

# DebateVis: Visualizing Political Debates for Non-Expert Users

Laura South \* Michail Schwab Nick Beauchamp Lu Wang John Wihbey Michelle A. Borkin

Northeastern University

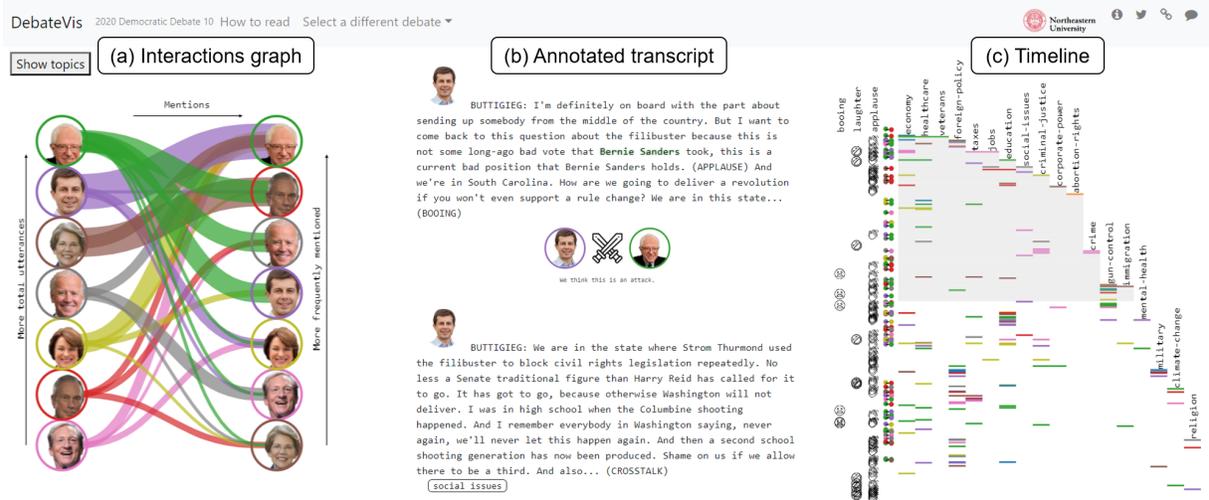


Figure 1: Transcript of the tenth debate of the 2020 U.S. Democratic primary election cycle, viewed with DEBATEVIS. The interactions graph (1) summarizes speaker behavior throughout the debate, including topics discussed and interactions with other speakers. The user can consult the annotated transcript (2) to read the full context of any statement in the debate. Rapid navigation through the transcript is possible by clicking on marks for the audience interactions, speaker interactions, and topic mentions automatically identified and visualized in the timeline (3). All three visualizations are connected through brushing and linking.

## ABSTRACT

Political debates provide an important opportunity for voters to observe candidate behavior, learn about issues, and make voting decisions. However, debates are generally broadcast late at night and last more than ninety minutes, so watching debates live can be inconvenient, if not impossible, for many potential viewers. Even voters who do watch debates may find themselves overwhelmed by a deluge of information in a substantive, issue-driven debate. Media outlets produce short summaries of debates, but these are not always effective as a method of deeply comprehending the policies candidates propose or the debate techniques they employ. In this paper we contribute reflections and results of an 18-month design study through an interdisciplinary collaboration with journalism and political science researchers. We characterize task and data abstractions for visualizing political debate transcripts for the casual user, and present a novel tool (DEBATEVIS) to help non-expert users explore and analyze debate transcripts.

**Index Terms:** Human-centered computing—Visualization—Visualization design and evaluation methods

## 1 INTRODUCTION

Imagine that an election is coming up in the next few months, and you are torn between two candidates. Perhaps between working full-time and taking care of your children, you do not have the time

\*Corresponding author. E-mails: [south.l, schwab.m, n.beauchamp, lu.wang, j.wihbey, m.borkin] @northeastern.edu

or energy to seek out resources to educate yourself on the candidates' stances on several key issues that are important to you. You read online that a debate is scheduled this week. Your first thought is to watch the debate live on television, but ultimately you decide that you cannot justify dedicating two hours to watching candidates discuss many topics, some of which may be of limited interest to you. You could seek out summary articles and videos compiled by journalists during post-debate coverage, but these will include only the parts of the debate that the creators found most interesting or most controversial.

Political debates are complex, dense artifacts of the political processes that influence our culture and our lives. Because debates are frequently broadcast late at night and last for several hours, keeping up with debate performances is inconvenient or impossible for many people. Even for people who do watch a debate live, it can be challenging to digest, summarize, and remember the events of a debate afterwards without re-viewing or re-reading the entire debate or relying on soundbites selected by others. A way to analyze and preserve the key aspects of the debate is essential, but debates involve complex interactions between many people, all with their own separate objectives and even their own ways of framing what happened (illustrated by the "spin room" afterwards). This means that it is inherently difficult to fully represent debates even in traditional journalistic narrative, let alone graphical representation.

Visualization can help solve this problem, but prior work has focused on creating visualization tools for in-depth expert analysis rather than the casual user. These tools often require users to have a strong grasp of complex visualization techniques, a method of gaining access to the tool from the developers, or specialized software to run the tool. In this paper, we describe the results of an 18-month design study to address the challenges inherent in visualizing debate transcripts for non-expert users, conducted with an interdisciplinary

team of researchers from computer science, political science, and journalism. We contribute a novel task analysis and abstraction for casual users exploring debate transcripts. This task analysis was used to inform the iterative development of DEBATEVIS, a novel visualization tool for exploring debate transcripts built for non-expert users.

## 2 RELATED WORK

### 2.1 Discourse and debate visualization

Most debate visualizations fall into one of two categories: *academic* and *non-academic*. *Academic* debate visualizations are built by researchers to be used by researchers. They may include a range of detailed views in response to specific domain goals uncovered in a task analysis phase, and often use ambitious visualization techniques unfamiliar to non-researchers, like node-link diagrams, heatmaps, and custom visualizations created for the application. Because these tools are built for a specific subset of researchers, they often require direct contact with the development team in order to gain access and are not freely available on the Internet. VisArgue [13], NEREx [14], and MultiConVis [17] are all examples of successful *academic* debate visualizations. Of these tools, only NEREx [14] incorporates speaker interactions.

*Non-academic* debate visualizations are targeted towards a more general audience. Media coverage of political debates and speeches tends to take one of three forms: full transcripts [1, 25], summary videos [23, 27], and data visualizations [2, 9, 11, 19]. Debate visualizations in this category are usually created by journalists or visual designers interested in politics and use simple visualization styles that are familiar to the average reader. Visualizations in this category are usually not connected to the full transcript of the debate [1, 2, 9, 11, 19, 27]. Several recent visualizations targeting general audiences have attempted to merge close reading of the transcript with higher level visual summaries, like Politico’s summary of the 2019 State of the Union address [25] or Twitter’s streamgraph of the most tweeted hashtags during the 2014 State of the Union address [4]. These tools are built to visualize an address given by one speaker, and do not scale to multi-speaker debate transcripts.

### 2.2 Casual information visualization

Casual information visualization, along with other subfields of visualization like data journalism and public visualizations, aims to understand how non-experts perceive and interact with data visualizations [26]. As a case study in creating text visualizations for non-experts, our work is closely related to casual information visualization. Segel and Heer reviewed the design space for data storytelling in online journalism [29], while Peck, Ayuso, & El-Etr explored how design guidelines for data visualization can exclude the “data poor” by interviewing residents of rural Pennsylvania with diverse economic and educational backgrounds [24]. In a treatise on ethical visualization research, Correll called for a shift towards visualization for the masses, arguing that “visualization work should be concerned with imbalances in power, and focus on distributing power in more equitable ways, and to more ethical ends” [12]. Visualizations like DEBATEVIS can help redistribute political power by encouraging non-experts to engage with political debates.

## 3 DATA

In our design study, the data visualized are transcripts of presidential debates. The transcripts are generated every time a public debate is aired on television. The transcripts used in our work come from the corporations broadcasting debates (e.g., NBC, CNN), media outlets (e.g., The Washington Post, The New York Times), or professional transcription services (e.g., Scribie). The transcripts are simple text files that require processing to extract relevant metadata. We automatically extract three types of data from these transcripts: topic labels, speaker interactions, and audience interactions. Code for

processing and visualizing transcript data, along with all transcripts, are available at <https://osf.io/6jefk/>.

### 3.1 Topic labels

We use a topic dictionary to assign topic labels to each utterance in a debate. The topic dictionary approach requires more manual effort than unsupervised or supervised topic modelling but produces more comprehensible results. Consider the following statement from Senator Bernie Sanders in the tenth 2020 Democratic primary debate:

“What the American people want, by the way, and a lot of the issues we’ll be discussing tonight are issues I raised four years ago: raising the minimum wage to a living wage, 15 bucks an hour. Making public colleges and universities tuition-free. And finally, doing what every other major country on Earth does, guaranteeing health care to all people as a human right through a Medicare for All, single-payer system.”

With the topic dictionary approach, this statement receives “healthcare”, “economy”, and “education” topic labels. With Latent Dirichlet Allocation (LDA), an unsupervised topic modelling algorithm, this statement is assigned less intuitive topic labels, such as 8 (“jobs”, “percent”, “money”), 0 (“going”, “want”, “get”), and 13 (“time”, “one”, “first”). LDA performance can be improved with semi-supervised topics [21] and hierarchical clustering [15]. However, the added benefit of automated methods can be limited on short documents comprised of individual debate utterances [16]. Topic distributions can also vary drastically between debates as the national conversation shifts, making it difficult to accurately learn topic labels from historical transcripts. We look forward to applying semi-supervised and hierarchical automated methods in future work.

### 3.2 Speaker interactions

Debates provide a rare opportunity for candidates to interact face to face: lesser known candidates can attack their more successful opponents to try to draw attention to their campaigns, creating an environment that is exciting to watch because it differs from the “stump speeches” voters are used to hearing. We define a speaker interaction as direct reference, by name, to another candidate in the debate. There are many kinds of speaker interactions that can take place in a debate, but we limit the scope of our speaker interactions to those that involve direct references. Direct reference interactions leave little room for interpretation, making it an ideal attribute to automatically extract from a raw transcript. In Benoit’s Functional Theory of Political Campaign Discourse, he classifies candidate utterances in political debates into three categories: acclaims, attacks, and defenses [5]. We take a similar approach in DEBATEVIS by dividing candidate interactions into one of four categories: agree, attack, defense, or neutral reference.

🗨️ *Agree*: Candidates *agree* when they want to align themselves with other speakers on stage. We consider an *agree* interaction to be a sentence that contains a direct reference and the word “agree” without any contradictory modifiers (e.g., “I do *not* agree”).

🗨️ *Attack*: Candidates *attack* other speakers on stage when they feel they can benefit from pointing out their opponent’s weaknesses. We use Textblob [20], a natural language processing library, to get sentiment scores for each sentence in the transcript. Utterances with a direct reference and sentiment below  $-0.1$  are labelled *attacks*.

🗨️ *Defense*: A *defense* can only occur after an attack, when the target of the prior attack chooses to respond. We consider an utterance to be a *defense* if it immediately follows an attack on the current speaker.

🗨️ *Neutral reference*: Finally, we classify all utterances that contain a direct reference to a fellow candidate but do not fall into one of

the first three categories as *neutral references*. Neutral references are less dramatic than blatant attacks or effusive praise, but by including them we provide another dimension of speaker interactions.

### 3.3 Audience interactions

Audience members voice their opinions of debate performances through applause, laughter, booing, or protesting. Transcribers often include audience interactions as actions occurring between candidate statements, such as “[APPLAUSE]” or “[LAUGHTER]”. By including this information, we can create a visualization that more closely echoes the traditional experience of watching a debate and hearing the crowd’s reaction firsthand.

## 4 DESIGN PROCESS

The final design of the DEBATEVIS interface, shown in Figure 1, was the result of three iterative design phases, including intermittent evaluations and usability tests with non-expert users through application of the Design Study Process Model [28]. All interviews and surveys as part of the design process were approved by our institution’s Internal Review Board.

### 4.1 Phase 1: (October 2018 - May 2019)

Our design process began with a series of informal interviews with faculty in political science and journalism at Northeastern University (Boston, MA, USA), along with a literature review of the presentation of debate transcripts in academic and non-academic settings. Phase 1 produced two key artifacts: an initial task analysis and an initial data abstraction. Our interviews made immediately clear that speakers and topics, and the interactions between the two categories, were the most important data points to be extracted from transcripts. **Artifact 1: Initial task analysis** Our initial task analysis identified three domain goals for non-expert users when exploring debate transcripts. We use Brehmer & Munzner’s visualization task typology [8] to associate each domain goal with an abstract task.

*T1: Overview* (Discover → Explore → Summarize) The *Overview* domain goal reflects users who do not have a specific interest in candidates or topics and want to explore the debate at large. A concrete example of this domain goal is “What did the candidates talk about in the last debate?”

*T2: Speaker Focus* (Discover → Locate → Compare) A user with the *Speaker Focus* domain goal is interested in the behavior of a specific candidate or range of candidates. A concrete example of this domain goal is “Who did Bernie Sanders attack in the first debate?”

*T3: Topic Focus* (Discover → Locate → Summarize) A user with the *Topic Focus* domain goal wants to know what has been said about a specific topic or range of topics, regardless of speaker. A concrete example of this domain goal is “What did the candidates have to say about healthcare in the debate?”

**Artifact 2: Initial data abstraction** Academic and non-academic debate visualization tools share an interest in the topics and issues discussed by candidates in debates. Both classes of visualizations often organize segments of the debate with automatically or manually generated topic labels. VisArgue [13] and NEREx [14] are examples of academic visualization tools with extensive topic modelling capabilities, while debate visualizations from non-academic sources often included manual topic annotations [10, 18, 22]. Our initial data extraction pipeline manually assigned topic annotations for each utterance.

### 4.2 Phase 2: (June 2019 - August 2019)

The second iterative development phase began in conjunction with the 2020 U.S. Democratic primary debates. The first two rounds of debates occurred on June 26-27 and July 30-31. Transcripts were visualized in DEBATEVIS immediately following broadcasts and informal usability testing was performed with five students at our university to assess the design and task analysis created in Phase 1. Although our evaluation had a small sample size and only included

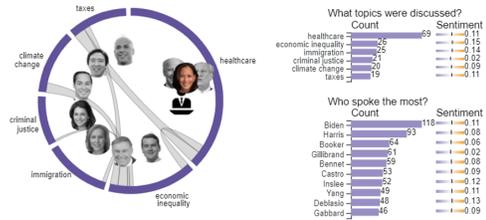


Figure 2: Chord diagram used to show speakers and topics in Phase 2. The chord diagram was difficult for non-expert users to understand and was replaced with a simpler network diagram in Phase 3.

students, our participants were reasonably representative of non-expert users: four out of five participants reported limited to no familiarity with the upcoming U.S. Presidential election and none reported experience with text analysis or text visualization.

**Artifact 1: Updated task analysis** Participants in our testing sample wanted to be able to share interesting moments they found in the debates with other people. As a result, we added a new domain goal to our task analysis:

*T4: Share* (Present) A user with the *Share* domain goal wants to share a specific moment in a debate with other users. A concrete example of this domain goal is “Look at what Elizabeth Warren said in her opening statement in the first debate!”

**Artifact 2: Updated data abstraction** Usability participants wanted to be able to see how the audience reacted to the candidates, so we added a new set of data attributes to our transcript processing pipeline: audience interactions (applause, laughter, booing, etc.).

### 4.3 Phase 3: (September 2019 - February 2020)

To begin Phase 3, we recruited five participants to participate in one-on-one semi-structured interviews about DEBATEVIS. The interviews lasted 30-60 minutes and were conducted by the paper’s lead author. The participants were university students (four graduate, one undergraduate) from varied academic background (one political science, one journalism, three computer science). All five participants were female. None of the participants reported expertise in political debates or text visualization. We chose to use this population for our study because they represented the non-expert, general audience we want to use DEBATEVIS. One of the study participants had seen a brief summary of the DEBATEVIS project, but had not used the tool herself. The rest had no prior exposure to DEBATEVIS.

**Artifact 1: Updated task analysis** Because our evaluation interviews lasted longer than previous usability testing, we were able to observe users as they adjusted to the interface and began asking deeper questions about the content of the debate. This led us to realize that our task analysis was missing one final domain goal, which lay at the intersection of the previously defined *Speaker Focus* and *Topic Focus* goals. Evaluation participants frequently mentioned wanting to know what a specific speaker said about a specific topic, so we added a new domain goal to our task analysis:

*T5: Topic-Speaker Focus* (Discover → Lookup → Identify) A user with the *Topic-Speaker Focus* domain goal wants to learn about the relationship between a given speaker and topic. A concrete example of this domain goal is “What did Beto O’Rourke say about gun control after the shooting in El Paso?”

**Artifact 2: Updated design** Following the Phase 2 evaluation, we refined the DEBATEVIS interface in response to user feedback. While most participants were interested in utterance counts split by topics and speakers, many struggled to interpret the chord diagram (Figure 2). In Phase 3, the chord diagram was replaced with a more straightforward network diagram encoding speaker interactions and topic mentions through edge thickness (Figure 3).

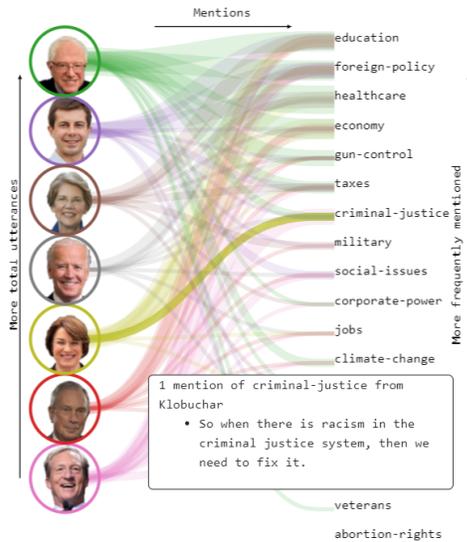


Figure 3: Topic-speaker view of the interactions graph, with hover detail showing Klobuchar's mentions of criminal justice in the 10th Democratic debate.

#### 4.4 Final design

The final DEBATEVIS interface has three components: an interactions graph, an annotated transcript, and a timeline. All three views are connected through brushing and linking and together correspond to the *Topic Focus*, *Speaker Focus*, *Topic-Speaker Focus*, and *Overview* domain goals. Users can share insights found in DEBATEVIS with a custom DEBATEVIS URL preserving their current location within a transcript. This URL can be shared with anyone to show the precise moment identified by the initial user, satisfying the *Share* domain goal. DEBATEVIS is implemented in JavaScript using D3 [7] and can be accessed at <https://web.northeastern.edu/debatevis>.

**Interactions graph** The interactions graph (Figure 1-1) summarizes how often each candidate spoke overall, mentioned other candidates, and discussed each topic. In the interactions graph, a node-link diagram displays speakers and topics as nodes connected by edges representing the number of mentions in the debate. Edges are colored according to the source speaker. The user can choose to view speaker interactions with fellow candidates or topic mentions for each speaker in the interactions graph. We use images of speakers to take advantage of memorability and perceptual benefits of recognizing human faces [3, 6]. Because the user can explore topic distributions and speaker behavior in the interactions graph, this visualization addresses the *Topic Focus*, *Speaker Focus* and *Topic-Speaker Focus* domain goals.

**Annotated transcript** The annotated transcript (Figure 1-2) shows the full text of the debate, annotated with automatically extracted topic labels and speaker interactions. An image of the current speaker is placed next to their last name at the beginning of each utterance to help the user track who is speaking.

**Timeline** The timeline (Figure 1-3) shows a high level overview of the entire debate at a glance. Every utterance occupies a row on the timeline, with marks indicating what audience interactions, speaker interactions, or topic mentions occurred at that moment in the debate. Speaker interactions appear on the timeline as connected pairs of circles, colored according to the source and target speakers. Audience interactions, like applause and laughter, are represented by black horizontal lines in columns labeled with emojis to indicate the type of audience interaction (e.g., applause 👏, laughter 😂).

Topic mentions are represented with horizontal lines placed in the corresponding topic columns, colored by speaker. The primary purpose of the timeline is navigation; the user can click any mark on the timeline to jump through the annotated transcript, rather than scrolling through the transcript line-by-line. Because the user can see broad patterns in topic distribution, audience interaction, and speaker interactions in the timeline, this visualization addresses the *Overview* domain goal.

**Evaluation** In future work, we plan to evaluate the final design of DEBATEVIS with a large audience of non-expert users. Although we conducted an evaluation with students in Phase 2, a broader evaluation with more diverse participants from our target audience will help us understand the impact of changes made during Phase 3.

## 5 DISCUSSION

Visual summaries of political debates have historically been built for two disparate groups: experts and non-experts. Debate visualizations created for experts use state-of-the-art techniques, like brushing and linking and connected views for distant and close reading, while visualizations for non-experts have too frequently been static and explanatory, without the same novel techniques developed by visualization researchers to improve text exploration. The task analysis and abstraction created through our design study characterizes how non-experts engage with debate transcripts, and could help explain differences between how experts and non-experts connect with and use visualizations. For example, we discovered in our first phase of development that many users wanted to share their observations with others via social media, leading us to create the *Share* domain goal. None of the expert-oriented visualizations encountered during our review of prior literature provided sharing functionality.

The data extraction pipeline we developed through iterative discussions with non-expert users also sheds light on the difference between what experts and non-experts look for in a visualization. For example, the non-expert users in our Phase 2 evaluation were fascinated by the audience interactions in our visualization. Two out of the five participants used audience laughter as an exploration mechanism by clicking on the timeline to learn what elicited each bout of laughter. This is an example of a data attribute that might not possess great scientific value, as audience annotations can be irregular and are subject to the judgements of the transcriber, but offers a familiar hook to help non-expert users engage with the visualization. In future visualizations for non-experts, we encourage developers to think outside the box and search for similar connection points to engage their users.

## 6 CONCLUSION

In this paper, we have presented DEBATEVIS, a visualization system developed through an 18-month collaborative design study to promote understanding and engagement with debate transcripts for non-expert users. We hope that this design study and the positive reactions to DEBATEVIS documented in our qualitative evaluation will encourage new debate visualizations with greater exploratory potential. We have shown that the same visualization techniques that have been successfully adopted by digital humanities researchers can be applied to general audiences. By enabling people to better understand the complex and nuanced political debates that influence their lives, we can use visualization to create a more informed public and therefore a more successful democracy.

## ACKNOWLEDGMENTS

The authors wish to thank the Visualization Lab at Northeastern University for their helpful feedback. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1451070. This work was supported by a Tier 1 Interdisciplinary Seed Grant and by the Khoury College of Computer Sciences at Northeastern University.

## REFERENCES

- [1] Transcript of the second debate. *The New York Times*, Oct 2016.
- [2] C. Alcantara, K. Rabinowitz, and A. Steckelberg. All the interruptions, alignments, and attacks of the democratic debates. *Washington Post*, Jun 2019.
- [3] W. A. Bainbridge, P. Isola, and A. Oliva. The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, 142(4):1323, 2013.
- [4] N. G. Belmonte. Extracting and visualizing insights from real-time conversations around public presentations. In *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 225–226. IEEE, 2014.
- [5] W. L. Benoit. *Political Election Debates: Informing Voters about Policy and Character*. Lexington Books, 2013.
- [6] M. A. Borkin, Z. Bylinskii, N. W. Kim, C. M. Bainbridge, C. S. Yeh, D. Borkin, H. Pfister, and A. Oliva. Beyond memorability: Visualization recognition and recall. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):519–528, 2015.
- [7] M. Bostock, V. Ogievetsky, and J. Heer. D<sup>3</sup> data-driven documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, 2011.
- [8] M. Brehmer and T. Munzner. A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, 2013.
- [9] W. Cai, J. Kao, J. C. Lee, A. Parlapiano, and J. K. Patel. Which candidates and topics got the most time during the democratic debate. *The New York Times*, Jun 2019.
- [10] W. Cai, L. Leatherby, and J. C. Lee. Which candidates got the most speaking time in the democratic debate. Nov. 2019.
- [11] A. Choi and E. Doherty. What went down on night two of the first democratic debates. *FiveThirtyEight*, Jun 2019.
- [12] M. Correll. Ethical dimensions of visualization research. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2019.
- [13] M. El-Assady, V. Gold, A. Hautli-Janisz, W. Jentner, M. Butt, K. Holzinger, and D. A. Keim. Visargue: A visual text analytics framework for the study of deliberative communication. In *PolText 2016-The International Conference on the Advances in Computational Analysis of Political Text*, pp. 31–36, 2016.
- [14] M. El-Assady, R. Sevastjanova, B. Gipp, D. Keim, and C. Collins. Nerex: Named-entity relationship exploration in multi-party conversations. In *Computer Graphics Forum*, vol. 36, pp. 213–225. Wiley Online Library, 2017.
- [15] T. L. Griffiths, M. I. Jordan, J. B. Tenenbaum, and D. M. Blei. Hierarchical topic models and the nested chinese restaurant process. In *Advances in neural information processing systems*, pp. 17–24, 2004.
- [16] L. Hong and B. D. Davison. Empirical study of topic modeling in twitter. In *Proceedings of the first workshop on social media analytics*, pp. 80–88, 2010.
- [17] E. Hoque and G. Carenini. Multiconvis: A visual text analytics system for exploring a collection of online conversations. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, pp. 96–107. ACM, 2016.
- [18] J. Kimelman. Full transcript: 2019 democratic debate night one, sortable by topic. *NBC News*, June 2019.
- [19] L. Leatherby, M. Rojanasakul, P. Murray, R. Best, Y. Qiu, D. Pogkas, A. Tartar, J. Gu, and A. McCartney. The democratic candidates who spoke the most at the debate — and what they said. *Bloomberg*, Jun 2019.
- [20] S. Loria, P. Keen, M. Honnibal, R. Yankovsky, D. Karesh, and E. Dempsey. Textblob: Simplified text processing, 2013.
- [21] J. D. McAuliffe and D. M. Blei. Supervised topic models. In *Advances in neural information processing systems*, pp. 121–128, 2008.
- [22] A. McCartney, J. Gu, C. Whiteaker, and M. Rojanasakul. Everything the candidates discussed at the atlanta debate. *Bloomberg*, Nov. 2019.
- [23] T. Parti. The moments that mattered in wednesday’s debate. *The Wall Street Journal*, Aug 2019.
- [24] E. M. Peck, S. E. Ayuso, and O. El-Etr. Data is personal: Attitudes and perceptions of data visualization in rural pennsylvania. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2019.
- [25] Politico. Trump’s state of the union address. *Politico*, Feb 2019.
- [26] Z. Pousman, J. Stasko, and M. Mateas. Casual information visualization: Depictions of data in everyday life. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1145–1152, 2007.
- [27] A. Rappeport, M. Haberman, A. Nagourney, and N. Confessore. Trump and Clinton’s second debate: Analysis. *The New York Times*, Oct 2016.
- [28] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization & Computer Graphics*, (12):2431–2440, 2012.
- [29] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, 2010.