Decoding Political Issue Framing in News Media using Word Embeddings

Term Project for PPOL 6801 Text as Data: Computational Linguistics

Zhiqiang Ji, Spring 2024

Motivation: Same news, different narratives

Dominion Voting Systems v. Fox News Network (***Dominion v. Fox***)

Fox News reached a $787.5 million settlement with Dominion on April 18, 2023

MSNBC TRUMP ON TRIAL DECISION 2024 RACHEL MADDOW LIVE TV

EVENT ENDED Last updated April 18, 2023, 5:47 PM EDT

Fox News settles with Dominion for $787.5 million after hourslong trial delay

Rupert Murdoch's media empire was expected to face a reckoning in court for broadcasting 2020 election lies. But the parties abruptly agreed to a hefty settlement.

REUTERS® World Business Markets Sustainability Legal Breakingviews More

Litigation | Government | Litigation | Corporate Counsel | Securities Enforcement

Fox settles Dominion lawsuit for $787.5 million over US election lies

By Helen Coster and Jack Queen

April 19, 2023 2:24 AM EDT · Updated a year ago

Aa

MEDIA

FOX News Media, Dominion Voting Systems reach agreement over defamation lawsuit

Dominion Voting Systems filed $1.6B lawsuit against network in 2021

By Joseph A. Wulfsohn · Fox News

Published April 18, 2023 5:37pm EDT

Settlement reached in Dominion v. Fox News case

Dominion sued Fox News for 1.6 billion in defamation case, Fox denied any wrongdoing.

Motivation (cont'd)

Research Questions:

1. How do news media frame the same political and policy issues and party leanings by attaching different sentiments?
2. How do news media establish positive associations between major issues and political parties/ideologies?
3. Has this politically-loaded narrative framing intensified in sync with the gradual polarization of the political climate in recent decades?

(Inspired by class readings on word embeddings and Rozado, David, and Musa al-Gharbi. "Using Word Embeddings to Probe Sentiment Associations of Politically Loaded Terms in News and Opinion Articles from News Media Outlets." *Journal of Computational Social Science* 5, no. 1 (May 1, 2022): 427–48. https://doi.org/10.1007/s42001-021-00130-y.)

Data

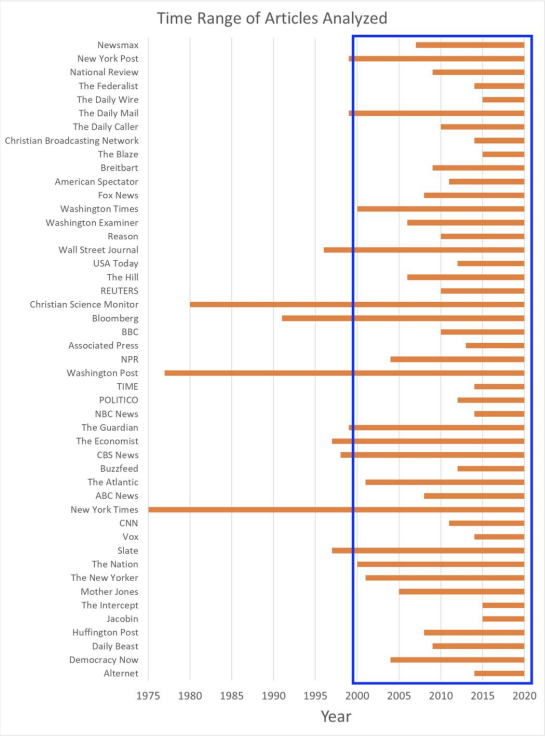
- From (Rozado & al-Gharbi 2022)
- Original: 47 news media outlets with the time range of 1975-2019, over 27 million articles were analyzed.
- News and opinion articles, headlines and main text.
- "Scrambled": the authors removed stop words and then randomly scrambled the remaining words within a sliding window of size 10 to avoid infringing on the articles' copyrights.
- My project: 6 news outlets (2 left, 2 central, 2 right), 2000-2019

	nyt	slate	reuters	csm	nyp	dailymail
2000-2004	421143	15313	6	42787	146136	64491
2005-2009	303699	26045	790883	44535	187012	170786
2010-2014	460506	61621	1322954	92239	257433	803932
2015-2019	290703	61284	1180341	48144	255737	1356851

Data (cont'd)

Choose news outlets:

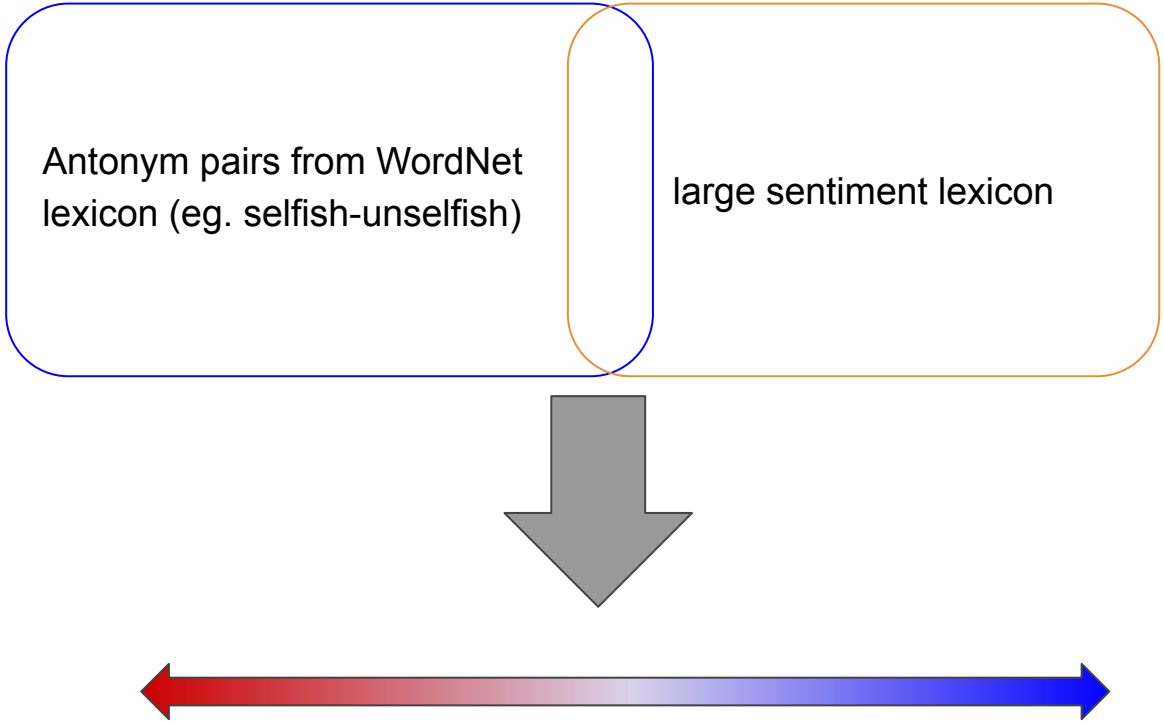
- Human-rated media bias chart
- Availability across the time span



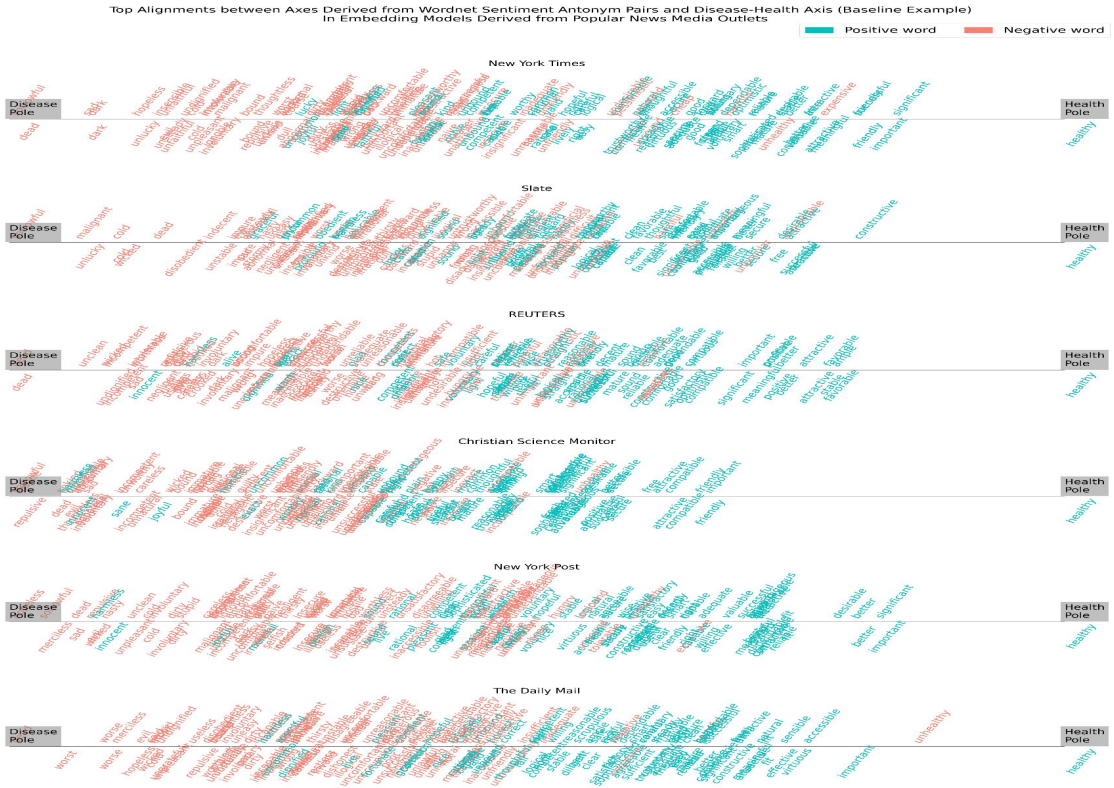
Methods: High-level 3 steps

- Train outlet/time-interval specific embedding models (6x4=24 models)
- Create axes
- Projecting words to the embedding space

Method (cont'd): Sentiment analysis

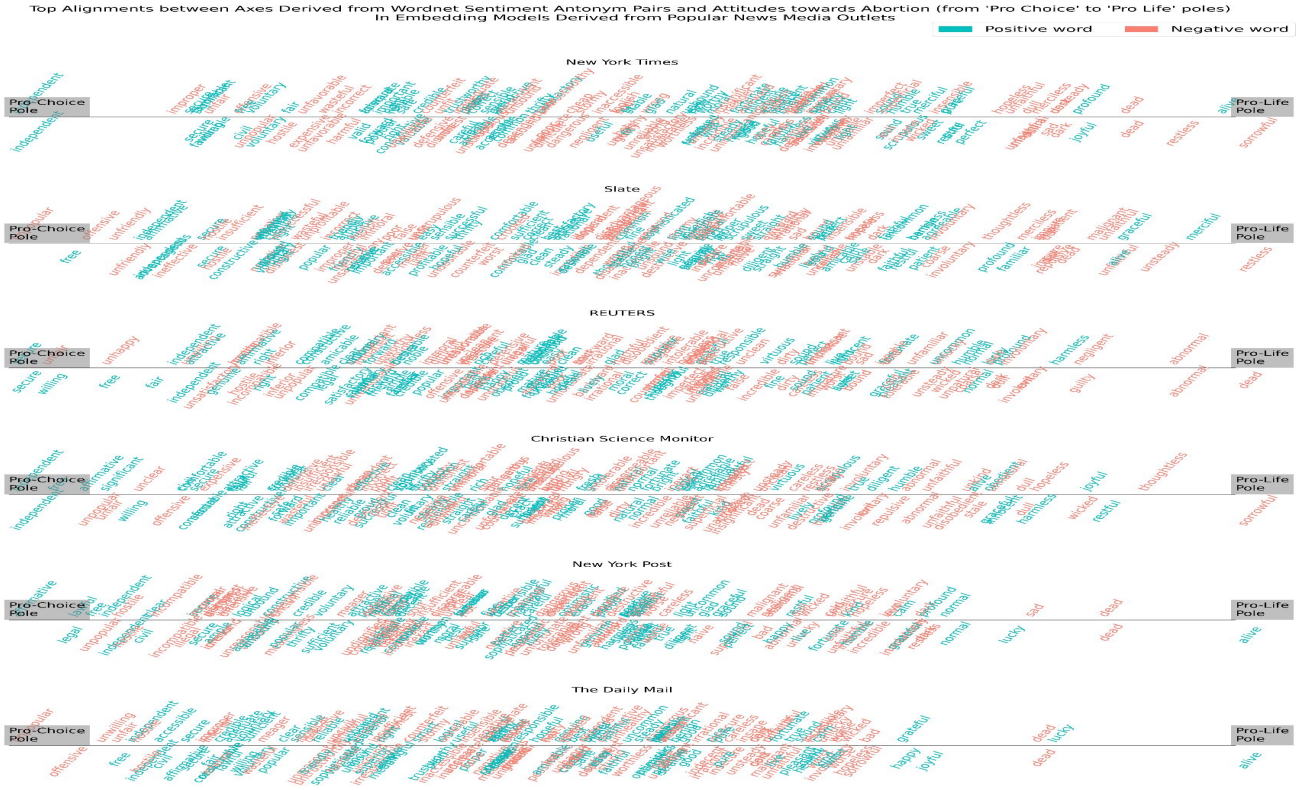


Results: Pos/Neg Sentiments (1. Disease→Health)

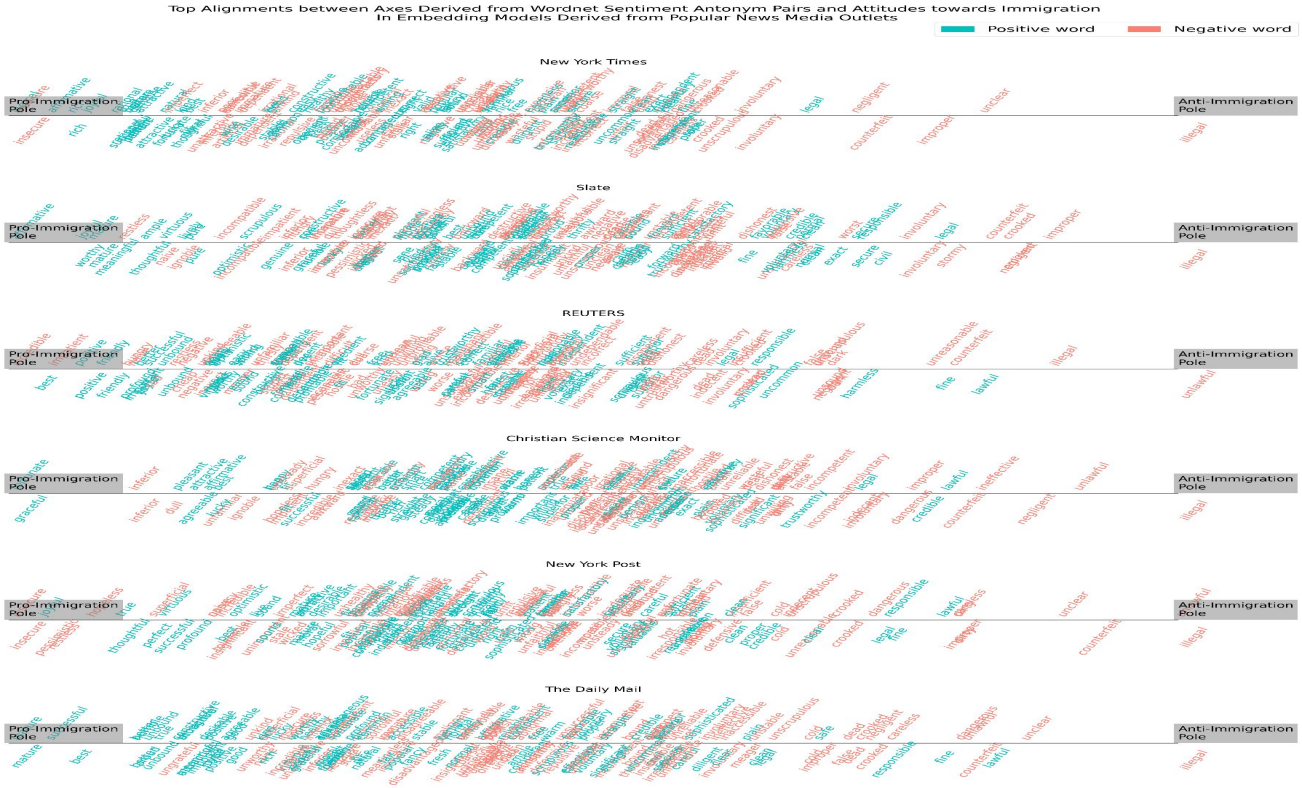


Results: Pos/Neg Sentiments (2. Ideology Con→Lib)

Results: Pos/Neg Sentiments (3. Abortion: Choice→Life)



Results: Pos/Neg Sentiments (4. Immigration: Pro→Con)



Results: Issue Ownership

Fagan, EJ. "Issue Ownership and the Priorities of Party Elites in the United States, 2004–2016." *Party Politics* 27, no. 1 (January 1, 2021): 149–60. https://doi.org/10.1177/1354068819839212.

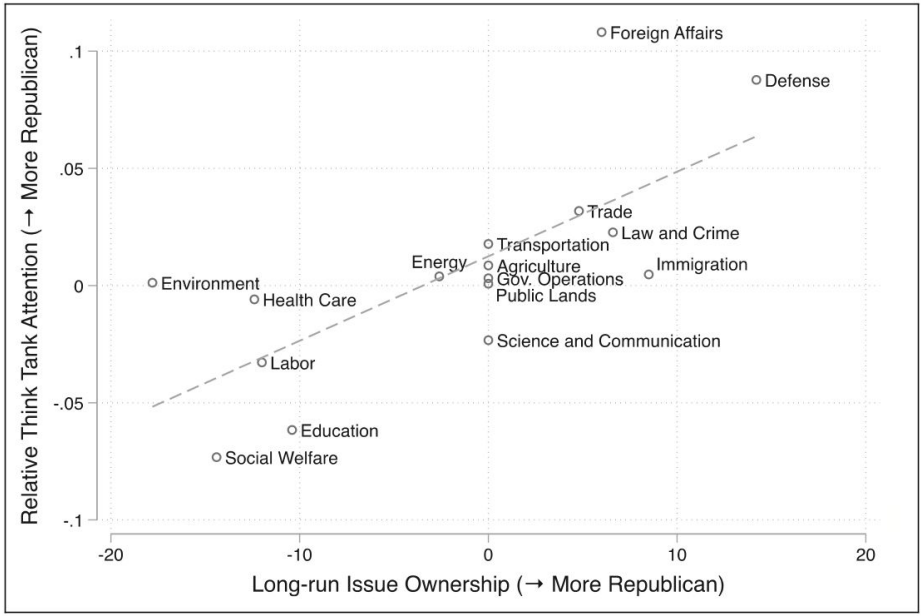
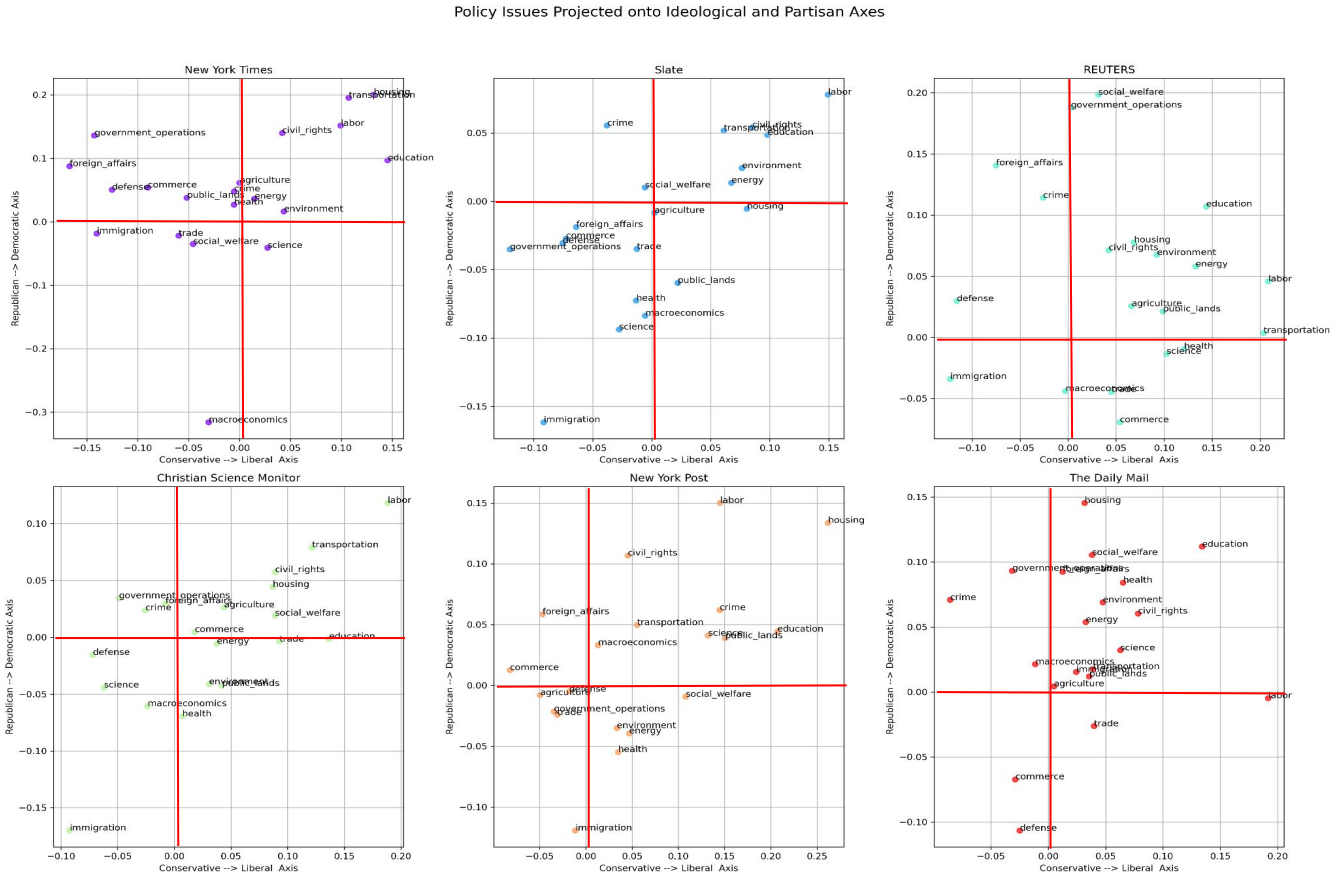


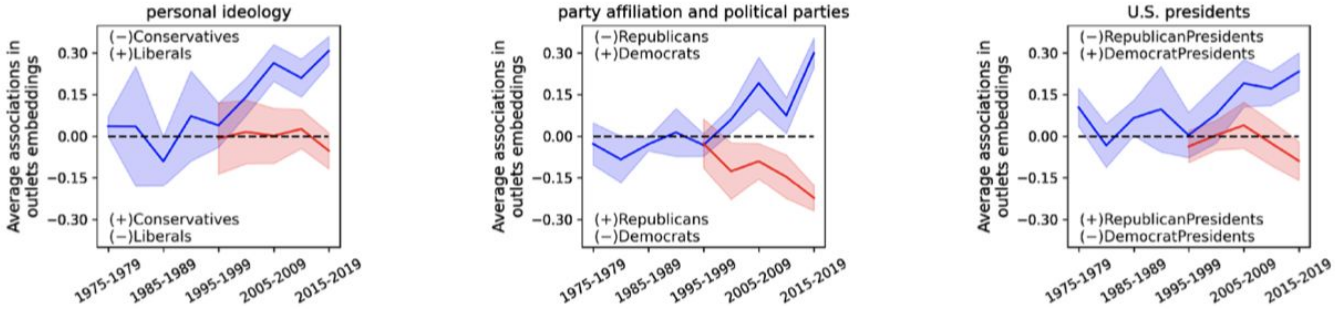
Figure 5. Relative attention of party-aligned think tanks and long-run issue ownership. The *Y*-axis represents the average proportion of issue attention from Republican-aligned think tanks from 2004 to 2016 minus the average proportion of issue attention from Democratic-aligned think tanks. Macroeconomics, civil rights, housing, and domestic commerce excluded. The *X*-axis represents long-run competency-based issue ownership as coded by Egan (2013), where a negative score represents more Democratic ownership, and a positive score represents more Republican ownership.

Results: Issue Ownership (cont'd)



Results: Chronological bias (still working on it)

Change in latent associations between words with political connotations and positive/negative terms in popular news media outlets (1975-2019)



Lesson Learned & Next Steps

- Issue axes are difficult to create
  - Each issue takes only a small portion of the corpus
  - Results can be murky
  - Hard to validify interpretations on results
- Choosing the effective lexicons is essential
- A more granular and comprehensive sentiment model would be beneficial
- Next steps:
  - Time series analysis
  - Use pre-trained Named Entity Recognition (NER) model to detect biased languages

Thank you! Enjoy your final weeks!