1 Natural Language Processing for Policymaking

Zhijing Jin[a] & Rada Mihalcea[b]
[a]Max Planck Institute for Intelligent Systems, Tübingen, Germany & ETH Zürich, Zürich, Switzerland [b]University of Michigan, Ann Arbor, USA

Language is the medium for many political activities, from campaigns to news reports. Natural language processing (NLP) uses computational tools to parse text into key information that is needed for policymaking. In this chapter, we introduce common methods of NLP, including text classification, topic modeling, event extraction, and text scaling. We then overview how these methods can be used for policymaking through four major applications including data collection for evidence-based policymaking, interpretation of political decisions, policy communication, and investigation of policy effects. Finally, we highlight some potential limitations and ethical concerns when using NLP for policymaking.

Keywords: natural language processing, text analysis, policymaking, artificial intelligence, machine learning

1 Introduction

Language is an important form of data in politics. Constituents express their stances and needs in text such as social media and survey responses. Politicians conduct campaigns through debates, statements of policy positions, and social media. Government staff needs to compile information from various documents to assist in decision-making. Textual data is also prevalent through the documents and debates in the legislation process, negotiations and treaties to resolve international conflicts, and media such as news reports, social media, party platforms, and manifestos.

Natural language processing (NLP) is the study of computational methods to automatically analyze text and extract meaningful information for subsequent analysis. The importance of NLP for policymaking has been highlighted since the

Zhijing Jin and Rada Mihalcea. *Natural Language Processing for Policymaking.*
Handbook of Computational Social Science for Policy, European Commission Joint Research Center, 2023.

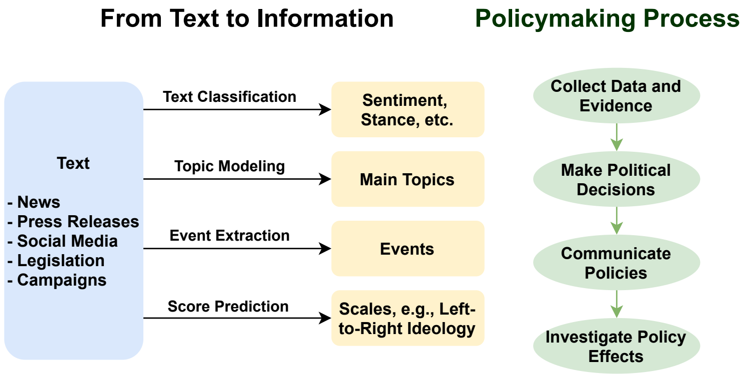

Figure 1: Overview of NLP for policymaking.

last century (Gigley 1993). With the recent success of NLP and its versatility over tasks such as classification, information extraction, summarization, and translation (Devlin et al. 2019; Brown et al. 2020), there is a rising trend to integrate NLP into the policy decisions and public administrations (Misuraca et al. 2020; Engstrom et al. 2020; Van Roy et al. 2021). Main applications include extracting useful, condensed information from free-form text (Engstrom et al. 2020), and analyzing sentiment and citizen feedback by NLP (Biran et al. 2022) as in many projects funded by EU Horizon projects (European Commission 2017). Driven by the broad applications of NLP (Jin, Chauhan, et al. 2021), the research community also starts to connect NLP with various social applications in the fields of computational social science (Lazer et al. 2009; Shah et al. 2015; Engel et al. 2021; Luz 2022) and political science in particular (Grimmer & Stewart 2013; Glavaš et al. 2019).

We show an overview of NLP for policymaking in Figure 1. According to this overview, the chapter will consist of three parts. First, we introduce in Section 2 NLP methods that are applicable to political science, including text classification, topic modeling, event extraction, and score prediction. Next, we cover a variety of cases where NLP can be applied to policymaking in Section 3. Specifically, we cover four stages: analyzing data for evidence-based policymaking, improving policy communication with the public, investigating policy effects, and interpreting political phenomena to the public. Finally, we will discuss limitations and ethical considerations when using NLP for policymaking in Section 4.

Table 1: Four common NLP methods, the type of information extracted by each of them, and example applications.

NLP Method	Information to Extract	Example Applications
Text classification	Category of text	Identify the sentiment, stance, etc.
Topic modeling	Key topics in text	Summarize topics in political agenda
Event extraction	List of events	Extract news events, international conflicts
Score prediction	Score	Text scaling

2 NLP for Text Analysis

NLP brings powerful computational tools to analyze textual data (Jurafsky & Martin 2000). According to the type of information that we want to extract from the text, we introduce four different NLP tools to analyze text data: text classification (by which the extracted information is the *category* of the text), topic modeling (by which the extracted information is the *key topics* in the text), event extraction (by which the extracted information is the list of *events* mentioned in the text), and score prediction (where the extracted information is a *score* of the text). Table 1 lists each method with the type of information it can extract and some example application scenarios, which we will detail in the following subsections.

2.1 Text Classification

As one of the most common types of text analysis methods, text classification reads in a piece of text and predicts its category using an NLP text classification model, as in Figure 2.

There are many off-the-shelf existing tools for text classification (Yin et al. 2019; Brown et al. 2020; Loria 2018) such as the implementation[1] using the Python package `transformers` (Wolf et al. 2020). A well-known subtask of text classification is sentiment classification (also known as sentiment analysis, or opinion mining), which aims to distinguish the subjective information in the text, such as positive or negative sentiment (Pang & Lee 2007). However, the existing tools only do well in categories that are easy to predict. If the categorization is customized and very specific to a study context, then there are two common solutions. One is to use dictionary-based methods, by a list of frequent keywords that correspond to a certain category (Albaugh et al. 2013) or using general linguistic dictionaries such as the Linguistic Inquiry and Word Count (LIWC) dictionary

[1] https://discuss.huggingface.co/t/new-pipeline-for-zero-shot-text-classification/681

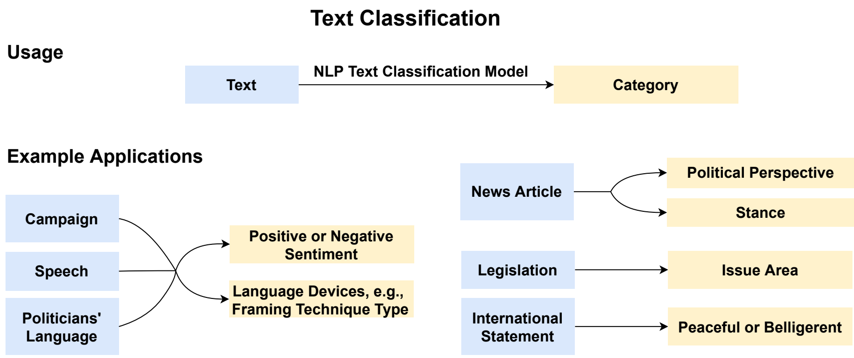

Figure 2: The usage and example applications of text classification on political text.

(Pennebaker et al. 2001). The second way is to adopt the data-driven pipeline, which requires human hand coding of documents into a predetermined set of categories, then train an NLP model to learn the text classification task (Sun et al. 2019), and verify the performance of the NLP model on a held-out subset of the data, as introduced in Grimmer & Stewart (2013). An example of adapting the state-of-the-art NLP models on a customized dataset is demonstrated in this guide.[2]

Using the text classification method, we can automate many types of analyses in political science. As listed in the examples in Figure 2, researchers can detect political perspective of news articles (Huguet Cabot et al. 2020), the stance in media on a certain topic (Luo et al. 2020), whether campaigns use positive or negative sentiment (Ansolabehere & Iyengar 1995), which issue area is the legislation about (Adler & Wilkerson 2011), topics in parliament speech (Albaugh et al. 2013; Osnabrügge et al. 2021), congressional bills (Hillard et al. 2008; Collingwood & Wilkerson 2012) and political agenda (Karan et al. 2016), whether the international statement is peaceful or belligerent (Schrodt 2000), whether a speech contains positive or negative sentiment (Schumacher et al. 2016), and whether a U.S. Circuit Courts case decision is conservative or liberal (Hausladen et al. 2020). Moreover, text classification can also be used to categorize the type of language devices that politicians use, such as what type of framing the text uses (Huguet Cabot et al. 2020), and whether a tweet uses political parody (Maronikolakis et al. 2020).

[2] https://skimai.com/fine-tuning-bert-for-sentiment-analysis/

2.2 Topic Modeling

Topic modeling is a method to uncover a list of frequent topics in a corpus of text. For example, news articles that are against vaccination might frequently mention the topic "autism," whereas news articles supporting vaccination will be more likely to mention "immune" and "protective." One of the most widely used models is the Latent Dirichlet Allocation (LDA) (Blei et al. 2001) which is available in the Python packages NLTK and Gensim, as in this guide.[3]

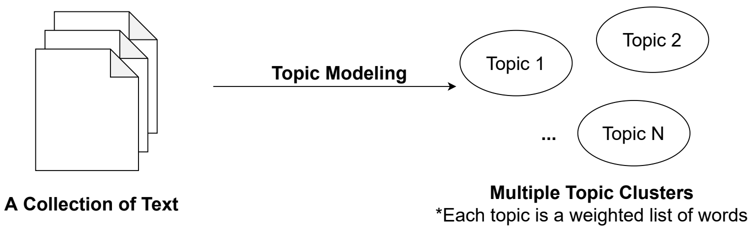

Figure 3: Given a collection of text documents, topic modeling generates a list of topic clusters.

Specifically, LDA is a probabilistic model that models each topic as a mixture of words, and each textual document can be represented as a mixture of topics. As in Figure 3, given a collection of textual documents, LDA topic modeling generates a list of topic clusters, for which the number N of topics can be customized by the analyst. In addition, if needed, LDA can also produce a representation of each document as a weighted list of topics. While often the number of topics is predetermined by the analyst, this number can also be dynamically determined by measuring the perplexity of the resulting topics. In addition to LDA, other topic modeling algorithms have been used extensively, such as those based on principal component analysis (PCA) (Chung & Pennebaker 2008).

Topic modeling, as described in this section, can facilitate various studies on political text. Previous studies analyzed the topics of legislative speech (Quinn et al. 2010; 2006), Senate press releases (Grimmer 2010a), and electorial manifestos (Menini et al. 2017).

2.3 Event Extraction

Event extraction is the task of extracting a list of events from a given text. It is a subtask of a larger domain of NLP called information extraction (Manning et al. 2008). For example, the sentence "Israel bombs Hamas sites in Gaza" expresses

[3] https://skimai.com/fine-tuning-bert-for-sentiment-analysis/

an event "*Israel* $\xrightarrow{bombs}$ *Hamas sites*" with the location "*Gaza*." Event extraction usually incorporates both entity extraction (e.g., Israel, Hamas sites, and Gaza in the previous example) and relation extraction (e.g., "bombs" in the previous example).

Event extraction is a handy tool to monitor events automatically, such as detecting news events (Walker et al. 2006; Mitamura et al. 2017), and detecting international conflicts (Azar 1980; Trappl 2006). To foster research on event extraction, there are tremendous efforts into textual data collection (McClelland 1976; Schrodt & Hall 2006; Merritt et al. 1993; Raleigh et al. 2010; Sundberg & Melander 2013), event coding schemes to accommodate different political events (Goldstein 1992; Bond et al. 1997; Gerner et al. 2002), and dataset validity assessment (Schrodt & Gerner 1994).

As for event extraction models, similar to text classification models, there are off-the-shelf tools such as the Python packages `stanza` (Qi et al. 2020) and `spaCy` (Honnibal et al. 2020). In case of customized sets of event types, researchers can also train NLP models on a collection of textual documents with event annotations (Hogenboom et al. 2011; Liu et al. 2020: *inter alia*).

2.4 Score Prediction

NLP can also be used to predict a score given input text. A useful application is political text scaling, which aims to predict a score (e.g., left-to-right ideology, emotionality, and different attitudes towards the European integration process) for a given piece of text (e.g., political speeches, party manifestos, and social media posts) (Laver et al. 2003; Lowe et al. 2011; Slapin & Proksch 2008; Gennaro & Ash 2021: *inter alia*).

Traditional models for text scaling include Wordscores (Laver et al. 2003) and WordFish (Slapin & Proksch 2008; Lowe et al. 2011). Recent NLP models represent the text by high-dimensional vectors learned by neural networks to predict the scores (Glavaš et al. 2017b; Nanni et al. 2019). One way to use the NLP models is to apply off-the-shelf general-purpose models such as InstructGPT (Ouyang et al. 2022) and design a prompt to specify the type of the scaling to the API,[4], or borrow existing, trained NLP models if the same type of scaling has been studied by previous researchers. Another way is to collect a dataset of text with hand-coded scales, and train NLP models to learn to predict the scale, similar to the practice in Slapin & Proksch (2008); Gennaro & Ash (2021), *inter alia*.

[4] https://beta.openai.com/docs/introduction

3 Using NLP for Policymaking

In the political domain, there are large amounts of textual data to analyze (NEUENDORF & KUMAR 2015), such as parliament debates (Van Aggelen et al. 2017), speeches (Schumacher et al. 2016), legislative text (Baumgartner et al. 2006; Bevan 2017), database of political parties worldwide (Döring & Regel 2019), and expert survey data (Bakker et al. 2015). Since it is tedious to hand-code all textual data, NLP provides a low-cost tool to automatically analyze such massive text.

In this section, we will introduce how NLP can facilitate four major areas to help policymaking: before policies are made, researchers can use NLP to analyze data and extract key information for evidence-based policymaking (Section 3.1); after policies are made, researchers can interpret the priorities among and reasons behind political decisions (Section 3.2); researchers can also analyze features in the language of politicians when communicating the policies to the public (Section 3.3); finally, after the policies have taken effect, researchers can investigate the effectiveness of the policies (Section 3.4).

3.1 Analyzing Data for Evidence-Based Policymaking

A major use of NLP is to extract information from large collections of text. This function can be very useful for analyzing the views and needs of constituents, so that policymakers can make decisions accordingly.

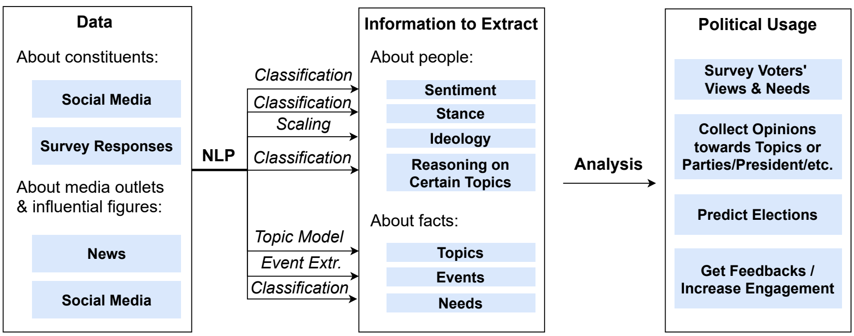

Figure 4: NLP to analyze data for evidence-based policymaking.

As in Figure 4, we will explain how NLP can be used to analyze data for evidence-based policymaking from three aspects: data, information to extract, and political usage.

Data. Data is the basis of such analyses. Large amounts of textual data can reveal information about constituents, media outlets, and influential figures. The data can come from a variety of sources, including social media such as Twitter and Facebook, survey responses, and news articles.

Information to Extract. Based on the large textual corpora, NLP models can be used to extract information that are useful for political decision-making, ranging from information about people, such as sentiment (Thelwall et al. 2011; Rosenthal et al. 2015), stance (Thomas et al. 2006; Gottipati et al. 2013; Stefanov et al. 2020; Luo et al. 2020), ideology (Hirst et al. 2010; Iyyer et al. 2014; Preoţiuc-Pietro et al. 2017), and reasoning on certain topics (Egami et al. 2018; Demszky et al. 2019; Camp et al. 2021), to factual information, such as main topics (Gottipati et al. 2013), events (Trappl 2006; Mitamura et al. 2017; Ding & Riloff 2018; Ding et al. 2019), and needs (Sarol et al. 2020; Crayton et al. 2020; Paul & Frank 2019) expressed in the data. The extracted information cannot only be about people, but also about political entities, such as the left-right political scales of parties and political actors (Slapin & Proksch 2008; Glavaš et al. 2017b), which claims are raised by which politicians (Blessing et al. 2019; Padó et al. 2019), and the legislative body's vote breakdown for state bills by backgrounds such as gender, rural-urban and ideological splits Davoodi et al. (2020).

To extract such information from text, we can often utilize the main NLP tools introduced in Section 2, including text classification, topic modeling, event extraction and score prediction (especially text scaling to predict left-to-right ideology). In NLP literature, social media, such as Twitter, is a popular source of textual data to collect public opinions (Thelwall et al. 2011; Paltoglou & Thelwall 2012; Pak & Paroubek 2010; Arunachalam & Sarkar 2013; Rosenthal et al. 2015).

Political Usage. Such information extracted from data is highly valuable for political usage. For example, voters' sentiment, stance, and ideology are important supplementary for traditional polls and surveys to gather information about the constituents' political leaning. Identifying the needs expressed by people is another important survey target, which helps politicians understand what needs they should take care of, and match the needs and availabilities of resources (Hiware et al. 2020).

Among more specific political uses is to understand the public opinion on parties/president, as well as on certain topics. The public sentiment towards parties (Pla & Hurtado 2014) and President (Marchetti-Bowick & Chambers 2012) can serve as a supplementary for the traditional approval rating survey, and stances towards certain topics (Gottipati et al. 2013; Stefanov et al. 2020; Luo et al. 2020) can be important information for legislators to make decisions on debatable issues such as abortion, taxes, and legalization of same-sex marriage. Many existing studies use NLP on social media text to predict election results (O'Connor et al. 2010; Beverungen & Kalita 2011; Unankard et al. 2014; Mohammad et al. 2015; Tjong Kim Sang & Bos 2012). In general, big-data-driven analyses can facilitate decision-makers to collect more feedback from people and society, enabling policymakers to be closer to citizens, and increase transparency and engagement in political issues (Arunachalam & Sarkar 2013).

3.2 Interpreting Political Decisions

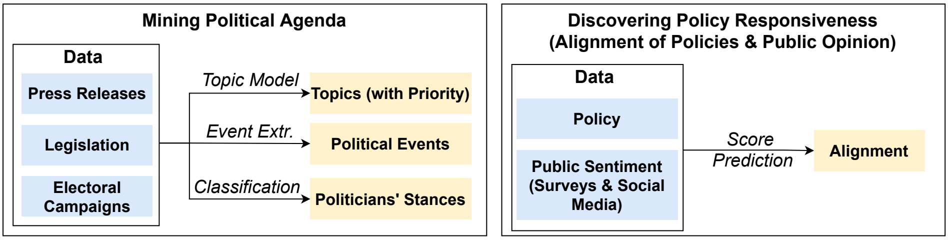

Figure 5: NLP to interpret political decisions.

After policies are made, political scientists and social scientists can use textual data to interpret political decisions. As in Figure 5, there are two major use cases: mining political agendas, and discovering policy responsiveness.

Mining Political Agendas. Researchers can use textual data to infer a political agenda, including the topics that politicians prioritize, political events, and different political actors' stances on certain topics. Such data can come from press releases, legislation, and electoral campaigns. Example of previous studies to analyze the topics and prioritization of political bodies include the research on the prioritization each Senator assigns to topics using press releases (Grimmer 2010b), topics in different parties' electoral manifestos (Glavaš et al. 2017a), topics in EU parliament speeches (Lauscher et al. 2016) and other various types of text (King & Lowe 2003; Hopkins & King 2010; Grimmer 2010a; Roberts et al. 2014), as well as political event detection from congressional text and news (Nanni et al. 2017).

Research on politicians' stances include identifying policy positions of politicians (Winter & Stewart 1977; Laver et al. 2003; Slapin & Proksch 2008; Lowe et al. 2011: *inter alia*), how different politicians agree or disagree on certain topics in electoral campaigns (Menini & Tonelli 2016), and assessment of political personalities (Immelman 1993).

Further studies look into how political interests affect legislative behavior. Legislators tend to show strong personal interest in the issues that come before their committees (Fenno 1973), and Mayhew (2004) identifies that Senators replying on appropriations secured for their state have a strong incentive to support legislations that allow them to secure particularistic goods.

Discovering Policy Responsiveness. Policy responsiveness is the study of how policies respond to different factors, such as how changes in public opinion lead to responses in public policy (Stimson et al. 1995). One major direction is that politicians tend to make policies that align with the expectations of their constituents, in order to run for successful re-election in the next term (Canes-Wrone et al. 2002). Studies show that policy preferences of the state public can be a predictor of future state policies (Caughey & Warshaw 2018). For example, Lax & Phillips (2009) show that more LGBT tolerance leads to more pro-gay legislation in response.

A recent study by Jin, Peng, et al. (2021) uses NLP to analyze over 10 million COVID-19-related tweets targeted at US governors; using classification models to obtain the public sentiment, they study how public sentiment leads to political decisions of COVID-19 policies made by US governors. Such use of NLP on massive textual data contrasts with the traditional studies of policy responsiveness which span over several decades and use manually collected survey results (Caughey & Warshaw 2018; Lax & Phillips 2009; 2012).

3.3 Improving Policy Communication with the Public

Policy communication is the study to understand how politicians present the policies to their constituents. As in Figure 6, common research questions in policy communication include how politicians establish their images (Fenno 1978) such as campaign strategies (Petrocik 1996; Simon 2002; Sigelman & Buell Jr 2004), how constituents allocate credit, what receives attention in Congress (Sulkin 2005), and what receives attention in news articles (Semetko & Valkenburg 2000; McCombs & Valenzuela 2004; Armstrong et al. 2006).

Based on data from press releases, political statements, electoral campaigns and news articles,[5] researchers usually analyze two types of information: the language techniques politicians use, and the contents such as topics and underlying moral foundations in these textual documents.

[5]Other data sources used in policy communication research include surveys of Senate staffers (Cook 1988), newsletters that legislators send to constituents (Lipinski 2009) and so on.

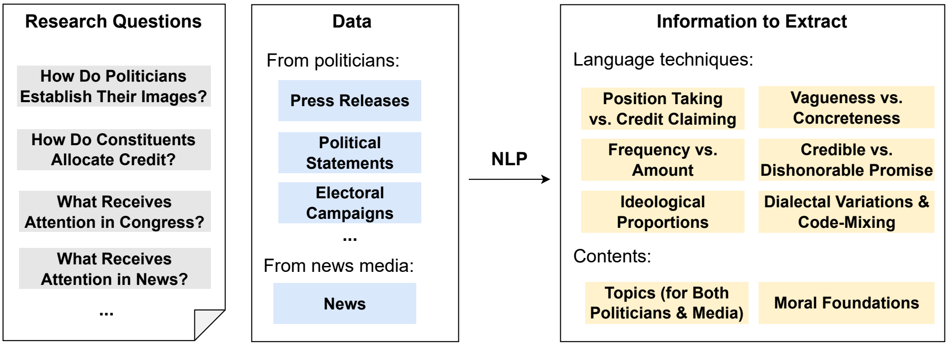

Figure 6: NLP to analyze policy communication.

Language Techniques. Policy communication largely focuses on the types of languages that politicians use. Researchers are interested in first analyzing the language techniques in political texts, and then, based on these techniques, researchers can dive into the questions of why politicians use them, and what are the effects of such usage.

For example, previous studies analyze what portions of political texts are position-taking versus credit-claiming (Grimmer et al. 2012; Grimmer 2013), whether the claims are vague or concrete (Baerg et al. 2018; Eichorst & Lin 2019), the frequency of credit-claiming messages versus the actual amount of contributions (Grimmer et al. 2012), and whether politicians tend to make credible or dishonorable promises (Grimmer 2010b). Within the political statements, it is also interesting to check the ideological proportions (Sim et al. 2013), and how politicians make use of dialectal variations and code-mixing (Sravani et al. 2021).

The representation styles usually affect the effectiveness of policy communication, such as the role of language ambiguity in framing the political agenda (Page 1976; Campbell 1983), and the effect of credit-claiming messages on constituents' allocation of credit (Grimmer et al. 2012).

Contents. The contents of policy communication include the topics in the political statements, such as what Senators discuss in floor statements (Hill & Hurley 2002), and what Presidents address in daily speeches (Lee 2008), and also the moral foundations used by politicians underlying their political tweets (Johnson & Goldwasser 2018).

Using the extracted content information, researchers can explore further questions such as whether competing politicians or political elites emphasize the same issues (Petrocik 1996; Gabel & Scheve 2007), and how the priorities politicians articulate co-vary with the issues discussed in the media (Bartels 1996). Another open research direction is to analyze the interaction between newspapers and politicians' messages, such as how often newspapers cover a certain politician's message and in what way, and how such coverage affects incumbency advantage.

Meaningful Future Work. Apart from analyzing the language of existing political texts that aims to maximize political interests, an advanced question that is more meaningful to society is how to improve policy communication to steer towards a more beneficial future for society as a whole. There is relatively little research on this, and we welcome future work on this meaningful topic.

3.4 Investigating Policy Effects

After policies are taken into effect, it is important to collect feedback or evaluate the effectiveness of policies. Existing studies evaluate the effects of policies along different dimensions: one dimension is the change in public sentiment, which can be analyzed by comparing the sentiment classification results before and after policies, following a similar paradigm in Section 3.1. There are also studies on how policies affect the crowd's perception of the democratic process (Miller et al. 1990).

Another dimension is how policies result in economic changes. Calvo-González et al. (2018) investigate the negative consequences of policy volatility that harm long-term economic growth. Specifically, to measure policy volatility, they first obtain main topics by topic modeling on presidential speeches, and then analyze how the significance of topics changes over time.

4 Limitations and Ethical Considerations

There are several limitations that researchers and policymakers need to take into consideration when using NLP for policymaking, due to the data-driven and black-box nature of modern NLP. First, the effectiveness of the computational models relies on the quality and comprehensiveness of the data. Although many political discourses are public, including data sources such as news, press releases, legislation, and campaigns, when it comes to surveying public opinions, social media might be a biased representation of the whole population. Therefore, when making important policy decisions, the traditional polls and surveys can provide more comprehensive coverage. Note that in the case of traditional polls, NLP can still be helpful in expediting the processing of survey answers.

The second concern is the black-box nature of modern NLP models. We do not encourage decision-making systems to depend fully on NLP, but suggest that NLP can assist human decision-makers. Hence, all the applications introduced in this chapter use NLP to compile information that is necessary for policymaking instead of directly suggesting a policy. Nonetheless, some of the models are hard to interpret or explain, such as text classification using deep learning models (Yin et al. 2019; Brown et al. 2020), which could be vulnerable to adversarial attacks by small paraphrasing of the text input (Jin et al. 2020). In practical applications, it is important to ensure the trustworthiness of the usage of AI. There could be a preference for transparent machine learning models if they can do the work well (e.g., LDA topic models, and traditional classification methods using dictionaries or linguistic rules), or tasks with well-controlled outputs such as event extraction to select spans of the given text that mention events. In cases where only the deep learning models can provide good performance, there should be more detailed performance analysis (e.g., a study to check the correlation of the model decisions and human judgments), error analysis (e.g., different types of errors, failure modes, and potential bias towards certain groups), and studies about the interpretability of the model (e.g., feature attribution of the model, visualization of the internal states of the model).

Apart from the limitations of the technical methodology, there are also ethical considerations arising from the use of NLP. Among the use cases introduced in this chapter, some applications of NLP are relatively safe as they mainly involve analyzing public political documents and fact-based evidence or effects of policies. However, others could be concerning and vulnerable to misuse. For example, although effective, truthful policy communication is beneficial for society, it might be tempting to overdo policy communication and by all means optimize the votes. As it is highly important for government and politicians to gain positive public perception, overly optimizing policy communication might lead to propaganda, intrusion of data privacy to collect more user preferences, and, in more severe cases, surveillance and violation of human rights. Hence, there is a strong need for policies to regulate the use of technologies that influence public opinions and pose a challenge to democracy.

5 Conclusions

This chapter provided a brief overview of current research directions in NLP that provide support for policymaking. We first introduced four main NLP tasks that are commonly used in text analysis: text classification, topic modeling, event extraction, and text scaling. We then showed how these methods can be used in policymaking for applications such as data collection for evidence-based policymaking, interpretation of political decisions, policy communication, and investigation of policy effects. We also discussed potential limitations and ethical considerations of which researchers and policymakers should be aware.

NLP holds significant promise for enabling data-driven policymaking. In addition to the tasks overviewed in this chapter, we foresee that other NLP applications, such as text summarization (e.g., to condense information from large documents), question answering (e.g., for reasoning about policies), and culturally-adjusted machine translation (e.g., to facilitate international communications), will soon find use in policymaking. The field of NLP is quickly advancing, and close collaborations between NLP experts and public policy experts will be key to the successful use and deployment of NLP tools in public policy.

References

Adler, E. Scott & John Wilkerson. 2011. Congressional bills project. *NSF 00880066 and 00880061*. http://www.congressionalbills.org/.

Albaugh, Quinn, Julie Sevenans, Stuart Soroka & Peter John Loewen. 2013. The automated coding of policy agendas: A dictionary-based approach. In *The 6th annual comparative agendas conference, Antwerp, Belgium*.

Ansolabehere, Stephen & Shanto Iyengar. 1995. *Going negative: How political advertisements shrink and polarize the electorate*. Vol. 95. New York: Simon & Schuster.

Armstrong, Elizabeth M, Daniel P Carpenter & Marie Hojnacki. 2006. Whose deaths matter? Mortality, advocacy, and attention to disease in the mass media. *Journal of Health Politics, Policy and Law* 31(4). 729–772.

Arunachalam, Ravi & Sandipan Sarkar. 2013. The new eye of government: citizen sentiment analysis in social media. In *Proceedings of the IJCNLP 2013 workshop on natural language processing for social media (SocialNLP)*, 23–28. Nagoya, Japan: Asian Federation of Natural Language Processing. https://www.aclweb.org/anthology/W13-4204.

Azar, Edward E. 1980. The conflict and peace data bank (COPDAB) project. *Journal of Conflict Resolution* 24(1). 143–152.

Baerg, Nicole, Dominik Duell & Will Lowe. 2018. Central bank communication as public opinion: experimental evidence. *Work in Progress*.

Bakker, Ryan, Catherine De Vries, Erica Edwards, Liesbet Hooghe, Seth Jolly, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen & Milada Anna Vachudova. 2015. Measuring party positions in Europe: The Chapel Hill expert survey trend file, 1999–2010. *Party Politics* 21(1). 143–152.

Bartels, Larry M. 1996. Politicians and the press: Who leads, who follows. In *Annual meeting of the american political science association*, 1–60.

Baumgartner, Frank R, Christoffer Green-Pedersen & Bryan D Jones. 2006. Comparative studies of policy agendas. *Journal of European public policy* 13(7). 959–974.

Bevan, Shaun. 2017. Gone fishing: The creation of the comparative agendas project master codebook. *Comparative Policy Agendas: Theory, Tools, Data*. http://sbevan.%20com/cap-master-codebook.html.

Beverungen, Gary & Jugal Kalita. 2011. Evaluating methods for summarizing twitter posts. *Proceedings of the 5th AAAI ICWSM*.

Biran, Ofer, Oshrit Feder, Yosef Moatti, Athanasios Kiourtis, Dimosthenis Kyriazis, George Manias, Argyro Mavrogiorgou, Nikitas M. Sgouros, Martim Taborda Barata, Isabella Oldani, María Angeles Sanguino & Pavlos Kranas. 2022. PolicyCLOUD: A prototype of a cloud serverless ecosystem for policy analytics. *CoRR* abs/2201.06077. https://arxiv.org/abs/2201.06077.

Blei, David M., Andrew Y. Ng & Michael I. Jordan. 2001. Latent dirichlet allocation. In Thomas G. Dietterich, Suzanna Becker & Zoubin Ghahramani (eds.), *Advances in neural information processing systems 14 [neural information processing systems: natural and synthetic, NIPS 2001, December 3-8, 2001, Vancouver, British Columbia, Canada]*, 601–608. MIT Press. https://proceedings.neurips.cc/paper/2001/hash/296472c9542ad4d4788d543508116cbc-Abstract.html.

Blessing, Andre, Nico Blokker, Sebastian Haunss, Jonas Kuhn, Gabriella Lapesa & Sebastian Padó. 2019. An environment for relational annotation of political debates. In *Proceedings of the 57th annual meeting of the association for computational linguistics: system demonstrations*, 105–110. Florence, Italy: Association for Computational Linguistics. DOI: 10.18653/v1/P19-3018. https://aclanthology.org/P19-3018.

Bond, Doug, J Craig Jenkins, Charles L Taylor & Kurt Schock. 1997. Mapping mass political conflict and civil society: issues and prospects for the automated development of event data. *Journal of Conflict Resolution* 41(4). 553–579.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever & Dario Amodei. 2020. Language models are few-shot learners. In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan & Hsuan-Tien Lin (eds.), *Advances in neural information processing systems 33: annual conference on neural information processing systems 2020, neurips 2020, december 6-12, 2020, virtual*. https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html.

Calvo-González, Oscar, Axel Eizmendi & German Jeremias Reyes. 2018. Winners never quit, quitters never grow: Using text mining to measure policy volatility and its link with long-term growth in latin america. *World Bank Policy Research Working Paper* (8310).

Camp, Nicholas P, Rob Voigt, Dan Jurafsky & Jennifer L Eberhardt. 2021. The thin blue waveform: racial disparities in officer prosody undermine institutional trust in the police. *Journal of Personality and Social Psychology*.

Campbell, James E. 1983. Ambiguity in the issue positions of presidential candidates: a causal analysis. *American Journal of Political Science*. 284–293.

Canes-Wrone, Brandice, David W Brady & John F Cogan. 2002. Out of step, out of office: electoral accountability and house members' voting. *American Political Science Review*. 127–140.

Caughey, Devin & Christopher Warshaw. 2018. Policy preferences and policy change: dynamic responsiveness in the american states, 1936–2014.

Chung, C K & J W Pennebaker. 2008. Revealing dimensions of thinking in Open-Ended Self-Descriptions: an automated meaning extraction method for natural language. en. *J. Res. Pers.* 42(1). 96–132.

Collingwood, Loren & John Wilkerson. 2012. Tradeoffs in accuracy and efficiency in supervised learning methods. *Journal of Information Technology & Politics* 9(3). 298–318.

Cook, Timothy E. 1988. Press secretaries and media strategies in the house of representatives: deciding whom to pursue. *American Journal of Political Science*. 1047–1069.

Crayton, Ancil, João Fonseca, Kanav Mehra, Michelle Ng, Jared Ross, Marcelo Sandoval-Castañeda & Rachel von Gnechten. 2020. Narratives and needs: analyzing experiences of cyclone amphan using twitter discourse. *CoRR* abs/2009.05560. https://arxiv.org/abs/2009.05560.

Davoodi, Maryam, Eric Waltenburg & Dan Goldwasser. 2020. Understanding the language of political agreement and disagreement in legislative texts. In *Proceedings of the 58th annual meeting of the association for computational linguistics*, 5358–5368. Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.476. https://aclanthology.org/2020.acl-main.476.

Demszky, Dorottya, Nikhil Garg, Rob Voigt, James Zou, Jesse Shapiro, Matthew Gentzkow & Dan Jurafsky. 2019. Analyzing polarization in social media: method and application to tweets on 21 mass shootings. In *Proceedings of the 2019 conference of the north American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, 2970–3005. Minneapolis, Minnesota: Association for Computational Linguistics. DOI: 10.18653/v1/N19-1304. https://aclanthology.org/N19-1304.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee & Kristina Toutanova. 2019. BERT: pre-training of deep bidirectional transformers for language understanding. In Jill Burstein, Christy Doran & Thamar Solorio (eds.), *Proceedings of the 2019 conference of the north american chapter of the association for computational linguistics: human language technologies, NAACL-HLT 2019, minneapolis, mn, usa, june 2-7, 2019, volume 1 (long and short papers)*, 4171–4186. Association for Computational Linguistics. DOI: 10.18653/v1/n19-1423. https://doi.org/10.18653/v1/n19-1423.

Ding, Haibo & Ellen Riloff. 2018. Human needs categorization of affective events using labeled and unlabeled data. In *Proceedings of the 2018 conference of the north American chapter of the association for computational linguistics: human language technologies, volume 1 (long papers)*, 1919–1929. New Orleans, Louisiana: Association for Computational Linguistics. DOI: 10.18653/v1/N18-1174. https://aclanthology.org/N18-1174.

Ding, Haibo, Ellen Riloff & Zhe Feng. 2019. Improving human needs categorization of events with semantic classification. In *Proceedings of the eighth joint conference on lexical and computational semantics (*SEM 2019)*, 198–204. Minneapolis, Minnesota: Association for Computational Linguistics. DOI: 10.18653/v1/S19-1022. https://aclanthology.org/S19-1022.

Döring, Holger & Sven Regel. 2019. Party facts: a database of political parties worldwide. *Party politics* 25(2). 97–109.

Egami, Naoki, Christian J. Fong, Justin Grimmer, Margaret E. Roberts & Brandon M. Stewart. 2018. How to make causal inferences using texts. *CoRR* abs/1802.02163. http://arxiv.org/abs/1802.02163.

Eichorst, Jason & Nick CN Lin. 2019. Resist to commit: concrete campaign statements and the need to clarify a partisan reputation. *The Journal of Politics* 81(1). 15–32.

Engel, Uwe, Anabel Quan-Haase, Sunny Xun Liu & Lars Lyberg. 2021. *Handbook of computational social science, volume 2*. Taylor & Francis.

Engstrom, David Freeman, Daniel E Ho, Catherine M Sharkey & Mariano-Florentino Cuéllar. 2020. Government by algorithm: artificial intelligence in federal administrative agencies. *NYU School of Law, Public Law Research Paper* (20-54).

European Commission. 2017. COM(2011) 808 Final: Horizon 2020 — the framework programme for research and innovation. In https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%5C%3A52011PC0809 (15 May, 2022).

Fenno, Richard F. 1973. Congressmen in committees. In Boston: Little Brown & Company.

Fenno, Richard F. 1978. *Home style: House members in their districts*. Boston: Addison Wesley.

Gabel, Matthew & Kenneth Scheve. 2007. Estimating the effect of elite communications on public opinion using instrumental variables. *American Journal of Political Science* 51(4). 1013–1028.

Gennaro, Gloria & Elliott Ash. 2021. Emotion and Reason in Political Language. *The Economic Journal* 132(643). 1037–1059. DOI: 10.1093/ej/ueab104. https://doi.org/10.1093/ej/ueab104.

Gerner, Deborah J, Philip A Schrodt, Omür Yilmaz & Rajaa Abu-Jabr. 2002. Conflict and mediation event observations (cameo): a new event data framework for the analysis of foreign policy interactions. *International Studies Association, New Orleans*.

Gigley, Helen M. 1993. Projected government needs in human language technology and the role of researchers in meeting them. In *Human Language Technology: proceedings of a workshop held at plainsboro, new jersey, march 21-24, 1993*. https://aclanthology.org/H93-1056.

Glavaš, Goran, Federico Nanni & Simone Paolo Ponzetto. 2017a. Cross-lingual classification of topics in political texts. In *Proceedings of the second workshop on NLP and computational social science*, 42–46. Vancouver, Canada: Association for Computational Linguistics. DOI: 10.18653/v1/W17-2906. https://aclanthology.org/W17-2906.

Glavaš, Goran, Federico Nanni & Simone Paolo Ponzetto. 2017b. Unsupervised cross-lingual scaling of political texts. In *Proceedings of the 15th conference of the European chapter of the association for computational linguistics: volume 2, short papers*, 688–693. Valencia, Spain: Association for Computational Linguistics. https://aclanthology.org/E17-2109.

Glavaš, Goran, Federico Nanni & Simone Paolo Ponzetto. 2019. Computational analysis of political texts: bridging research efforts across communities. In *Proceedings of the 57th annual meeting of the association for computational linguistics: tutorial abstracts*, 18–23. Florence, Italy: Association for Computational Linguistics. DOI: 10.18653/v1/P19-4004. https://aclanthology.org/P19-4004.

Goldstein, Joshua S. 1992. A conflict-cooperation scale for weis events data. *Journal of Conflict Resolution* 36(2). 369–385.

Gottipati, Swapna, Minghui Qiu, Yanchuan Sim, Jing Jiang & Noah A. Smith. 2013. Learning topics and positions from Debatepedia. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, 1858–1868. Seattle, Washington, USA: Association for Computational Linguistics. https://aclanthology.org/D13-1191.

Grimmer, Justin. 2010a. A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in Senate press releases. *Political Analysis* 18(1). 1–35.

Grimmer, Justin. 2013. Appropriators not position takers: the distorting effects of electoral incentives on congressional representation. *American Journal of Political Science* 57(3). 624–642.

Grimmer, Justin, Solomon Messing & Sean J Westwood. 2012. How words and money cultivate a personal vote: the effect of legislator credit claiming on constituent credit allocation. *American Political Science Review* 106(4). 703–719.

Grimmer, Justin & Brandon M Stewart. 2013. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis* 21(3). 267–297.

Grimmer, Justin Ryan. 2010b. *Representational style: The central role of communication in representation*. Harvard University.

Hausladen, Carina I., Marcel H. Schubert & Elliott Ash. 2020. Text classification of ideological direction in judicial opinions. *International Review of Law and Economics* 62. 105903. DOI: https://doi.org/10.1016/j.irle.2020.105903. https://www.sciencedirect.com/science/article/pii/S0144818819303667.

Hill, Kim Quaile & Patricia A Hurley. 2002. Symbolic speeches in the us senate and their representational implications. *Journal of Politics* 64(1). 219–231.

Hillard, Dustin, Stephen Purpura & John Wilkerson. 2008. Computer-assisted topic classification for mixed-methods social science research. *Journal of Information Technology & Politics* 4(4). 31–46.

Hirst, Graeme, Yaroslav Riabinin & Jory Graham. 2010. Party status as a confound in the automatic classification of political speech by ideology. In *Proceedings of the 10th international conference on statistical analysis of textual data (jadt 2010)*, 731–742.

Hiware, Kaustubh, Ritam Dutt, Sayan Sinha, Sohan Patro, Kripa Ghosh & Saptarshi Ghosh. 2020. NARMADA: need and available resource managing assistant for disasters and adversities. In *Proceedings of the eighth international workshop on natural language processing for social media*, 15–24. Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.socialnlp-1.3. https://aclanthology.org/2020.socialnlp-1.3.

Hogenboom, Frederik, Flavius Frasincar, Uzay Kaymak & Franciska de Jong. 2011. An overview of event extraction from text. In Marieke van Erp, Willem Robert van Hage, Laura Hollink, Anthony Jameson & Raphaël Troncy (eds.), *Proceedings of the workhop on detection, representation, and exploitation of events in the semantic web (derive 2011), bonn, germany, october 23, 2011*, vol. 779 (CEUR Workshop Proceedings), 48–57. CEUR-WS.org. http://ceur-ws.org/Vol-779/derive2011%5C_submission%5C_1.pdf.

Honnibal, Matthew, Ines Montani, Sofie Van Landeghem & Adriane Boyd. 2020. spaCy: Industrial-strength Natural Language Processing in Python. DOI: 10.5281/zenodo.1212303.

Hopkins, Daniel J & Gary King. 2010. A method of automated nonparametric content analysis for social science. *American Journal of Political Science* 54(1). 229–247.

Huguet Cabot, Pere-Lluís, Verna Dankers, David Abadi, Agneta Fischer & Ekaterina Shutova. 2020. The Pragmatics behind Politics: Modelling Metaphor, Framing and Emotion in Political Discourse. In *Findings of the association for computational linguistics: emnlp 2020*, 4479–4488. Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.findings-emnlp.402. https://aclanthology.org/2020.findings-emnlp.402.

Immelman, Aubrey. 1993. The assessment of political personality: a psychodiagnostically relevant conceptualization and methodology. *Political Psychology*. 725–741.

Iyyer, Mohit, Peter Enns, Jordan Boyd-Graber & Philip Resnik. 2014. Political ideology detection using recursive neural networks. In *Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 1: long papers)*, 1113–1122. Baltimore, Maryland: Association for Computational Linguistics. DOI: 10.3115/v1/P14-1105. https://aclanthology.org/P14-1105.

Jin, Di, Zhijing Jin, Joey Tianyi Zhou & Peter Szolovits. 2020. Is BERT really robust? A strong baseline for natural language attack on text classification and entailment. In *The thirty-fourth AAAI conference on artificial intelligence, AAAI 2020, the thirty-second innovative applications of artificial intelligence conference, IAAI 2020, the tenth AAAI symposium on educational advances in artificial intelligence, EAAI 2020, new york, ny, usa, february 7-12, 2020*, 8018–8025. AAAI Press. https://aaai.org/ojs/index.php/AAAI/article/view/6311.

Jin, Zhijing, Geeticka Chauhan, Brian Tse, Mrinmaya Sachan & Rada Mihalcea. 2021. How good is NLP? A sober look at NLP tasks through the lens of social impact. In *Findings of the association for computational linguistics: acl-ijcnlp*

2021, 3099–3113. Online: Association for Computational Linguistics. DOI: 10.18653/v1/2021.findings-acl.273. https://aclanthology.org/2021.findings-acl.273.

Jin, Zhijing, Zeyu Peng, Tejas Vaidhya, Bernhard Schoelkopf & Rada Mihalcea. 2021. Mining the cause of political decision-making from social media: A case study of COVID-19 policies across the US states. In *Findings of the 2021 conference on empirical methods in natural language processing, EMNLP 2021*. Association for Computational Linguistics.

Johnson, Kristen & Dan Goldwasser. 2018. Classification of moral foundations in microblog political discourse. In *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: long papers)*, 720–730. Melbourne, Australia: Association for Computational Linguistics. DOI: 10.18653/v1/P18-1067. https://aclanthology.org/P18-1067.

Jurafsky, Daniel & James H. Martin. 2000. *Speech and language processing - An introduction to natural language processing, computational linguistics, and speech recognition* (Prentice Hall series in artificial intelligence). Prentice Hall.

Karan, Mladen, Jan Šnajder, Daniela Širinić & Goran Glavaš. 2016. Analysis of policy agendas: Lessons learned from automatic topic classification of Croatian political texts. In *Proceedings of the 10th SIGHUM workshop on language technology for cultural heritage, social sciences, and humanities*, 12–21. Berlin, Germany: Association for Computational Linguistics. DOI: 10.18653/v1/W16-2102. https://aclanthology.org/W16-2102.

King, Gary & Will Lowe. 2003. An automated information extraction tool for international conflict data with performance as good as human coders: A rare events evaluation design. *International Organization* 57(3). 617–642.

Lauscher, Anne, Pablo Ruiz Fabo, Federico Nanni & Simone Paolo Ponzetto. 2016. Entities as topic labels: Combining entity linking and labeled lda to improve topic interpretability and evaluability. *IJCoL. Italian Journal of Computational Linguistics* 2(2-2). 67–87.

Laver, Michael, Kenneth Benoit & John Garry. 2003. Extracting policy positions from political texts using words as data. *American political science review* 97(2). 311–331.

Lax, Jeffrey R & Justin H Phillips. 2009. Gay rights in the states: public opinion and policy responsiveness. *American Political Science Review* 103(3). 367–386.

Lax, Jeffrey R & Justin H Phillips. 2012. The democratic deficit in the states. *American Journal of Political Science* 56(1). 148–166.

Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, et al. 2009. Computational social science. *Science* 323(5915). 721–723.

Lee, Frances E. 2008. Dividers, not uniters: presidential leadership and senate partisanship, 1981-2004. *The Journal of Politics* 70(4). 914–928.

Lipinski, Daniel. 2009. *Congressional communication: Content and consequences*. University of Michigan Press.

Liu, Kang, Yubo Chen, Jian Liu, Xinyu Zuo & Jun Zhao. 2020. Extracting events and their relations from texts: A survey on recent research progress and challenges. *AI Open* 1. 22–39. DOI: https://doi.org/10.1016/j.aiopen.2021.02.004. https://www.sciencedirect.com/science/article/pii/S266665102100005X.

Loria, Steven. 2018. TextBlob documentation. *Release 0.15* 2.

Lowe, Will, Kenneth Benoit, Slava Mikhaylov & Michael Laver. 2011. Scaling policy preferences from coded political texts. *Legislative studies quarterly* 36(1). 123–155.

Luo, Yiwei, Dallas Card & Dan Jurafsky. 2020. Detecting stance in media on global warming. In *Findings of the association for computational linguistics: emnlp 2020*, 3296–3315. Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.findings-emnlp.296. https://aclanthology.org/2020.findings-emnlp.296.

Luz, Saturnino. 2022. Computational Linguistics and Natural Language Processing. English. In Federico Zanettin & Christopher Rundle (eds.), *The Routledge Handbook of Translation and Methodology*. United States: Routledge.

Manning, Christopher D., Prabhakar Raghavan & Hinrich Schütze. 2008. *Introduction to information retrieval*. Cambridge University Press. DOI: 10.1017/CBO9780511809071. https://nlp.stanford.edu/IR-book/pdf/irbookprint.pdf.

Marchetti-Bowick, Micol & Nathanael Chambers. 2012. Learning for microblogs with distant supervision: political forecasting with Twitter. In *Proceedings of the 13th conference of the European chapter of the association for computational linguistics*, 603–612. Avignon, France: Association for Computational Linguistics. https://www.aclweb.org/anthology/E12-1062.

Maronikolakis, Antonis, Danae Sánchez Villegas, Daniel Preotiuc-Pietro & Nikolaos Aletras. 2020. Analyzing political parody in social media. In *Proceedings of the 58th annual meeting of the association for computational linguistics*, 4373–4384. Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.403. https://aclanthology.org/2020.acl-main.403.

Mayhew, David R. 2004. *Congress: The electoral connection*. Yale university press.

McClelland, Charles A. 1976. *World event/interaction survey codebook*.

McCombs, Maxwell & Sebastián Valenzuela. 2004. *Setting the agenda: Mass media and public opinion*. John Wiley & Sons.

Menini, Stefano, Federico Nanni, Simone Paolo Ponzetto & Sara Tonelli. 2017. Topic-based agreement and disagreement in US electoral manifestos. In *Proceedings of the 2017 conference on empirical methods in natural language processing*, 2938–2944. Copenhagen, Denmark: Association for Computational Linguistics. DOI: 10.18653/v1/D17-1318. https://aclanthology.org/D17-1318.

Menini, Stefano & Sara Tonelli. 2016. Agreement and disagreement: comparison of points of view in the political domain. In *Proceedings of COLING 2016, the 26th international conference on computational linguistics: technical papers*, 2461–2470. Osaka, Japan: The COLING 2016 Organizing Committee. https://aclanthology.org/C16-1232.

Merritt, Richard L, Robert G Muncaster & Dina A Zinnes. 1993. *International event-data developments: DDIR phase II*. University of Michigan Press.

Miller, William Lockley, Harold D Clarke, Martin Harrop, Lawrence LeDuc, Paul F Whiteley, et al. 1990. *How voters change: the 1987 british election campaign in perspective*. Oxford University Press.

Misuraca, Gianluca, Colin van Noordt & Anys Boukli. 2020. The use of AI in public services: Results from a preliminary mapping across the EU. In Yannis Charalabidis, Maria Alexandra Cunha & Demetrios Sarantis (eds.), *ICEGOV 2020: 13th international conference on theory and practice of electronic governance, athens, greece, 23-25 september, 2020*, 90–99. ACM. DOI: 10.1145/3428502.3428513. https://doi.org/10.1145/3428502.3428513.

Mitamura, Teruko, Zhengzhong Liu & Eduard H. Hovy. 2017. Events detection, coreference and sequencing: what's next? overview of the TAC KBP 2017 event track. In *Proceedings of the 2017 text analysis conference, TAC 2017, gaithersburg, maryland, usa, november 13-14, 2017*. NIST. https://tac.nist.gov/publications/2017/additional.papers/TAC2017.KBP%5C_Event%5C_Nugget%5C_overview.proceedings.pdf.

Mohammad, Saif M., Xiaodan Zhu, Svetlana Kiritchenko & Joel D. Martin. 2015. Sentiment, emotion, purpose, and style in electoral tweets. *Inf. Process. Manag.* 51(4). 480–499. DOI: 10.1016/j.ipm.2014.09.003. https://doi.org/10.1016/j.ipm.2014.09.003.

Nanni, Federico, Goran Glavas, Simone Paolo Ponzetto & Heiner Stuckenschmidt. 2019. Political text scaling meets computational semantics. *CoRR* abs/1904.06217. http://arxiv.org/abs/1904.06217.

Nanni, Federico, Simone Paolo Ponzetto & Laura Dietz. 2017. Building entity-centric event collections. In *2017 ACM/IEEE joint conference on digital libraries, JCDL 2017, toronto, on, canada, june 19-23, 2017*, 199–208. IEEE Computer Society. DOI: 10.1109/JCDL.2017.7991574. https://doi.org/10.1109/JCDL.2017.7991574.

NEUENDORF, Kimberly A & Anup KUMAR. 2015. Content analysis. *The International Encyclopedia of Political Communication* 8. 1–10.

O'Connor, Brendan, Ramnath Balasubramanyan, Bryan R. Routledge & Noah A. Smith. 2010. From tweets to polls: linking text sentiment to public opinion time series. In William W. Cohen & Samuel Gosling (eds.), *Proceedings of the fourth international conference on weblogs and social media, ICWSM 2010, washington, dc, usa, may 23-26, 2010*. The AAAI Press. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1536.

Osnabrügge, Moritz, Elliott Ash & Massimo Morelli. 2021. Cross-domain topic classification for political texts. *Political Analysis*. 1–22. DOI: 10.1017/pan.2021.37.

Ouyang, Long, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F. Christiano, Jan Leike & Ryan Lowe. 2022. Training language models to follow instructions with human feedback. *CoRR* abs/2203.02155. DOI: 10.48550/arXiv.2203.02155. https://doi.org/10.48550/arXiv.2203.02155.

Padó, Sebastian, Andre Blessing, Nico Blokker, Erenay Dayanik, Sebastian Haunss & Jonas Kuhn. 2019. Who sides with whom? towards computational construction of discourse networks for political debates. In *Proceedings of the 57th annual meeting of the association for computational linguistics*, 2841–2847. Florence, Italy: Association for Computational Linguistics. DOI: 10.18653/v1/P19-1273. https://aclanthology.org/P19-1273.

Page, Benjamin I. 1976. The theory of political ambiguity. *American Political Science Review* 70(3). 742–752.

Pak, Alexander & Patrick Paroubek. 2010. Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the seventh international conference on language resources and evaluation (LREC'10)*. Valletta, Malta: European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2010/pdf/385_Paper.pdf.

Paltoglou, Georgios & Mike Thelwall. 2012. Twitter, myspace, digg: unsupervised sentiment analysis in social media. *ACM Trans. Intell. Syst. Technol.* 3(4). 66:1–66:19. DOI: 10.1145/2337542.2337551. https://doi.org/10.1145/2337542.2337551.

Pang, Bo & Lillian Lee. 2007. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.* 2(1-2). 1–135. DOI: 10.1561/1500000011. https://doi.org/10.1561/1500000011.

Paul, Debjit & Anette Frank. 2019. Ranking and selecting multi-hop knowledge paths to better predict human needs. In *Proceedings of the 2019 conference of the north American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, 3671–3681. Minneapolis, Minnesota: Association for Computational Linguistics. DOI: 10.18653/v1/N19-1368. https://aclanthology.org/N19-1368.

Pennebaker, James W, Martha E Francis & Roger J Booth. 2001. Linguistic inquiry and word count: liwc 2001. *Mahway: Lawrence Erlbaum Associates* 71(2001). 2001.

Petrocik, John R. 1996. Issue ownership in presidential elections, with a 1980 case study. *American journal of political science*. 825–850.

Pla, Ferran & Lluís-F. Hurtado. 2014. Political tendency identification in Twitter using sentiment analysis techniques. In *Proceedings of COLING 2014, the 25th international conference on computational linguistics: technical papers*, 183–192. Dublin, Ireland: Dublin City University & Association for Computational Linguistics. https://www.aclweb.org/anthology/C14-1019.

Preoţiuc-Pietro, Daniel, Ye Liu, Daniel Hopkins & Lyle Ungar. 2017. Beyond binary labels: political ideology prediction of Twitter users. In *Proceedings of the 55th annual meeting of the association for computational linguistics (volume 1: long papers)*, 729–740. Vancouver, Canada: Association for Computational Linguistics. DOI: 10.18653/v1/P17-1068. https://aclanthology.org/P17-1068.

Qi, Peng, Yuhao Zhang, Yuhui Zhang, Jason Bolton & Christopher D. Manning. 2020. Stanza: a Python natural language processing toolkit for many human languages. In *Proceedings of the 58th annual meeting of the association for computational linguistics: system demonstrations*. https://nlp.stanford.edu/pubs/qi2020stanza.pdf.

Quinn, Kevin M, Burt L Monroe, Michael Colaresi, Michael H Crespin & Dragomir R Radev. 2006. An automated method of topic-coding legislative speech over time with application to the 105th-108th US Senate. In *Midwest political science association meeting*, 1–61.

Quinn, Kevin M, Burt L Monroe, Michael Colaresi, Michael H Crespin & Dragomir R Radev. 2010. How to analyze political attention with minimal assumptions and costs. *American Journal of Political Science* 54(1). 209–228.

Raleigh, Clionadh, Andrew Linke, Håvard Hegre & Joakim Karlsen. 2010. Introducing ACLED-Armed Conflict Location and Event Data. *Journal of Peace Research* 47(5). 651–660. https://journals.sagepub.com/doi/10.1177/0022343310378914.

Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson & David G Rand. 2014. Structural topic models for open-ended survey responses. *American Journal of Political Science* 58(4). 1064–1082.

Rosenthal, Sara, Preslav Nakov, Svetlana Kiritchenko, Saif Mohammad, Alan Ritter & Veselin Stoyanov. 2015. Semeval-2015 task 10: sentiment analysis in twitter. In Daniel M. Cer, David Jurgens, Preslav Nakov & Torsten Zesch (eds.), *Proceedings of the 9th international workshop on semantic evaluation, semeval@naacl-hlt 2015, denver, colorado, usa, june 4-5, 2015*, 451–463. The Association for Computer Linguistics. DOI: 10.18653/v1/s15-2078. https://doi.org/10.18653/v1/s15-2078.

Sarol, M. Janina, Ly Dinh, Rezvaneh Rezapour, Chieh-Li Chin, Pingjing Yang & Jana Diesner. 2020. An empirical methodology for detecting and prioritizing needs during crisis events. In *Findings of the association for computational linguistics: emnlp 2020*, 4102–4107. Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.findings-emnlp.366. https://aclanthology.org/2020.findings-emnlp.366.

Schrodt, Philip A. 2000. Pattern recognition of international crises using hidden markov models. *Political complexity: Nonlinear models of politics*. 296–328.

Schrodt, Philip A & Deborah J Gerner. 1994. Validity assessment of a machine-coded event data set for the middle east, 1982-92. *American Journal of Political Science*. 825–854.

Schrodt, Philip A & Blake Hall. 2006. Twenty years of the kansas event data system project. *The political methodologist* 14(1). 2–8.

Schumacher, Gijs, Martijn Schoonvelde, Denise Traber, Tanushree Dahiya & Erik De Vries. 2016. EUSpeech: A new dataset of EU elite speeches. In *Proceedings of the international conference on the advances in computational analysis of political text (poltext 2016)*, 75–80.

Semetko, Holli A & Patti M Valkenburg. 2000. Framing European politics: A content analysis of press and television news. *Journal of communication* 50(2). 93–109.

Shah, Dhavan V, Joseph N Cappella & W Russell Neuman. 2015. Big data, digital media, and computational social science: possibilities and perils. *The ANNALS of the American Academy of Political and Social Science* 659(1). 6–13.

Sigelman, Lee & Emmett H Buell Jr. 2004. Avoidance or engagement? issue convergence in us presidential campaigns, 1960–2000. *American Journal of Political Science* 48(4). 650–661.

Sim, Yanchuan, Brice D. L. Acree, Justin H. Gross & Noah A. Smith. 2013. Measuring ideological proportions in political speeches. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, 91–101. Seattle, Washington, USA: Association for Computational Linguistics. https://aclanthology.org/D13-1010.

Simon, Adam F. 2002. *The winning message: Candidate behavior, campaign discourse, and democracy*. Cambridge University Press.

Slapin, Jonathan B & Sven-Oliver Proksch. 2008. A scaling model for estimating time-series party positions from texts. *American Journal of Political Science* 52(3). 705–722.

Sravani, Dama, Lalitha Kameswari & Radhika Mamidi. 2021. Political discourse analysis: a case study of code mixing and code switching in political speeches. In *Proceedings of the fifth workshop on computational approaches to linguistic code-switching*, 1–5. Online: Association for Computational Linguistics. DOI: 10.18653/v1/2021.calcs-1.1. https://aclanthology.org/2021.calcs-1.1.

Stefanov, Peter, Kareem Darwish, Atanas Atanasov & Preslav Nakov. 2020. Predicting the topical stance and political leaning of media using tweets. In *Proceedings of the 58th annual meeting of the association for computational linguistics*, 527–537. Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.50. https://aclanthology.org/2020.acl-main.50.

Stimson, James A, Michael B MacKuen & Robert S Erikson. 1995. Dynamic representation. *American political science review*. 543–565.

Sulkin, Tracy. 2005. *Issue politics in Congress*. Cambridge University Press.

Sun, Chi, Xipeng Qiu, Yige Xu & Xuanjing Huang. 2019. How to fine-tune BERT for text classification? In Maosong Sun, Xuanjing Huang, Heng Ji, Zhiyuan Liu & Yang Liu (eds.), *Chinese computational linguistics - 18th china national conference, CCL 2019, kunming, china, october 18-20, 2019, proceedings*, vol. 11856 (Lecture Notes in Computer Science), 194–206. Springer. DOI: 10.1007/978-3-030-32381-3_16. https://doi.org/10.1007/978-3-030-32381-3%5C_16.

Sundberg, Ralph & Erik Melander. 2013. Introducing the ucdp georeferenced event dataset. *Journal of Peace Research* 50(4). 523–532.

Thelwall, Mike, Kevan Buckley & Georgios Paltoglou. 2011. Sentiment in twitter events. *J. Assoc. Inf. Sci. Technol.* 62(2). 406–418. DOI: 10.1002/asi.21462. https://doi.org/10.1002/asi.21462.

Thomas, Matt, Bo Pang & Lillian Lee. 2006. Get out the vote: determining support or opposition from congressional floor-debate transcripts. In *Proceedings of the 2006 conference on empirical methods in natural language processing*, 327–335. Sydney, Australia: Association for Computational Linguistics. https://aclanthology.org/W06-1639.

Tjong Kim Sang, Erik & Johan Bos. 2012. Predicting the 2011 Dutch senate election results with Twitter. In *Proceedings of the workshop on semantic analysis in social media*, 53–60. Avignon, France: Association for Computational Linguistics. https://www.aclweb.org/anthology/W12-0607.

Trappl, Robert. 2006. *Programming for peace: computer-aided methods for international conflict resolution and prevention*. Vol. 2. Springer Science & Business Media.

Unankard, Sayan, Xue Li, Mohamed A. Sharaf, Jiang Zhong & Xueming Li. 2014. Predicting elections from social networks based on sub-event detection and sentiment analysis. In Boualem Benatallah, Azer Bestavros, Yannis Manolopoulos, Athena Vakali & Yanchun Zhang (eds.), *Web information systems engineering - WISE 2014 - 15th international conference, thessaloniki, greece, october 12-14, 2014, proceedings, part II*, vol. 8787 (Lecture Notes in Computer Science), 1–16. Springer. DOI: 10.1007/978-3-319-11746-1_1. https://doi.org/10.1007/978-3-319-11746-1%5C_1.

Van Aggelen, Astrid, Laura Hollink, Max Kemman, Martijn Kleppe & Henri Beunders. 2017. The debates of the European Parliament as linked open data. *Semantic Web* 8(2). 271–281.

Van Roy, V, F Rossetti, K Perset & L Galindo-Romero. 2021. *AI watch - national strategies on artificial intelligence: A european perspective, 2021 edition*. Scientific analysis or review, Policy assessment, Country report KJ-NA-30745-EN-N (online). Luxembourg (Luxembourg). DOI: 10.2760/069178(online).

Walker, Christopher, Stephanie Strassel, Julie Medero & Kazuaki Maeda. 2006. Ace 2005 multilingual training corpus. *Linguistic Data Consortium, Philadelphia* 57.

Winter, David G & Abigail J Stewart. 1977. Content analysis as a technique for assessing political leaders. *A psychological examination of political leaders*. 27–61.

Wolf, Thomas, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest & Alexander Rush. 2020. Transformers: state-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, 38–45. Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.emnlp-demos.6. https://aclanthology.org/2020.emnlp-demos.6.

Yin, Wenpeng, Jamaal Hay & Dan Roth. 2019. Benchmarking zero-shot text classification: datasets, evaluation and entailment approach. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (emnlp-ijcnlp)*, 3914–3923. Hong Kong, China: Association for Computational Linguistics. DOI: 10.18653/v1/D19-1404. https://aclanthology.org/D19-1404.