Data Innovation for International Development

Broniecki, Philipp; Hanchar, Anna; Mikhaylov, Slava J.

Data Innovation for International Development: An overview of natural language processing for qualitative data analysis

Philipp Broniecki
School of Public Policy
University College London
Email: philipp.broniecki.14@ucl.ac.uk

Anna Hanchar
The Data Atelier
London, UK
Email: anna.h@thedataatelier.com

Slava J. Mikhaylov
Institute for Analytics and Data Science
Department of Government
University of Essex
Email: s.mikhaylov@essex.ac.uk

***Abstract*—Availability, collection and access to quantitative data, as well as its limitations, often make qualitative data the resource upon which development programs heavily rely. Both traditional interview data and social media analysis can provide rich contextual information and are essential for research, appraisal, monitoring and evaluation. These data may be difficult to process and analyze both systematically and at scale. This, in turn, limits the ability of timely data driven decision-making which is essential in fast evolving complex social systems. In this paper, we discuss the potential of using natural language processing to systematize analysis of qualitative data, and to inform quick decision-making in the development context. We illustrate this with interview data generated in a format of micro-narratives for the UNDP Fragments of Impact project.**

I. Introduction

Practitioners in the development sector have long recognized the potential of qualitative data to inform programming and gain a better understanding of values, behaviors and attitudes of people and communities affected by their efforts. Some organizations mainly rely on interview or focus group data, some also consider policy documents and reports, and others have started tapping into social media data. Regardless of where the data comes from, analyzing it in a systematic way to inform quick decision-making poses challenges, in terms of high costs, time, or expertise required.

The application of natural language processing (NLP) and machine learning (ML) can make feasible the speedy analysis of qualitative data on a large scale.

We start with a brief description of the main approaches to NLP and how they can be augmented by human coding. We then move on to the issue of working with multiple languages and different document formats. We then provide an overview of the recent application of these techniques in a United Nations Development Program (UNDP) study.

II. Supervised and Unsupervised Learning

There are two broad approaches to NLP - supervised learning and unsupervised learning [1]. Supervised learning assumes that an outcome variable is known and an algorithm is used to predict the correct variable. Classifying email as spam based on how the user has classified previous mail is a classic example. In social science, we may want to predict voting behavior of a legislator with the goal of inferring ideological positions from such behavior. In development, interest may center around characteristics that predict successful completion of a training program based on a beneficiary's previous experience or demographic characteristics.

Supervised learning requires human coding - data must be read and labelled correctly. This can require substantial resources. At the same time, the advantage is that validation of a supervised learning result is relatively straightforward as it requires comparing prediction results with actual outcomes. Furthermore, there is no need to label all text documents (or interview data from each respondent) prior to analyzing them. Rather, a sufficiently large set of documents can be labelled to train an algorithm and then used to classify the remaining documents.

In unsupervised learning, the outcome variable is unknown. The exercise is, therefore, of a more exploratory nature. Here the purpose is to reveal patterns in the data that allow us to distinguish distinct groups whose differences are small, while variations across groups are large. In a set of text documents, we may be interested in the main topics about which respondents are talking. In social sciences, we may look for groups of nations within the international system that use a similar language or that describe similar issues, like small-island states prioritizing climate change. Identifying such groups is often referred to as 'dimension reduction' of data.

Validation of unsupervised learning results is less straightforward than with supervised learning. We use data external to our analysis to validate the findings [2].

A complementary approach to unsupervised and supervised learning is the use of crowdsourced human coders. [3] show that crowdsourcing text analysis is a way to achieve reliable and replicable results quickly and inexpensively through the CrowdFlower platform. This approach can work well in supporting a supervised approach where outcomes need to be labelled. For example, [3] use this technique to produce party positions on the economic left-right and liberal-conservative dimensions from party manifestos. Online coders receive small

specific tasks to reduce individual biases. Their individual responses are then aggregated to create an overall measure of party positions.

III. Working with Multiple Languages

A common obstacle to analyzing textual data in many fields, including the international development sector, is the plethora of languages that practitioners and researchers need to consider – each with subtle but meaningful differences. Fortunately, significant commercial interest in being able to translate large quantities of text inexpensively has led to major advances in recent years driven by Microsoft, Google, and Yandex with the introduction of neural machine translation [4]. They provide services that are free of charge that can be easily integrated into standard programming languages like Python and R.[1] Open source neural machine translation systems are also being made available [5].

In a recent application, [6] estimate the policy preferences of Swiss legislators using debates in the federal parliament. With speeches delivered in multiple languages, the authors first translate from German, French, and Italian into English using Google Translate API. They then estimate the positions of legislators using common supervised learning methods from text and compare to estimates of positions from roll-call votes.

[7] evaluate the quality of automatic translation for social science research. The authors utilize the *europarl* dataset [8] of debate transcripts in the European Parliament and compare English, Danish, German, Spanish, French, and Polish official versions of the debates with their translations performed using Google Translate. [7] find that features identified from texts are very similar between automatically translated documents and official manual translations. Furthermore, topic model estimates are also similar across languages when comparing machine and human translations of EU Parliament debates.

IV. Working with Documents

In recent years, great strides have been made into leveraging information from text documents. For example, researchers have analyzed speeches, legislative bills, religious texts, press communications, newspaper articles, stakeholder consultations, policy documents, and regulations. Such documents often contain many different dimensions or aspects of information and it is usually impossible to manually process them for systematic analysis. The analytical methods used to research the content of such documents are similar. We introduce prominent applications from the social sciences to provide an intuition about what can be done with such data.

A. Open-ended survey questions

Open-ended questions are a rich source of information that should be leveraged to inform decision-making. We could be interested in several aspects of such a text document. One useful approach would be to find common, recurring topics across multiple respondents. This is an unsupervised learning task because we do not know what the topics are. Such models are known as topic models. They summarize multiple text documents into a number of common, semantic topics. [9] use a structural topic model (STM) that allows for the evaluation of the effects of structural covariates on topical structure, with the aim of analyzing several survey experiments and open-ended questions in the American National Election Study.

B. Religious statements

[10] analyze Islamic fatwas to determine whether Jihadist clerics write about different topics to those by non-Jihadists. Using an STM model, they uncover fifteen topics within the collection of religious writings. The model successfully identifies characteristic words in each topic that are common within the topic but occur infrequently in the fourteen other topics. The fifteen topics are labelled manually where the labels are human interpretations of the common underlying meaning of the characteristic words within each topic. Some of the topics deal with fighting, excommunication, prayer, or Ramadan. While the topics are relatively distinct, there is some overlap, i.e. the distance between them varies. This information can be leveraged to map out rhetorical network.

C. Public debates

[11] uses topic modeling to link the content of parliamentary speeches in the UK's House of Commons with the number of signatures of constituency-level petitions. He then investigates whether the signatures have any bearing on the responsiveness of representatives, i.e. whether Members of Parliament take up an issue if more people sign a petition on that issue. Also using speeches from the UK House of Commons, [12] produces evidence for the female role-model hypothesis. He shows that the appointment of female ministers leads to more speaking time and speech centrality of female backbenchers.

D. News reports

[13] uses supervised machine learning to generate data on UN peacekeeping activities in Cote d'Ivoire. The text data inputs are news articles from the website of the UN peacekeeping mission in Cote d'Ivoire. Based on a manually classified subset of articles, an algorithm is trained to classify terms into activity categories. Based on this algorithm, the remaining articles are then categorized. Analyzing these data yields new micro-level insights into the activities of peacekeepers on the ground and their effects.

[14] take a similar approach to researching the effectiveness of self-promotion strategies of politicians. They analyze 170,000 press releases from the U.S. House of Representatives. First, 500 documents are classified by hand into five categories of credit claiming, next the supervised learning algorithm, ReadMe [15], is used to code the remaining documents automatically. Using this data, they show that the number of times legislators claim credit generates more support than whether or not the subject they claim credit for amounts to much.

[1] Bing, Google, and Yandex place character limits on single document translations, with the latter allowing for slightly longer text chunks. For longer documents, commercial alternatives would have to be used.

E. Sentiment analysis

Instead of uncovering topics, we may want to know of a positive or negative tone of any given document. In an open-ended survey response, we could be interested in how the respondent rates the experience with the program. Sentiment analysis is a common tool for such a task. It is based on dictionaries of words that are associated with positive or negative emotions. The sentiment of a document such as an open-ended question would then be based on relative word counts and the sentiment scores with which these words are associated. [16] find positive and negative keywords and count their frequency in Chinese newspaper articles that mention the United States. With this data, they identify attitudes towards the United States in China.

F. Text reuse

Another application is to study text reuse in order to trace the flow of ideas. [17] analyze bill sections to identify whether two sections of a bill propose similar ideas. The algorithm they use was devised to trance gene sequencing and takes the frequency and order of words into account. Using this technique, they can measure the influence of one bill on another. Similarly, [18] studies policy diffusion by shedding light on how the American Supreme Court and Courts of Appeals influence diffusion of state policies. He uses plagiarism software to quantify the exact degree to which an existing law is reflected in a new proposal. Tracing ideas or influence over time and space can generate insights into the sources of information, the degree of spillover, and the influence of certain actors, ideas, or policies. It can shed light on network structures and long-term effects that would otherwise be hidden to us.

G. Estimating preferences of actors

In the social sciences, text is often used to infer preferences. Various scaling techniques have been developed and refined over recent years. Wordfish is a scaling algorithm that enables us to estimate policy positions based on word frequencies [19]. Researchers have used this approach to measure policy positions on European integration in the European Parliament [20], on austerity preferences in Ireland [21], and intra-party preferences in the energy debate in Switzerland [22]. Recent developments in the field allow researchers to estimate attitudes in multiple issue dimensions and, therefore, allow for more fine-grained preference estimates [23].

H. Taking context into account

More recent developments in NLP depart from frequencies of single words or groupings of multiple words. Instead, each word is an observation and its variables are other words or characters. Thus, each word is represented by a vector that describes words and their frequencies in the neighborhood. This approach allows for the capture of text semantics [24].

[25] apply this to evaluating real estate in the U.S. by comparing property descriptions with words that are associated with high quality. [26] collected all country statements made during the United Nations General Debate where heads of state and government make statements that they consider to be important for their countries. Using this data, [27] run a neural network. They construct an index of similarity between nations and policy themes that allows us to identify preference alliances. This enables them to identify major players using network centrality and show that speeches contain information that goes beyond mere voting records.

In a large exploratory effort, [28] use dynamic topic modeling which captures the evolution of topics over time, along with topological analytical techniques that allow for the analysis of multidimensional data in a scale invariant way. This enables them to understand political institutions, in this case through the use of speeches from the UK House of Commons. They classify representatives into groups according to speech content and verbosity, and identify a general pattern of political cohesion. They further show that it is feasible to track the performance of politicians with regard to specific issues using text. Topological techniques are especially useful to discover networks of relations using text. [27] apply this to uncover ideological communities in the network of states in the international system using UN General Assembly speeches.

V. Working with Short Text, Micro-Blogs, Social Media

Social media networks such as Twitter, the microblogging service, or the social network, Facebook, connect a vast amount of people in most societies. They generally contain shorter text excerpts compared to the sources of text previously discussed. However, their size and dynamic nature make them a compelling source of information. Furthermore, social networks online reflect social networks offline [29]. They provide a rare and cheap source of information on dynamic micro-level processes.

A. Twitter

Similar to our discussion above, topic models can be used to analyze social media data. [10] use such a model to analyze how the United States is viewed in China and in Arabic-speaking countries in response to the disclosure of classified information by Edward Snowden. They collect tweets containing the word “Snowden” in both languages. The tweets are then translated to English using machine translation. [10] show that Chinese posts are concerned more about attacks in terms of spying, while Arabic posts discuss human rights violations.

We can use social media to analyze networks and sentiments. Similar to word counts, volume of posts can carry information. [30] collect tweets originating from and referring to political actors around the 2014 elections to the European Parliament. They consider the language and national distribution as well as the dynamics of social media usage. Using network graphs depicting the conversations within and between countries, they identify topics debated nationally, and also find evidence for a Europe-wide debate around the EP elections and the European Union generally. Using

sentiment analysis, they further show that positive statements were correlated with pro-integration attitudes whereas negative debates were more national and anti-integration.

This EU example translates well to national conversations involving multiple ethnic or linguistic groups elsewhere. Moreover, we can learn how information spreads from social networks. Consequently, within ethical boundaries, we may also be able to target information more efficiently. An analysis of Twitter data from the Arab Spring suggests that coordination that originated from the periphery of a network rather than the center sparked more protest [31]. Coordination was measured as a Gini index of Hashtags while centrality was measured by a count of followers of an account.

B. Facebook

Social media has been used to estimate preferences as well. The advantage of social media compared to speeches or any other preference indicator is coverage. [32] use endorsement of official pages on Facebook to scale ideological positions of politicians from different levels of government and the public into a common space. Their method extends to other social media such as Twitter where endorsements and likes could be leveraged.

C. Weibo, RenRen, and Chinese microblogs

The most prominent example of supervised classification with social media data involves the first large scale study of censorship in China. [33] automatically downloaded Chinese blogposts as they appeared online. Later they returned to the same posts and checked whether or not they had been censored. Furthermore, they analyzed the content of the blog posts and showed that rather than banning critique directed at the government, censorship efforts concentrate on calls for collective expression, such as demonstrations.

Further investigations of Chinese censorship were made possible by leaked correspondence from the Chinese Zhanggong District. The leaks are emails in which individuals claim credit for propaganda posts in the name of the regime. The emails contain social media posts and account names. [34] used the leaked posts as training data for a classification algorithm that subsequently helped them to identify more propaganda posts. In conjunction with a follow-up survey experiment they found that most content constitutes cheerleading for the regime rather than, for example, critique of foreign governments.

In the next section we discuss an application of natural language processing in international development research.

VI. UNDP Fragments of Impact Initiative

In 2015, the United Nations Development Programme (UNDP) Regional Hub for Europe and CIS launched a Fragments of Impact Initiative (FoI) that helped to collect qualitative (micro-narratives) and quantitative data from multiple countries.

Within a six-month period, around 10,000 interviews were conducted in multiple languages. These covered the perception of the local population in countries including Tajikistan, Yemen, Serbia, Kyrgyzstan and Moldova on peace and reconciliation, local and rural development, value chain, female entrepreneurship and empowerment, and youth unemployment issues. The micro-narratives were collected using SenseMaker(r), a commercial tool for collecting qualitative and quantitative data. The micro-narratives were individual responses to context-tailored questions. An example of such a question is: "Share a recent example of an event that made it easier or harder to support how your family lives."

While the analysis and visualization of quantitative data was not problematic, systematic analysis and visualization of qualitative data, collected in a format of micro-narratives, would have been impossible.

To find a way to deal with the expensive body of micro-narrative data, UNDP engaged a group of students from the School of Public Policy, University College London, under the supervision of Prof Slava Mikhaylov (University of Essex) and research coordination of Dr Anna Hanchar (The Data Atelier). The objective of this work was to explore how to systematize the analysis of country-level qualitative data, visualize the data, and inform quick decision-making and timely experiment design. The results of the project were presented at the Data for Policy 2016 [35].

The UCL team had access to micro-narratives, as well as context specific meta-data such as demographic information and project details. For a cross-national comparison for policy-makers, the team translated the responses in multiple languages into English using machine translation, in this case Translate API (Yandex Technologies). As a pre-processing step, words without functional meaning (e.g. 'I'), rare words that occurred in only one narrative, numbers, and punctuation were all removed. The remaining words were stemmed to remove plural forms of nouns or conjugations of verbs.

As part of this exploration exercise, and guided by UNDP country project leads, the UCL team applied structural topic modeling [9] as an NLP approach and created an online dashboard containing data visualization per country. The dashboard included descriptive data, as well as results. Figure 1 illustrates an example of the dashboard. The analysis also allowed for the extraction of general themes described by respondents in the micro-narratives, and looked for predictors such as demographics that correlated with these themes. In Moldova, the major topic among men was rising energy prices. Among women the main topic was political participation and protest, which suggests that female empowerment programs could potentially be fruitful.

In Kyrgyzstan, the team found that the main topics revolved around finding work, access to resources and national borders. Using the meta-data on urbanization, it became clear that rural respondents described losing livestock that had crossed the border to Tajikistan, or that water sources were located across the border. The urban population was concerned about being able to cross the border to Russia for work. Figure 2

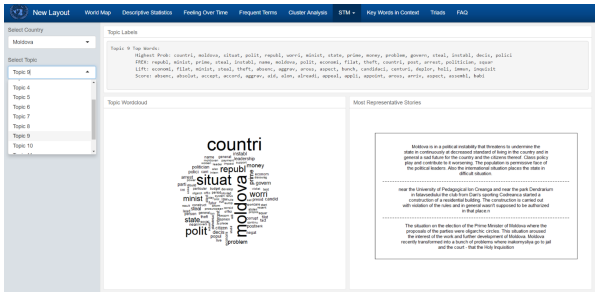

Fig. 1. *Dashboard Interface* Example of topic modeling results for Moldova showing the highest probability words for one of the topics and corresponding most representative micro-narratives.

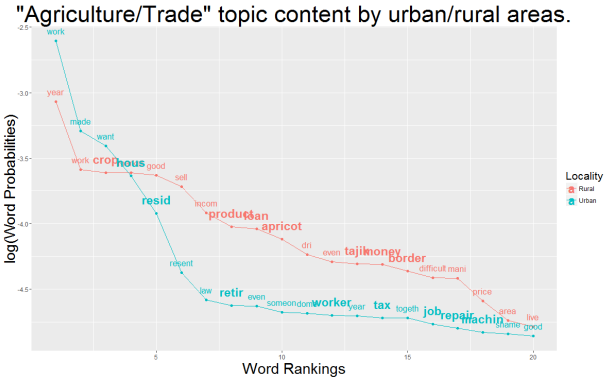

Fig. 2. *Comparing topic prevalence and topical content for urban and rural areas in Kyrgyzstan.*

shows word probabilities from the main "agriculture/trade" topic across respondents from urban and rural communities.

For Serbia, the analysis compared issues faced by Roma populations in areas of high and low Roma concentration. Figure 3 shows the relationship between topics discussed by Roma respondents in areas of high concentration.

[35] found that Roma respondents identified education

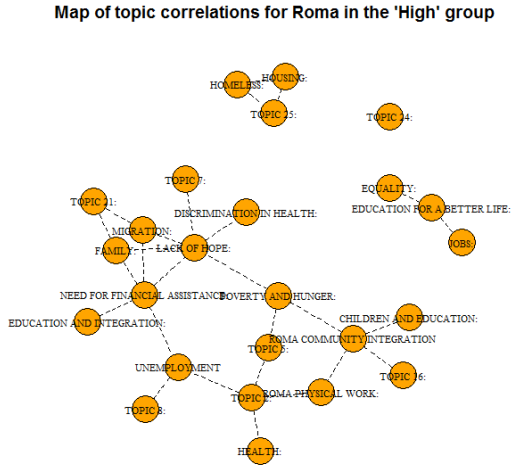

Fig. 3. *Topic correlations in areas of high Roma concentration in Serbia.*

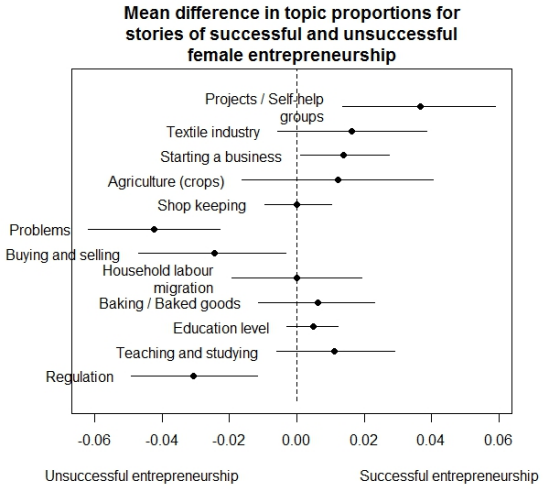

Fig. 4. *Identifying constraints on female entrepreneurship in Tajikistan*

as the overarching main topic, independent of the density of the Roma population. Differences were found between respondents across the level of integration with society. In areas of high Roma concentration, respondents were aware of available channels for inclusion. In low Roma density areas, respondents were mainly concerned with severe poverty and discrimination preventing societal inclusion.

In Tajikistan, [35] investigated the relationship between household labor migration and female entrepreneurship success. They found strong regional differences between the Sughd and Khatlon regions where topics in less successful Khatlon revolved around red tape. Moreover, successful entrepreneurship was very much related to receiving remittances. Figure 4 illustrates topics that correlate with success.

Analysis of micro-narratives from Yemen showed that the most recurrent themes focused on family issues (Figure 5). There are significant differences in terms of engagement in civil society between young people and the older population. Young respondents emphasized pro-active behavior, political engagement, and interest in community-driven initiatives fostering political development.

VII. Conclusion

In this overview, our aim has been to demonstrate how new forms of data can be leveraged to complement the work of practitioners in international development. We have demonstrated that a wide variety of questions can be asked.

Exploratory work can be performed to systematize large quantities of text. Additionally, we can learn about the sentiment that specific groups express towards specific topics. Networks can be uncovered, the spread of information or ideas can be traced, and influential actors identified. We can classify documents based on human coding of a subset of documents and establish which topics predict/correlate with predefined outcomes such as successful employment or completion of a program.

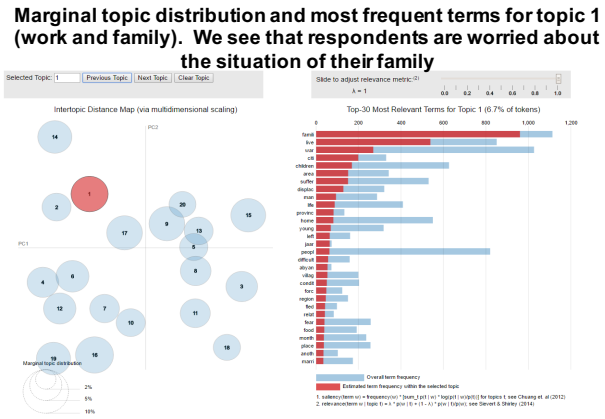

Fig. 5. *Exploring differences in perspectives by age in Yemen*

While the application used here to illustrate the discussion focuses on texts in the form of open-ended questions, social networks can be used and their coverage and topicality can be leveraged.

Natural language processing has the potential to unlock large quantities of untapped knowledge that could enhance our understanding of micro-level processes and enable us to make better context-tailored decisions.

ACKNOWLEDGMENT

Authors' names are listed in alphabetical order. Authors have contributed equally to all work.

REFERENCES

[1] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*. Springer, 2013, vol. 112.

[2] J. Grimmer and B. M. Stewart, "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts," *Political Analysis*, vol. 21, no. 3, pp. 267–297, 2013.

[3] K. Benoit, D. Conway, B. E. Lauderdale, M. Laver, and S. J. Mikhaylov, "Crowd-sourced Text Analysis: Reproducible and Agile Production of Political Data," *American Political Science Review*, vol. 110, no. 2, pp. 278–295, 2016.

[4] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey *et al.*, "Google's neural machine translation system: Bridging the gap between human and machine translation," *arXiv preprint arXiv:1609.08144*, 2016.

[5] G. Klein, Y. Kim, Y. Deng, J. Senellart, and A. M. Rush, "Opennmt: Open-source toolkit for neural machine translation," *arXiv preprint arXiv:1701.02810*, 2017.

[6] K. Benoit, D. Schwarz, and D. Traber, "The sincerity of political speech in parliamentary systems: A comparison of ideal points scaling using legislative speech and votes," in *2nd Annual Conference of EPSA, Berlin*, 2012, pp. 19–21.

[7] E. de Vries, M. Schoonvelde, and G. Schumacher, "Lost in translation? evaluating the usefulness of machine translation for bag-of-words text models," 2017, Open Science Framework.

[8] P. Koehn, "Europarl: A parallel corpus for statistical machine translation," in *MT summit*, vol. 5, 2005, pp. 79–86.

[9] M. E. Roberts, B. M. Stewart, D. Tingley, C. Lucas, J. Leder-Luis, S. K. Gadarian, B. Albertson, and D. G. Rand, "Structural Topic Models for Open-Ended Survey Responses," *American Journal of Political Science*, vol. 58, no. 4, pp. 1064–1082, 2014.

[10] C. Lucas, R. A. Nielsen, M. E. Roberts, B. M. Stewart, A. Storer, and D. Tingley, "Computer assisted text analysis for comparative politics," *Political Analysis*, vol. 23, no. 2, pp. 254–277, 2015.

[11] J. Blumenau, "Are E-Petitions Pointless? Evidence from the House of Commons," 2017, unpublished.

[12] ——, "Legislative Role Models: Female Ministers , Participation , and Influence in the UK House of Commons," *unpublished*, 2017.

[13] H. Smidt, "Peacekeeping activities in cote d'ivoire," unpublished.

[14] J. Grimmer, S. Messing, and S. J. Westwood, "How Words and Money Cultivate a Personal Vote: The Effect of Legislator Credit Claiming on Constituent Credit Allocation," *American Political Science Review*, vol. 106, no. 4, pp. 703–719, 2012.

[15] D. J. Hopkins and G. King, "A Method of Automated Nonparametric Content Analysis for Social Science," *American Journal of Political Science*, vol. 54, no. 1, pp. 229–247, 2010.

[16] A. I. Johnston and D. Stockmann, "Chinese attitudes toward the United States and Americans," *Anti-Americanisms in World Politics*, pp. 157–195, 2007.

[17] J. Wilkerson, D. Smith, and N. Stramp, "Tracing the Flow of Policy Ideas in Legislatures: A Text Reuse Approach," *American Journal of Political Science*, vol. 59, no. 4, pp. 1002–1021, 2015.

[18] R. K. Hinkle, "Into the Words: Using Statutory Text to Explore the Impact of Federal Courts on State Policy Diffusion," *American Journal of Political Science*, vol. 52, no. 3, pp. 1002–1021, 2015.

[19] J. B. Slapin and S.-O. Proksch, "A Scaling Model for Estimating Time-Series Party Positions from Texts," *American Journal of Political Science*, vol. 52, no. 3, pp. 705–722, 2008.

[20] S.-O. Proksch and J. B. Slapin, "Position Taking in European Parliament Speeches," *British Journal of Political Science*, vol. 40, no. 3, pp. 587–611, 2010.

[21] A. Herzog and K. Benoit, "The Most Unkindest Cuts: Speaker Selection and Expressed Government Dissent during Economic Crisis," *Journal of Politics*, vol. 77, no. 4, pp. 1157–1175, 2015.

[22] D. Schwarz, D. Traber, and K. Benoit, "Estimating Intra-Party Preferences: Comparing Speeches to Votes*," *Political Science Research and Methods*, vol. 5, no. 2, pp. 379–396, 2017.

[23] B. E. Lauderdale and A. Herzog, "Measuring Political Positions from Legislative Speech," *Political Analysis*, vol. 24, no. 3, pp. 374–394, 2016.

[24] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.

[25] M. Shahbazi, J. R. Barr, V. Hristidis, and N. N. Srinivasan, "Estimation of the investability of real estate properties through text analysis," in *Semantic Computing (ICSC), 2016 IEEE Tenth International Conference on*. IEEE, 2016, pp. 301–306.

[26] A. Baturo, N. Dasandi, and S. J. Mikhaylov, "Understanding state preferences with text as data: Introducing the UN General Debate corpus," *Research and Politics*, vol. 4, no. 2, 2017.

[27] S. Gurciullo and S. Mikhaylov, "Detecting policy preferences and dynamics in the un general debate with neural word embeddings," *arXiv preprint arXiv:1707.03490*, 2017.

[28] S. Gurciullo, M. Smallegan, M. Pereda, F. Battiston, A. Patania, S. Poledna, D. Hedblom, B. T. Oztan, A. Herzog, P. John *et al.*, "Complex politics: A quantitative semantic and topological analysis of uk house of commons debates," *arXiv preprint arXiv:1510.03797*, 2015.

[29] J. Bisbee and J. M. Larson, "Testing Social Science Network Theories with Online Network Data: An Evaluation of External Validity," *American Political Science Review*, vol. 111, no. 3, pp. 502–521, 2017.

[30] P. Nulty, Y. Theocharis, S. A. Popa, O. Parnet, and K. Benoit, "Social media and political communication in the 2014 elections to the European Parliament*," *Electoral Studies*, vol. 44, pp. 429–444, 2016.

[31] Z. C. Steinert-Threlkeld, "Spontaneous Collective Action: Peripheral Mobilization During the Arab Spring," *American Political Science Review*, vol. 111, no. 2, pp. 379–403, 2017.

[32] R. Bond and S. Messing, "Quantifying Social Media's Political Space : Estimating Ideology from Publicly Revealed Preferences on Facebook," *American Political Science Review*, vol. 109, no. 1, pp. 62–78, 2015.

[33] G. King, J. Pan, and M. E. Roberts, "How Censorship in China Allows Government Criticism but Silences Collective Expression," *American Political Science Review*, vol. 107, no. 2, pp. 326–343, 2013.

[34] ——, "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument," *American Journal of Political Science*, vol. 111, no. 3, pp. 484–501, 2017.

[35] C. Amato, A. Karmhus, E. C. Magrini, C. Mealings, and B. Vigreux, "Generating policy insights from micro-narratives," in *Data for Policy 2016*. 2nd Annual International Conference, September 2016.