MA Thesis - Leon Siefken

For the Title: Master of Political Science

Applying Large Langue Models for the Measure of Populism:

Assessing the use of Simplistic Language in Election Manifestos

---

Abstract

In recent years, populist parties have witnessed a surge in popularity within Western democracies. Populism, a multi-dimensional and latent construct that is often merely implied, requires extensive measurement methods. Given these attributes of populism, a valid, fully automated measure that can capture its multi-dimensionality, latency, and non-compensatory aspects is yet to be proposed. This thesis aims to evaluate the applicability of Large Language Models (LLMs) for the automated text-analysis of populist rhetoric. More specifically, I employ Google's Gemini AI to code the paragraphs of German election manifestos from five national-level elections. This method of measurement holds the potential to account for populisms multi-dimensionality, non-compensatory aspect, and even its latency, due to Geminis ability to make subtle and human-like interpretations. The findings are forthcoming.

1 Introduction

With the recent upsurge of populist parties in Western democracies, scholars of political science have attempted to explain their appeal, success, and strategies (Mudde, 2004; Hawkins and Rovira Kaltwasser, 2017; Meijers and Zaslove, 2021). While many publications have focused on defining populism and what makes it appealing to the electorate, there is only a handful of research dedicated to operationalizing the concept in a way that aligns with its latent nature and non-compensatory characteristic (Wuttke, Schimpf and Schoen, 2020; Pauwels, 2011). Populism's latent nature often requires some degree of interpretation of what has been said or written. The antagonism between an elite and the citizens can be strategically subtle or in implication only, much like a dog whistle. This, in turn, requires sophisticated ways of measurement, making populism difficult to measure in an automated and reproducible way. Meijers and Zaslove (2021) deem expert surveys to be the gold standard for the measurement of populism, while other scholars have made use of text as data, hand-coding speeches, or election manifestos (Hawkins and Silva, 2016). Both ways are often costly and time-consuming. Other scholars have attempted to employ fully automated text analysis to measure populism. Di Cocco and Monechi (2021) have set up a machine learning model, trained on populist texts, to classify text as populist or not. However, this model was later found to produce invalid scores for populism (Jankowski and Huber, 2023). Another proposed computer-assisted measure is dictionary analysis by assigning scores to the text using a populism dictionary. This way of measuring populism in texts appears to have become the most common method to use, next to hand-coding (Breyer, 2023; Gründl, 2022; Pauwels, 2011; Rooduijn and Pauwels, 2011*a*). However, the application of dictionaries for the measure of ideologies, sentiments, and attitudes has been criticized and was deemed the least valid method of text analysis by Van Atteveldt, Van der Velden and Boukes (2021). These drawbacks potentially intensify when applied to a latent and multi-dimensional concept such as populism, which can require some degree of interpretation. Something that a dictionary analysis can not do or account for.

This thesis aims to explore the potential of Large Language Models (LLM) as a new method of measurement and assess their applicability for the automated analysis of text for populist discourse. Specifically, I apply Google's novel Gemini AI to analyze

German election manifestos from the national-level elections between 2005 and 2021. LLMs have the advantage of being able to do interpretations of texts, potentially being able to account for populism's latent nature. The LLM was tasked with coding each paragraph of each manifesto for people-centrism and anti-elitism, using a slightly adjusted version of the coding scheme by Hawkins and Silva (2016) as Gemini's coding prompt. The LLM's performance is then assessed by comparing its coding to hand-coding of the same paragraphs. Scores are then compared to those generated by a dictionary analysis to assess if LLMs offer a more accurate measurement. To better understand Gemini's coding, topic models are estimated on the paragraphs deemed populist. If valid, one or more topics should reflect populism and its dimensions. Lastly, the data is used to reproduce the findings by Bischof and Senninger (2018). The authors find that populist rhetoric correlates with the use of simplistic language in election manifestos. Answering this question using all three measurements (LLM, hand-coding, and dictionary) will help to assess whether potential differences in coding affect the results of an analysis or if this is merely a technicality.

2 Populism in Political Texts

2.1 Ideational and Non-Compensatory

The thesis will rely on the so-called ideational definition of populism as a thin-centered ideology (Mudde, 2004). This understanding of populism was first proposed by Mudde (2004) and built upon definitions of populism as a political strategy and a style of discourse by arguing that both are symptoms of an underlying ideology (Gidron and Bonikowski, 2013; Pauwels, 2011). Populism is understood as

> "[...] an ideology that considers society to be ultimately separated into two homogeneous and antagonistic groups, the pure people versus the corrupt elite, and which argues that politics should be an expression of the volonté générale (general will) of the people."(Mudde, 2004)

The definition implies two major actors related to each other in a Manichean world-view. On the one side, there are *the pure people*, a homogenous group that defines a silent majority within society, striving for the achievement of the *general will*. On the other side, there is *the elite*, the ones in power. Also, a homogenous group with

opposing interests to the people, described as corrupt and exclusively following selfish interests. These two groups are seen as being in an antagonistic relationship with each other, a *pure good* versus *pure evil*, giving populism a moralistic and normative character (Hawkins and Rovira Kaltwasser, 2017). For populists, politics should exclusively be an expression of what they define to be the will of the people. However, populism is not understood as an ideology that can exist independently. The "thin-centered" aspect of populism implies that populism as an ideology has a chameleonic character that can be bent to be combined with any other ideology. This could be e.g. nationalism, liberalism, or socialism, which act as so-called "host ideologies". The host ideology defines who is understood to be part of the corrupt and who defines the pure people (Meijers and Zaslove, 2021; Mudde, 2004). Generally, this definition implies that populism has at least two dimensions, anti-elitism, and people-centrism, with the antagonistic aspect being a part of both. Meijers and Zaslove (2021) argue that something or someone can only be described as populist when both dimensions of its definition are present. This non-compensatory understanding of populism is shared by Wuttke, Schimpf and Schoen (2020), who argue that one dimension of (attitudinal) populism can not substitute the other and that this way of measuring populism is necessary to justify it being an original concept and not simply an alternative measure for what is essentially political trust or external political efficacy (Geurkink et al., 2020).

Using this definition, populism becomes measurable through its symptoms, so, among others, through specific policies, the type of leadership, and most importantly, the strategic employment of populist language. This in turn makes populism and its dimensions measurable in party literature (Pauwels, 2011).

2.2 Populist Discourse in Party Literature

Party literature, and specifically election manifestos, make for an interesting case for analysis. In nature, they are a document that comprises a party's ideas, plans, and goals, agreed upon by the party. This makes them more suitable for analyzing and comparing parties' stances, strategies, and ideology than e.g. a candidate's speech (Volkens et al., 2021; Rooduijn, de Lange and van der Brug, 2014). This is regularly done for different concepts in the Manifesto Project by Volkens et al. (2021). The expectation of strategic use of specific language stems from the fact that election manifestos are directed toward the electorate, expressing how where they see room for improvement

in politics as well as their stances on different issues. For populism specifically, this expectation has had mixed empirical findings (Di Cocco and Monechi, 2021; Meijers and Zaslove, 2021; Pauwels, 2017; Hawkins and Rovira Kaltwasser, 2019*a*). Scholars have observed election manifestos to be strategically sober in tone when compared to e.g. speeches. As only a few voters are likely to read a party's full manifesto, parties have little incentive to employ populist language and instead keep the tone moderate if a coalition seems possible to them (Rooduijn, de Lange and van der Brug, 2014; Krause and Wagner, 2021). On the other hand, Siefken (2024) has found that the use of populist discourse in manifestos is dependent on the manifestos, with introductions, conclusions, and chapters about the state of democracy showing the highest chance for populist language. This is likely linked to the fact that introductions and conclusions often summarize the parties' identity and goals, also likely making them the most appealing chapters to potential voters. This finding could also be made by Hawkins and Rovira Kaltwasser (2019*b*). Chapters about democracy, on the other hand, lend themselves to including a high amount of populism as it can be seen as an issue over which populist parties might claim ownership (not regarding the host ideology) (Siefken, 2024; Green-Pedersen, 2007). Thus, election manifestos might be kept more moderate in tone, but are not immune to strategically including populism.

Thus, it should be expected that populist scores calculated using manifestos might be rather low, but still vary between parties. This makes populist scores based on manifestos not as suitable for an in-depth understanding of how populist a party and its members are, as e.g. the POPPA data (Meijers and Zaslove, 2021) would be, but still offer a valuable score for comparison of the degree of populism within different parties. And, as argued before, the nature of manifestos as a regularly published document agreed upon by the party members makes them more valuable for comparing parties than speeches of single candidates would. Thus, manifestos were chosen as the unit of analysis in the paper.

3 Measuring Populism in Political Texts

Populism's latent nature requires sophisticated measures, often requiring some degree of interpretation. Over the years, scholars have proposed different methods for the measurement of populism in texts, which Rooduijn and Pauwels (2011*b*) categorizes

into being either classic content analysis (e.g. hand-coding) or computer-based content analysis. Three of these measures and their benefits and drawbacks will briefly be discussed.

Hand-coding is often understood as the most reliable way of text analysis (Van Atteveldt, Van der Velden and Boukes, 2021). It requires two or more coders, trained using a codebook and exemplary data, to read through the text and code it, its paragraphs, or sentences manually (Rooduijn and Pauwels, 2011*b*). It is the coders' choice if the read text contains no, one, or both dimensions of populism, while sufficient inter-coder reliability assures that all coders are measuring the same concept, assuring the reliability of the measurement. Populism often requires interpretation, being rather chameleonic, taking different forms depending on the host ideology, makes coding done by human coders, fluent in the language of the texts, the ideal way to measure populism (and other concepts) in texts (Van Atteveldt, Van der Velden and Boukes, 2021; March, 2019; Engesser et al., 2017). While scholars like Hawkins (2009), Hawkins and Rovira Kaltwasser (2019*b*), or Jagers and Walgrave (N.d.) have applied holistic classification of texts, scholars like Rooduijn and Pauwels (2011*b*) and Rooduijn, de Lange and van der Brug (2014) focused on analyzing texts on the paragraph level. Both approaches classified a text or paragraph as populist when both, anti-elitism and people-centrism, were present. However, the coding on a paragraph level allows for an aggregation e.g. to relative frequencies, resulting in a continuous measure of populism, in line with the definition of populism in this thesis. However, while this classic hand-coding can be considered the gold standard for text analysis it comes with the the drawback of being heavily time consuming. Not only do the texts have to be read and classified, but human coders also have to be trained beforehand to reliably execute the classification in line with the operationalization. Additionally, the coding is based on human interpretation and its biases, making the outcome potentially slightly different when applied to the same texts with other coders.

Computer-assisted measures attempt to fix these drawbacks by offering reproducible, standardized, and time-efficient methods for analyzing texts. When it comes to computer-assisted measures for populism in texts, mainly two methods have been proposed. The first is the dictionary analysis, which can be understood as counting the relative frequency of words from a pre-defined dictionary in the respective texts. The resulting score is used as a proxy for the prevalence of the concept of interest. (Grimmer

and Stewart, 2013). For populism, different dictionaries have been proposed. Gründl (2022) recently published a dictionary for measuring populism in social media posts specifically, focusing on the demand side of populism and populist attitudes. Pauwels (2011) proposed a dictionary for the analysis of party literature, specifically excluding election manifestos due to their moderate tone. While both scholars claim to make valuable findings using the dictionaries, other scholars have called the method into question (Pauwels, 2017; Bruycker and Rooduijn, 2021; Van Atteveldt, Van der Velden and Boukes, 2021). It appears to produce a significant number of false positives and negatives, as it lacks any understanding of the words and their context, making interpretations impossible. This problem becomes especially prominent when analyzing a theoretical concept as latent and in need of interpretation such as populism. If the text refers to the people, the government, or the system positively or negatively can hardly be measured with a dictionary analysis. Additionally, the findings are highly dependent on the researchers' choice of words to include in the dictionary. Thus Van Atteveldt, Van der Velden and Boukes (2021) conclude that the dictionary approach is the least valid method of text analysis. Lastly, dictionaries fail to account for the multidimensional and non-compensatory aspects, as defined in the ideational definition of populism. If a manifesto includes a high frequency of words tied to anti-elitism, the dimension can compensate e.g. a complete lack of people-centrism, falsely giving the text a high populism score. However, as of now, it is the most commonly applied method for measuring populism in texts (??????????????????).

Another computer-assisted method for measuring populism has been proposed by Di Cocco and Monechi (2021). They use supervised machine learning in an attempt to develop an algorithm trained on election manifestos for classifying texts as populist or not. Compared to the dictionary approach, the advantage of this method is that the algorithm can account for the word's context and can be seen as a decent middle-ground between hand-coding and dictionary analysis Van Atteveldt, Van der Velden and Boukes (2021). While Di Cocco and Monechi (2021) claim sufficient findings, as their scores correlate with the classification in the PopuList data, Hawkins and Rovira Kaltwasser (2019*b*) find that this method struggles with classifying texts that, by a hand-coder, would be placed as in between non-populist and strongly populist. In a later paper by Jankowski and Huber (2023), the paper by Di Cocco and Monechi (2021) has been heavily criticized for its validity checks. Exploring the proposed method

further, they find that instead of measuring populism, the algorithm heavily relies on other contents such as the party names or the policy positions within the manifestos for its classification. While Hawkins and Rovira Kaltwasser (2019*b*) see potential in this method for the future, it appears to be insufficient as of now.

Generally, it appears that hand-coding is the only fully valid measure for populism in texts. But as described earlier, it is heavily time-consuming and will become costly when having to employ coders. While supervised machine learning has the potential to be improved in the future, it still needs further refinement until it reaches the point of sufficiency. However, the recent rise and fast development of Large Language Models (LLM) might offer a novel way of computer-assisted classification by having the artificial intelligence act out as a trained hand-coder.

4 The State of LLMs for Text-Analysis

Transformer-based models such as GPT or Goolge's Gemini have seen quick development in recent years. Their advantage compared to e.g. the dictionary approach is that these models do not rely on the bag-of-words approach but will account for the order of the words, and the context they are in, and can understand and interpret the content of the texts (Terechshenko et al., 2020). With this, scholars have tried to assess their potential for helping with research. However, the application within the social sciences is sparse for now, with numerous papers being under review.

Scholars like Argyle, Bail, Busby, Gubler, Howe, Rytting, Sorensen and Wingate (2023), Argyle, Busby, Fulda, Gubler, Rytting and Wingate (2023), and Bisbee et al. (2024) have used LLMs in social sciences for simulating samples and survey participants and as an intervention in an experiment. OpenAI offers versions of GPT specifically for helping with programming, finding the consensus in a research strain, or data analysis (OpenAI, 2024). Other papers explore the coding capabilities of LLMs, in line with this paper (Gilardi, Alizadeh and Kubli, 2023; Ornstein, Blasingame and Truscott, 2023; Terechshenko et al., 2020; Wu et al., 2023). In a paper by Terechshenko et al. (2020), LLMs are used for the classification of a variety of political texts, with the findings being compared to the coding of other methods of classification. As the transformer models can interpret the context of each word, the authors find that LLMs outperform other automated models for classification such as embeddings or n-grams, showing higher

accuracy in its coding. Wu et al. (2023) use GPT-3.5 for the estimation of US senators' ideological positions by having GPT pairwise compare the senators and having the model tell which of them is more conservative when it comes to abortion, gun control, and general ideology. While ChatGPT is not trained specifically for these kinds of tasks and we can't be sure what GPT's decision is based on, their findings highly correlate with the DW-Nominate scores of said senators, signaling some possible validity of the measure (Wu et al., 2023). A paper by Ornstein, Blasingame and Truscott (2023) comes to similar conclusions and shows evidence that LLMs also outperform supervised machine learning when it comes to a variety of tasks such as sentiment analysis or topic modeling. They go as far as arguing that the performance of GPT is on par with expert- and crowd-coding. Gilardi, Alizadeh and Kubli (2023) directly compare ChatGPT's performance to MTurk crowd-coding for annotating political texts for stance, framing, topics, and relevance. What they find is that the LLM is not only cheaper and faster than crow-coding, it also outperforms it in most tasks by about 25%. On the other hand, Koco et al. (2023) test GPT on 25 different tasks and find that it struggles with more difficult tasks, leading to a high loss of information when compared to state-of-the-art methods of said task. One of those problematic tasks was analyzing texts for emotions, which is key for measuring populist discourse and its Manichean good people versus evil elites character.

Thus, in conclusion, the previous application of LLM by social scientists has shown generally positive and promising findings in regard to annotating political texts, however, there seems to be significant losses when applied to difficult tasks. This makes the attempt to measure ideational populism with LLM especially interesting, as not only computer-based approaches but also human coders can struggle to reliably, and validly perform this task due to populism's latency and chameleonic character.

5 Data and Methods

References

Argyle, Lisa P., Christopher A. Bail, Ethan C. Busby, Joshua R. Gubler, Thomas Howe, Christopher Rytting, Taylor Sorensen and David Wingate. 2023. "Leveraging AI for democratic discourse: Chat interventions can improve online political conversations at scale." *Proceedings of the National Academy of Sciences* 120(41):e2311627120. **URL:** *https://www.pnas.org/doi/abs/10.1073/pnas.2311627120*

Argyle, Lisa P., Ethan C. Busby, Nancy Fulda, Joshua R. Gubler, Christopher Rytting and David Wingate. 2023. "Out of One, Many: Using Language Models to Simulate Human Samples." *Political Analysis* 31(3):337351.

Bisbee, James, Joshua D. Clinton, Cassy Dorff, Brenton Kenkel and Jennifer M. Larson. 2024. "Synthetic Replacements for Human Survey Data? The Perils of Large Language Models." *Political Analysis* p. 116.

Bischof, Daniel and Roman Senninger. 2018. "Simple politics for the people? Complexity in campaign messages and political knowledge." *European Journal of Political Research* 57(2):473–495.
URL: *https://ejpr.onlinelibrary.wiley.com/doi/abs/10.1111/1475-6765.12235*

Breyer, Magdalena. 2023. "Populist positions in party competition: Do parties strategically vary their degree of populism in reaction to vote and office loss?" *Party Politics* 29(4):672–684.
URL: *https://doi.org/10.1177/13540688221097082*

Bruycker, Iskander De and Matthijs Rooduijn. 2021. "The Peoples Champions? Populist Communication as a Contextually Dependent Political Strategy." *Journalism & Mass Communication Quarterly* 98(3):896–922.

Di Cocco, J. and B. Monechi. 2021. "How Populist are Parties? Measuring Degrees of Populism in Party Manifestos Using Supervised Machine Learning." *Political Analysis* pp. 1–17.

Engesser, Sven, Nicole Ernst, Frank Esser and Florin Büchel. 2017. "Populism and social media: how politicians spread a fragmented ideology." *Information, Communication & Society* 20(8):1109–1126.

Geurkink, Bram, Andrej Zaslove, Roderick Sluiter and Kristof Jacobs. 2020. "Populist Attitudes, Political Trust, and External Political Efficacy: Old Wine in New Bottles?" *Political Studies* 68(1):247–267.
URL: *https://doi.org/10.1177/0032321719842768*

Gidron, Noam and Bart Bonikowski. 2013. Varieties of Populism: Literature Review and Research Agenda. Weatherhead Working Paper Series, No. 13-0004 Weatherhead Working Paper Series, No. 13-0004.

Gilardi, Fabrizio, Meysam Alizadeh and Maël Kubli. 2023. "ChatGPT outperforms crowd workers for text-annotation tasks." *Proceedings of the National Academy of Sciences* 120(30):e2305016120.
URL: *https://www.pnas.org/doi/abs/10.1073/pnas.2305016120*

Green-Pedersen, C. 2007. "The Growing Importance of Issue Competition: The Changing Nature of Party Competition in Western Europe." *Political Studies* 55(3):607–628.

Grimmer, Justin and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21(3):267297.

Gründl, Johann. 2022. "Populist ideas on social media: A dictionary-based measurement of populist communication." *New Media & Society* 24(6):1481–1499.

Hawkins, K.A. and B.C. Silva. 2016. "A Head-to-Head Comparison of Human-Based and Automated Text Analysis for Measuring Populism in 27 Countries." *Explaining Populism: Team Populism January Conference* Provo-UT(January):28–30.

Hawkins, Kirk A. 2009. "Is Chávez Populist?: Measuring Populist Discourse in Comparative Perspective." *Comparative Political Studies* 42(8):1040–1067.

Hawkins, Kirk A. and Cristóbal Rovira Kaltwasser. 2017. "The Ideational Approach to Populism." *Latin American Research Review* 52(4):513528.

Hawkins, Kirk. A. and Cristóbal Rovira Kaltwasser. 2019*a*. Introduction: The ideational approach. In *The Ideational Approach to Populism: Concept, Theory, and Analysis*, ed. Kirk A. Hawkins, Ryan E. Carlin, Levente Littvay and Cristóbal Rovira Kaltwasser. New York, USA: Routledge pp. 1–24.

Hawkins, Kirk. A. and Cristóbal Rovira Kaltwasser. 2019*b*. Textual analysis: Big data Approaches. In *The Ideational Approach to Populism: Concept, Theory, and Analysis*, ed. Kirk A. Hawkins, Ryan E. Carlin, Levente Littvay and Cristóbal Rovira Kaltwasser. New York, USA: Routledge pp. 27–48.

Jagers, Jan and Stefaan Walgrave. N.d. "Populism as political communication style: An empirical study of political parties' discourse in Belgium." *European Journal of Political Research*. Forthcoming.

Jankowski, Michael and Robert A. Huber. 2023. "When Correlation Is Not Enough: Validating Populism Scores from Supervised Machine-Learning Models." *Political Analysis* 31(4):591605.

Koco, Jan, Igor Cichecki, Oliwier Kaszyca, Mateusz Kochanek, Dominika Szydo, Joanna Baran, Julita Bielaniewicz, Marcin Gruza, Arkadiusz Janz, Kamil Kanclerz, Anna Koco, Bartomiej Koptyra, Wiktoria Mieleszczenko-Kowszewicz, Piotr Mikowski, Marcin Oleksy, Maciej Piasecki, ukasz Radliski, Konrad Wojtasik, Stanisaw Woniak and Przemysaw Kazienko. 2023. "ChatGPT: Jack of all trades, master of none." *Information Fusion* 99:101861.
URL: *https://www.sciencedirect.com/science/article/pii/S156625352300177X*

Krause, Werner and Aiko Wagner. 2021. "Becoming part of the gang? Established and nonestablished populist parties and the role of external efficacy." *Party Politics* 27(1):161–173.

March, Luke. 2019. Textual Analysis. The UK Party System. In *The Ideational Approach to Populism: Concept, Theory, and Analysis*, ed. Kirk A. Hawkins, Ryan E. Carlin, Levente Littvay and Cristóbal Rovira Kaltwasser. New York, USA: Routledge pp. 49–66.

Meijers, M.J. and A. Zaslove. 2021. "Measuring populism in political parties: appraisal of a new approach." *Comparative Political Studies* 54(2):378–390.

Mudde, C. 2004. "The Populist Zeitgeist." *Government and Opposition* 39:541–563.

OpenAI. 2024. "Introducing GPTs.". [Online; accessed 20-May-2023].
URL: *https://openai.com/index/introducing-gpts/*

Ornstein, Joseph T., Elise N. Blasingame and Jake S. Truscott. 2023. "How to Train Your Stochastic Parrot: Large Language Models for Political Texts.". [Online; accessed 24-May-2024].
URL: *https://joeornstein.github.io/publications/ornstein-blasingame-truscott.pdf*

Pauwels, Teun. 2011. "Measuring Populism: A Quantitative Text Analysis of Party Literature in Belgium." *Journal of Elections, Public Opinion and Parties* 21(1):97–119.

Pauwels, Teun. 2017. CHAPTER 6: MEASURING POPULISM: A REVIEW OF CURRENT APPROACHES: A Handbook. In *Political Populism: A Handbook*, ed. Reinhard Heinisch, Christina Holtz-Bacha and Oscar Mazzoleni. International studies on populism, Bd. 3 Baden-Baden, Germany: Nomos pp. 123–136.

Rooduijn, M., S. L. de Lange and W. van der Brug. 2014. "A populist Zeitgeist? Programmatic contagion by populist parties in Western Europe." *Party Politics* 20(4):563–575.

Rooduijn, M. and T. Pauwels. 2011*a*. "Measuring Populism: Comparing Two Methods of Content Analysis." *West European Politics* 34(2):372–407.

Rooduijn, Matthijs and Teun Pauwels. 2011*b*. "Measuring Populism: Comparing Two Methods of Content Analysis." *West European Politics* 34(6):1272–1283.

Siefken, Leon. 2024. "When are Populist Parties Populist? A Study of Populist Sentiment in Election Manifestos of Germanys AfD and Die Linke." *German Politics* 0(0):1–24.

Terechshenko, Zhanna, Fridolin Linder, Vishakh Padmakumar, Fengyuan Liu, Jonathan Nagler, Joshua Aaron Tucker and Richard Bonneau. 2020. "A Comparison of Methods in Political Science Text Classification: Transfer Learning Language Models for Politic.".
URL: *https://ssrn.com/abstract=3724644*

Van Atteveldt, W., M. A. C. G. Van der Velden and M. Boukes. 2021. "The Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary Approaches, and Machine Learning Algorithms." https://doi.org/10.1080/19312458.2020.1869198. [Online; accessed 18-May-2024].

Volkens, Andrea, Tobias Burst, Werner Krause, Pola Lehmann, Theres MatthieSS, Nicolas Merz, Sven Regel, Bernhard WeSSels and Lisa Zehnter. 2021. "The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR). Version 2021a.".
URL: *https://doi.org/10.25522/manifesto.mpds.2021a*

Wu, Patrick Y., Jonathan Nagler, Joshua A. Tucker and Solomon Messing. 2023. "Large Language Models Can Be Used to Estimate the Latent Positions of Politicians.". [Online; accessed 24-May-2024].
URL: *https://www.patrickywu.com/PatrickYWu$_J$MP1$_L$aMPscores.pdf*

Wuttke, Alexander, Christian Schimpf and Harald Schoen. 2020. "When the Whole Is Greater than the Sum of Its Parts: On the Conceptualization and Measurement of Populist Attitudes and Other Multidimensional Constructs." *American Political Science Review* 114(2):356374.