Analyzing Positions and Topics in Political Discussions of the German Bundestag

Cäcilia Zirn
Data and Web Science Group
University of Mannheim
Germany
caecilia@informatik.uni-mannheim.de

Abstract

We present ongoing doctoral work on automatically understanding the positions of politicians with respect to those of the party they belong to. To this end, we use textual data, namely transcriptions of political speeches from meetings of the German Bundestag, and party manifestos, in order to automatically acquire the positions of political actors and parties, respectively. We discuss a variety of possible supervised and unsupervised approaches to determine the topics of interest and compare positions, and propose to explore an approach based on topic modeling techniques for these tasks.

1 Introduction

The Bundestag is the legislative institution of Germany. In its plenary sessions, the members discuss the introduction and formulation of bills. Subjects under discussion include a wide spectrum of issues, ranging from funding of public transport through fighting right-wing extremism, or the deployment of German troops in Afghanistan. For each issue, a few selected members give a speech stating their opinion towards the topic, while the audience is allowed to interact: by questions, heckles, applause or even laughter. Transcriptions of the Bundestag's sessions provide us with a gold-mine of political speech data, encoding heterogeneous political phenomena such as, for instance, the prominence or engagement of the different politicians with respect to the current political situation, or their interest for specific topics.

In our work, we propose to leverage these data to enable the analysis of the speakers' positions with respect to the party they belong to, on the basis of the content of their speech. Questions we investigate include: which party's views do different politicians support? How much are their political views aligned with those of their party? Although we know *a-priori* which party a speaker belongs to, we view their positions on different topics with respect to their party's official lines as degrees of alignment, and measure them based on the content of their speeches. There are several circumstances under which a speaker might deviate from his or her party's opinion. For instance, he might stem from an election district where membership of a particular party increases his chances of being elected. Moreover, it might just happen that a politician who generally supports his party's lines personally has a different view on one particular topic. If we are able to measure positions from text, we allow for methods of analyzing adherence to party lines, which is an important issue in political science (cf. (Clinton et al., 2004), (Ceron, 2013) and (Ansolabehere et al., 2001)).

At its heart, our work aims at modeling politicians' positions towards a specific topic, as inferred from their speech. To estimate a position, in turn, we need a statement of the party's opinion towards the topic of interest, which can be then used for comparison against the speech. Various work in political science suggests to take this from party manifestos like (Keman, 2007) and (Slapin and Proksch, 2008). Research in political science has previously focused on analyzing political positions within text, for instance (Laver and Garry, 2000), (Laver et al., 2003), (Keman, 2007) or (Sim et al., 2013). However, most of previous work focused on the general position of a party or a person, like (Slapin and Proksch, 2008), as opposed to fine-grained positions towards specific topics. In our research, we address the two following tasks:

1. *Determine the speeches' topics* – namely develop methods to determine the topic(s) covered by a political speech, such as those given in the Bundestag.

Proceedings of the ACL 2014 Student Research Workshop, pages 26–33,
Baltimore, Maryland USA, June 22-27 2014.

2. *Quantify adherence to party lines* – namely estimate the speaker's position relatively to his party's opinion towards the respective topic(s).

In the following thesis proposal we present a variety of approaches that we plan to investigate in order to address these tasks, as well as discuss their limitations and challenges.

The first task, determining the topics, could be in principle addressed using well-studied supervised approaches like state-of-the-art machine learning algorithms. However, we cannot rely on the fact that all topics are covered in the training data. Consequently, we propose to explore an unsupervised approach that integrates information from an external resource. We suggest to use a variant of topic models which allows us to influence the creation of the topics.

The second task, determining the positions, is a bigger challenge, given the current state of the art. Some previous research looked at the related field of opinion mining, also on political discussion, as in (Abu-Jbara et al., 2012), (Anand et al., 2011) or (Somasundaran and Wiebe, 2009). These methods, however, are hardly applicable to the complex data of plenary meetings. In our scenario, we have to deal with a very specific kind of text, since the discussions do not consist of spontaneous dialogues, but rather formal statements. Consequently, we are forced to deal with a type of language which lies in-between dialogue and text. More concretely, within these speeches speakers roughly assume what positions the parties have and also have expectations about their opponents' opinions. Besides, as opposed to full-fledged dialogues, our data shows a very limited amount of interaction between the speaker and the audience, solely consisting of a few questions, heckles, laughter or applause. Further, as it is the goal of the discussions to constructively develop laws and agree on formulations, the speakers do not just state reasons pro or contra some issue. They rather illustrate different aspects of the discussed items. Furthermore, they try to convince others by emphasizing what their party has achieved in the past or criticize decisions taken in the past. To address these complex problems, we propose to start by using manually annotated party manifestos in order to provide us with an upper bound. Next, we propose to investigate the applicability of topic models to provide us, again, with a flexible unsupervised approach.

2 Data

The German Bundestag meets about 60 times a year, and discusses various items in each plenary session. There are various types of items on the agenda: they can be discussions about bills, but also question times or government's statements. We are interested in the first type only. Each bill has a unique identifier which is also mentioned by the session chair. By looking it up in a database provided by the Bundestag, it is possible to filter the bill discussions from other forms of items.

For each discussed item, a few selected members are permitted to give a speech. Most of the members belong to a party and their affiliation is publicly known.

The Bundestag releases the transcripts of its sessions as plain text documents. OffenesParlament[1] is a project run by volunteers that processes these documents and publishes them in a structured form on the web as HTML documents. The data distinguishes between parts of a given speech, utterances by the chairman and heckles, each annotated with its speaker. OffenesParlament makes the attempt to divide each session's transcript into parts containing a single item of the agenda only. This is not trivial, as it is the chairman who leads over using a non-standardized formulations, and thus contains many mistakes.

We collected a number of regular expressions and hope to improve the segmentation of the items. We will evaluate the performance of this heuristic by checking a sample with human judges.

Our extracted dataset covers the time period between March 2010 and December 2012 and consists of 182 meetings.

3 Determining topics in speeches

We aim at comparing the positions stated within the speeches to the general positions of the parties represented in the Bundestag. The parties' positions can be found in their manifestos, and are commonly used as a source by scholars, as in (Keman, 2007) or (Slapin and Proksch, 2008). In order to being able to compare speakers' and parties positions, we need to address two different tasks, namely: i) identifying the topic of a speech, and ii) locating that very same topic within the party manifesto or some further resource. The latter task depends on how the comparison is done. In this

[1] `http://offenesParlament.de`

section, we will focus on the first task: determining the topic of the speech.

There are two general approaches to classify the topics of text: either the topics are known in advance and constitute a static set of categories, for example (Hillard et al., 2008), or they are unknown in advance and dynamically created depending on the data, as in (Quinn et al., 2010) (see also (Grimmer and Stewart, 2013) and (Sebastiani, 2002) for an overview). In our scenario, we assume a common set of topics over several data sources, namely the party manifestos and transcripts of speeches in our case. Therefore, we opt for a fixed set of topic categories.

3.1 Definition of topical categories

In political science, there are various schemes to categorize political topics. A well-known and important project is the Comparative Manifesto Project (Budge et al., 2001), in which party manifestos are hand-coded on sentence level with a scheme of 560 categories. A similar project is the Comparative Agendas Project[2], which uses 21 top level categories further divided into fine-grained subcategories.

An alternative approach is to use the ministries as definition of the available categories, which inspired the category scheme used in (Seher and Pappi, 2011). In our work, we develop a category scheme for our particular task on the basis of the responsibilities of committees of the Bundestag, as suggested by internal discussions with scholars of political science. Similar to the ministries in government, the responsibilities for political areas are divided among various committees (see Table 1 for a list of committees). Each item discussed in the Bundestag is assigned to all committees who investigate the issues in more detail. For instance, in our data we find that a discussion about continuing the German participation in the International Security Assistance Force in Afghanistan has been assigned to the following committees: Foreign Affairs, Internal Affairs, Legal Affairs, Defense, Human Rights and Humanitarian Aid, Economic Cooperation and Development. For each issue, one of the committees is appointed as the leading one (German: federführende Ausschuss), the Committee of Foreign Affairs in this case.

Note that, crucially for our work, this assignment process provides us with human-annotated topic labels: in fact, not only can we use the committees as category definitions, but we can also use these very same assignments as a gold standard. Consequently, we use the definitions describing the responsibilities of the committees as our category scheme for political topics. We exclude three committees from the experiments namely: a) the Committee on Scrutiny of Elections, Immunity and the Rules of Procedure, b) the Committee on Petitions, and c) the Committee of Legal Affairs. This is because these committees are not directly responsible for a particular political domain, but perform meta functions.

Affairs of the European Union
Labour and social Affairs
Food, Agriculture and Consumer Protection
Family Affairs, Senior Citizens, Women and Youth
Health
Cultural and Media Affairs
Committee on Human Rights and Humanitarian Aid
Tourism
Environment, Nature Conservation and Nuclear Safety
Transport, Building and Urban Development
Scrutiny of Elections, Immunity and the Rules of Procedure
Economics and Technology
Economic Cooperation and Development
Foreign Affairs
Finance
Budget
Internal Affairs
Petitions
Legal Affairs
Sports
Defense
Education, Research and Technology Assessment

Table 1: Committees of the 17th German Bundestag.

Descriptions of the particular committees including their responsibilities and tasks as well as concrete examples of their work, accomplished by lists of current members, can be found in flyers released by the Bundestag[3].

Given this definition of political categories on the basis of the committees, we can create a gold standard for our topic classification scenario: to label a speech, we take the item it is given about, and use the committees the item has been assigned to as labels. The committee responsible, in turn, can be seen as the most important (i.e., primary) topic label[4]. Topic assignments are automatically harvested from a freely available source of infor-

[2] http://www.comparativeagendas.info

[3] https://www.btg-bestellservice.de/index.php?navi=1&subnavi=52

[4] Henceforth, we refer to the committees as labels for our topic classification task as “category” or “class”

mation, namely a public database offered by the German Bundestag[5]. Each item discussed in the Bundestag is associated with a printed document (*Drucksache*) tagged with a unique identifier, by which it can be tracked in the database and where the list of assigned committees can be queried.

Given these topic assignments, we aim at acquiring a model to classify the speeches with their assigned categories. To this end, we could focus on predicting the main label only (i.e. the committee responsible), or rather perform a multi-class labeling task predicting all labels (all committees the item is assigned to). We now overview a supervised and unsupervised approach to address these classification problems.

3.2 Supervised approach

Given that we have labeled data, a first solution is to opt for a supervised approach to text classification, which has been successfully used for many tasks like topic detection ((Diermeier et al., 2012), (Husby and Barbosa, 2012), or sentiment analysis (Bakliwal et al., 2013), to name a few. Consequently, in our case we could represent the speeches as a word vector and train state-of-the-art machine learning algorithms like Support Vector Machines, using the assigned committees as labels.

3.3 Unsupervised approach

In order to develop a generally applicable approach that can easily be applied to other resources such as speeches given in a context different from that of the Bundestag, we are interested to explore an unsupervised approach and compare it to the supervised one.

External definition of categories. The particular issues that fall into the responsibility of a committee are broad and might not be completely covered when using the speeches themselves as training data. As mentioned in Section 3.1, we have a clear definition of the tasks of each committee provided within the flyers. We will use them as a basis for the category definitions, and extend them with political issues discussed in party manifestos. We will explain this further in Section 3.3.

Known set of categories. Techniques such as LDA (Blei et al., 2003) create the topics dynamically during the classification process. Recently they became quite popular in political science, c.f. (Grimmer, 2010), (Quinn et al., 2010) or (Gerrish and Blei, 2011). As discussed in Section 3, we prefer to have a fixed set of categories. This allows for comparison between applications of the classification on different sources and domains separately. But while topic models do not fit this requirement, they have one property that corresponds quite well to our task: rather than assigning the text one single label, they return a distribution over topics contained by it. The items discussed in the speeches touch a range of political topics, and are assigned to various committees. There are variations of topic models that allow for influencing the creation of the topics, such as the systems of (Ramage et al., 2009) (Labeled LDA), (Andrzejewski and Zhu, 2009) or (Jagarlamudi et al., 2012). Labeled LDA is trained on a corpus of documents. In contrast to standard topic model approaches, it needs as input the information which labels (topics) are contained by the document, though not their proportions, thus uses a fixed set of categories.

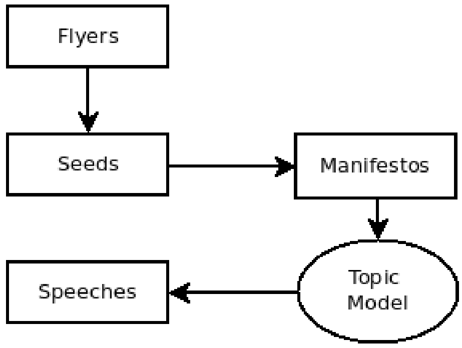

Figure 1: Approach overview

We illustrate our methodology in Figure 1. Our proposed approach starts by extracting seed words for the categories from the flyers about the committees. These seed words are then used to label training data for labeled LDA. As training data, we take an external resource: the manifestos[6] of all parties. Finally, we apply the trained model to the speeches to infer the labels. The output can be evaluated by comparing the predicted categories to the committees the issue is actually assigned to. In the following, we will explain each step in more detail.

[5] `dipbt.bundestag.de/dip21.web/bt`

[6] We combine the general party programs and the current election programs of each party

1) Extraction of seed words. We first download the flyers provided by the Bundestag. Then, we filter for nouns and calculate their TF-IDF values for the committee, by which we rank them. In a final step, we ask a scholar of political science to clean them, i.e. to delete nouns that are not necessarily important for the particular committee or are too ambiguous, and to cut the tail of low-ranked nouns. To give an example, we finally receive the following keywords for the committee of Labour and Social Affairs: *age-related poverty, labour-market policy, employee, social affairs, social security, labour, work, pension, basic social security, regulated rates, partial retirement, social standard, subcontracted labour*.

2) Automatically generating training data. We take the manifestos of all parties in the Bundestag to train our labeled LDA model. While topic models expect a whole collection of documents as input, we only provide a handful of them: accordingly, we generate a pseudo document collection by cutting the documents into snippets, following our previous work in (Zirn and Stuckenschmidt, 2013), and treating each of them as single documents. If a keyword for a committee is found within a snippet, we add the corresponding category to the documents labels. We finally run labeled LDA using standard configurations on the so labeled data.

3) Applying labeled LDA. Finally, we can apply the trained model on our transcribed speech data: we do this by inferring, for each speech, the distribution of topics, i.e. of categories. To evaluate the model, we check that the committee responsible corresponds to the highest probable topic inferred for the speech, and the other n assigned committees to the n most probable topics.

Currently, in our work, we are in the final stages of creating the gold standard, and evaluating our method. However, we have already implemented the proposed system as prototype, and accordingly show a part of the created topic model in Table 2 to give the reader an impression.

ENCNS	**LSA**	TBUD
consumer *(male)*	labour	mobility
consumer *(female)*	employee *male*	research
environment	employees *female*	infrastructure
protection	salary	railway
products	pension	traffic
farming	labour market	investments
nature	old-age provision	development
variety	unemployment	future
raw materials	employment	rails
transparency	percentage	streets

Table 2: Top 10 terms for the committees on Environment, Nature Conservation and Nuclear Safety (ENCNS), on Labour and social affairs (LSA) and on Transport, Building and Urban Development (TBUD).

4 Detecting positions

The overall goal of our work is to analyze the positions expressed by the speakers towards the debated item. As we aim at performing a fine-grained analysis, approaches merely classifying pro or contra (like those of (Walker et al., 2012) or (Somasundaran and Wiebe, 2009) are not applicable in our case. The same applies to the task of subgroup detection (as done by (Abu-Jbara et al., 2012), (Anand et al., 2011) or (Thomas et al., 2006)).

In order to produce a finer-grained model of positions, we want to develop a model that places positions stated in text along a one-dimensional scale, as done by (Slapin and Proksch, 2008) with their system called Wordfish, (Gabel and Huber, 2000),(Laver and Garry, 2000), (Laver et al., 2003) or (Sim et al., 2013). Wordfish places party manifestos on a left-right-scale, what visualizes very well which parties are close to each other and which ones are distant. This is similar in spirit to the purpose of our work, since we are interested primarily in estimating closeness and distances between the speakers' and the parties' positions. However, in contrast to their work, we are interested in positions towards specific topics, as opposed to general parties' positions.

We define our task as follows: we want to analyze the distance between the position towards a topic expressed in a speech and the position towards the same topic stated in a party manifesto. In the previous section, we described an approach to determine the topic of the speech. We now move on and present how we can retrieve the segments of the manifestos that correspond to the topic(s) addressed within the speeches, as well as how to compare these positions.

4.1 Approach A: Hand-coding of manifestos

Extract positions As part of a larger collaboration project with scholars of political science we decided to start with hand-coding a set of manifestos on sentence-level in order to have a gold standard for further work. To facilitate the manual work, we use a computer-assisted method based using the seed words created in Section 3.3. In more detail, we first use occurrences of the seed words to assign them the corresponding category label. Then, a human annotator validates these assignments, optionally adding missing labels.

If the sentence-wise labeled data proofs successful and necessary for the further analysis of political positions, we will investigate approaches to automate this process, for example with supervised learning or bootstrapping techniques starting with our seed words. For each topic, we can then accumulate the sentences assigned to its corresponding category and use this data as the party's opinion towards this topic.

Compare positions The comparison between the speech and the parties' opinions can then be performed as follows: for each party, we extract the sentences from the manifesto that are tagged with the topic covered in the speech. We then represent the extracted sentences and the speeches as word vectors, and compare them with a distance metric, e.g., a standard measure like cosine similarity, which gives us the closeness of the speech to each party's position.

4.2 Approach B: Topic Models

Extract positions Instead of selecting sentences from the manifesto that cover a topic, the position could be extracted from the manifesto using topic models, as shown in (Thomas et al., 2006) and (Gerrish and Blei, 2011). To extract the topics from the manifestos, we run labeled LDA separately on each manifesto, following the technique described in Section 3, yet with an important difference. In Section 3, we trained one common topic model on all manifestos, in order to have a broad coverage over all topics. Here, we are interested in the positions carried by the particular words chosen by the party to describe a topic. Accordingly, we train a separate topic model on each manifesto. The result is a distribution over terms for each committee, hence for each topic.

Compare positions As a result of the process to determine the topic of a speech (Section 3), the speeches also have a representation of the discussed topics as a distribution over terms. This way we can directly compare the distributions for the most probable topics in the speech with the corresponding topic in the party manifestos. This can be done using measures to estimate the distance between probability distributions like, for instance, Kullback-Leibler distance or Jensen-Shannon divergence.

5 Conclusions and Future Work

In this paper, we presented an overview of our thesis proposal on comparing positions found within political speeches against those expressed in party manifestos. To the best of our knowledge, this is the first work of this kind to aim at providing a fine-grained analysis of speakers' positions on political data. Arguably, the most exiting aspect of this work is that it grounds a variety of Natural Language Processing topics – e.g., polarity detection, topic modeling, among others – within a concrete, multi-faceted application scenario.

Being this a proposal, the first step in the future will be to complete the implementation of all above described methods and evaluate them. In our dataset, we are provided with additional information apart from the speech text: we know about heckles, laughter and applause and even know their origin. This knowledge can be used to estimate a network of support or opposition. This knowledge is also used in (Strapparava et al., 2010) to predict persuasiveness of sentences, which could constitute another source of information for our model. Another idea would be to make use of the speaker's given party affiliations and bootstrap an approach to analyze their positions: if we assume that a majority of the speakers actually does follow their parties' lines, we can train a classifier for each party for each topic, and apply it to the same data to detect outliers. Besides, a big research question would be to see how much we can complement our topic models with additional supervision in the form of symbolic knowledge sources like wide-coverage ontologies, e.g., DBpedia. Finally, while we do focus in this work on German data, we are interested in extending our model to other languages, including resource-rich ones like English as well as resource-poor ones.

Acknowledgements

We thank Google for travel and conference support for this paper.

References

Amjad Abu-Jbara, Mona Diab, Pradeep Dasigi, and Dragomir Radev. 2012. Subgroup detection in ideological discussions. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers - Volume 1*, ACL '12, pages 399–409, Stroudsburg, PA, USA. Association for Computational Linguistics.

Pranav Anand, Marilyn Walker, Rob Abbott, Jean E. Fox Tree, Robeson Bowmani, and Michael Minor. 2011. Cats rule and dogs drool!: Classifying stance in online debate. In *Proceedings of the 2Nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis*, WASSA '11, pages 1–9, Stroudsburg, PA, USA. Association for Computational Linguistics.

David Andrzejewski and Xiaojin Zhu. 2009. Latent dirichlet allocation with topic-in-set knowledge. In *Proceedings of the NAACL HLT 2009 Workshop on Semi-Supervised Learning for Natural Language Processing*, pages 43–48. Association for Computational Linguistics.

Stephen Ansolabehere, James M Snyder, and Charles Stewart III. 2001. The effects of party and preferences on congressional roll-call voting. *Legislative Studies Quarterly*, 26(4):533–572.

Akshat Bakliwal, Jennifer Foster, Jennifer van der Puil, Ron O'Brien, Lamia Tounsi, and Mark Hughes. 2013. Sentiment analysis of political tweets: Towards an accurate classifier. In *Proceedings of the Workshop on Language Analysis in Social Media*, pages 49–58, Atlanta, Georgia, June. Association for Computational Linguistics.

D.M. Blei, A.Y. Ng, and M.I. Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research (JMLR)*, 3:993–1022.

Ian Budge, Hans"=Dieter Klingemann, Andrea Volkens, Judith Bara, and Eric Tanenbaum. 2001. *Mapping Policy Preferences. Estimates for Parties, Electors, and Governments 1945-1998*. Oxford University Press, Oxford u. a.

Andrea Ceron. 2013. Brave rebels stay home: Assessing the effect of intra-party ideological heterogeneity and party whip on roll-call votes. *Party Politics*, page 1354068812472581.

Joshua Clinton, Simon Jackman, and Douglas Rivers. 2004. The statistical analysis of roll call data. *American Political Science Review*, 98(02):355–370.

Daniel Diermeier, Jean-Franois Godbout, Bei Yu, and Stefan Kaufmann. 2012. Language and ideology in congress. *British Journal of Political Science*, 42:31–55, 1.

Matthew J. Gabel and John D. Huber. 2000. Putting parties in their place: Inferring party left-right ideological positions from party manifestos data. *American Journal of Political Science*, 44(1):pp. 94–103.

Sean Gerrish and David M Blei. 2011. Predicting legislative roll calls from text. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pages 489–496.

Justin Grimmer and Brandon M Stewart. 2013. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*.

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

Dustin Hillard, Stephen Purpura, and John Wilkerson. 2008. Computer assisted topic classification for mixed methods social science research. *Journal of Information Technology and Politics*.

Stephanie Husby and Denilson Barbosa. 2012. Topic classification of blog posts using distant supervision. In *Proceedings of the Workshop on Semantic Analysis in Social Media*, pages 28–36, Avignon, France, April. Association for Computational Linguistics.

Jagadeesh Jagarlamudi, Hal Daumé, III, and Raghavendra Udupa. 2012. Incorporating lexical priors into topic models. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, EACL '12, pages 204–213, Stroudsburg, PA, USA. Association for Computational Linguistics.

Hans Keman. 2007. Experts and manifestos: Different sources - same results for comparative research. *Electoral Studies*, 26:76–89.

Michael Laver and John Garry. 2000. Estimating policy positions from political texts. *American Journal of Political Science*, pages 619–634.

Michael Laver, Kenneth Benoit, and John Garry. 2003. Extracting policy positions from political texts using words as data. *American Political Science Review*, 97(02):311–331.

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

Daniel Ramage, David Hall, Ramesh Nallapati, and Christopher D Manning. 2009. Labeled lda: A supervised topic model for credit attribution in multi-labeled corpora. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1*, pages 248–256. Association for Computational Linguistics.

Fabrizio Sebastiani. 2002. Machine learning in automated text categorization. *ACM computing surveys (CSUR)*, 34(1):1–47.

Nicole Michaela Seher and Franz Urban Pappi. 2011. Politikfeldspezifische positionen der landesverbände der deutschen parteien. Working Paper 139, Mannheimer Zentrum für Europäische Sozialforschung (MZES).

Yanchuan Sim, Brice D. L. Acree, Justin H. Gross, and Noah A. Smith. 2013. Measuring ideological proportions in political speeches. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 91–101, Seattle, Washington, USA, October. Association for Computational Linguistics.

Jonathan B. Slapin and Sven-Oliver Proksch. 2008. A Scaling Model for Estimating Time-Series Party Positions from Texts. *American Journal of Political Science*, 52(3):705–722, July.

Swapna Somasundaran and Janyce Wiebe. 2009. Recognizing stances in online debates. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1 - Volume 1*, ACL '09, pages 226–234, Stroudsburg, PA, USA. Association for Computational Linguistics.

Carlo Strapparava, Marco Guerini, and Oliviero Stock. 2010. Predicting persuasiveness in political discourses. In *LREC*. European Language Resources Association.

Matt Thomas, Bo Pang, and 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*, EMNLP '06, pages 327–335, Stroudsburg, PA, USA. Association for Computational Linguistics.

Marilyn A Walker, Pranav Anand, Robert Abbott, and Ricky Grant. 2012. Stance classification using dialogic properties of persuasion. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 592–596. Association for Computational Linguistics.

Cäcilia Zirn and Heiner Stuckenschmidt. 2013. Multidimensional topic analysis in political texts. *Data & Knowledge Engineering*.