Semantic Navigator: Query Driven Active Learning for Historical Narrative Understanding

EVA MAXFIELD BROWN, University of Washington Information School
MADELEINE GRUNDE-MCLAUGHLIN, University of Washington Computer Science & Engineering
ISABELLA PESTOVSKI, University of Washington Mechanical Engineering
LANYI ZHU, University of Washington Political Science
NICHOLAS WEBER, University of Washington Information School

Open legislative debate is a cornerstone of contemporary democracy. At the local level, city and county council meetings are where lawmakers and the general public engage most directly in legislative debate. Digital archives of these events enable access and transparency - they allow journalists and community members to meaningfully understand discussion around various policy issues over time. However, these digital archives scale exponentially - making it infeasible, for example, to review tens to hundreds of hours of footage to understand the context of one particular issue. To help overcome issues of scale, including the search and discovery of topics in council meetings, the Council Data Project (CDP) stores videos and generates transcripts to enable keyword-based search against the timestamped transcripts of meeting discussion. Searching by keyword, however, has limited power when users wish to refine searches by a conceptual and semantic meaning. We present Semantic Navigator, an interactive Active Learning approach in which users can better refine their search query to complex topics. Semantic Navigator is implemented as a web application to support an iterative refinement process in which the search model surfaces potentially relevant parts of meetings, and the user marks each text as relevant or irrelevant to further direct the search model. To test the utility of this search refinement process, we conduct a prototype evaluation with five participants. While preliminary, we find that participants considered the semantic searching and refinement process highly useful, and have provided multiple suggestions of interface improvement to reduce the learning curve of successfully interacting with our system.

CCS Concepts: • **Information systems → Search interfaces**; *Probabilistic retrieval models*; • **Applied computing → Document searching**; *E-government*; *Digital libraries and archives.*

Additional Key Words and Phrases: active learning, archival search, journalism technology, search, prototype evaluation

ACM Reference Format:
Eva Maxfield Brown, Madeleine Grunde-McLaughlin, Isabella Pestovski, Lanyi Zhu, and Nicholas Weber. 2023. Semantic Navigator: Query Driven Active Learning for Historical Narrative Understanding. 1, 1 (September 2023), 10 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Authors' addresses: Eva Maxfield Brown, University of Washington Information School, evamxb@uw.edu; Madeleine Grunde-McLaughlin, University of Washington Computer Science & Engineering; Isabella Pestovski, University of Washington Mechanical Engineering; Lanyi Zhu, University of Washington Political Science; Nicholas Weber, University of Washington Information School.

XXXX-XXXX/2023/9-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

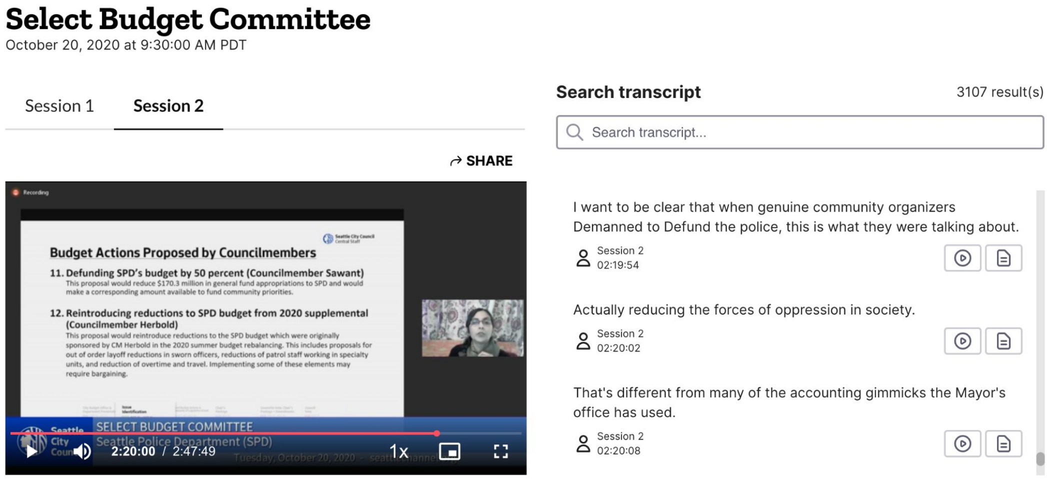

Fig. 1. Our approach applies Active Learning to sort through data from the Council Data Project [2], the site of which is shown above. This data includes chunks of text from Seattle city council meeting transcripts, as well as a link to the Council Data Project site to explore further. As there are thousands of hours of content in the archive, our search method helps people better surface sections of the transcripts relevant to their search query.

1 INTRODUCTION

Council Data Project (CDP) is an organization which produces software to gather, process, archive, and republish municipal legislative meeting information. In total, CDP has archived more than 5,000 municipal legislative meetings (city councils, county councils, state legislatures, etc.). Typically, these archives are used by journalists, researchers, and members of the public to search for specific meetings which are relevant to a specific legislative policy or discussion item. Recently however, journalists and members of the public have asked for a method for viewing the "narrative arc" of a policy issue across time within a municipal council – the ability to search across meetings. Finding, tracking, and researching these narrative arcs is currently quite a time consuming and difficult process. Furthermore, defining the policy issue with a user's intended definition and specificity can be difficult through keyword search alone.

The problem may be best described through an example: after the murder of George Floyd, many city councils across the United States pledged to activists and protesters that they would defund their police departments. In many cases these were empty promises with little to no action ever being taken. In a few cases, city councils across the United States took minor actions towards defunding police departments by moving their budgets to other departments and services [8]. While CDP provides the data and tools needed to begin to construct and understand these policy pushes, they currently do not have a tool to enable search across events focusing on a single topic. Members of the public already have a difficult time following the news and actions of city councils, and tracking their actions over extended periods of time is even more difficult.

Understanding the beginning, middle, and end, of a push for reform or policy in this narrative arc form, may help residents understand and become more engaged in the process of municipal legislation. In order to gain this narrative arc, residents must be able to filter and search many hundreds of hours of meeting transcripts for a specific, and potentially novel, definition of that reform or policy.

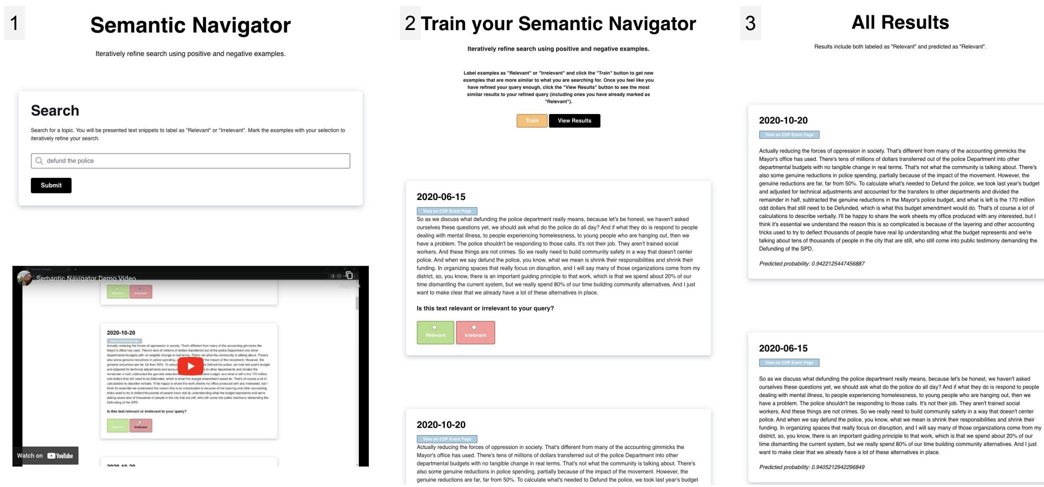

Fig. 2. Our tool for searching for complex concepts involves three steps. **Step 1:** Users enter a search query using natural language. A demo video is available if needed. **Step 2:** Users refine their search through multiple iterations in which they label text as relevant or irrelevant, then press “Train.” When they are happy with the search results, they click “View Results.” **Step 3:** Users can interact with and view the results of their search query, as well as the probability that the model would classify the chunk as "relevant" to their search.

2 RELATED WORK

2.1 Active Learning and Annotation

There is a small but growing area of research for quick, iterative, fine-tuning of Machine Learning (ML) models through user data annotation [4, 9, 14]. Active learning approaches use iterative labeling to train a model from scratch by selecting the most important data points to label [15]. While such systems are typically used to assist in annotating and training a gold-standard model, we instead use an active learning approach in a search capacity to allow users to make many models for their own individual search objectives.

Most relevant to our work is Lee’s "Newspaper Navigator" which takes advantage of a digitized dataset to allow researchers to conduct a more meaningful search through historical newspaper images [10]. Lee created an AI navigator that returns images from within historical newspapers for a given search query and allows the user to annotate results. The user interacts with these results by up-voting some as relevant and down-voting others as irrelevant. After building a small collection, the user can then train their AI navigator to search for other examples. The user’s choice affects which AI-selected results should be used as positive and negative examples in the next round of the training process. We apply this iterative search and refinement process to long-form text rather than visual data.

2.2 Computational Tools for Journalists

Algorithms assisting journalists in monitoring, tracking, and even partially generating stories has become a major part for newsrooms large and small [5, 6]. Diakopoulos et al. built a prototype "news discovery tool" called Algorithm Tips, which was designed to help journalists find newsworthy leads about algorithmic decision-making systems used by various levels of the US government [5]. These types of "news detection software" are common. The Big Local News project from Stanford University has been building the Agenda Watch platform to assist city hall and local government reporters stay up-to-date as legislative documents and meeting agendas are released. Agenda Watch collects and indexes hundreds of thousands of government documents to then notify journalists of events, budget items, and legislative actions that may be considered newsworthy.

Automatic tracking and notification of newsworthy events aren't the only approach that technologists have brought to journalists. Franks et al., explored the INJECT system which aimed to "support the creative capabilities of time and resource poor journalists" by helping them find new angles on stories [6]. Their work utilized a combination of natural language processing approaches to extract topical information and named entities from articles to perform automatic data linkage between entities found across multiple articles. In doing so, their system aims to provide a possible overview of different aspects and narrative angles prescribed to previously written about figures.

While automatic tracking and notifications are built for "beat reporters", there is additional technology designed to assist investigative reporters. Brehmer et al. introduced "Overview," a system for visual document mining [1]. It included visualization methods for analyzing large document collections using document clustering and tagging. They similarly approached large document exploration through these means because of the understanding that standard keyword-based text search is useful, but there are plenty of situations in investigative journalism where it may not be possible to formulate a keyword based query. Their work on "Overview" was continued by Stray in 2016 with an extensive ethnography of journalists using their system in practice [16]. They recorded a series of case studies which ultimately led them to six practices for natural language processing researchers to follow in creating software to assist journalists: robust import of documents, robust analysis methods, search (not exploration), quantitative summaries, interactive methods, and clarity and accuracy in results.

Our work is most similar to the tools built for investigative journalism which enable search and discovery of information through a large number of documents. Prior work typically has utilized keyword based search for finding documents, and while Brehmer et al. explored methods closer to semantic search, they approached the problem through visualization. Visualization can be incredibly useful, however, visualization of semantic embeddings typically involves reducing the dimensions of the embeddings which ultimately may reduce the users ability to differentiate documents effectively. Our active learning approach additionally follows the guidance from Stray's follow up study in which they stated that interactive methods are incredibly useful for allowing the user to guide the session themselves.

2.3 Archival Search and Discovery

Similar to Brehmer et al. work on "Overview" as an analysis and visualization system for assisting investigative journalists search through large document collections, ClioQuery from Handler et al., was similar but tailored more towards archival use cases [7]. ClioQuery incorporated both keyword-based search and additional analysis-during search techniques, such as auto-summarization of documents and analyzing and visualizing word clusters. Further, as many archival datasets are from historical archives, ClioQuery had a built-in a timeline plot for visualizing Ngram usage over time using the documents dates similar to that of Google's Ngram Viewer [11]. Demonstrated through the results of their user study, their work was an improvement in archival search against a baseline keyword-based search alone. Further, they provide guidance to future work by highlighting that what was particularly successful was that their work enabled interactive corpus investigation organized around the analysis of query mentions in context.

Lee et al. must be discussed as well; Newspaper Navigator builds off of these interactive best practices and discussions of the drawbacks of keyword only search methods by demonstrating how active learning can be used for iterative refinement of a query [10]. Newspaper Navigator demonstrated our active learning approach for large image collections and we are studying if such successes hold for text corpora. Specifically we question if users will spend more time annotating each item in the dataset and potentially feel that the active learning process is too time consuming to be considered useful.

3 METHODS

3.1 Data Collection and Structure

For testing our application and use case, we utilize the data from Brown and Weber's Councils in Action dataset [3]. Specifically, we process their transcripts of Seattle City Council meetings for one-hundred randomly sampled meetings between 2020 and 2023 collected and generated by Council Data Project [2]. As shown in Figure 1, these meetings are typically long, with meetings commonly over two hours in duration. To enable search for fine-grained queries, we process each transcript into multiple chunks of text, with each portion being cut off at the nearest sentence past the 1,024 character threshold. We then store the CDP provided metadata from which each text portion was created and additionally generate a unique identifier for the text portion itself for storage. Finally, we generate a semantic embedding for each chunk of text using the sentence transformers library and the "all-MiniLM-L12-v2" model which produces a 384 dimension semantic embedding [12, 13, 17]. Once we have created the semantic embedding for each text portion, we store the text to a unique file on Google Cloud Storage. The completed dataset then contains, the metadata for the portion of text, including the unique identifier we created, the semantic embedding, and the path for remotely reading the text file from cloud storage.

By pre-computing the semantic embeddings for each example and storing each text portion to it's own file, we are able to create a dataset which can be kept in memory on the server while containing all of the metadata and embedding data needed for querying and training, while leaving the full text of each snippet to be loaded as needed.

Once all processing is complete, our small test dataset contained 8,282 unique portions of text from the randomly sampled one-hundred Seattle City Council meetings from between 2020 and 2023.

We provide a few examples of the produced text portions below:

Councilmember Kshama Sawant speaking during a March 15, 2021 Full Council Meeting: Timestamped Link

> And we've had a lot of discussion. And we've had a lot of evidence about how. In cities that have right to counsel, especially in cities that have strong right to counsel. It has actually had a positive impact, not only on how it prevents eviction. But also on how addiction filings themselves have gone down. Because landlords now know that they face not just a vulnerable tenant. But also a vulnerable tenant. And so they are able to connect the tenant to. Other sources of help like renter assistance. But as we know, renter assistance, rental assistance is not. Cannot work by itself. It's a rental assistance and eviction defense are compliments to each other. Preventing the eviction in court legally is a necessary component of the eviction defense. And it's not just, we know it's women and especially specifically black women who are deeply affected by evictions. And we see from New York city, for example, that 86% of tenants who are represented by council. Are remaining in their home. This is just staggering evidence of what. Good.

Public Commenters speaking during an October 30, 2020 Select Budget Committee Meeting: Timestamped Link

> There's a lot of crossing points. A number of them are not great. You have a good experience down in the South and really solid experience up North. 408Th and my 50th aren't great experiences. Light rail going in, it will be the most central connection in the U-District. Both of those are not great. We need a connection to come light rail. Thank you very much. Next person is Valerie. Good morning. Good morning. This is Valerie from district 2. agenda item 28 is $65,000 to the OIG for a sentinel review of police interaction during the customer's George Floyd is demonstrations. I listened to a virtual community listening form hosted by the Oig in off. It was a dramatic example of everything that's wrong with the police accountability system in Seattle. The event was not publicized. A few protesters called in to provide evidence of violent and Traumatizing police behavior at demonstrations over the summer. The event continued and appeared to be a set up. Once again our accountability system provided illusion of an official response with no real consequences for abusive policing.

3.2 Interactive Querying Method

The interactive learning process is shown in Figure 3. A user can train their own Semantic Navigator for each topic they search for. The process begins when the user submits their search query. Using the same embedding model which created semantic embeddings for the dataset ("all-MiniLM-L12-v2"), we get an embedding the user's query. The query embedding is compared to the transcript portions embeddings using cosine similarity and the twenty most similar chunks are returned and displayed on the annotation page as a form.

The user annotates each portion as "Relevant" or "Irrelevant" using on screen buttons. When the user is done annotating and wants to re-train the model, the dataset is updated to reflect the user's choices. First, the form data is sorted into two lists containing positive (relevant) and negative (irrelevant) examples. To improve the robustness of the model, we utilize negative sampling and label four hundred additional randomly sampled embeddings as negative examples. If the user does not pick any irrelevant examples, then these random embeddings are used solely as the negative examples, otherwise they are concatenated with the user-selected negative examples. The positive and negative examples, and the matching embeddings, are then used to train a logistic regression classification model. After removing the previously annotated examples from the CDP dataset, the classifier is used calculate the probabilities of the remaining transcript portions being classified as "relevant." The twenty transcript portions with the highest probability are returned and displayed for the user to annotate for the next iteration in the active learning cycle.

This iterative training and annotation process can continue as long as the user wants. The user can at any time, choose instead to view their results. In doing so, the model is retrained a final time according to the process described above. The top fifty results with the highest probability of being classified as "relevant" are returned and displayed on the results page.

4 PROTOTYPE EVALUATION

We conducted a prototype evaluation to assess the effectiveness of the Semantic Navigator in assisting users and obtained their feedback. We selected five participants, all of whom possess extensive knowledge of city council meetings and are our target users. In order to carry out the usability testing, we initially introduced the project to them and encouraged them to provide feedback and comments while thinking aloud. Subsequently, we posed several initial questions, including their current role and the frequency with which they engage with city council meetings

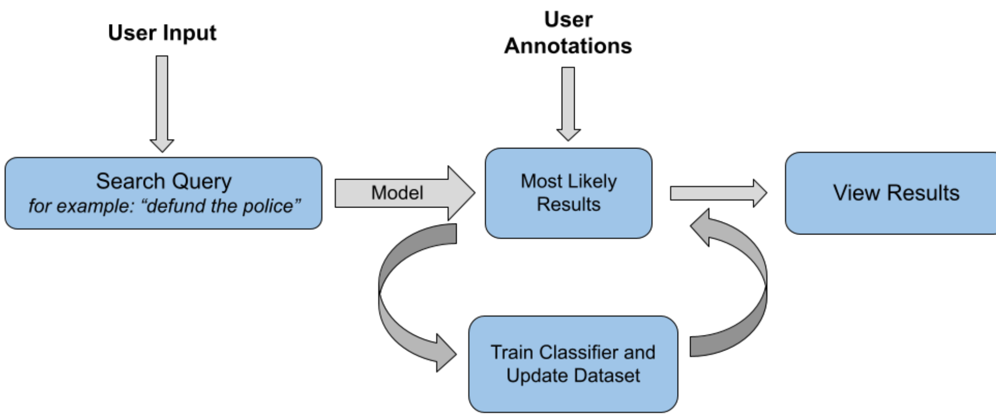

Fig. 3. This flowchart demonstrates the steps taken in the interactive training process. Note that the middle section can be repeated as often as the user deems necessary and that their final results will incorporate their previously annotated examples as well as the next most likely examples.

and government documents, in order to evaluate their existing knowledge of City Council meetings. The specifics of their roles and exposure to City Council meetings are presented in Table 1.

Description	**Exposure to City Council Meetings**
City Hall Reporter	High, review multiple meetings per week
PM for News Tech	High, prev. reporting job and now tool dev.
Labor Researcher	Low, only follows specific policies
Community Member	Medium, reviews a few meetings per week
Civic Tech Dev	Low, engages rarely to catch up on events

Table 1. User descriptions and exposure to City Council meetings

Afterwards, we provided all participants with a specific task to evaluate the usability and functionality of the Semantic Navigator application. We instructed them to envision themselves as individuals preparing to write a report, or more generally research, Seattle police reform process. We instructed them to find portions of transcripts regarding "defunding the police." With this standardized scenario, we requested participants to demonstrate where and how they would initiate their search, and we carefully observed their search behaviors.

Upon completion of their search and training with the models, we invited users to assess the final results they obtained and gauge the usefulness of the Semantic Navigator. This was done through a Likert scale ranging from 1 to 5, with 1 representing "not useful", and 5 representing "incredibly useful". Additionally, we monitored the users' time spent during the search process and asked for their perception on whether the process was longer or shorter in duration than expected. Furthermore, we encouraged users to share any suggestions they had, such as desired additional functions, and provide feedback on the overall process.

5 USER STUDY RESULTS

Through the prototype evaluation, we gained five main insights. First, most users were initially unfamiliar with the process of training a search and recommendation model themselves. Some participants overlooked the instructions on the page and became confused about the meaning of "relevant" and "irrelevant." Initially using the annotation buttons to inform the model that the result was "generally relevant" for the overall query rather than using those buttons to narrow and guide the model towards something more specific. However, once they understood how the application could be used, they found it to be highly beneficial for their search and result refinement.

Second, two out of the five users expressed uncertainty about the number of examples they needed to annotate before training or viewing the results. This feedback is valuable, and we are considering implementing a feature that provides users with information about the model's current performance within the web application.

Third, two out of the five users suggested adding a function to sort and filters results by date and highlight or bold their initial search query in the transcript portion itself. We also recognized the merit in this suggestion and have plans to incorporate it in future versions.

Fourth, all users spent approximately ten minutes on the search and model training. Four out of the five users reported that their time spent searching met their expectations or was even shorter than expected, while one user felt it was slightly longer due to not reading the instructions and struggling with understanding how to use the application effectively. These findings provide strong reassurance that the user journey is time-efficient, and the model is genuinely effective for users.

Lastly, when we asked the users to rate the the Semantic Navigator application on a 1-5 Likert scale for the usefulness of the application in searching across time and meetings, the average rating was 4.2. We feel that this shows the Semantic Navigator is helpful for the users.

6 DISCUSSION

In its current state, Semantic Navigator functions as a proof-of-concept for a human-in-the-loop, active learning approach to help researchers and reporters search through thousands of city council meetings. The results from our prototype evaluation tell us that this type of search utilizing human-machine-learning collaboration is relatively unfamiliar to our target users and represents an untapped space for possible productivity tools.

Due to this unfamiliarity with the human-in-the-loop training process, it was unclear to several of our test users how many relevant/irrelevant examples would be needed to successfully train the underlying model. This prompted a suggestion to include features that would allow a user to visualize how their Semantic Navigator is improving as they annotate. This will help users conduct a more efficient search, but will also give them more experience and understanding of using machine learning algorithms by grounding it in a real-life context. The user feedback on the speed of searching and training confirms that a meaningful semantic search can be conducted using transcript text portions and their matching embeddings. While Semantic Navigator is lacking in some useful features for more thorough searching and narrative arc construction, this first iteration accomplishes the main task of creating a successful training algorithm and presenting the results in a user-friendly way.

Overall, the feedback and suggestions from our test users is encouraging and will motivate the future versions of Semantic Navigator.

7 LIMITATIONS

We want to emphasize again that much of this work is a prototype. With only five users in our study to evaluate our application, all of whom were generally positive about the Semantic Navigator application, we feel that there is serious promise in our approach, but more research is needed to have a greater understanding of the potential pros and cons of this type of query pattern.

Further, many of our user study participants are potentially biased as all of the participants of our study have previously discussed artificial intelligence and machine learning for local journalism.

Future studies should resolve this problem by recruiting a larger, unacquainted, participant group to test not only the application but how the application supports various types of users, from journalists to activists to researchers and more.

8 FUTURE WORK

There are many routes to continue on with this work. First, recalling from the literature how tools for journalists tend to support either beat-reporters or investigative journalists, we believe that there is a large middle ground to explore. While our initial implementation and feature set tended more towards the investigative reporter and archival use case, including a notifications feature that enables users to be alerted when newly available items are above a threshold for classification probability could be very useful for beat-reporters who have previously trained a model for their specific beat.

Continuing to build in the direction of tooling for investigative reporters and archival use cases, we can additionally add some of the "analysis and visualization in-context" content that many search tools have incorporated previously. For example, including a basic timeline visualizations or adding additional dimensions to such a timeline which include processed summaries of the selected content, sentiment analysis, and outlier detection, over time.

9 CONCLUSION

To the growing area of iterative, user-refined machine learning models, we contribute Semantic Navigator, a tool for creating refined and personalized searches through city council meeting data. Underlying our custom interface is an active learning algorithm that is updated through multiple rounds of user annotations. To test the utility of our approach, we have conducted a prototype evaluation with five users. We find that this search approach, though initially unintuitive, can help users across a variety of disciplines better understand their government's discussions around particular issues of interest. To make content from the council meetings even more digestible, we hope to apply this method in future work to automatically create timelines of the discussions surfaced by the search model.

REFERENCES

[1] Matthew Brehmer, Stephen Ingram, Jonathan Stray, and Tamara Munzner. 2014. Overview: The Design, Adoption, and Analysis of a Visual Document Mining Tool for Investigative Journalists. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2271–2280. https://doi.org/10.1109/TVCG.2014.2346431

[2] Eva Maxfield Brown, To Huynh, Isaac Na, Brian Ledbetter, Hawk Ticehurst, Sarah Liu, Emily Gilles, Katlyn M. f. Greene, Sung Cho, Shak Ragoler, and Nicholas Weber. 2021. Council Data Project: Software for Municipal Data Collection, Analysis, and Publication. *Journal of Open Source Software* 6, 68 (2021), 3904. https://doi.org/10.21105/joss.03904

[3] Eva Maxfield Brown and Nicholas M. Weber. 2022. Councils in Action: Automating the Curation of Municipal Governance Data for Research. *Proceedings of the Association for Information Science and Technology* 59 (2022).

[4] Simon Bruns, André Calero Valdez, Christoph Greven, Martina Ziefle, and Ulrik Schroeder. 2015. What should i read next? a personalized visual publication recommender system. In *Human Interface and the Management of Information. Information and Knowledge in Context: 17th International Conference, HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part II.* Springer, 89–100.

[5] Nicholas A. Diakopoulos, Daniel Trielli, and Grace Lee. 2021. Towards Understanding and Supporting Journalistic Practices Using Semi-Automated News Discovery Tools. *Proceedings of the ACM on Human-Computer Interaction* 5 (2021), 1 – 30.

[6] Suzanne Franks, Rebecca Wells, Neil A. M. Maiden, and Konstantinos Zachos. 2021. Using computational tools to support journalists' creativity. *Journalism* 23 (2021), 1881 – 1899.

[7] Abram Handler, Narges Mahyar, and Brendan T. O'Connor. 2022. ClioQuery: Interactive Query-oriented Text Analytics for Comprehensive Investigation of Historical News Archives. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 12 (2022), 1 – 49.

[8] Bai Linh Hoang and Andrea Benjamin. 2023. "Defund" or "Refund" the Police?: City Council Responsiveness to the Black Lives Matter Protests. *Urban Affairs Review* (2023). https://api.semanticscholar.org/CorpusID:260015888

[9] Benjamin Charles Germain Lee, Doug Downey, Kyle Lo, and Daniel S Weld. 2020. LIMEADE: From AI Explanations to Advice Taking. *ACM Transactions on Interactive Intelligent Systems* (2020).

[10] Benjamin Charles Germain Lee, Jaime Mears, Eileen Jakeway, Meghan Ferriter, Chris Adams, Nathan Yarasavage, Deborah Thomas, Kate Zwaard, and Daniel S Weld. 2020. The newspaper navigator dataset: extracting and analyzing visual content from 16 million historic newspaper pages in chronicling America. *Conference on Information and Knowledge Management* (2020).

[11] Jean-Baptiste Michel, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak, and Erez Lieberman Aiden. 2011. Quantitative Analysis of Culture Using Millions of Digitized Books. *Science* 331, 6014 (2011), 176–182. https://doi.org/10.1126/science.1199644 arXiv:https://www.science.org/doi/pdf/10.1126/science.1199644

[12] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. https://arxiv.org/abs/1908.10084

[13] Nils Reimers and Iryna Gurevych. 2020. Making Monolingual Sentence Embeddings Multilingual using Knowledge Distillation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. https://arxiv.org/abs/2004.09813

[14] Patrick Schramowski, Wolfgang Stammer, Stefano Teso, Anna Brugger, Franziska Herbert, Xiaoting Shao, Hans-Georg Luigs, Anne-Katrin Mahlein, and Kristian Kersting. 2020. Making deep neural networks right for the right scientific reasons by interacting with their explanations. *Nature Machine Intelligence* 2, 8 (2020), 476–486.

[15] Burr Settles. 2009. Active learning literature survey. (2009).

[16] Jonathan Stray. 2016. What do Journalists do with Documents ? Field Notes for Natural Language Processing Researchers.

[17] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. 2020. Transformers: State-of-the-Art Natural Language Processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Association for Computational Linguistics, Online, 38–45. https://www.aclweb.org/anthology/2020.emnlp-demos.6