Bridging Dictionary: AI-Generated Dictionary of Partisan Language Use

Hang Jiang*
hjian42@mit.edu
Massachusetts Institute of Technology
Cambridge, MA, USA

Doug Beeferman
dougb5@mit.edu
Massachusetts Institute of Technology
Cambridge, MA, USA

William Brannon
wbrannon@mit.edu
Massachusetts Institute of Technology
Cambridge, MA, USA

Andrew Heyward
aheyward@media.mit.edu
Massachusetts Institute of Technology
Cambridge, MA, USA

Deb Roy
dkroy@mit.edu
Massachusetts Institute of Technology
Cambridge, MA, USA

Abstract

Words often carry different meanings for people from diverse backgrounds. Today's era of social polarization demands that we choose words carefully to prevent miscommunication, especially in political communication and journalism. To address this issue, we introduce the Bridging Dictionary, an interactive tool designed to illuminate how words are perceived by people with different political views. The Bridging Dictionary includes a static, printable document featuring 796 terms with summaries generated by a large language model. These summaries highlight how the terms are used distinctively by Republicans and Democrats. Additionally, the Bridging Dictionary offers an interactive interface that lets users explore selected words, visualizing their frequency, sentiment, summaries, and examples across political divides. We present a use case for journalists and emphasize the importance of human agency and trust in further enhancing this tool. The deployed version of Bridging Dictionary is available at https://dictionary.ccc-mit.org/.

CCS Concepts

• Human-centered computing → Empirical studies in HCI; **Web-based interaction**; **Visualization toolkits**; **• Computing methodologies → Natural language processing**.

Keywords

natural language processing, social polarization, computational journalism, text analysis, civic media, trust and human agency

ACM Reference Format:
Hang Jiang, Doug Beeferman, William Brannon, Andrew Heyward, and Deb Roy. 2024. Bridging Dictionary: AI-Generated Dictionary of Partisan Language Use. In *Companion of the 2024 Computer-Supported Cooperative Work and Social Computing (CSCW Companion '24), November 9–13, 2024, San Jose, Costa Rica.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3678884.3681820

*HJ and DB contributed equally to this work. HJ led the evaluation and writing and DB led the development.

CSCW Companion '24, November 9–13, 2024, San Jose, Costa Rica

ACM ISBN 979-8-4007-1114-5/24/11
https://doi.org/10.1145/3678884.3681820

1 Introduction

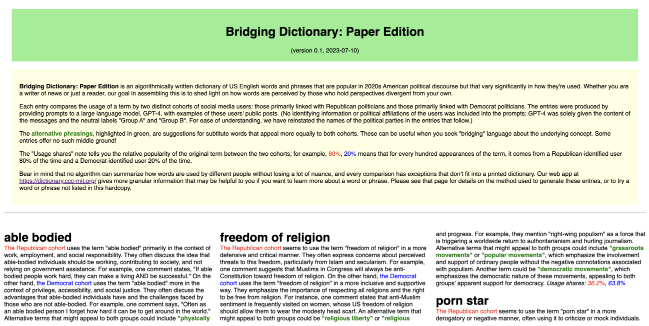
Polarization is a significant feature of the political landscape in the United States [8, 15, 19]. Previous research has shown that Republicans and Democrats often use and interpret words differently, even when speaking the same language [12, 20]. This linguistic divide poses considerable challenges to public discourse, particularly in journalism, where the use of words without an understanding of their varying connotations across political communities can have serious consequences. Journalists face the added difficulty of manually reading and editing news content, a process that is not only time-consuming but also prone to errors due to the constantly evolving connotations of words online. To address this problem, we introduce the Bridging Dictionary (BD), a tool designed to automatically identify controversial terms across the political divides and to summarize their different usages. We provide not only a useful resource for the academic community, journalists, and wider audiences but also highlight the importance of considering human agency and trust in developing human-AI systems.

2 Related Work

Polarized language use in NLP. Researchers in political science and Natural Language Processing (NLP) have found that there is a partisan difference in language understanding [12]. R. KhudaBukhsh et al. [20] used modern machine-translation methods to show that the Republican and Democratic communities use English words differently. For instance, there are different connotations when partisans use "undocumented workers" or "illegal aliens" to discuss the same group of people. To address this issue, Webson et al. [22] developed an NLP method to mitigate the political bias of text representations, showing that it improves the viewpoint diversity of document rankings. Recently, NLP researchers have also proposed novel methods to quantify and debias the political bias of language models [13, 14]. However, previous work focuses on measuring and debiasing NLP models instead of facilitating humans in writing less biased content. Our work fills the gap by leveraging NLP to help humans understand the language bias and facilitate them in writing and editing through an interactive tool.

NLP for qualitative analysis and sensemaking. NLP has been used to develop computational tools for qualitative analysis and sensemaking [3, 7]. Among these tools, topic models are particularly prevalent for text analysis across various domains [9, 10, 24]. However, traditional NLP models often lack world knowledge, resulting in limited insights and interpretability [1]. The modern large language models (LLMs) have enabled users to interact with unstructured data through simple queries to extract more nuanced and interpretable insights. Recent studies have explored the application of LLMs to automatically extract and summarize valuable information from texts [2, 5, 18, 23]. Despite these advancements, there is a lack of research focusing on how LLMs can assist journalists by providing qualitative insights for writing and editing. This study aims to bridge this gap by introducing an interactive tool to summarize the varying usage of terms across political divides, thereby guiding journalists in their word choices in news writing.

3 System Overview



Bridging Dictionary: Paper Edition

(version 0.1, 2023-07-10)

Bridging Dictionary: Paper Edition is an algorithmically written dictionary of US English words and phrases that are popular in 2020s American political discourse but that vary significantly in how they're used. Whether you are a writer of news or just a reader, our goal in assembling this is to shed light on how words are perceived by those who hold perspectives divergent from your own.

Each entry compares the usage of a term by two distinct cohorts of social media users: those primarily linked with Republican politicians and those primarily linked with Democrat politicians. The entries were produced by providing prompts to a large language model, GPT-4, with examples of these users' public posts. (No identifying information or political affiliations of the users was included into the prompts; GPT-4 was solely given the content of the messages and the neutral labels "Group A" and "Group B". For ease of understanding, we have reinstated the names of the political parties in the entries that follow.)

The **alternative phrasings**, highlighted in green, are suggestions for subtitute words that appeal more equally to both cohorts. These can be useful when you seek "bridging" language about the underlying concept. Some entries offer no such middle ground!

The "Usage shares" note tells you the relative popularity of the original term between the two cohorts; for example, 80%, 20% means that for every hundred appearances of the term, it comes from a Republican-identified user 80% of the time and a Democrat-identified user 20% of the time.

Bear in mind that no algorithm can summarize how words are used by different people without losing a lot of nuance, and every comparison has exceptions that don't fit into a printed dictionary. Our web app at https://dictionary.ccc-mit.org/ gives more granular information that may be helpful to you if you want to learn more about a word or phrase. Please see that page for details on the method used to generate these entries, or to try a word or phrase not listed in this hardcopy.

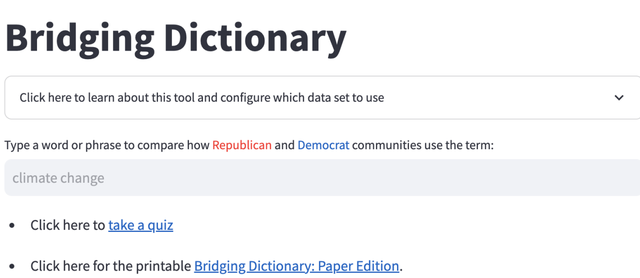
able bodied
The Republican cohort uses the term "able bodied" primarily in the context of work, employment, and social responsibility. They often discuss the idea that able-bodied individuals should be working, contributing to society, and not relying on government assistance. For example, one comment states, "If able bodied people work hard, they can make a living AND be successful." On the other hand, the Democrat cohort uses the term "able bodied" more in the context of privilege, accessibility, and social justice. They often discuss the advantages that able-bodied individuals have and the challenges faced by those who are not able-bodied. For example, one comment says, "Often as an able bodied person I forget how hard it can be to get around in the world." Alternative terms that might appeal to both groups could include **"physically**

freedom of religion
The Republican cohort seems to use the term "freedom of religion" in a more defensive and critical manner. They often express concerns about perceived threats to this freedom, particularly from Islam and secularism. For example, one comment suggests that Muslims in Congress will always be anti-Constitution toward freedom of religion. On the other hand, the Democrat cohort uses the term "freedom of religion" in a more inclusive and supportive way. They emphasize the importance of respecting all religions and the right to be free from religion. For instance, one comment states that anti-Muslim sentiment is frequently visited on women, whose US freedom of religion should allow them to wear the modesty head scarf. An alternative term that might appeal to both groups could be **"religious liberty"** or **"religious**

and progress. For example, they mention "right-wing populism" as a force that is triggering a worldwide return to authoritarianism and hurting journalism. Alternative terms that might appeal to both groups could include **"grassroots movements"** or **"popular movements"**, which emphasize the involvement and support of ordinary people without the negative connotations associated with populism. Another term could be **"democratic movements"**, which emphasizes the democratic nature of these movements, appealing to both groups' apparent support for democracy. *Usage shares: 36.2%, 63.8%*

porn star
The Republican cohort seems to use the term "porn star" in a more derogatory or negative manner, often using it to criticize or mock individuals.

(a) The front page of "Bridging Dictionary: Paper Edition" with GPT-generated summaries for representative terms.



(b) The front page of the interactive Bridging Dictionary demo. The user can type any term they are interested in.

Figure 1: Selected views of the Bridging Dictionary: a static paper page and an interactive front page with user text input.

The Bridging Dictionary (BD) comprises two main components: (1) a paper dictionary and (2) an interactive demo. As illustrated in Figure 1, the paper dictionary presents 796 representative terms in a print-ready format, supplemented with summaries generated by an LLM. The interactive demo, on the other hand, enables users to explore the usage of a given term within Republican and Democratic communities in greater detail. BD leverages `gpt-3.5-turbo`, a widely-recognized LLM, via OpenAI's API to generate these summaries. The term usages are sampled from a Twitter dataset [11], which includes 4.7 million partisan-generated tweets (amounting to 100 million tokens) from each side during the 2020 American election. Users can customize both the dataset and the available LLMs from OpenAI.

3.1 Paper Dictionary

We generate a static printable document called "Bridging Dictionary: Paper Edition" that comprises 796 terms with LLM-generated summaries. These terms are curated algorithmically: we identify words and multi-word phrases that (1) occur sufficiently often within both partisan communities and (2) have significant differences between the two communities, either in sentiment score or in usage frequency. The parameters for these operations (i.e., the thresholds for "sufficient" and "significant") are adjusted manually by an editor.

3.2 Interactive Demo

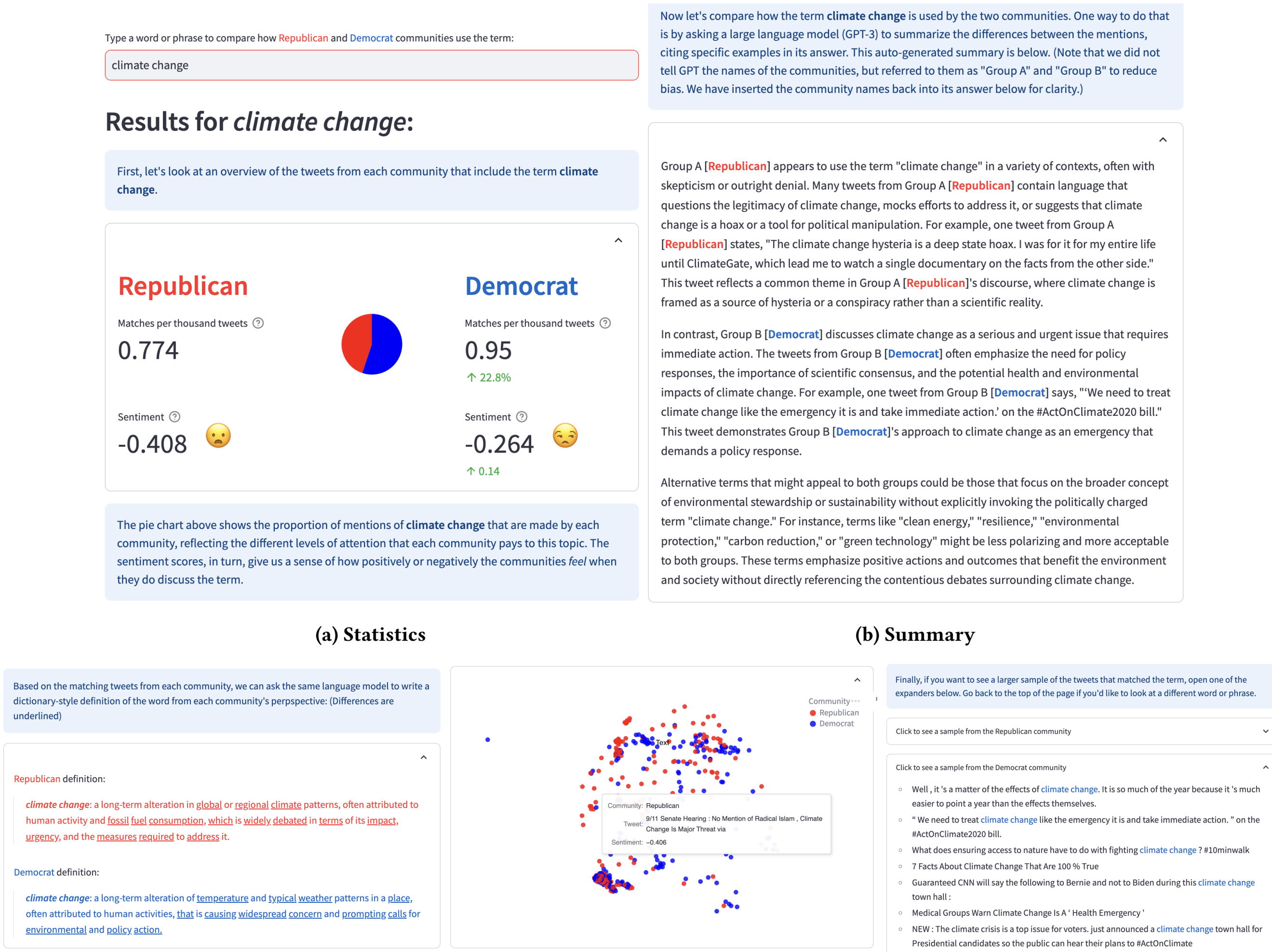
The interactive demo is a web-based generative dictionary providing greater detail and flexibility than the print version. It is implemented with the Streamlit framework [21]. Whenever a user types a phrase, the system provides a few functions to explore how two communities (Republicans and Democrats) use this term, with alternative suggestions and visualizations of the input data.

Statistics. This section offers an overview of tweets containing the specified term from each community. The column "Matches per Thousand Tweets" indicates the frequency of tweets that include the term (e.g., "climate change") per 1,000 tweets within a community. The accompanying pie chart illustrates the proportion of term usage between two partisan communities. Sentiment scores represent the average sentiment for tweets from a community that matched the term, with higher scores indicating a more positive sentiment. Colored text highlights the comparative scores between the two communities, helping users to interpret the data easily.

Summary. This section features LLM-generated summaries that explain the usage of the term across divides and propose alternative terms. The generation follows a standard retrieval-augmented generation (RAG) procedure, as outlined by Gao et al. [4]. Initially, the system randomly samples up to 50 tweets containing the term from a specific community, creates a prompt using these tweets, and prompts the LLM to produce summaries with a simple query. Importantly, the LLM is unaware of the community identity and only uses the sampled tweets for its summarization.

Definition. This section generates a dictionary-style definition of a term from each community's perspective. This follows the same RAG process as the Summary section but asks the model for a definition instead of a summary.

Topic scatterplot. This section presents a two-dimensional interactive scatterplot that organizes individual tweets by topic, grouping similar topics closely together. Users can explore specific tweets by hovering over the corresponding points, enabling them to review the sampled tweets used for summary and definition generation. The process follows a well-established pipeline (e.g., BERTopic [6])



(a) Statistics (b) Summary

(c) Definition (d) Topic scatterplot (e) Sample list

Figure 2: Six core sections in the interactive demo to help users explore how two communities use a term differently.

transforming raw text into a scatterplot. It involves computing the embedding for each tweet using a sentence embedding model, then projecting these embeddings into two dimensions. By default, the points are clustered and color-coded based on the outcome of a clustering algorithm applied to these embeddings. For projection, we employ UMAP [17], and for clustering, we use HDBSCAN [16].

Sample list. This section lists the sampled tweets from each group, allowing users to read the information source.

4 Discussion and Evaluation

After deploying the Bridging Dictionary, we interviewed a professional journalist from *Frontline*, PBS's flagship investigative journalism series known for its in-depth reporting on critical social and political issues, and received positive feedback. Based on the interview, the journalist particularly appreciate the statistics and AI-generated explanations of how words are perceived across political lines. However, they prefer selecting alternative words themselves rather than relying on AI-suggested alternatives. The topic scatterplot and sample list features help the journalist make informed word choices by increasing transparency and trust in AI-generated output. During our interview, two promising directions emerged: (1) enhancing the connection between LLM-generated content and information sources by considering human agency and trust in human-AI interaction, and (2) broadening information sources beyond Twitter and regularly updating the dataset to reflect evolving language trends. We intend to further develop the tool based on these suggestions, conduct a more comprehensive field study involving more professional journalists and other users, and assess the tool's impact on writing and editing.

Acknowledgments

We express our sincere gratitude to the PBS Frontline team for their support and to our colleagues at the MIT Center for Constructive Communication for their valuable input.

References

[1] Bhagyashree Vyankatrao Barde and Anant Madhavrao Bainwad. 2017. An overview of topic modeling methods and tools. In *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE, 745–750.

[2] Robert Chew, John Bollenbacher, Michael Wenger, Jessica Speer, and Annice Kim. 2023. LLM-assisted content analysis: Using large language models to support deductive coding. *arXiv preprint arXiv:2306.14924* (2023).

[3] Kevin Crowston, Eileen E Allen, and Robert Heckman. 2012. Using natural language processing technology for qualitative data analysis. *International Journal of Social Research Methodology* 15, 6 (2012), 523–543.

[4] Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Qianyu Guo, Meng Wang, and Haofen Wang. 2023. Retrieval-Augmented Generation for Large Language Models: A Survey. *ArXiv* abs/2312.10997 (2023). https://api.semanticscholar.org/CorpusID:266359151

[5] Katy Ilonka Gero, Chelse Swoopes, Ziwei Gu, Jonathan K Kummerfeld, and Elena L Glassman. 2024. Supporting Sensemaking of Large Language Model Outputs at Scale. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–21.

[6] Maarten Grootendorst. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794* (2022).

[7] Timothy C Guetterman, Tammy Chang, Melissa DeJonckheere, Tanmay Basu, Elizabeth Scruggs, and VG Vinod Vydiswaran. 2018. Augmenting qualitative text analysis with natural language processing: methodological study. *Journal of medical Internet research* 20, 6 (2018), e231.

[8] Gordon Heltzel and Kristin Laurin. 2020. Polarization in America: Two possible futures. *Current opinion in behavioral sciences* 34 (2020), 179–184.

[9] Nan Hu, Ting Zhang, Baojun Gao, and Indranil Bose. 2019. What do hotel customers complain about? Text analysis using structural topic model. *Tourism Management* 72 (2019), 417–426.

[10] Karoliina Isoaho, Daria Gritsenko, and Eetu Mäkelä. 2021. Topic modeling and text analysis for qualitative policy research. *Policy Studies Journal* 49, 1 (2021), 300–324.

[11] Hang Jiang, Doug Beeferman, Brandon Roy, and Deb Roy. 2022. CommunityLM: Probing Partisan Worldviews from Language Models. In *Proceedings of the 29th International Conference on Computational Linguistics*. International Committee on Computational Linguistics, Gyeongju, Republic of Korea, 6818–6826. https://aclanthology.org/2022.coling-1.593

[12] Ping Li, Benjamin Schloss, and D Jake Follmer. 2017. Speaking two "Languages" in America: A semantic space analysis of how presidential candidates and their supporters represent abstract political concepts differently. *Behavior research methods* 49, 5 (2017), 1668–1685.

[13] Ruibo Liu, Chenyan Jia, Jason Wei, Guangxuan Xu, and Soroush Vosoughi. 2022. Quantifying and alleviating political bias in language models. *Artificial Intelligence* 304 (2022), 103654.

[14] Ruibo Liu, Chenyan Jia, Jason Wei, Guangxuan Xu, Lili Wang, and Soroush Vosoughi. 2021. Mitigating political bias in language models through reinforced calibration. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 14857–14866.

[15] Nolan McCarty, Keith T Poole, and Howard Rosenthal. 2016. *Polarized America: The dance of ideology and unequal riches*. mit Press.

[16] Leland McInnes, John Healy, Steve Astels, et al. 2017. hdbscan: Hierarchical density based clustering. *J. Open Source Softw.* 2, 11 (2017), 205.

[17] Leland McInnes, John Healy, and James Melville. 2018. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* (2018).

[18] Cassandra Overney, Belén Saldías, Dimitra Dimitrakopoulou, and Deb Roy. 2024. SenseMate: An Accessible and Beginner-Friendly Human-AI Platform for Qualitative Data Analysis. In *Proceedings of IUI '24*. ACM, Greenville, 922–939. https://doi.org/10/gtt392

[19] Keith T Poole and Howard Rosenthal. 1984. The polarization of American politics. *The journal of politics* 46, 4 (1984), 1061–1079.

[20] Ashiqur R. KhudaBukhsh, Rupak Sarkar, Mark S. Kamlet, and Tom Mitchell. 2021. We Don't Speak the Same Language: Interpreting Polarization through Machine Translation. *Proceedings of the AAAI Conference on Artificial Intelligence* 35, 17 (May 2021), 14893–14901. https://ojs.aaai.org/index.php/AAAI/article/view/17748

[21] Streamlit. 2021. Streamlit – A faster way to build and share data apps. https://streamlit.io/

[22] Albert Webson, Zhizhong Chen, Carsten Eickhoff, and Ellie Pavlick. 2020. Are "Undocumented Workers" the Same as "Illegal Aliens"? Disentangling Denotation and Connotation in Vector Spaces. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Bonnie Webber, Trevor Cohn, Yulan He, and Yang Liu (Eds.). Association for Computational Linguistics, Online, 4090–4105. https://doi.org/10.18653/v1/2020.emnlp-main.335

[23] Ziang Xiao, Xingdi Yuan, Q Vera Liao, Rania Abdelghani, and Pierre-Yves Oudeyer. 2023. Supporting qualitative analysis with large language models: Combining codebook with GPT-3 for deductive coding. In *Companion proceedings of the 28th international conference on intelligent user interfaces*. 75–78.

[24] Xiaohui Yan, Jiafeng Guo, Yanyan Lan, and Xueqi Cheng. 2013. A biterm topic model for short texts. In *Proceedings of the 22nd international conference on World Wide Web*. 1445–1456.