Councils in Action

Eva Maxfield Brown, TODO: Council Data Project Contributors, Nic Weber

TODO

Introduction

Municipal government meetings are a place where policy is discussed by elected officials, a place for the community to provide last minute comment on policy, and the place where policy is ultimately enacted. The rapid growth and adoption of remotely available, and audio or video recorded, municipal government meetings enables the study of urban politics possible on new scales.

- Scale is hard
- Understanding what is available for each municipality
- Choosing what to store and how
- The data itself

This paper introduces a new dataset called: "Councils in Action" which can be used to study municipal government political deliberations and the enactment of public policy. Using Councils in Action, we have created tools for exploratory data analysis to support new forms of urban political scholarship.

- Tooling shouldn't just be limited to academics
- Such data can be used to create search engines for journalists and members of the community
- Querying, filtering, and more statistical methods are all available to remove barriers to exploring the dataset.

Exploratory data analysis is the process of using summary statistics and data visualization to allow an individual to get a better understanding of the data and potentially help form hypotheses (Tukey, 1980).

- As municipal government meetings become more fully digitized and available, we argue that exploratory data analysis can be used to remove barriers that are created with scale.

- EDA is especially useful at scale when evaluating hypotheses that may be computationally complex prior to working on the problem in earnest.

Our recent work with Council Data Project produced data attempts to solve two problems of urban politics scholarship: municipal government data used in urban political scholarship is typically limited to a single or few municipalities, and such data is typically not machine-readable for use in large-scale quantitative study.

- Councils in Action is a dataset composed of high-quality transcripts, video, audio, voting data, and policy documents.
- The tools we have created allow for large-scale, comparative analysis of municipal meetings.
- Unlike prior datasets which enable urban political scholarship, Councils in Action has both transcripts and policy outcomes together.

This paper uses exploratory data analysis to explore how meeting transcripts and other text-as-data can be combined with traditional municipal government meeting data (i.e. voting records and policy documents) to enable comparative, longitudinal, urban political scholarship.

- We discuss the broader use of text-as-data in the larger field of political science
- We demonstrate a few methods which we argue are a start for enabling longitudinal-comparative urban political research.

TODO: Reference limitations? (I.e. "recordings are common in typically large, wealthy municipalities", our dataset naturally tends towards the larger municipalities)

Background on Municipal Government and Urban Politics

TODO

Related Work

Text-as-Data in Political Scholarship

Text-as-data has been the basis for many forms of political science and legal research:

- Political polarization (Németh, 2022)
- Partisan affiliation with word embeddings (Abercrombie & Batista-Navarro, 2022; Bayram et al., 2019; Bergam et al., 2022; Khanna, 2020; Rao & Spasojevic, 2016; Yan et al., 2017)
- Debate dynamics (Abbott et al., 2011; Ahmadalinezhad & Makrehchi, 2018; Gupta et al., 2022; Jacobi & Schweers, 2017; Jalali & Sadeghi, 2014; Parthasarathy et al., 2019)

- Sentiment (Abercrombie & Batista-Navarro, 2019, 2020; Burfoot, 2008; Burfoot et al., 2011)
- Policy Classification (Medina, 2019)
- In many cases, the text being processed originates from legislative meeting transcripts (federal legislature floor speeches, UK parliament floor speeches, judicial transcripts, etc.).
- Quantitative urban political science scholarship leans heavily on access to publicly available meeting minutes. (Einstein et al., 2018, 2022; Sahn, 2022; Yoder, 2020)
- Specifically these studies heavily rely on states and cities which release public comments with their meeting minutes.
- Further, quantitative urban political science scholarship tends to tie the available public comments to policy documents or decisions that were discussed during the same meeting.
- Quantitative urban political scholarship as diverse as the larger field of political science is possible when provided access to a large corpus of transcript data. (Else, n.d.)

Prior Datasets of Municipal Meetings

There are problems in simply getting data in the same structure across many municipalities. (Davis et al., 1998; Stewart, 2010)

Each municipality may have its own regulations for distribution in addition to the distribution format and access method There are some companies, and teams, GovTech, CivicTech, academic which have created systems for archiving this data

TODO: TABLE

- Councilmatic
- Minutes
- Big Local News
- YouTube transcripts analysis paper (http://soubhikbarari.com/research/localgov.pdf)
- Open Civic Data
- Swagit
- Legistar
- Internet Archive
- Block Party (https://blockparty.studio)

TODO: TABLE None of these currently meet all the conditions and data structures Councils in Action does.

- Transcripts
- Meeting Minutes

- Documents
- Openly and easily available
- Reusable

Exploratory Data Analysis of Large Text Corpora

Compare and reference Google Trends and (Aiden & Michel, 2011; Organisciak et al., 2021)

TODO: "Summary":

- Text-as-data is well utilized in the parent field of political science
- Meeting minutes, and specifically recorded public comments, are heavily utilized by leading researchers in urban political scholarship
  - Minutes are especially useful because comments can be tied directly to policy outcome
- There is a need for this data to not only be made available but to made easy to use
- Such a dataset would include thousands of hours of meeting data
- EDA and "distant reading" approaches are useful for parsing, filtering, and gaining an understanding of the data

The Councils in Action Dataset

Data Model

Access and Use

Web User Interface

Programmatic API

Comparative, Longitudinal, Municipal Meeting Data

Text-as-Data

Ngram Usage

Semantic Representations

Voting Blocs

Discussion

Conclusion

References

Abbott, R., Walker, M. A., Anand, P., Tree, J. E. F., Bowmani, R., & King, J. (2011). *Recognizing disagreement in informal political argument*.

Abercrombie, G., & Batista-Navarro, R. T. (2019). Sentiment and position-taking analysis of parliamentary debates: A systematic literature review. *Journal of Computational Social Science*, *3*, 245–270.

Abercrombie, G., & Batista-Navarro, R. T. (2020). ParlVote: A corpus for sentiment analysis of political debates. *International Conference on Language Resources and Evaluation*.

Abercrombie, G., & Batista-Navarro, R. T. (2022). Policy-focused stance detection in parliamentary debate speeches. *Northern European Journal of Language Technology*.

Ahmadalinezhad, M., & Makrehchi, M. (2018). Detecting agreement and disagreement in political debates. *International Conference on Social, Cultural, and Behavioral Modeling*.

Aiden, E. L., & Michel, J.-B. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, *331*, 176–182.

Bayram, U., Pestian, J. P., Santel, D., & Minai, A. (2019). What's in a word? Detecting partisan affiliation from word use in congressional speeches. *2019 International Joint Conference on Neural Networks (IJCNN)*, 1–8.

Bergam, N., Allaway, E., & McKeown, K. (2022). Legal and political stance detection of SCOTUS language. *ArXiv*, *abs/2211.11724*.

Burfoot, C. (2008). Using multiple sources of agreement information for sentiment classification of political transcripts. *Australasian Language Technology Association Workshop*.

Burfoot, C., Bird, S., & Baldwin, T. (2011). Collective classification of congressional floor-debate transcripts. *Annual Meeting of the Association for Computational Linguistics*.

Davis, C. N., Chance, S. F., & Chamberlin, B. F. (1998). Guardians of access: Local prosecutors and open meetings laws. *Communication Law and Policy*, *3*, 35–54.

Einstein, K. L., Glick, D. M., Puig, L. G., & Palmer, M. (2022). Still muted: The limited participatory democracy of zoom public meetings. *Urban Affairs Review*.

Einstein, K. L., Palmer, M., & Glick, D. M. (2018). Who participates in local government? Evidence from meeting minutes. *Perspectives on Politics*, *17*, 28–46.

Else, D. C. M. T. A. (n.d.). *The asymmetric nationalization of local politics: Partisanship aligns with politics in*.

Gupta, A., Blodgett, S. L., Gross, J. H., & O'Connor, B. T. (2022). Examining political rhetoric with epistemic stance detection. *ArXiv*, *abs/2212.14486*.

Jacobi, T., & Schweers, D. (2017). Justice, interrupted: The effect of gender, ideology and seniority at supreme court oral arguments. *Virginia Law Review*, *103*, 1379–1496.

Jalali, M. S., & Sadeghi, B. (2014). A critical discourse analysis of political speech of four candidates of rasht city council elections in 2013, with a view to fairclough approach. *European Journal of Social Sciences Education and Research*, *1*, 8.

Khanna, S. (2020). *Guess the stance: Predicting political orientation from congressional statements*.

Medina, S. R. (2019). *Multi-label text classification with transfer learning for policy documents*.

Németh, R. (2022). A scoping review on the use of natural language processing in research on political polarization: Trends and research prospects. *Journal of Computational Social Science*, 1–25.

Organisciak, P., Schmidt, B. M., & Downie, J. S. (2021). Giving shape to large digital libraries through exploratory data analysis. *Journal of the Association for Information Science and Technology*, *73*, 317–332.

Parthasarathy, R., Rao, V., & Palaniswamy, N. (2019). Deliberative democracy in an unequal world: A text-as-data study of south india's village assemblies. *American Political Science Review*, *113*, 623–640.

Rao, A., & Spasojevic, N. (2016). Actionable and political text classification using word embeddings and LSTM. *ArXiv*, *abs/1607.02501*.

Sahn, A. (2022). *Public comment and public policy*.

Stewart, D. R. (2010). Let the sunshine in, or else: An examination of the "teeth" of state and federal open meetings and open records laws. *Communication Law and Policy*, *15*, 265–310.

Tukey, J. W. (1980). Exploratory data analysis. *ACM SIGSPATIAL International Workshop on Advances in Geographic Information Systems*.

Yan, H., Lavoie, A., & Das, S. (2017). The perils of classifying political orientation from text. *LINKDEM@IJCAI*.

Yoder, J. (2020). Does property ownership lead to participation in local politics? Evidence from property records and meeting minutes. *American Political Science Review*, *114*, 1213–1229.