

Politicization or Medicalization? Partisan Framing of COVID-19 Vaccines on Twitter Using Language Models and Embeddings

Shuti Zhao

McCourt School of Public Policy, Georgetown University
sz734@georgetown.edu

Abstract

COVID-19 vaccines represented a major scientific breakthrough, yet in the United States vaccination quickly became politicized. This study examines how U.S. Members of Congress discussed COVID-19 vaccines on Twitter between 2020 and 2022, focusing on partisan differences in framing, stance, and narrative signals associated with misinformation. Using a corpus of 16,857 vaccine-related tweets, the analysis combines structural topic modeling and large language model (LLM) classification to identify dominant issue frames and vaccine stances, including Medical/Safety, Mandate/Rights, and Conspiracy/Misinformation narratives. An embedding-based semantic analysis is then applied to quantify the alignment of tweets with key concepts such as “freedom,” “mandate,” “safety,” and “hoax,” and to assess how these linguistic cues predict anti-vaccine messaging. Results reveal stark partisan asymmetries: Democratic lawmakers overwhelmingly frame vaccines as a public health issue grounded in science, while Republican lawmakers more frequently employ politicized and misinformation-linked narratives, particularly following the introduction of vaccine mandates. Embedding-based measures capture these patterns most effectively when using multi-anchor and data-driven concept vectors rather than single-word anchors.

Keywords: large language models, COVID-19 vaccines, political framing, misinformation, Twitter, text embeddings

1. Introduction

The development of effective COVID-19 vaccines represented a major scientific breakthrough, yet in the United States vaccination quickly became entangled in partisan politics. From the outset of the pandemic, COVID-19 was framed not only as a public health emergency but also as a political issue (Gollust et al., 2020). Partisan identity soon emerged as a strong predictor of pandemic-related attitudes and behaviors (Gadarian et al., 2021), and early 2021 surveys documented substantial gaps in vaccine intent between Republicans and Democrats (Kaiser Family Foundation, 2021). These divides were amplified by politicized narratives and misinformation, including claims that downplayed COVID-19’s severity or characterized public concern as a “hoax,” contributing to disagreement over vaccine safety and necessity (Motta et al., 2020). In contrast, public health authorities consistently emphasized scientific consensus regarding vaccine safety and efficacy. This polarized environment set the stage for sharply divergent vaccine discourse in the public sphere.

Against this backdrop, this study examines how members of the U.S. Congress communicated about COVID-19 vaccines on Twitter and which narratives were associated with misinformation-oriented messaging. Prior research shows that vaccine communication on social media became increasingly polarized along ideological lines (Jiang et al., 2021), and that Democratic and Republican lawmakers employed distinct themes and rhetorical strategies when discussing COVID-19 vaccines online (Engel-Rebitzer et al., 2022). Building on this literature, we distinguish between *politicized framing*—which situates vaccines within political conflict or ideology (e.g., mandates, individual rights, government overreach, or conspiracy and “hoax” claims)—and *medicalized framing*, which treats vaccines as a public health issue grounded in science (e.g., safety, efficacy, and medical guidance).

1.1. Research Questions

This study addresses the following research question:

How did U.S. Members of Congress communicate about COVID-19 vaccines on Twitter, and what narratives signaled misinformation?

We further investigate four related sub-questions:

- **Framing:** How did Republican and Democratic lawmakers differ in their use of Medical/Safety, Mandate/Rights, and Conspiracy/Misinformation frames?
- **Stance over time:** How did pro-vaccine versus anti-vaccine stances evolve from 2020 to 2022 across parties?
- **Brand narratives:** Were specific vaccines (Pfizer, Moderna, AstraZeneca) discussed differently in terms of politicization versus medicalization?
- **Language signals:** Which semantic cues (e.g., “freedom,” “mandate,” “safety,” “hoax”) were most predictive of anti-vaccine messaging?

2. Data & Methods

2.1. Data Collection

This study analyzes tweets authored by U.S. Members of Congress that discussed COVID-19 vaccines between 2020 and 2022. Tweets were obtained from the public *Tweets from Congress* archive and filtered for vaccine-related content using keyword rules (e.g., “vaccine,” vaccine brand names, and common misinformation terms). After removing simple retweets, the analytic sample contains approximately 16,857 English-language tweets from lawmakers in both major parties. Each tweet record includes the author’s party affiliation (Democrat, Republican, or Independent) and additional metadata, including whether a specific vaccine brand was mentioned.

2.2. Frame Identification via Topic Modeling

To characterize narrative structure without imposing ex ante labels, Structural Topic Modeling (STM) was first applied to the vaccine-related tweets. STM is an unsupervised probabilistic topic model that identifies latent themes by clustering co-occurring words and documents. Party metadata was incorporated as a document-level covariate to assess whether specific topics were differentially prevalent across partisan groups. Model output (top words and representative tweets per topic) yielded multiple coherent topics.

To translate these latent topics into interpretable frames, a human-in-the-loop labeling step was implemented using a GPT-4-based large language model (LLM). For each STM topic—summarized by its highest-probability terms and example tweets—the LLM was prompted to provide a concise substantive label. Related topics were subsequently aggregated into four higher-level frames:

- **Medical/Safety:** health information, efficacy, and safety (e.g., “safe,” “effective,” “doses,” “health,” “science”).
- **Mandate/Rights:** mandates, personal liberty, and government overreach (e.g., “mandate,” “freedom,” “rights,” “unconstitutional”).
- **Conspiracy/Misinformation:** conspiratorial narratives and false claims (e.g., “hoax,” “lies,” “fake,” “truth,” “media”).
- **Other:** residual content not captured by the three dominant narratives (e.g., procedural updates or tangential mentions).

Tweets were assigned to one of these frames based on topic membership; ambiguous cases were resolved via manual inspection when necessary. This STM+LLM pipeline yields frame categories that are both data-driven and substantively interpretable.

2.3. LLM-Based Classification (Stance & Frame)

In parallel to the STM workflow, a GPT-4 LLM classifier was applied at the tweet level to obtain context-sensitive labels for (i) vaccine stance and (ii) issue frame. Prompts provided operational definitions and examples to standardize labeling. Each tweet was assigned:

- **Vaccine stance:** *Pro-vaccine*, *Anti-vaccine*, or *Mixed/Unclear*, depending on whether the tweet endorsed vaccination, opposed it (e.g., skepticism, fear appeals, or misinformation), or was neutral/ambiguous.
- **Issue frame:** *Medical/Safety*, *Mandate/Rights*, *Conspiracy/Misinformation*, or *Other*, using the definitions above.

A subset of classifications was manually reviewed for quality control. LLM frame assignments were broadly consistent with STM topic structure, supporting reliability of the combined coding scheme. Rare sarcasm or highly coded language was flagged for cautious interpretation.

2.4. Semantic Embeddings & Concept Vectors

To complement discrete stance and frame labels with continuous semantic measures, each tweet was represented as a dense vector embedding using a pretrained OpenAI text embedding model. Let $\mathbf{e}_i \in \mathbb{R}^d$ denote the embedding of tweet i , where d is the embedding dimension. These embeddings encode semantic similarity such that tweets with similar linguistic content are located closer together in the embedding space.

Single-anchor prototype approach. As an initial embedding-based strategy, prototype vectors were constructed for four theoretically motivated concept anchors: *freedom*, *mandate*, *safety*, and *hoax*. For each concept, a short prototypical text was manually specified to capture its core semantic meaning (e.g., individual liberty for *freedom*, public health protection for *safety*, and conspiratorial narratives for *hoax*). Each prototypical text was embedded using the same language model, yielding one prototype vector per concept, denoted $\mathbf{c}_k \in \mathbb{R}^d$.

For each tweet, semantic alignment with a given concept was quantified using cosine similarity between the tweet embedding and the corresponding concept prototype:

$$\text{sim}(\mathbf{e}_i, \mathbf{c}_k) = \frac{\mathbf{e}_i^\top \mathbf{c}_k}{\|\mathbf{e}_i\| \|\mathbf{c}_k\|}.$$

These similarity scores were rescaled to the interval $[0, 1]$ to facilitate interpretation and comparison across concepts. Higher values indicate stronger semantic alignment between a tweet and the corresponding framing concept.

Multi-anchor concept vectors. Because single anchors can be sensitive to vocabulary overlap and contextual ambiguity, a second strategy used multi-anchor concept vectors. For each narrative (e.g., mandate/rights, hoax/misinformation, medical/safety), multiple semantically related words or phrases were embedded and averaged to form a more stable concept vector. This reduces dependence on any single lexical choice and better captures distributed semantics.

Data-driven politicization–medicalization axis. A third strategy constructed a corpus-derived axis by contrasting average embeddings across LLM-coded frames. Let $\bar{\mathbf{e}}_{\text{pol}}$ be the mean embedding of tweets labeled as politicized frames (Mandate/Rights or Conspiracy/Misinformation) and $\bar{\mathbf{e}}_{\text{med}}$ be the mean embedding of Medical/Safety tweets. The politicization direction is:

$$\mathbf{v} = \bar{\mathbf{e}}_{\text{pol}} - \bar{\mathbf{e}}_{\text{med}}.$$

Each tweet receives a continuous politicization score via projection:

$$s_i = \frac{\mathbf{e}_i^\top \mathbf{v}}{\|\mathbf{e}_i\| \|\mathbf{v}\|}.$$

Higher values indicate language more aligned with politicized (rights/hoax) discourse; lower values indicate more medicalized language.

2.5. Regression Analysis

To quantify which semantic cues predict misinformation-oriented messaging and to evaluate partisan effects conditional on language, a logistic regression model was fit with

LLM-coded stance as the outcome. The dependent variable Y_i equals 1 if tweet i was labeled *Anti-vaccine* and 0 otherwise. Predictors include concept similarity features and party affiliation:

- **Hoax similarity:** cosine similarity between tweet embedding and the hoax/misinformation concept vector.
- **Mandate/Freedom similarity:** similarity to the freedom/mandate concept vector.
- **Medical/Safety similarity:** similarity to the safety/medical concept vector.
- **Party indicator:** Republican = 1, Democrat = 0 (Independents were handled separately due to small counts).

The model is:

$$\Pr(Y_i = 1) = \text{logit}^{-1} \left(\beta_0 + \beta_1 \text{HoaxSim}_i + \beta_2 \text{MandateSim}_i + \beta_3 \text{SafetySim}_i + \beta_4 \text{Republican}_i \right).$$

where coefficient magnitudes quantify how language cues and partisanship shift the odds of anti-vaccine stance.

2.6. Summary of Analytic Pipeline

The overall workflow proceeds as follows:

1. **STM + LLM frame labeling:** discover topics via STM and map topics to interpretable frames using GPT-4.
2. **LLM tweet-level labels:** classify each tweet's stance (pro/anti/mixed) and frame (medical/mandate/conspiracy/other).
3. **Embeddings:** compute dense semantic vectors for each tweet using a pretrained embedding model.
4. **Concept similarity scoring:** compute cosine similarity to (i) prototype single-anchor vectors, (ii) multi-anchor vectors, and (iii) a data-driven politicization axis.
5. **Regression modeling:** predict anti-vaccine stance using similarity features and party affiliation.
6. **Validation:** cross-check STM themes against LLM frame labels and manually review ambiguous cases; agreement across methods was high.

3. Results

3.1. Partisan Differences in Vaccine Stance

Figure 1 shows the distribution of stance labels by party for congressional tweets about COVID-19 vaccines. A partisan divide in expressed sentiment toward the vaccines is immediately apparent. Among Democratic members of Congress, an overwhelming majority of vaccine-related tweets were classified as *pro-vaccine*—roughly 65–70% of

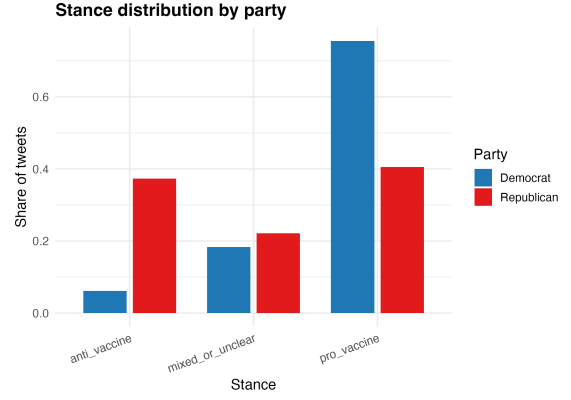


Figure 1: Distribution of stance labels by party for congressional tweets about COVID-19 vaccines.

Democrats' tweets endorsed vaccination or conveyed positive sentiment about vaccines, with essentially zero Democratic tweets in our sample classified as explicitly *anti-vaccine*. The remainder (approximately 30% of Democratic tweets) were mostly neutral or ambiguous in stance (e.g., simple informational posts or references to vaccines without clear approval or disapproval).

In stark contrast, Republican lawmakers' tweets were far more mixed and skewed toward skepticism. Only about 40% of Republican vaccine-related tweets were clearly *pro-vaccine*, while roughly one-third (about 30%) of GOP tweets were coded as *anti-vaccine*. The remaining approximately 30% fell into the *mixed/unclear* category. In practical terms, Republican members of Congress were much more likely than Democrats to voice skepticism or opposition regarding vaccines on Twitter. Virtually one in three Republican vaccine tweets contained anti-vaccine sentiment, compared to almost zero among Democratic tweets.

3.2. Partisan Framing Differences in Narrative

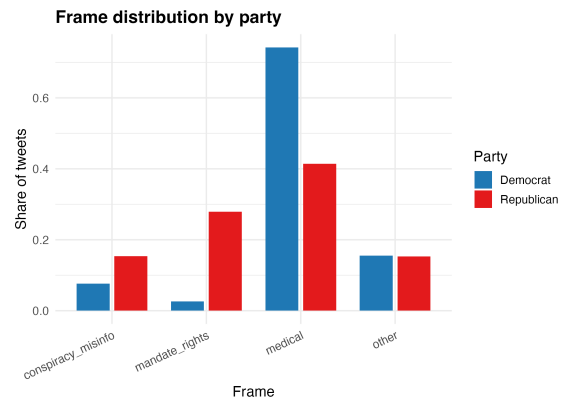


Figure 2: Issue frame distribution by party for vaccine-related tweets.

Beyond simple stance, Republicans and Democrats emphasized very different narratives or frames when discussing vaccines. Figure 2 shows the share of each party's tweets that fell into four frame categories: Medical/Safety, Mandate/Rights, Conspiracy/Misinformation, and Other. The

contrast is striking and aligns with the stance divergence described above.

Democrats: Medical frame dominance. Democratic lawmakers overwhelmingly employed a Medical/Safety frame in their vaccine tweets. About 70% of Democratic tweets highlighted public health and scientific aspects—for example, stressing vaccine efficacy and safety, urging constituents to get vaccinated to protect their health, celebrating the success of vaccines in curbing the pandemic, or sharing factual information about vaccine availability. The Mandate/Rights narrative was virtually absent in Democrats’ tweets (essentially 0%): Democrats almost never framed vaccines as an issue of government overreach or personal liberty versus authority. Similarly, very few Democratic tweets engaged with the Conspiracy/Misinformation frame except to debunk it—only a handful (under 5%) of Democratic tweets even mentioned common misinformation themes, and when they did, it was typically to refute false claims or emphasize trust in science. The remaining Democratic tweets fell into the Other category when they were not explicitly medical, but this was a relatively small portion.

Republicans: Politicized frames. Republican lawmakers, by contrast, showed a far more diversified framing of the vaccine discourse, including a significantly higher uptake of political and conspiratorial narratives. Only about 45% of Republican vaccine tweets used the Medical/Safety frame—far less than Democrats’ 70%. Instead, a substantial portion of GOP tweets pivoted to politicized frames. Roughly 25% of Republican tweets were categorized as Mandate/Rights, explicitly discussing government mandates, personal freedom, individual rights, or resistance to perceived government overreach. These tweets often cast the vaccine issue as one of choice versus coercion. In addition, about 10–15% of Republican tweets centered on Conspiracy/Misinformation themes, including content that spread or lent credibility to dubious claims and conspiracy theories about vaccines. The remainder of GOP tweets fell into the Other category. In sum, Democrats focused on health and science, whereas Republicans more frequently politicized the issue, framing vaccines in terms of personal freedom, government overreach, and conspiratorial overtones.

3.3. Predicted Effects of Framing by Party

To assess whether these framing differences translate into substantively different vaccine positions across parties, Figure 3 plots the predicted probability of an anti-vaccine stance as a function of issue frame and party affiliation. The figure reveals a strong interaction between party and frame. Mandate/Rights and Conspiracy/Misinformation frames are associated with sharply higher probabilities of anti-vaccine messaging among Republican lawmakers, while these frames are virtually absent or non-predictive among Democrats. In contrast, Medical/Safety framing corresponds to consistently low predicted probabilities of anti-vaccine stance for both parties. This interaction helps explain why partisan differences in vaccine stance closely track differences in narrative framing.

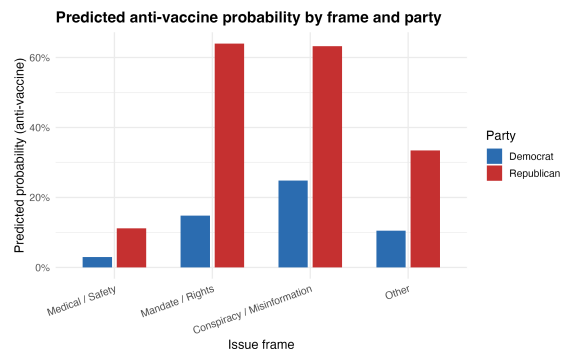


Figure 3: Predicted probability of anti-vaccine stance by issue frame and party.

3.4. Trends in Vaccine Discourse Over Time (2020–2022)

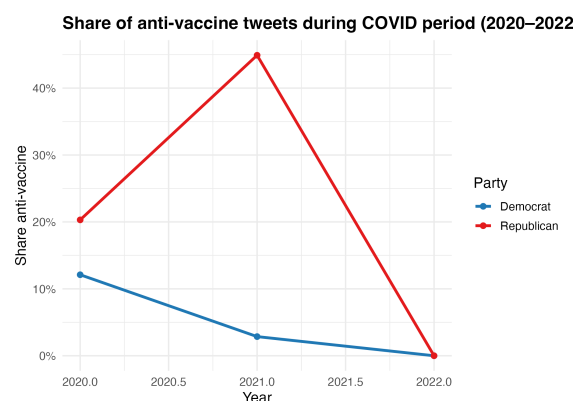


Figure 4: Share of anti-vaccine tweets by party over time (2020–2022).

Figure 4 shows that the partisan gap in vaccine stance was not static but widened dramatically as the pandemic progressed, largely driven by changes in Republican rhetoric. In 2020, tweets from both parties were largely supportive or neutral. By 2021, following widespread vaccine availability and debates over mandates, Republican opposition surged on Twitter while Democrats remained uniformly pro-vaccine. In 2022, vaccine-related tweeting declined overall, but Republicans continued to post a substantial share of anti-vaccine content, whereas Democrats posted virtually none.

3.5. Narratives by Vaccine Brand

Figure 5 plots the average semantic characteristics of tweets mentioning specific vaccine brands. Tweets referencing Pfizer and Moderna cluster in the medical/scientific region, indicating high alignment with safety and efficacy language and low politicization. AstraZeneca-related tweets occupy a more intermediate position, reflecting neutral or informational discourse. By contrast, tweets referring to vaccines generically (without brand names) are the most politicized and least medical in tone. When lawmakers spoke about “the vaccine” in abstract terms, they were far more likely to inject political narratives or misinformation, whereas

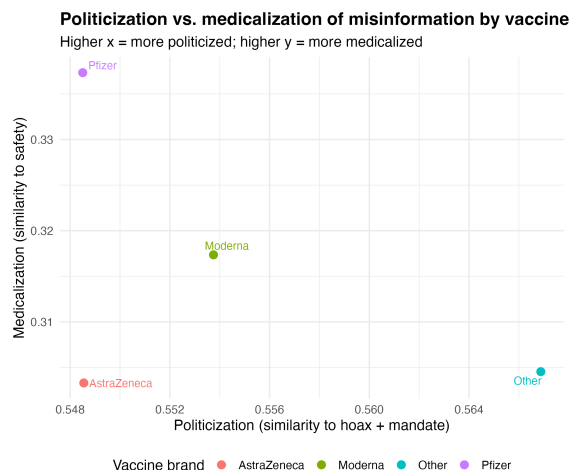


Figure 5: Average politicization versus medicalization of vaccine-related tweets by brand.

brand-specific references tended to anchor discussion in concrete scientific details.

3.6. Language Cues Predicting Anti-Vax Messaging

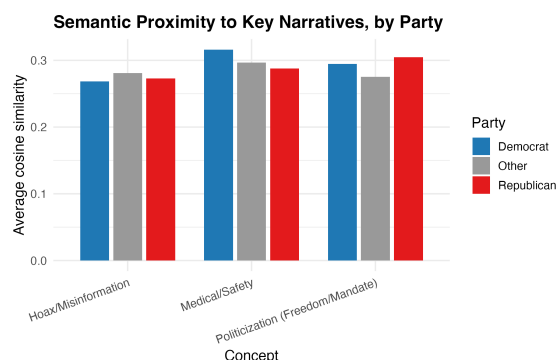


Figure 6: Average semantic similarity of Democratic vs. Republican tweets to selected framing concept anchors (single-word approach).

Figure 6 presents the average semantic similarity of Democratic and Republican tweets to selected framing concept anchors using an initial single-word embedding approach. Given the strong partisan and framing differences observed above, this analysis seeks to identify which linguistic signals in tweets best predict an anti-vaccine stance. Using the embedding-based features described in the Methods section, each tweet’s cosine similarity to the Freedom/Mandate, Hoax/Misinformation, and Medical/Safety narratives was computed.

Somewhat surprisingly, these simple measures reveal only modest differences between the parties. For instance, Democratic and Republican tweets exhibit nearly identical average similarity scores to the “freedom” concept (approximately 0.28–0.30 for both parties), with Democrats, if anything, scoring slightly higher. Intuitively, one might expect Republicans to invoke freedom-related language

more frequently in the vaccine context, given their emphasis on opposition to mandates. However, the embedding results suggest that Democrats also employed freedom-related rhetoric, potentially in contexts such as “freedom through vaccination” or the restoration of normal life following widespread vaccine uptake.

Similarly, similarity to the “hoax” concept shows no dramatic partisan gap. Both parties display low and largely overlapping average similarity scores (roughly 0.22–0.25), indicating that overt conspiracy-related keywords were relatively uncommon in congressional vaccine tweets overall. Even among Republicans, only a subset of lawmakers employed explicit “hoax” language, while some Democratic tweets referenced misinformation solely to refute it. Because embedding similarity does not capture negation or intent, such refutational uses are still registered as semantic proximity to the concept.

The single anchor that exhibits a clearer partisan difference is “mandate.” Republican tweets are somewhat more similar to the mandate concept than Democratic tweets, consistent with Republicans’ greater emphasis on mandates and personal rights. Conversely, Democratic language shows slightly higher similarity to safety-related terms, reflecting a stronger alignment with scientific and public health messaging. Overall, however, the magnitude of these differences remains modest and is substantially smaller than the stark contrasts observed in the categorical frame distributions. This suggests that a naïve single-word embedding approach understates the true extent of partisan divergence in vaccine framing.

To better capture nuanced framing differences, the embedding strategy was refined in two ways. First, multi-anchor concept vectors were constructed by averaging embeddings across multiple related terms for each narrative, as described in the Methods section. This richer semantic representation yields clearer separation between partisan discourse.

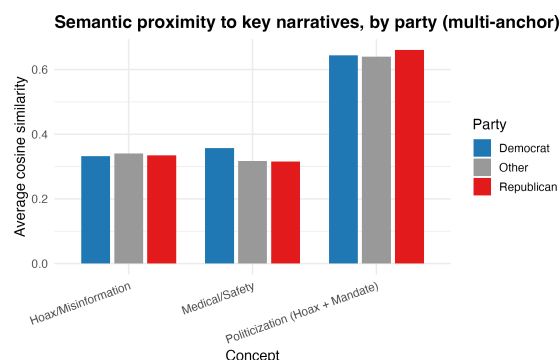


Figure 7: Semantic proximity to key narratives by party using multi-anchor concept vectors.

As shown in Figure 7, Republican tweets score higher on the aggregate Freedom/Rights vector (e.g., {freedom, rights, liberty, mandate}), while Democratic tweets score higher on the composite Medical/Science vector (e.g., {safe, effective, health, science}). These differences align

more closely with the qualitative patterns identified earlier, with Republicans leaning toward politicized rhetoric and Democrats emphasizing medical framing. Nonetheless, even with multi-anchor concepts, the absolute gaps remain moderate.

Second, a data-driven politicization axis was derived directly from the corpus, positioning each tweet along a continuous politicized–medicalized spectrum.

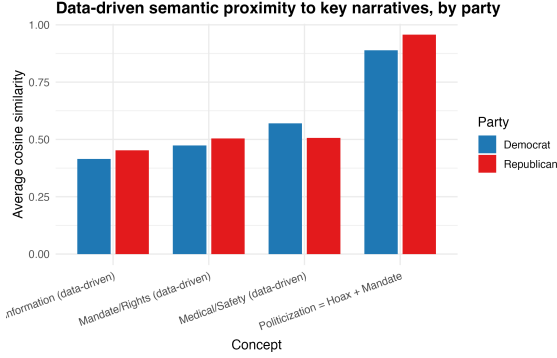


Figure 8: Data-driven semantic proximity to politicized versus medicalized narratives, by party.

Figure 8 reveals a clear partisan gap using this composite measure. Republican tweets score higher on average (approximately 0.47), indicating greater alignment with politicized language, while Democratic tweets score lower (approximately 0.32), reflecting more medicalized discourse. Although this gap is smaller than that observed in discrete frame classifications (e.g., Mandate/Rights framing), it nonetheless confirms a statistically meaningful difference in narrative style. The attenuation of the gap likely reflects the fact that many tweets blend multiple themes, causing continuous embedding measures to average across competing signals.

Finally, to directly link linguistic cues to misinformation stance, a logistic regression model was estimated predicting whether a tweet expresses an anti-vaccine stance.

Table 1: Logistic regression predicting anti-vaccine tweet stance from language similarities and party affiliation.

	(1)
Intercept	-4.318*** (0.152)
Republican (vs. Democrat)	1.612*** (0.061)
Hoax / Misinformation similarity	16.995*** (0.583)
Mandate / Freedom similarity	16.956*** (0.500)
Medical / Safety similarity	-28.146*** (0.575)
Num. Obs.	16,833

Table 1 shows that similarity to Hoax/Misinformation language is the strongest predictor of an anti-vaccine stance. Tweets containing such language are dramatically more likely to be anti-vaccine, even after controlling for party affiliation. Similarity to Mandate/Freedom language is also positively associated with anti-vaccine messaging, while similarity to Medical/Safety language strongly reduces the likelihood of an anti-vaccine stance.

Party affiliation remains statistically significant after accounting for language, though its effect size is notably smaller than those of the content-based predictors. This suggests that most of the partisan gap in anti-vaccine messaging is mediated by narrative choice rather than party identity alone. In practical terms, Republican tweets are more likely to be anti-vaccine largely because they more frequently employ politicized and misinformation-linked rhetoric. When Republican tweets adopt medical or scientific language, they are typically pro-vaccine despite the author’s party.

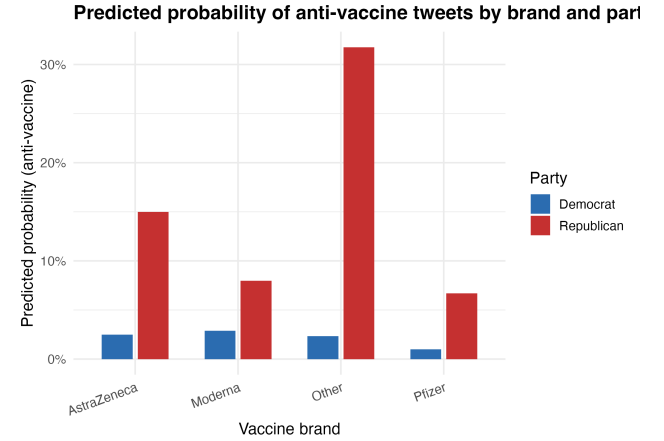


Figure 9: Predicted probability of anti-vaccine tweets by vaccine brand and party affiliation. Predicted values are derived from the logistic regression model in Table 1, holding embedding-based language similarity measures at their sample means.

To aid interpretation of the logistic regression coefficients in Table 1, Figure 9 presents predicted probabilities of an anti-vaccine stance by party and vaccine brand. These predictions translate the log-odds estimates into substantively meaningful probability differences, holding all embedding-based language similarity measures constant at their sample means.

The predicted probabilities reveal a pronounced partisan divide across all vaccine categories. Republican-authored tweets consistently exhibit substantially higher predicted probabilities of being anti-vaccine than Democratic-authored tweets, regardless of the vaccine brand mentioned. Among Democratic lawmakers, predicted anti-vaccine probabilities remain near zero across all brands, typically below 3%, reflecting their overwhelmingly pro-vaccine or neutral messaging.

In contrast, predicted probabilities for Republican tweets vary considerably by brand. The highest predicted prob-

ability appears in tweets referencing vaccines generically (“Other”), where the likelihood of an anti-vaccine stance exceeds 30%. This pattern aligns with earlier descriptive findings that non-branded vaccine discourse is more politicized and more likely to invoke mandates, skepticism, or misinformation. Brand-specific discussions, particularly those referencing Pfizer and Moderna, are associated with lower predicted anti-vaccine probabilities, suggesting that explicit references to named vaccines tend to anchor discourse in more factual or medical contexts.

Overall, Figure 9 reinforces the regression results by showing that the partisan gap in vaccine messaging is not only statistically significant but also substantively large. Even after accounting for linguistic cues captured by embedding similarities, party affiliation and the framing context implied by vaccine branding jointly shape the probability that a tweet expresses anti-vaccine sentiment.

4. Discussion

4.1. Partisan Divergence in Vaccine Narratives

This analysis demonstrates that Republican and Democratic members of Congress presented sharply different public narratives about COVID-19 vaccines on Twitter. Using a classification-based content analysis, we found that Republicans were far more likely to express skepticism or opposition to vaccines and to frame the issue in terms of personal liberty or misinformation, whereas Democrats almost uniformly promoted vaccines with a public health rationale. These findings align with expectations from political communication research on partisan responses to health crises: once the topic became politicized, Republican elites largely adopted a critical stance toward the pro-vaccine position championed by Democrats and public health authorities.

This partisan divergence in messaging is likely to have had real-world consequences. Prior work suggests that elite cues play an important role in shaping public vaccine attitudes and behaviors, particularly in highly polarized informational environments. The stark contrast in congressional messaging documented here may therefore have contributed to broader polarization in public perceptions of COVID-19 vaccination.

4.2. Limits of Naive Embedding-Based Measures

When we attempted to quantify these narratives using embedding-based semantic measures, we observed divergent results that carry important methodological implications. The initial embedding approach—measuring cosine similarity to single-word concept anchors—largely failed to distinguish between Republican and Democratic discourse. For example, considering only the frequency or semantic proximity of words such as “freedom” or “hoax” produced a misleading picture in which the two parties appeared more similar than they actually were.

This result highlights a key limitation of simplistic computational metrics: nuanced framing differences can elude basic embedding approaches even when they are readily apparent to human interpretation—or, in our case, to an ad-

vanced LLM acting as a proxy for human judgment. Embeddings capture surface-level semantic similarity, but they often fail to encode the pragmatic intent behind word usage.

4.3. Context, Intent, and Framing Ambiguity

There are several reasons why embedding-based similarity struggled to match the clarity of the classification results. First, framing differences often emerge through context and emphasis rather than exclusive vocabulary. Both pro-vaccine and anti-vaccine tweets may invoke concepts such as “freedom,” but with opposing meanings: for example, “Vaccination will bring us freedom from COVID” versus “We demand freedom from vaccine mandates.” A generic embedding model treats both usages as similar, even though their political implications are fundamentally different.

Similarly, references to a “vaccine hoax” may appear in tweets either to debunk misinformation or to propagate it. Embedding similarity to the word “hoax” is high in both cases, despite the divergent intent. In short, semantic similarity does not equate to semantic intent. Embeddings register which words appear in a text, but not why they are used or how they function within a broader argumentative structure.

Second, there is substantial vocabulary overlap across frames. Lawmakers from both parties routinely used common terms such as “vaccine,” “COVID,” “virus,” and “safety,” albeit embedded in different narratives. The concept-vector approach necessarily compresses a complex linguistic space into a limited number of dimensions, which inevitably discards information. Our initial single-anchor approach was especially noisy: Democrats and Republicans appeared equally close to the “hoax” concept, not because they shared conspiratorial beliefs, but because mentioning misinformation (even to refute it) registers similarly in embedding space.

4.4. Refining Embeddings: Multi-Anchor and Data-Driven Approaches

Refining the embedding approach improved its alignment with the classification-based findings. The multi-anchor method, which combined several related terms for each narrative, provided a richer semantic target and reduced noise from individual words. The data-driven politicization axis further enhanced separation by learning the composite distinction between politicized and medicalized language directly from the data.

These refinements produced results consistent with the qualitative patterns identified earlier. For example, Republicans scored higher on the data-driven politicization scale, while Democrats clustered closer to medical and scientific language (Appendix Table A1). Nonetheless, even the best-performing embedding measure revealed a smaller partisan gap than the categorical frame analysis.

One explanation is that Republican vaccine discourse often blended different rhetorical elements within the same tweet. A single message might praise vaccine development while simultaneously criticizing mandates. Embeddings reflect this mixture as an intermediate score, whereas the classifier assigns a dominant frame, yielding a starker contrast.

Thus, classification highlights the presence of politicized elements, while embeddings treat framing as a continuum. Both approaches capture distinct but complementary aspects of political communication.

4.5. Implications of the Regression Analysis

The logistic regression analysis illustrates the value of combining classification and embedding-based features. Even after controlling for party affiliation, semantic similarity to conspiracy-oriented and anti-mandate narratives strongly predicted anti-vaccine stance. This finding suggests that degrees of misinformation-oriented language matter, even when tweets are not overtly labeled as anti-vaccine. Embedding-based features therefore add explanatory power by capturing latent linguistic signals that categorical labels alone may miss.

4.6. Limitations

Several limitations should be noted. First, our dataset is confined to the 2020–2022 period and relies on keyword filtering, which may miss indirect references to vaccines or evolving rhetoric in later stages of the pandemic. Patterns may differ during booster campaigns, under new variant conditions, or across other platforms such as Facebook or congressional floor speeches.

Second, the LLM-generated frame and stance labels, while generally accurate, are not infallible. Sarcasm, parody, and subtle dog-whistle rhetoric may be misclassified, and borderline cases depend on model interpretation. Different prompts or models could yield slightly different labels. On the embedding side, we relied on general-purpose language embeddings; models fine-tuned on political or public health discourse might better capture domain-specific nuances. Our construction of concept vectors was also somewhat ad hoc, and more systematic approaches—such as supervised dimension reduction—could improve performance.

4.7. Broader Implications and Future Research

Despite these limitations, our findings carry important substantive and methodological implications. Substantively, they underscore how COVID-19 vaccination became deeply politicized in elite discourse, with likely downstream effects on public opinion and behavior. Future work could link elite messaging trends to vaccination uptake or public trust measures to better assess causal influence.

Methodologically, this study shows the risks of relying on naive text similarity measures to study nuanced framing. Off-the-shelf embeddings may suggest similarity where meaningful differences exist. Incorporating context—through advanced models such as GPT-4 or carefully designed hybrid approaches—is essential. Future research might explore LLM-assisted feature engineering, human-in-the-loop concept refinement, or interactive modeling frameworks to better capture political meaning.

4.8. Conclusion

In conclusion, this study demonstrates how U.S. members of Congress diverged sharply in their portrayal of COVID-19 vaccines on Twitter. Classification-based content analysis revealed clear partisan narrative splits, while

embedding-based approaches required refinement to approximate those patterns. Together, these methods illustrate both the promise and the pitfalls of modern NLP tools for studying political communication. We argue for cautious optimism: with thoughtful design and hybrid methodologies, computational text analysis can meaningfully map the evolving tension between science and politics in public discourse.

A Appendix: Additional Embedding-Based Results

A1. Data-Driven Politicization Scores by Party

Table 2 reports the average politicization scores for Democratic and Republican tweets using the data-driven embedding approach described in the Methods section. These scores place each tweet on a continuous semantic spectrum ranging from medical/scientific language (0) to politicized language (1), where politicization reflects proximity to Mandate/Rights and Conspiracy/Misinformation narratives.

Table 2: Average Politicization vs. Medicalization Semantic Scores by Party (Data-Driven Embedding Approach)

Party	Average Politicization Score
Democrat	0.32
Republican	0.47

The results show a clear partisan difference in narrative style. On average, Republican tweets exhibit substantially higher politicization scores (0.47) than Democratic tweets (0.32), indicating greater alignment with politicized frames emphasizing government mandates, individual rights, and conspiratorial themes. Democratic tweets, by contrast, cluster closer to the medical/scientific end of the spectrum, reflecting language focused on safety, efficacy, and public health.

Importantly, this partisan gap emerges despite the fact that the embedding measure averages over all words in each tweet rather than isolating a single dominant frame. As a result, the magnitude of the difference in politicization scores is more moderate than the stark contrasts observed in the categorical frame coding, where Republican tweets frequently fell into Mandate/Rights or Conspiracy frames and Democratic tweets almost never did. The continuous embedding measure therefore captures the *degree* of politicization rather than its presence or absence.

This finding helps reconcile the classification-based and embedding-based results. While categorical frame labels highlight sharp narrative divergence by assigning each tweet a dominant frame, the data-driven embedding approach reveals that many tweets—particularly among Republicans—blend medical and political language within the same message. For example, a tweet may acknowledge vaccine development while opposing mandates, yielding an intermediate politicization score. The embedding therefore reflects hybrid rhetoric that categorical approaches must collapse into a single label.

Taken together, Table 2 provides quantitative support for the broader claim that Republican vaccine discourse was more politicized on average than Democratic discourse, even when accounting for overlapping vocabulary and mixed framing. This convergence across methods strengthens confidence in the substantive conclusion that partisan elites differed not only in their stance toward vaccines but also in the underlying narrative logic through which they communicated about them.

B Acknowledgements

Place all acknowledgements (including those concerning research grants and funding) in a separate section at the end of the paper.

C Bibliographical References

- Engel-Rebitzer, P., Lee, J. K., and Fowler, E. F. (2022). Elite messaging and vaccine hesitancy: Congressional communication on twitter during covid-19. *Health Communication*, pages 1–13.
- Gadarian, S. K., Goodman, S. W., and Pepinsky, T. B. (2021). Partisanship, health behavior, and policy attitudes in the early stages of the covid-19 pandemic. *PLoS ONE*, 16(4):e0249596.
- Gollust, S. E., Nagler, R. H., and Fowler, E. F. (2020). The emergence of covid-19 in the u.s. media: A public health crisis framed as a political issue. *American Journal of Public Health*, 110(9):1372–1379.
- Jiang, J., Fan, L., and Yang, S. (2021). Political polarization in vaccine discourse on social media. *Journal of Medical Internet Research*, 23(11):e26472.
- Kaiser Family Foundation. (2021). Kff covid-19 vaccine monitor: January 2021. <https://www.kff.org>. Accessed 2024-10-01.
- Motta, M., Stecula, D., and Farhart, C. (2020). How right-leaning media coverage of covid-19 facilitated the spread of misinformation in the early stages of the pandemic. *Canadian Journal of Political Science*, 53(2):335–342.