoutov A, Ooghe-Tabanou B, et al. What colours are the yellow vests? An ideological scaling of Facebook groups. Statistique et Société. 2020.
18. Jensen J, Naidu S, Kaplan E, Wilse-Samson L, Gergen D, Zuckerman M, et al. Political polarization and the dynamics of political language: Evidence from 130 years of partisan speech [with comments and discussion]. Brookings Papers on Economic Activity. 2012:1–81.
19. Kwak H, An J, Jing E, Ahn YY. FrameAxis: characterizing microframe bias and intensity with word embedding. PeerJ Computer Science. 2021;7:e644.
20. Liu B, Zhang L. A survey of opinion mining and sentiment analysis. In: Mining text data. Springer; 2012. p. 415–463.
21. Ravi K, Ravi V. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowledge-based systems. 2015;89:14–46.
22. van Vliet L, Törnberg P, Uitermark J. The Twitter parliamentarian database: Analyzing Twitter politics across 26 countries. PLoS one. 2020;15(9):e0237073.
23. Krause W, Lehmann P, Theres M, Merz N, Regel S, Wessels B. The Manifesto Data Collection: South America. Version 2020b. Berlin: Wissenschaftszentrum Berlin für Sozialforschung; 2020. Available from: https://doi.org/10.25522/manifesto.mpdssa.2020b.
24. Bakker R, Hooghe L, Jolly S, Marks G, Polk J, Rovny J, et al. 2019 Chapel Hill Expert Survey. Chapel Hill. 2020. Www.chesdata.eu.
25. Mikhaylov S, Laver M, Benoit K. Coder reliability and misclassification in comparative manifesto project codings. In: 66th MPSA annual national conference. vol. 3. Citeseer; 2008. p. 3.
26. Compact Language Detector v3 (CLD3);. Accessed: 2021-09-08. https://github.com/google/cld3.
27. Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:181004805. 2018.
28. Grossman E, Sauger N. Economic internationalization and the decline of the left–right dimension. Party Politics. 2019;25(1):36–49.

29. Ramaciotti Morales P, Lamarche-Perrin R, Fournier-S'Niehotta R, Poulain R, Tabourier L, Tarissan F. Measuring diversity in heterogeneous information networks. Theoretical Computer Science. 2021;859:80–115.