Paid Voices vs. Public Feeds: Interpretable Cross-Platform Theme-Based Analysis of Climate Discourse

Anonymous ACL submission

Abstract

Climate discourse online shapes public understanding of climate change and informs political and policy debate, yet it unfolds across structurally different environments: paid advertising platforms host targeted, institutionally produced messaging, while public social media reflects largely organic, user-driven discussion. We present a comparative analysis of climate discourse across paid advertisements on Meta (previously Facebook) and public posts on Bluesky from July 2024 to September 2025. To support it, we develop an interpretable thematic discovery pipeline that clusters texts by semantic similarity and uses large language models (LLMs) to label clusters with concise, human-interpretable themes, requiring **no** predefined topic inventory or seed set. Using these themes, we find the two environments diverge systematically: paid advertising centers on strategic promotion of specific solutions in a formal, forward-looking register, whereas organic discourse centers on systemic critique in a crisis-oriented, scientifically grounded one. We also evaluate the utility of the discovered themes through downstream *stance prediction* and *theme-guided retrieval* tasks. While our analysis focuses on climate communication, the framework generalizes to comparative thematic analysis across heterogeneous communication environments.

1 Introduction

Online platforms have become a primary arena for climate communication, shaping how climate change is discussed, debated, and acted upon by the public (Pera and Aiello, 2024; Islam et al., 2023b; Bloomfield and Tillery, 2019; Pearce et al., 2019; Abbar et al., 2016). Climate narratives circulated through social media influence political attitudes, policy preferences, and perceptions of scientific credibility (Liu et al., 2025; Vivion et al., 2024; Holder et al., 2023; Treen et al., 2020; Lejano and Nero, 2020; Walter et al., 2018; Gupta et al., 2018), making climate discourse a central object of study for computational social science (CSS) and natural language processing (NLP).

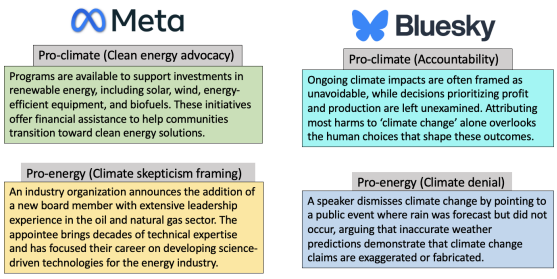


Figure 1: Examples of climate discourse across paid and public platforms (paraphrased and anonymized).

However, climate communication does not unfold in a uniform online environment. Paid advertising ecosystems such as Meta's advertising platform (Meta Platforms, Inc., 2025) are designed for strategic, targeted persuasion in service of political or economic goals (Islam et al., 2023b,a; Islam and Goldwasser, 2022; Capozzi et al., 2021; Ribeiro et al., 2019), while public social media platforms such as Bluesky (Bluesky Public Social Network, 2025) host largely user-driven discourse shaped by personal expression and grassroots interaction (Quelle and Bovet, 2024). We treat paid advertising as a proxy for institutionally produced, resource-backed messaging and public posting as a proxy for organic, user-driven discourse; Fig. 1 illustrates the contrast (motivation only—all empirical claims come from later analyses). Prior computational works have examined climate discourse within individual platforms, focusing on tasks such as topic modeling (Dahal et al., 2019), stance detection (Upadhyaya et al., 2023; Luo et al., 2020), and targeted messaging analysis (Islam and Goldwasser, 2025b,c,a; Islam et al., 2023b), leaving unclear whether thematic patterns are shared across communication environments or specific to where they emerge.

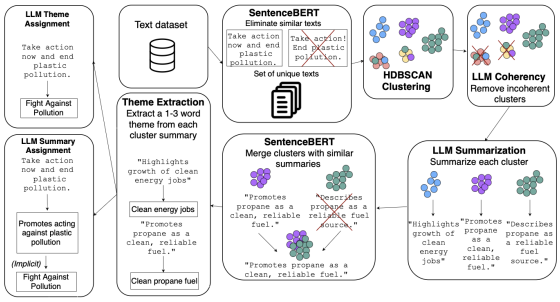


Figure 2: Overview of our framework.

In this paper, we conduct a comparative analysis of climate discourse across paid advertisements on Meta and public posts on Bluesky. Rather than proposing a new topic modeling algorithm, we develop a measurement framework that supports cross-platform thematic analysis despite differences in vocabulary, style, and communicative norms. Our approach first discovers latent semantic clusters and then uses large language models (LLMs) (Brown et al., 2020) to generate concise, human-interpretable theme labels. By separating semantic discovery from linguistic interpretation, the framework keeps themes grounded in the data while making them accessible for substantive analysis and downstream use. It also avoids reliance on predefined topic inventories or seed sets (Fig. 2). We evaluate the resulting themes against standard topic modeling baselines—LDA (Blei et al., 2003), BERTopic (Grootendorst, 2022), and TopicGPT (Pham et al., 2024)—using both human judgments and an LLM-based evaluator, and probe their utility on two downstream tasks, *stance prediction* and *theme-guided retrieval*, as validation probes for whether the themes are coherent semantic abstractions. We then use the themes to examine thematic prevalence across platforms and how discourse shifts around major real-world events. Our goal is not to make causal claims about platform effects or speaker intent, but to descriptively characterize how climate discourse differs across two communication environments.

Across our analyses, we find that the two environments reflect distinct communicative orientations. Paid advertising concentrates on the strategic promotion of specific solutions, whereas organic discourse concentrates on the systemic critique of corporate and political actors. This contrast recurs across thematic structure, rhetorical register, and stance alignment, and the two platforms respond to major events in opposite directions: advertising contracts following the 2024 U.S. election while public discourse expands.

In summary, this paper makes three primary contributions[1]:

(1) We construct and release the first multi-platform dataset enabling direct comparison between paid climate advertising and public climate discourse.

(2) We develop an interpretable thematic discovery pipeline for cross-platform thematic analysis that does not require predefined topic inventories or seed themes.

(3) We conduct quantitative and qualitative analysis. Our findings show that paid and public climate discourse differ systematically in thematic structure, stance alignment, and temporal responsiveness, illustrating the value of comparative thematic analysis across heterogeneous communication environments.

2 Related Work

Climate communication and political advertising. Prior works examine how climate and political campaigns operate in paid advertising ecosystems. Studies analyze messaging strategies, stance, and targeting in climate-related ads using weak supervision, graph-based advertiser–ad models, and LLM-based analysis (Islam and Goldwasser, 2025b; Rowlands et al., 2024; Islam et al., 2023b,a). The most closely related concurrent work is (Anonymous, 2026), which analyzes cross-platform climate discourse by extracting and classifying claims using a Toulmin-inspired argumentative framework (Toulmin, 1958). In contrast, our work takes a coarser-grained, theme-level perspective, examining how climate messaging strategies rather than individual argumentative moves differ across platform contexts. These studies motivate treating paid advertising as a distinct, strategically and institutionally produced communication environment.

LLMs for theme and narrative discovery. Recent work uses LLMs to discover interpretable latent structure in large corpora (Islam, 2026a; Brady and Islam, 2025; Nakshatri et al., 2025). Islam and Goldwasser (2025a) cluster documents from an initial seed and use LLMs to summarize clusters and label themes; others uncover finer-grained units such as arguments from a predefined theme set (Islam and Goldwasser, 2025c), and concept-induction methods like LLooM (Lam et al., 2024) incorporate human-provided seed concepts. In

[1] https://osf.io/csp3x/overview?view_only=85708d98978b4c7094e47f8604dd5e1a

contrast, our pipeline performs thematic discovery **without** any seed set.

Framing analysis of climate discourse. Framing theory (Entman, 1993) describes how communicators emphasize certain aspects of reality to promote particular framing functions. Computational work operationalizes framing through cross-issue codebooks and annotated corpora (Boydstun et al., 2014; Card et al., 2015) and (weakly) supervised frame classification (Roy and Goldwasser, 2020; Field et al., 2018). In the climate domain, framing analyses examine climate and energy policy narratives (Vikström et al., 2023) and how multimodal climate communication shapes public understanding (Dancygier, 2023). We adopt Entman's four functions across two communicative environments.

Linguistic and psycholinguistic analysis. Beyond framing, computational social science has used psycholinguistic features—particularly LIWC (Boyd et al., 2022)—to characterize how climate discourse varies across regional and organizational contexts (Wuraola et al., 2023; Gulliver et al., 2021). We extend this work by applying framing lexicons and LIWC variables to enable theme-level rhetorical-register comparison across paid and organic communication.

Stance detection. Stance detection is widely used to analyze climate narratives and ideological framing (Luo et al., 2020), including detection of denial and neutralization techniques in climate skepticism (Bhatia et al., 2021). This motivates our stance-based validation: meaningful themes should align systematically with stance labels.

Cross-platform discourse analysis. Recent datasets enable comparison of political advertising across platforms such as Meta and Google (Arenas et al., 2025), and other work tracks how climate communication shifts across ecosystems such as YouTube and TikTok (Pera and Aiello, 2024) or within decentralized platforms like Bluesky (Quelle and Bovet, 2024). However, paid advertising and organic discourse are typically studied separately. We bridge this gap by jointly analyzing themes, framing, register, and temporal dynamics in Meta ads and Bluesky posts.

3 Dataset

We analyze climate discourse across two structurally distinct platforms. The Meta Ad Library contains paid advertisements designed for strategic persuasion and includes disclosed advertiser identities and metadata (e.g., impressions, spend, targeting). In contrast, Bluesky is a decentralized social media platform where climate-related content is generated organically by users without monetary promotion or targeting, reflecting grassroots discussion and personal expression.

We collect[2] a set of over $113K$ climate-related ads from Meta Ad Library and $1.3M$ Bluesky climate posts from July 2024 to September 2025 using a curated set of climate-related keywords shown in Table 4 App. A. The climate keywords are taken from a selection constructed by Islam et al. (2023b) and modified to reflect recent shifts in climate-change related topics. Using the Meta Ad Library, a total of $113,493$ climate related ads are collected, which is then reduced to a dataset of size $17,026$ upon using SentenceBERT (Reimers and Gurevych, 2019) with a cosine similarity threshold of 0.80 to eliminate highly similar ads.

Bluesky, due to its nature as a social networking platform, has many more posts (≈ 1.3 million) that match our keywords within the same time frame. From the obtained post database, we randomly sample $20,000$ posts to serve as a similarly-sized comparison to the retrieved Meta dataset. We furthermore apply the same deduplication procedure, yielding $19,182$ unique Bluesky posts. We apply a greedy deduplication strategy for both Meta and Bluesky: for each text, all subsequent texts exceeding the similarity threshold are removed, retaining the first occurring text. The Bluesky sample maintains temporal representativeness of the full corpus. Across all months, the absolute difference between population and sample proportions $\leq 0.39\%$. Furthermore, human validation on 500 ads and 500 posts confirms 94.6% are climate-relevant (App. A).

4 Methodology

We develop a multi-stage pipeline for discovering the thematic structure of climate-related social-media posts and advertisements (Fig. 2). This process is done independently per-platform as described below.

We first embed texts into a shared semantic space using SentenceBERT (all-MiniLM-L6-v2) with L2 normalization, followed by dimensionality reduction via Principal Component Analysis (PCA) and Uniform Manifold Approximation and Projection

[2]The social media data used in this study may contain offensive, biased, toxic, or harmful language.

Method	Meta		Bluesky	
	Human	LLM	Human	LLM
LDA	0.25	0.43	0.35	0.44
BERTopic	0.25	0.37	0.26	0.25
TopicGPT	0.95	0.96	0.93	0.88
LLM Summary	0.45	0.60	0.62	0.46
LLM Theme	0.67	0.46	0.67	0.51

Table 1: Accuracy comparison across methods for Meta and Bluesky data under human and LLM evaluations.

(UMAP) (McInnes et al., 2018). Density-based clustering (HDBSCAN) (McInnes et al., 2017) is applied to group semantically similar texts. To characterize each cluster, we rank members by their probability scores and extract the top-k (k=5) texts from each cluster as representatives for downstream processing. Because automatically discovered clusters may be noisy, we filter them for semantic coherence with an LLM, prompting it to judge whether each cluster's representative texts reflect a single underlying theme (few-shot details in App. B.1). Coherence filtering retains 69.1% of the 97 initial Meta clusters and 52.6% of the 97 Bluesky clusters.

For each retained cluster we generate a concise LLM summary capturing its shared theme and intent, using the five texts with highest probability score as few-shot examples (App. B.4 Fig. 10). We then merge clusters whose summaries exceed a tuned embedding-similarity threshold and assign each merged cluster a short (1–3 word), human-interpretable theme label (merging and labeling details in App. B.2, App. B.3).

Finally, texts are assigned to themes via two complementary strategies: (i) matching a text to its most similar cluster summary and inheriting that summary's theme, and (ii) directly assigning the text to one of the generated theme labels. This yields **47** themes for Meta (after de-duplication; App. B.3) and **46** for Bluesky (Table 6, App. B.5).

All LLM steps (coherence filtering, summarization, theme labeling) use Mistral-Large-Instruct-2407 with few-shot prompting, chosen for its strong performance on structured annotation tasks among comparably sized open-weight models within our compute budget (Islam and Goldwasser, 2025b). Full prompts and implementation details are in App. B.4, App. B. We compare our two assignment variants against three topic-modeling baselines: LDA, BERTopic, and TopicGPT. Assignments are evaluated by both human annotators (App. C.3) and an LLM judge (Fig. 14, App. B.4).

5 Results

5.1 Results: Meta

For each ad, we compare theme assignments produced by our two assignment strategies of Text-to-Theme and Text-to-Summary against keywords inferred by three baseline methods: LDA, BERTopic, and TopicGPT. We randomly sample 500 Meta ads and assess the correctness of each method's assigned theme or keyword set. Each sample is independently evaluated by (1) human annotators (App. C.3) and (2) an LLM-based judge (Qwen3-235b-a22b) (Yang et al., 2025). The prompt given to the LLM judge is shown in Fig. 14 App. B.4. Annotators determine whether the assigned theme or keyword set accurately reflects the central topic and framing of the ad (App. C.3). Representative annotation examples are shown in Tables 7 and 8 App. C.1. We quantify the level of agreement between the two judges by computing Cohen's Kappa (Cohen, 1960) across both evaluation conditions. For summary-based judgments, we observe 79.23% raw agreement ($\kappa = 0.58$) between the human and LLM annotator. For theme-based judgments, we see 77.42% agreement ($\kappa = 0.50$). Both Cohen's Kappa scores fall in the moderate agreement range, which is consistent with what is typically reported for semantic-level annotation tasks of this nature. Results are shown in Table 1.

TopicGPT achieves high surface-level accuracy (96% for the LLM judge) However, this result is largely driven by the generality of the topics it produces (e.g., "Environment," "Energy," "Climate Change"), which broadly match most posts in a climate-focused corpus. This is reflected in a low Cohen's κ of 0.19, indicating only slight agreement once chance overlap is accounted for. However, there is 93.19% raw agreement for this dataset. The near-collapse of Kappa to Slight agreement despite high raw agreement score is attributable to the highly skewed label distribution: 94.99% of human annotations and 96.19% of LLM annotations are marked "True". In contrast to TopicGPT, our methods produce more specific and semantically targeted themes ("Water Resilience," "Climate Legal Accountability," and "Capitalism's Climate Doom") while achieving higher accuracy than LDA and BERTopic across both evaluation settings. App. C.1 provides additional details.

5.2 Results: Bluesky

We evaluate theme assignment on a randomly sampled set of 500 Bluesky posts using both human and LLM-based judges. Across both evaluation settings (Table 1), our methods outperform both LDA and BERTopic, with our Text-to-Theme variant achieving the highest accuracy among our methods. TopicGPT again achieves high surface-level accuracy (88% for LLM judge) due to its broader topic labels. App. C.2 provides a detailed discussion of systematic differences between human and LLM-based judgments on the Bluesky dataset.

6 Analysis

Through our analysis, we explore how paid advertising on Meta and organic discourse on Bluesky reflect fundamentally different communicative orientations: strategic persuasion and systemic critique. We see these orientations manifest themselves consistently across thematic structure (§6.1), rhetorical register (§6.2), stance alignment (§6.3), and responsiveness to events (§6.4).

6.1 Thematic Structure

To enable direct cross-platform comparison, we re-run our pipeline on the joint combination of all Meta and Bluesky texts (36,206 total texts). Modified experimental details are provided in App. D. The resulting 354 coherent themes split into three categories based on within-theme platform composition (70% threshold): 145 Meta-dominant, 112 Bluesky-dominant, and 97 shared themes.The top Meta-dominant themes by cluster size are *Solar incentive urgency* (296 Meta texts), *Offshore wind momentum* (248), and *Grassroots solar empowerment* (259)—all renewable-energy promotion (Fig. 3). Meta texts are distributed relatively evenly across the top themes, forming a diverse set of topics for campaign.

The Bluesky-dominant distribution differs greatly from the Meta-dominant distribution. A single theme, *Corporate climate sabotage*, accounts for 5,599 Bluesky texts; This number is larger than the entire top-four Meta-dominant cluster size combined (Fig. 4). The remaining top Bluesky-dominant themes (*Greenwashing exposed*, *Anti-capitalist resistance*, etc.) are organized around systemic critique of corporate, political, and economic actors. The full theme-size distribution is found in App. D.1 Fig. 19. Topically, we see that Meta-dominant clusters focus on a diverse set of promotional topics such as *Affordable solar adoption*, whereas Bluesky topics are heavily concentrated on a few critique-oriented themes. We include a list of shared themes for Bluesky and Meta in App. D.1 (Fig. 18), where each cluster in this category is neither predominantly comprised of Meta or Bluesky texts. A supplementary named-entity analysis (App. E.2) further reinforces this topical contrast: Meta-dominant clusters emphasize U.S. states and named individuals consistent with targeted promotional campaigns, whereas Bluesky-dominant clusters emphasize countries and collective actors consistent with larger-scale critique and structural framing.

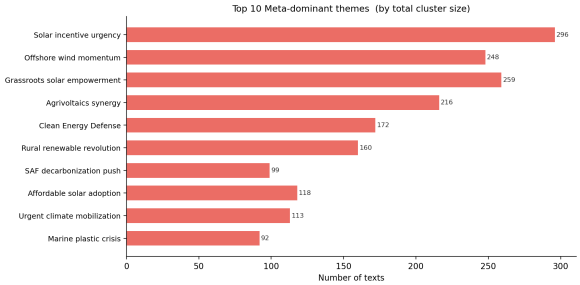


Figure 3: Top 10 themes (joint-clustering) which are Meta-Dominant, meaning that over 70% of texts in each given theme cluster is a Meta theme.

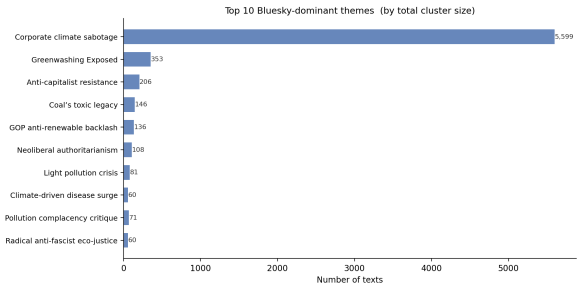


Figure 4: Top 10 themes (joint-clustering) which are Bluesky-Dominant, meaning that over 70% of texts in each given theme cluster is a Bluesky theme.

6.2 Rhetorical Register

We next characterize *how* the two platforms talk about climate using two different lenses: an Entman framing lexicon (Entman, 1993) operationalizing the four framing functions across eight subcategories (crisis/solution, institutional/individual actor, economic/moral, scientific authority/populist skepticism), and the Linguistic Inquiry and Word Count (LIWC-22) framework (Boyd et al., 2022) summary variables. Lexicon construction, cluster-

and demographic-level breakdowns, and theme-level LIWC profiles are provided in App. F. Together, these lenses show an asymmetric register. Paid communication leans formal, authoritative, and solution-oriented; Organic discourse leans personal, crisis-oriented, and scientifically grounded.

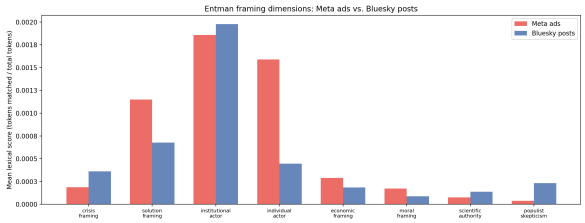
Framing. Fig. 5 (a) presents mean lexical framing scores across the eight subcategories. Meta ads score substantially higher on solution and individual-actor framing, consistent with calls-to-action targeting the individual viewer. Bluesky posts score higher on crisis framing, scientific authority, and (apparent) populist skepticism—the last likely reflecting criticism *of* skeptics rather than skepticism itself. Cluster-level analysis (App. F) shows Bluesky-dominant clusters peak on emotional urgency (0.79) and institutional actor (0.71) framing, while Meta-dominant clusters score lowest on scientific framing (0.33) and highest on emotional hope (0.75) and future framing (0.70), suggesting paid climate advertising favors emotive and future-focused persuasion as opposed to evidence-heavy rhetoric.

LIWC. Fig. 5 (b) shows Mann-Whitney U comparisons across four LIWC summary variables, with effect sizes via rank-biserial correlation (r). Meta ads score significantly higher on *Analytic* ($r = -0.189$, $p < .001$) and *Clout* ($r = -0.440$, $p < .001$), indicating a more formal, categorical, and authoritative register. The *Clout* difference is the largest effect observed. By contrast, Bluesky posts score higher on *Authentic* ($r = +0.103$, $p < .001$). Meta ads also exhibit more positive *Tone* ($r = -0.310$, $p < .001$), with a median near 75 compared to roughly 52 for Bluesky.

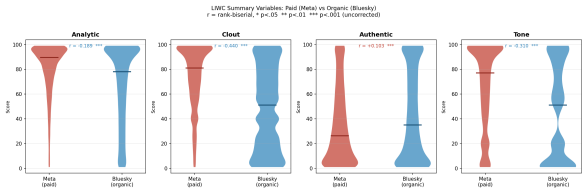
6.3 Stance Alignment

If induced themes capture meaningful discourse structure, they should be stance-aligned. We manually annotate 500 Meta ads and 500 Bluesky posts as *Pro-Climate*, *Pro-Energy*, or *Neutral*, with breakdowns found in App. G.2.

Fig. 6 presents theme–stance correlations on Meta for LDA and our Text-to-Theme variant; full panels for BERTopic, Text-to-Summary, and Bluesky are provided in App. G.2. The LDA baseline produces diffuse, weak correlations across stances. Our themes produce markedly more concentrated correlations: Pro-Climate themes align with Pro-Climate stance, and Pro-Energy themes with Pro-Energy stance. Bluesky correlations are

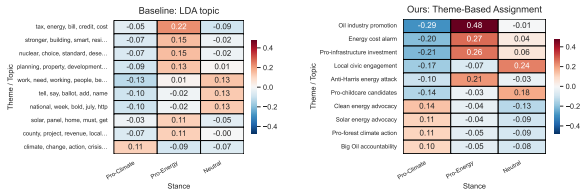


(a) Mean Entman framing scores by platform.



(b) LIWC summary variables by platform.

Figure 5: Rhetorical register comparison: paid climate communication is forward-looking, authoritative, and emotionally positive; organic discourse is crisis-oriented, scientifically grounded, and affectively mixed.



(a) **Baseline**: LDA. (b) **Ours**: LLM Theme.

Figure 6: Theme–stance correlation on Meta. Our themes produce concentrated, interpretable stance alignment, while baselines yield diffuse correlations.

constrained by the platform's stance imbalance but remain directionally consistent, whereas LDA baseline correlations are more sensitive to noise. Overall, we find the themes encode stance-relevant structure, and the platforms differ not only what they discuss, but also in stance-balance of their discourse.

6.4 Temporal Responsiveness

Finally, we examine how the two platforms respond to external climate-relevant events: the *November 5, 2024 U.S. presidential election* and the *January 2025 Southern California wildfires*. The two platforms respond to the same events in opposite directions: Paid climate advertising contracts post-election and shifts toward attribution-heavy, less future-oriented framing; Organic discourse expands. These divergent responses are consistent with the platform orientations established in §6.1. We report text-volume shifts here; corresponding spend, impression, and detailed LIWC dynamics

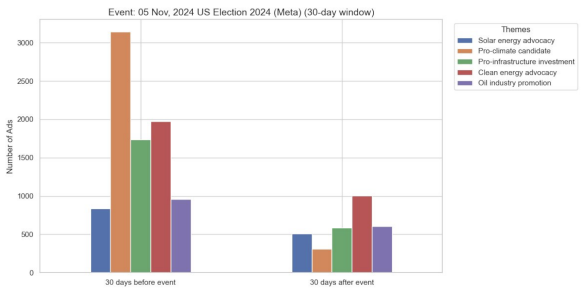
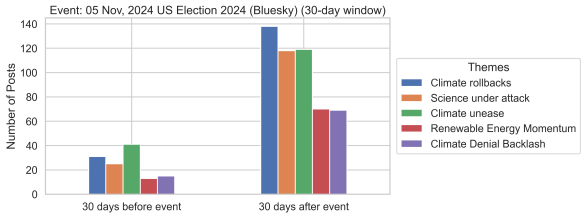


Figure 7: Top Meta themes 30 days before and after the November 2024 U.S. election. All five top-shift themes contract post-election.

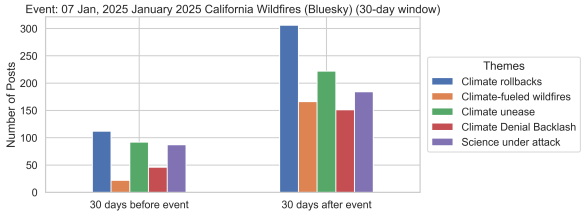
are provided in App. G.4 and App. G.5.

Meta contracts. As shown in Fig. 7, the 2024 U.S. election is followed by a sharp decline in climate-related advertising across the five themes with the largest pre/post shifts. These five themes span both Pro-Climate topics, such as clean-energy advocacy, and Pro-Energy topics, such as oil-industry promotion. This suggests that the post-election pullback is not necessarily stance-specific. Linguistically, we see that *Cause* language increases significantly post-election, whereas *moral*, *risk*, and *focusfuture* all decline (App. G.4 Fig. 33). The urgency and consequence-driven appeals used to promote candidates pre-election are dampened once the electoral cycle closes. Interestingly, a decreased *power* and increased *tentat* suggest uncertainty post-election. The 2025 wildfire window produces a different linguistic signature: *power* language characteristic of heightened institutional action rises sharply, while *Tone* becomes more negative (App. G.5).

Bluesky amplifies. Fig. 8 illustrates the inverse pattern. Within the ±30-day window around the 2024 U.S. election, every one of the top-shift themes—*Climate rollbacks*, *Science under attack*, *Climate unease*, *Renewable Energy Momentum*, and *Climate Denial Backlash*—rises sharply post-election. The same expansion pattern recurs around the 2025 California wildfires, with *Climate rollbacks* and *Science under attack* again among the steepest risers. Linguistically, the two Bluesky events diverge (App. G.3 Fig. 31). Around the election, LIWC register variables (*Analytic*, *Clout*, *Authentic*, *Tone*) remain stable, and only *cause* shifts significantly. Around the wildfires, *Tone* declines and *cause* increases, indicating that organic discourse around this acute disaster is both more negatively valenced and more oriented toward attribution and responsibility.



(a) Top Bluesky themes ±30 days around U.S. election.



(b) Top Bluesky themes ±30 days around CA wildfires.

Figure 8: Thematic shifts on Bluesky surrounding major events: rapid emergence of critique-oriented themes that were absent pre-event.

7 Downstream Tasks

We evaluate the utility of generated themes on two downstream tasks:
(i) A classification task where models predict stance (e.g., *Pro-Climate*, *Pro-Energy*, and *Neutral*) based on content. We compare performance using (a) text-only, (b) theme-only, and (c) combined features.
(ii) A retrieval task where themes act as queries to retrieve semantically relevant posts using dense embedding similarity. Precision@k is used to assess how often retrieved posts reflect theme-relevant content, using automatic keyword overlap for evaluation.

7.1 Stance Prediction

Stance Prediction is designed as a three-way classification task. We use **stratified** sampling to construct train, validation, and test splits (80/20 train–test, then 80/20 train–validation), preserving stance label distributions. Across model families and platforms, incorporating themes provides complementary signals and yields stronger gains in more balanced settings (e.g., Meta), while remaining competitive under severe class imbalance (Bluesky), as shown in Table 2. Theme-only models perform competitively, indicating that themes capture stance-relevant signal beyond surface text. On Bluesky, where class imbalance is more severe, improvements are smaller but remain consistent

MODEL	META		BLUESKY	
	ACC	F1	ACC	F1
$LR_{\text{tf-idf}txt}$	**0.78**	0.53	0.89	0.46
$LR_{\text{tf-idf}thm}$	0.72	0.59	0.90	0.47
$LR_{\text{tf-idf}txt+thm}$	0.75	**0.61**	**0.91**	**0.52**
$RoBERTa_{txt}$	0.73	0.63	0.88	0.82
$RoBERTa_{thm}$	0.72	0.60	0.88	0.82
$RoBERTa_{txt+thm}$	**0.75**	**0.67**	0.88	0.82
$Kimi_{txt}$	0.85	0.74	0.84	**0.52**
$Kimi_{thm}$	0.80	0.63	0.83	0.41
$Kimi_{txt+thm}$	**0.86**	**0.75**	**0.86**	0.51
$Llama3_{txt}$	0.76	0.58	0.66	0.41
$Llama3_{thm}$	0.66	0.54	0.73	0.36
$Llama3_{txt+thm}$	**0.78**	**0.63**	**0.81**	**0.42**

Table 2: Contribution of theme to stance classification on test data. txt: text, thm: theme.

Platform	P@1	P@5
Meta	0.783	0.691
Bluesky	0.818	0.736

Table 3: Precision@k for retrieving texts.

for simpler models. Full experimental details are provided in the App. H.1.

7.2 Theme-Guided Retrieval

We assess the semantic alignment between LLM-derived themes and content (texts) via a retrieval-based evaluation. For each unique theme, we embed both the theme and all ads/posts using SBERT. To qualitatively assess the alignment between theme and text embeddings, we project their vector representations into 2D using t-SNE (van der Maaten and Hinton, 2008) (Fig. 37 App. H.2). Across both datasets, theme embeddings (×) cluster near semantically related texts (•), visually supporting their utility for semantic retrieval. Meta shows a denser, centralized theme-ad structure, suggesting more coordinated messaging; while Bluesky reveals broader dispersion, reflecting diverse, user-driven discourse (Fig. 37 App. H.2).

We then compute cosine similarity and retrieve the top-k most similar posts per theme. Relevance is evaluated by checking whether retrieved posts contain keywords extracted from the theme (simple frequency-based filtering over normalized tokens). Table 3 reports Precision@1 and Precision@5 as aggregate metrics. These results indicate that theme embeddings effectively retrieve semantically relevant content, even when evaluated with simple unsupervised metrics. Retrieved examples illustrate stylistic contrasts, i.e., Meta ads are coordinated and campaign-driven (Table 15 App. H.2), while Bluesky posts reflect diverse, user-driven discourse (Table 16 App. H.2).

We additionally perform small-scale human relevance assessment on a random sample of 20 themes (top-5 retrieved posts per theme) for each platform. Using a rubric (2 = clearly relevant, 1 = partially relevant, 0 = not relevant), the resulting P@5 (binary ≥ 1) is **0.60** for Meta and **0.73** for Bluesky, with mean relevance scores of **0.68** and **0.88**, respectively. For Meta, the result is comparable but more conservative than the keyword-based estimate reported in Table 3. Notably, the semantic P@5 for Bluesky closely matches the keyword-based estimate reported in Table 3, indicating that theme-guided retrieval captures semantic alignment beyond keyword overlap.

8 Conclusion

We present a comparative analysis of climate discourse across paid advertising and public social media, examining Meta advertisements and Bluesky posts from July 2024 to September 2025 through an interpretable thematic discovery framework that labels semantic clusters with concise, human-interpretable themes. Under both human and LLM-based evaluation, the induced themes are more coherent and stance-consistent than standard topic-modeling baselines, and downstream stance prediction and retrieval confirm they capture meaningful semantic structure rather than surface lexical patterns. Applying these themes, we find the two environments reflect fundamentally different communicative orientations. Paid advertising concentrates on the strategic promotion of specific solutions in a formal, future-focused register, while organic discourse concentrates on the systemic critique of corporate and political actors in a crisis-oriented, scientifically grounded one. This contrast propagates consistently across thematic structure and responsiveness to events—most strikingly, the two move in opposite directions around the 2024 U.S. election, where advertising contracts and public discourse expands. More broadly, this work demonstrates the value of interpretable theme modeling for comparative thematic analysis across heterogeneous communication environments. While our analysis focuses on climate discourse, the framework can support cross-platform studies of narratives in other domains.

9 Limitations

While our framework enables interpretable cross-platform analysis of climate discourse, several limitations remain. First, our study focuses on two platforms: Meta's advertising ecosystem and Bluesky as representative cases of paid and organic communication environments. Although these platforms exhibit clear structural differences, our findings may not directly generalize to other advertising systems or social media platforms due to different user bases, moderation policies, or affordances. We acknowledge the existence of confounding factors, and that the two platforms differ across many dimensions beyond simply paid versus organic discussion. These dimensions include user demographics, moderation policies, platform age, and more. To account for these compounds, we include supplemental demographic-level analyses on the Meta dataset, as well as note confounds for which no comparable data is available—such as per-user demographics on Bluesky—limit the interpretability of cross-platform comparisons, and that observed differences should not be attributed to the paid versus organic distinction alone.

Second, our thematic modeling pipeline relies on pre-trained large language models for cluster coherence filtering, summarization, and theme labeling. While we employ open-source LLMs to promote reproducibility and reduce cost barriers, these models may encode biases present in their training data (Islam, 2026b; Blodgett et al., 2020; Brown et al., 2020), which could influence theme interpretation. We do not explicitly measure or mitigate such biases in this work. Besides, we only used pre-trained LLMs and did not consider fine-tuning due to the resource constraints. Relatedly, our coherence filtering combines LLM-based judgments with quantitative cluster cohesion (mean pairwise cosine similarity), but the underlying model choices and thresholds were calibrated on our climate corpus; because LLM-based coherence judgments are model- and prompt-dependent, the procedure may require re-tuning to generalize to other domains, platforms, or datasets.

Lastly, **our analysis is descriptive rather than causal**. Although we identify systematic differences in thematic structure and temporal dynamics across platforms, we do not attribute intent to individual actors or infer causal effects of platform incentives on discourse. Future work could combine our framework with causal or experimental designs to further investigate these mechanisms.

10 Ethics Statement

To the best of our knowledge, we did not violate any ethical code while conducting the research work described in this paper. We report the technical details for the reproducibility of the results, as well as release the datasets and codes publicly. The author's personal views are not represented in any results we report, as it is solely outcomes derived from machine learning or AI models.

The social media data used in this study may contain offensive, biased, toxic, or harmful language. Such content reflects user-generated discourse and does not represent the views of the authors or institutions. All analyses were conducted for research purposes only.

We used LIWC-22 under a research license. Our use is consistent with its intended academic research purpose. Any derivatives or outputs involving LIWC features should not be used outside of research contexts.

References

Sofiane Abbar, Tahar Zanouda, Laure Berti-Equille, and Javier Borge-Holthoefer. 2016. Using twitter to understand public interest in climate change: The case of qatar. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 10, pages 168–177.

Anonymous. 2026. Cross-platform analysis of claims in climate discourse. Under review.

Andreu Arenas, Josep Domingo-Ferrer, Marta C. González, and 1 others. 2025. Comparable 2022 general election advertising datasets from meta and google. *Scientific Data*, 12(1):1–15.

Shraey Bhatia, Jey Han Lau, and Timothy Baldwin. 2021. Automatic classification of neutralization techniques in the narrative of climate change scepticism. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2167–2175.

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022.

Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. 2020. Language (technology) is power: A critical survey of "bias" in nlp. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5454–5476.

Emma Frances Bloomfield and Denise Tillery. 2019. The circulation of climate change denial online: Rhetorical and networking strategies on facebook. *Environmental Communication*, 13(1):23–34.

Bluesky Public Social Network. 2025. Bluesky social platform. https://bsky.app/.

R. L. Boyd, A. Ashokkumar, S. Seraj, and J. W. Pennebaker. 2022. *LIWC-22 Manual*.

Amber Boydstun, Dallas Card, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2014. Tracking the development of media frames within and across policy issues.

Alexander Brady and Tunazzina Islam. 2025. Latent topic synthesis: Leveraging llms for electoral ad analysis. *arXiv preprint arXiv:2510.15125*.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, and 1 others. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Arthur Capozzi, Gianmarco De Francisci Morales, Yelena Mejova, Corrado Monti, André Panisson, and Daniela Paolotti. 2021. Clandestino or rifugiato? anti-immigration facebook ad targeting in italy. In *CHI*.

Dallas Card, Amber Boydstun, Justin H Gross, Philip Resnik, and Noah A Smith. 2015. The media frames corpus: Annotations of frames across issues. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 438–444.

Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.

David R Cox. 1958. The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2):215–232.

Biraj Dahal, Sathish AP Kumar, and Zhenlong Li. 2019. Topic modeling and sentiment analysis of global climate change tweets. *Social network analysis and mining*, 9(1):24.

Barbara Dancygier. 2023. Multimodal media: Framing climate change. *Discourse Studies*, 25(2):220–236.

Robert M Entman. 1993. Framing: Toward clarification of a fractured paradigm. *Journal of communication*, 43(4):51–58.

Anjalie Field, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky, and Yulia Tsvetkov. 2018. Framing and agenda-setting in russian news: a computational analysis of intricate political strategies. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3570–3580.

Maarten Grootendorst. 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*.

Robyn Gulliver, Kelly S. Fielding, and Winnifred R. Louis. 2021. Assessing the mobilization potential of environmental advocacy communication. *Journal of Environmental Psychology*, 74:101563.

Kuhika Gupta, Joseph Ripberger, and Wesley Wehde. 2018. Advocacy group messaging on social media: Using the narrative policy framework to study twitter messages about nuclear energy policy in the united states. *Policy studies journal*, 46(1):119–136.

Faye Holder, Sanober Mirza, Jake Carbone, and Ruth E McKie. 2023. Climate obstruction and facebook advertising: how a sample of climate obstruction organizations use social media to disseminate discourses of delay. *Climatic Change*, 176(2):16.

Tunazzina Islam. 2026a. Reasoning-based refinement of unsupervised text clusters with llms.

Tunazzina Islam. 2026b. Who gets which message? auditing demographic bias in llm-generated targeted text. *arXiv preprint arXiv:2601.17172*.

Tunazzina Islam and Dan Goldwasser. 2022. Understanding covid-19 vaccine campaign on facebook using minimal supervision. In *2022 IEEE International Conference on Big Data (Big Data)*, pages 585–595. IEEE.

Tunazzina Islam and Dan Goldwasser. 2025a. Discovering latent themes in social media messaging: A machine-in-the-loop approach integrating llms. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 19, pages 859–884.

Tunazzina Islam and Dan Goldwasser. 2025b. Post-hoc study of climate microtargeting on social media ads with llms: Thematic insights and fairness evaluation. In *Findings of the Association for Computational Linguistics: EMNLP 2025*.

Tunazzina Islam and Dan Goldwasser. 2025c. Uncovering latent arguments in social media messaging by employing llms-in-the-loop strategy. In *Findings of the Association for Computational Linguistics: NAACL 2025*, pages 7397–7429.

Tunazzina Islam, Shamik Roy, and Dan Goldwasser. 2023a. Weakly supervised learning for analyzing political campaigns on facebook. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 17, pages 411–422.

Tunazzina Islam, Ruqi Zhang, and Dan Goldwasser. 2023b. Analysis of climate campaigns on social media using bayesian model averaging. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and*

Society, AIES '23, page 15–25, New York, NY, USA. Association for Computing Machinery.

Michelle S Lam, Janice Teoh, James A Landay, Jeffrey Heer, and Michael S Bernstein. 2024. Concept induction: Analyzing unstructured text with high-level concepts using lloom. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–28.

Raul P Lejano and Shondel J Nero. 2020. *The power of narrative: Climate skepticism and the deconstruction of science*. Oxford University Press.

Qiaoyi Liu, Yuheun Kim, and Jeff Hemsley. 2025. Scientists, but deny science? climate change sceptics networks on youtube led by scientists. *Information Research an international electronic journal*, 30(iConf):741–751.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Yiwei Luo, Dallas Card, and Dan Jurafsky. 2020. Detecting stance in media on global warming. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 3296–3315, Online. Association for Computational Linguistics.

Leland McInnes, John Healy, and Steve Astels. 2017. hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11):205.

Leland McInnes, John Healy, and James Melville. 2018. Umap: Uniform manifold approximation and projection for dimension reduction. *JOSS*.

Meta Platforms, Inc. 2025. Meta ad library. `https://www.facebook.com/ads/library/`.

Nishanth Nakshatri, Nikhil Mehta, Siyi Liu, Sihao Chen, Daniel J. Hopkins, Dan Roth, and Dan Goldwasser. 2025. Talking point based ideological discourse analysis in news events. *Preprint*, arXiv:2504.07400.

Warren Pearce, Sabine Niederer, Suay Melisa Özkula, and Natalia Sánchez Querubín. 2019. The social media life of climate change: Platforms, publics, and future imaginaries. *Wiley interdisciplinary reviews: Climate change*, 10(2):e569.

Arianna Pera and Luca Maria Aiello. 2024. Shifting climates: Climate change communication from youtube to tiktok. In *Proceedings of the 16th ACM Web Science Conference*, pages 376–381.

Chau Minh Pham, Alexander Hoyle, Simeng Sun, Philip Resnik, and Mohit Iyyer. 2024. Topicgpt: A prompt-based topic modeling framework. *Preprint*, arXiv:2311.01449.

Dorian Quelle and Alexandre Bovet. 2024. Bluesky: Network topology, polarization, and algorithmic curation. *CoRR*.

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.

Filipe N Ribeiro, Koustuv Saha, Mahmoudreza Babaei, Lucas Henrique, Johnnatan Messias, Fabricio Benevenuto, Oana Goga, Krishna P Gummadi, and Elissa M Redmiles. 2019. On microtargeting socially divisive ads: A case study of russia-linked ad campaigns on facebook. In *ACM FAccT*.

Harri Rowlands, Gaku Morio, Dylan Tanner, and Christopher D Manning. 2024. Predicting narratives of climate obstruction in social media advertising. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 5547–5558.

Shamik Roy and Dan Goldwasser. 2020. Weakly supervised learning of nuanced frames for analyzing polarization in news media. In *EMNLP*.

Kimi Team, Yifan Bai, Yiping Bao, Guanduo Chen, Jiahao Chen, Ningxin Chen, Ruijue Chen, Yanru Chen, Yuankun Chen, Yutian Chen, and 1 others. 2025. Kimi k2: Open agentic intelligence. *arXiv preprint arXiv:2507.20534*.

Stephen E Toulmin. 1958. *The Uses of Argument*. Cambridge University Press.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, and 1 others. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Kathie M d'I Treen, Hywel TP Williams, and Saffron J O'Neill. 2020. Online misinformation about climate change. *Wiley Interdisciplinary Reviews: Climate Change*, 11(5):e665.

Apoorva Upadhyaya, Marco Fisichella, and Wolfgang Nejdl. 2023. Intensity-valued emotions help stance detection of climate change twitter data. In *IJCAI*, pages 6246–6254.

Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605.

Suvi Vikström, Erkki Mervaala, Hanna-Liisa Kangas, and Jari Lyytimäki. 2023. Framing climate futures: the media representations of climate and energy policies in finnish broadcasting company news. *Journal of Integrative Environmental Sciences*, 20.

Maryline Vivion, Valérie Trottier, Ève Bouhêlier, Isabelle Goupil-Sormany, Thierno Diallo, and 1 others. 2024. Misinformation about climate change and related environmental events on social media: Protocol for a scoping review. *JMIR Research Protocols*, 13(1):e59345.

Stefanie Walter, Michael Brüggemann, and Sven Engesser. 2018. Echo chambers of denial: Explaining user comments on climate change. *Environmental Communication*, 12(2):204–217.

Ifeoluwa Wuraola, Nina Dethlefs, and Daniel Marciniak. 2023. Linguistic pattern analysis in the climate change-related tweets from UK and Nigeria. In *Proceedings of the 2023 CLASP Conference on Learning with Small Data (LSD)*, pages 90–97, Gothenburg, Sweden. Association for Computational Linguistics.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, Chujie Zheng, Dayiheng Liu, Fan Zhou, Fei Huang, Feng Hu, Hao Ge, Haoran Wei, Huan Lin, Jialong Tang, and 41 others. 2025. Qwen3 technical report. *Preprint*, arXiv:2505.09388.

A Curated Keywords

The curated set of climate-related keywords shown in Table 4 was collected from (Islam et al., 2023b) and modified to reflect recent shifts in climate-change related topics.

Climate-Related Keywords
climate change, climate denial, sustainable, fossil fuel, green new deal, oil and gas industry, solar energy, renewable energy, greenhouse gas, pollution, clean energy, green energy, global warming, deforestation, carbon emission, drilling, fracking, coal mining, junk science, climate misinformation, greenwashing

Table 4: Climate change related keywords for Meta and Bluesky data collection.

To evaluate resulting relevance of the datasets obtained through keyword extraction, we conduct human validation on a random sample of 500 Meta ads and 500 Bluesky posts. Among the Meta sample, 23 items (4.6%) are judged to be irrelevant to climate-related content. Among the Bluesky sample, 31 items (6.2%) are judged irrelevant. As such, 94.6% of sampled items across both platforms are confirmed to be substantively climate-related. Note that for Meta, the estimated relevance rate of 95.4% yields a 95% CI of (93.6%, 97.2%). For Bluesky, the estimated relevance rate of 93.8% yields a 95% CI of (91.7%, 95.9%). These intervals are tight around the observed proportions, and so the high climate-relevance rates observed in our samples are reliable estimates of the relevance rates in the full datasets.

B Experimental Details

In this section, we now describe the experimental configuration used to instantiate and evaluate the proposed framework. We compute sentence-level embeddings using SentenceBERT (all-MiniLM-L6-v2) and then applying L2-normalization on each row. We then reduce dimensionality of the original embeddings by first applying Principle Component Analysis (PCA) and retaining the top 100 components, and then applying Uniform Manifold Approximation and Projection (UMAP) to project the data to 20 dimensions. Through this two-step dimensionality reduction, we suppress noise and outlier sensitivity, allowing for a more easily separable representation to be used for clustering. We perform density-based clustering using HDBSCAN using a minimum cluster size of 20 and a minimum number of samples of five. Clusters are characterized using membership probabilities, and the *top-5* texts per cluster are selected as representatives.

B.1 Cluster Coherency and Summarization

Cluster coherency checking and cluster summarization are performed using Mistral-Large-Instruct-2407. We provide few-shot examples for coherence evaluation. We look at the texts with the five highest probabilities per cluster and use them to determine whether the text is coherent or incoherent. A few-shot example given to LLM is shown in Table 5. We provide the LLM with three coherent examples and three incoherent examples. Then LLM labels cluster coherence based on the prompt shown in Fig. 9. Upon removing incoherent examples, we prompt the LLM to generate a summary from the top five texts from each coherent cluster (Fig. 10).

B.2 Cluster Merging

We use cosine similarity with SBERT (all-mpnet-base-v2) to determine which clusters are semantically similar and should be merged. We perform a grid search with threshold ($\tau \in \{0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90\}$) and use Silhouette Score (S) and Davies-Bouldin Index (DBI) as guideline metrics for determining best threshold value upon merging. For the Meta data, we find that a threshold value of 0.8 (S=0.161, DBI=0.665) obtains the best merging results. We began with 67 total clusters and merged 16 clusters for a total of 51 clusters. This clustering resulted in 3 duplicate themes, which are also removed (see

B.3). For the Bluesky data, we also find that a threshold value of 0.8 (S=0.484, DBI=0.444) obtains the best merging results. We began with 51 total clusters and merged 5 of them, for a final total of 46 clusters.

B.3 Theme Labeling and Assignment

Theme labels are generated from cluster summaries using few-shot prompting with Mistral-Large-2407. The prompt used to generate each theme is the following: `"Produce a short theme label in 1-3 words for the given summary. This theme label should capture the central idea, stance, and/or topic of the text. Examples: 'Against climate policy'....."` shown in Fig. 11. As a final preprocessing step, clusters with duplicate themes are removed. This yields no removals for Bluesky and eliminates 3 clusters from Meta. We have a total of **47** themes for Meta and **46** themes for Bluesky. Our final themes for Meta and Bluesky datasets are shown in Table 6. Once we obtain summaries and themes for each coherent cluster, we assign each text to its best-fitting theme through the following two methods:

(i) Text to Summary where we prompt LLM with the full set of cluster-level summaries. The model selects the single summary which best-fits the text. Because each summary was previously mapped to a theme during the Theme Label Generation step, the ad/post automatically inherits the corresponding theme once its summary is chosen. For efficiency, assignments are processed in batches of ten texts per LLM call. The exact prompt used for this procedure is shown in Fig. 12.

(ii) Text to Theme where we assign each ad to a theme directly by providing Mistral-Large-2407 the full set of cluster themes. This step is analogous to the summary assignment but uses cluster themes rather than cluster summaries. Again, a batch size of ten ads per LLM call is used. The prompt given to the LLM for this step is shown in Fig. 13.

We assign by semantic fit rather than original cluster membership, since geometric proximity in embedding space does not always match the most appropriate theme, particularly for boundary-spanning texts.

B.4 Prompting

Table 5 shows the few-shot example given to LLM as a prompt for the coherency check. We provide the LLM with three coherent examples and three incoherent examples. LLM then labels cluster coherence based on the prompt shown in Fig. 9.

Upon removing incoherent examples, we prompt Mistral-Large-Instruct-2407 with the prompt shown in Fig. 10. Each one-sentence summary is generated from the bodies of the top five texts from each coherent cluster. The prompt for generating themes is shown in Fig. 11, and the prompt used for assigning ads to corresponding summaries is shown in Fig. 12. Fig. 13 shows the prompt to assign each text to its corresponding theme. Lastly, Fig. 14 shows the prompt to evaluate results using an LLM-based judge.

Prompt:

Label each of the 5 paragraphs below. If the labels are all the same, the cluster is *coherent*. If the labels differ, the cluster is *incoherent*. Please state whether the cluster is coherent or incoherent, and explain your reasoning.

1: ** Paragraph 1 **
2: ** Paragraph 2 **
3: ** Paragraph 3 **
2: ** Paragraph 4 **
3: ** Paragraph 5 **

Response:

The cluster is <coherent/incoherent> because the labels are <consistent/inconsistent>.

Figure 9: Coherency prompt with example response from LLM.

Prompt:

Summarize the following five texts by writing 1–3 sentences (<= 100 words) capturing the common message and intent:

1. ** Text 1 **
2. ** Text 2 **
3. ** Text 3 **
4. ** Text 4 **
5. ** Text 5 **

Figure 10: Prompt used for generating cluster summaries.

B.5 Discovered Themes

Table 6 shows the resulting themes using our method on Meta and Bluesky datasets. To analyze redundancy of the derived theme sets, we compute pairwise cosine similarities between theme labels using SentenceBERT (sentence-transformers/all-mpnet-base-v2) and we find mean cosine similarity of 0.31 for the Meta theme set and 0.26 for the

#	Paragraph	Label
1	Copper has become the world's third most consumed metal, but uncertain demand levels and supply constraints have created a volatile outlook.	Against copper mining
2	Highlights widespread support for copper mining in Arizona, emphasizing its role in renewable energy, national security, and supply chain resilience.	Supports copper mining
3	Describes the Kennecott Mine's long history, production levels, and contribution to sustainable technologies like wind turbines and electric vehicles.	Supports copper mining
4	Discusses Alaska's Estelle Gold Project and its potential for gold and antimony production.	Promotes gold mining
5	Promotes mining in Minnesota as a modern STEM career, focusing on innovation in iron mining.	Supports iron mining
Cluster Coherency: Incoherent		
Reasoning: The cluster is incoherent because the labels are inconsistent.		

Table 5: Few-shot example for coherency prompting.

Prompt:

Produce a short theme label in 1–3 words for the given summary. This theme label should capture the central idea, stance, and/or topic of the text.

Examples:

'Against climate policy'
'Alternative energy'
'Oil lobby messaging'

*Summary to label: ** summary ***

Figure 11: Prompt used for generating cluster themes from cluster summaries.

Prompt:

Classify each of the 10 texts with its corresponding best-fitting summary from the list of provided summaries. Do not invent new summaries. Each text should have its own assigned summary from the summary set.

Summaries:

1. Summary 1
2. Summary 2
…
52. Summary 52

Texts:

1. ** Text 1 **
…
10. ** Text 10 **

Figure 12: Prompt used for assigning text to corresponding summaries.

Bluesky theme set. From this evaluation, we find that the vast majority of theme pairs are semantically distinct. In total, only 3 theme pairs out of 2116 total possible pairs combined across the two theme sets are flagged for high similarity above a threshold value of 0.80, which is a redundancy rate of 0.14%. The distribution of pairwise similarities and correlation matrices for both datasets can be found in Fig. 15.

In Fig. 27 and Fig. 28, the distribution of texts into each theme from the full theme sets is shown for Meta and Bluesky, respectively. For Meta, the theme that receives the least number of assignments is "Donziger vs. Big Oil" with 29 assignments, whereas the theme with the greatest number of assignments is "Local civic engagement: with 1365 assignments. We see a mean of 354.4 assignments per theme. 23 themes are needed to cover 80% of the Meta dataset, whereas 9 themes are needed to cover 50% of the dataset. Respectively, we see a minimum number of assignments of 78 for the Bluesky dataset with the theme "Marine Climate Crisis" and a maximum of 2051 with "Climate rollbacks". For this theme set, the average number of texts per theme is 416.9, where 23 themes are needed to cover 80% of the Bluesky dataset and 10 themes are needed to cover 50%.

C Evaluations

C.1 Evaluation on Meta Dataset

Table 8 shows representative annotation examples. Table 1 reports quantitative results across both evaluation settings (human and LLM judges). Un-

PLATFORMS	THEMES
Meta	Agriculture policy updates, Anti-coal advocacy, Anti-fracking activism, Anti-Harris energy attack, Anti-offshore drilling, Big Oil accountability, Clean air advocacy, Clean energy advocacy, Clean energy boom, Clean water advocacy, Climate crisis advocacy, Climate education advocacy, Climate voter mobilization, Data center reform, Donziger vs. Big Oil, Energy cost alarm, Ethanol advocacy, Ethical fashion advocacy, Farmland renewables boost, Green building advocacy, Green construction reform, Hydrogen as solution, Local civic engagement, Local food advocacy, Methane regulation push, Ocean conservation call, Oil industry promotion, Pro-BESS advocacy, Pro-childcare candidates, Pro-clean energy, Pro-climate candidate, Pro-forest climate action, Pro-infrastructure investment, Pro-nuclear advocacy, Pro-recycling advocacy, Pro-solar advocacy, Pro-wind energy, Propane advocacy, Public climate input, SAF advocacy push, Save the Arctic, Save the pollinators, Solar emergency power, Solar energy advocacy, Sustainable farming progress, Wildfire recovery advocacy, Women-led climate action.
Bluesky	Anti-renewable hypocrisy, Aviation decarbonization, Capitalism's climate doom, China's green leap, Clean Energy Push, Cleaning fatigue, Climate Denial Backlash, Climate action collaboration, Climate injustice exposed, Climate justice action, Climate legal accountability, Climate migration crisis, Climate rollbacks, Climate unease, Climate-fueled wildfires, Coffee climate crisis, Corporate Climate Accountability, Deadly Heat Crisis, Deforestation & Wildlife Crisis, EV transition barriers, End fossil fuels, Fascism's fatal flaws, Fossil Fuel Overreach, Fossil fuel corruption, Fossil fuel deception, Green Energy Sovereignty, Greenwashing deception, Independent investigative journalism, Insurance Climate Urgency, MAGA Climate Deception, Marine climate crisis, Oil's Deceptive Greed, Plastic treaty failure, Renewable Energy Momentum, Renewable energy surge, Resource expansion drilling, Responsible tourism future, Science under attack, Sound's dual impact, Starry night preservation, Sustainable luxury fashion, Trees for Climate, Water pollution crisis, Water resilience, Wealth vs. Welfare, Wildlife climate crisis

Table 6: Resulting themes on Meta and Bluesky datasets.

Prompt:

Classify each of the 10 texts with its corresponding best-fitting theme from the list of provided themes. Do not invent new themes. Each text should have its own assigned theme from the theme set.

Themes:

1. Theme 1
2. Theme 2
   …
52. Theme 52

Texts:

1. ** Text 1 **
   …
10. ** Text 10 **

Figure 13: Prompt to assign each text to its corresponding theme.

Prompt:

Evaluate whether the given theme/keywords fit the text.
If the theme/keywords are missing or do not reflect the main content output 'No'.
If the theme/keywords do reflect the text, output 'Yes'. Output only Yes or No.

text: <post/ad content>

theme/keywords: <theme>

Output: Yes

Figure 14: Prompt for LLM-based judge to annotate results based on the correctness of assigned keywords/theme.

der human evaluation, Text-to-Theme achieves the highest accuracy (0.673) of our methods, substantially outperforming Text-to-Summary (0.452), LDA (0.25), and BERTopic (0.25). While LDA attains comparatively higher accuracy under LLM-based evaluation, its performance drops under human judgment. Because LDA, BERTopic, and TopicGPT each assign multiple keywords or topics per document, LLM-based evaluation scores may vary depending on how the judge interprets partial correctness — for instance, whether a label set is marked correct only when all assigned topics are relevant, or when at least one is. This potential difference in interpretability should be considered when examining the results. BERTopic performs poorly under both evaluation settings, likely due to its high proportion of outlier assignments. TopicGPT performs best overall; however, much of this accuracy might be attributed to the generic themes derived from the TopicGPT assignments, as shown in Tables 11, 12.

In contrast to the human evaluation results, the LLM judge assigns the highest accuracy to Text-to-Summary (0.598), followed by Text-to-Theme (0.456), despite both methods ultimately assigning labels from the same fixed theme set. Consequently, the observed difference in LLM-based accuracy does not stem from lexical differences in the presented labels, but rather from differences in the upstream assignment process.

Specifically, Text-to-Summary relies on an intermediate and implicit summary-to-theme mapping that is influenced by large context windows during summary assignment. This design can degrade assignment accuracy, as shown by the LLM and human evaluation results for Bluesky, as well as for the human annotation results for Meta. However, the Text-to-Summary method may also introduce patterns that align more closely with the LLM judge's internal decision heuristics, leading to higher apparent accuracy under LLM evaluation. As such, the LLM judge may be sensitive to systematic properties of the assignment pipeline rather than to true semantic alignment between the text and the assigned theme. For this reason, we treat human annotation as the primary and more reliable evaluation signal.

C.2 Evaluation on Bluesky Dataset

Table 1 shows the human and LLM judge evaluation results of theme-assignment on the Bluesky dataset. According to both the human and LLM judgment settings, our Text-to-Theme method outperforms Text-to-Summary, as well as the LDA and BERTopic baseline methods of assignment. This result mirrors the core finding that we obtain from our Meta dataset that concise, theme-level labels and prompting often provide a more robust mapping than the full cluster summaries.

Another key point is that BERTopic's lower (and in some cases comparative) accuracy to LDA is often a result of its high rate of outliers or clusters that do not map cleanly to topics. One such topic that exemplifies this is: ['alt', 'alt text', 'note', 'text', 'spicy', 'plan fix', 'ouch', 'aim make', 'salty', 'thank doing']. In total, 9418 posts were assigned as outliers using BERTopic, whereas for our summary-based assignment, only 1 post is left unassigned, and for the theme-based assignment, only 6 posts are unassigned. For LDA, each post gets assigned a set of keywords, even if this set is not necessar-

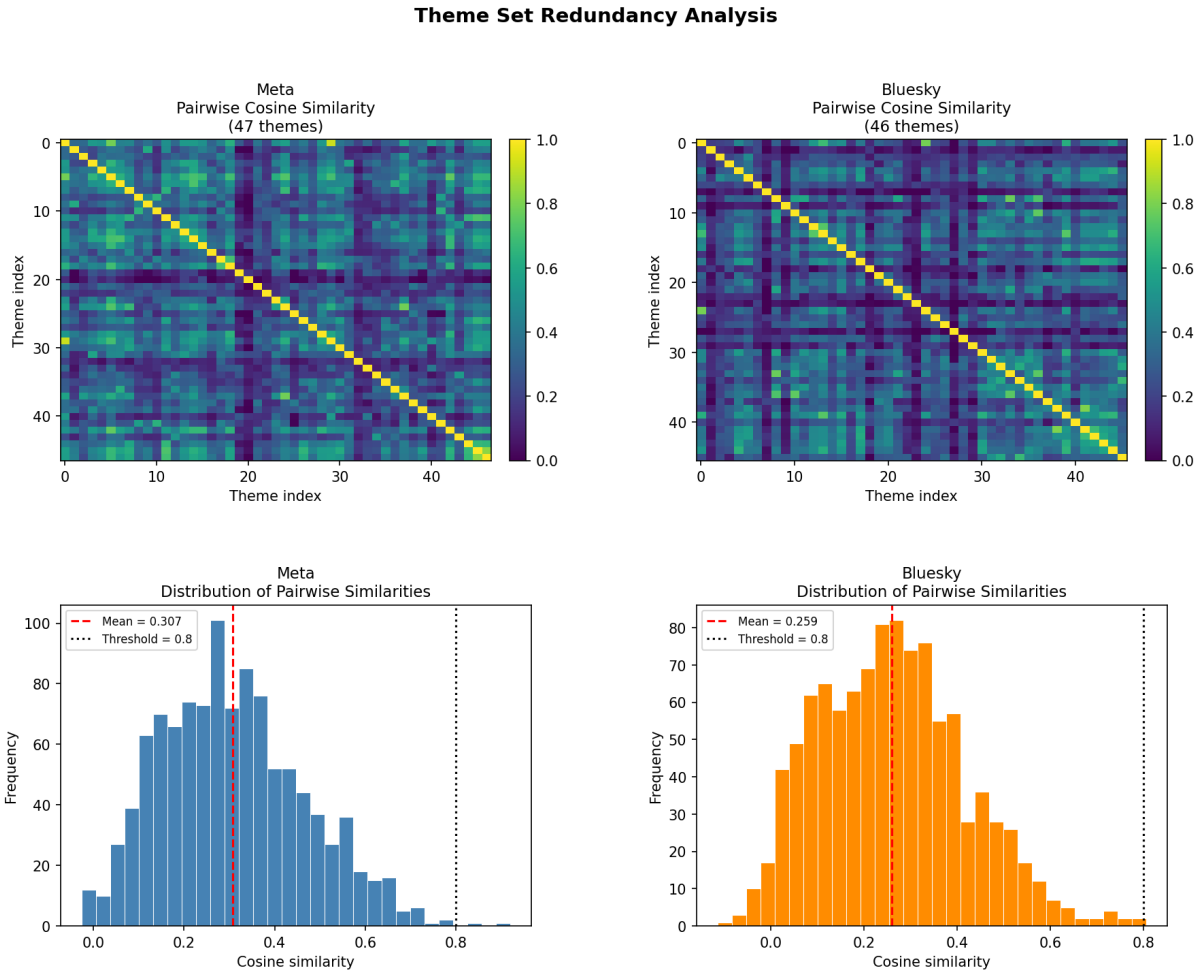


Figure 15: Correlation between and distribution of pairwise similarities for themes of the Meta and Bluesky datasets.

ad_text	lda_keywords	bert_keywords	LLM-Assigned Themes		LLM Judge Assignments			
			Theme-Based	Summary-Based	lda_correct	bert_correct	Theme Correct	Summary Correct
Data can't tell the whole story, but it's a powerful tool for understanding changes on our planet. Grab this article from our free collection of resources to teach your students essential data literacy skills.	free, call, tool, generation, receive	students, teaching, teaching climate, climate education, classroom, climate project, teach, educators, lesson, prep	Climate education advocacy	Climate education advocacy	No	Yes	Yes	Yes
Stop Kamala's Radical Plan for America Before It's Too Late. Kamala Harris is dangerously liberal, and her extreme policies threaten to fundamentally weaken our nation. She advocates for abolishing fracking, vote harris, patsy, liberal.	american, job, economy, america, michigan	kamala, kamala harris, harris, fracking, fighter, abolishing, ban fracking, vote harris, patsy, liberal	Anti-Harris energy attack	Anti-Harris energy attack	No	No	Yes	Yes
Whether you've been considering solar panels, harnessing the power of the wind, upgrading to energy-efficient equipment, or taking advantage of federal rebate programs, learn how clean energy can benefit your household.	program, wind, provide, grant, farming	infrastructure, investment, clean energy projects, rural america, infrastructure investment and jobs act, bipartisan infrastructure law, rural clean energy, credits, investment, projects	Clean energy advocacy	Clean energy advocacy	Yes	No	Yes	Yes
Utah families are being charged for lawmakers' fossil fuel favoritism. Contact your legislator today before it's too late.	tax, energy, bill, credit, cost	utah, investinamerica, iija, act utah, legislation, legislature, lawmakers, legislature's decision, act, bill	Energy cost alarm	Energy cost alarm	Yes	No	Yes	Yes
Already struggling with your electrical bill? It might get worse if lawmakers roll back clean energy incentives. Call your legislator today to protect American energy innovation.	tax, energy, bill, credit, cost	credits, rural energy, energy tax, legislation, credit, act, taxes, utah, bill, change	Solar energy advocacy	Energy cost alarm	Yes	No	No	Yes

Table 7: Evaluation of topic-model keywords and LLM theme assignments using LLM judge (Meta).

ily fitting to the post text. Therefore, the overall accuracy for LDA might appear higher than for BERTopic with both the human and LLM judges, but this is due to the large number of unassigned posts with BERTopic.

An additional observation is the overall increase in accuracy for summary-based assignment on Bluesky compared to Meta. Although both theme-based methods show around 67% accuracy according to the human judge (66.6% for Bluesky, 67.3% for Meta), there is a discrepancy in summary-based accuracy. This discrepancy is also very clear with the LLM evaluation results, where Bluesky Text-to-Summary shows 45.8% accuracy and Meta Text-to-Summary shows 59.8% accuracy. Bluesky has 46 total themes, whereas Meta has 47, and as such the corresponding summary sets differ in both size and informational density. Because summary-based assignment requires presenting the LLM with a list of candidate summaries during text-to-summary matching, a smaller summary set effectively reduces the prompt context window and the number of competing semantic options the model must consider. This reduction appears to benefit Bluesky in particular, where fewer summaries lead to less semantic overlap and lower cognitive load for the LLM during assignment. The difference in summary-based accuracy for Meta may also be affected by the type of dataset: Meta's dataset contains a larger number of more narrowly defined themes, often driven by campaign-specific messaging and targeted issue framing. As a result, summary-based assignment on Meta may be more sensitive to semantic mismatches between individual posts and their corresponding cluster summaries, leading to lower overall accuracy relative to Bluesky.

When comparing human judgments with those produced by the LLM judge, we observe systematic differences in the relative accuracy ranges across methods. LDA and BERTopic produce compact keyword lists, which align closely with the token-level evidence an LLM judge is optimized to evaluate. As a result, the LLM judge may systemati-

ad_text	lda_keywords	bert_keywords	LLM-Assigned Themes		Human Judge Assignments			
			Theme-Based	Summary-Based	lda_correct	bert_correct	Theme Correct	Summary Correct
Data can't tell the whole story, but it's a powerful tool for understanding changes on our planet. Grab this article from our free collection of resources to teach your students essential data literacy skills.	free, call, tool, generation, receive	students, teaching, teaching climate, climate education, classroom, climate project, teach, educators, lesson, prep	Climate education advocacy	Climate education advocacy	No	Yes	Yes	Yes
Stop Kamala's Radical Plan for America Before It's Too Late. Kamala Harris is dangerously liberal, and her extreme policies threaten to fundamentally weaken our nation. She advocates for abolishing fracking, vote harris, patsy, liberal.	american, job, economy, america, michigan	kamala, kamala harris, harris, fracking, fighter, abolishing, ban fracking, vote harris, patsy, liberal	Anti-Harris energy attack	Anti-Harris energy attack	No	Yes	Yes	Yes
Whether you've been considering solar panels, harnessing the power of the wind, upgrading to energy-efficient equipment, or taking advantage of federal rebate programs, learn how clean energy can benefit your household.	program, wind, provide, grant, farming	infrastructure, investment, clean energy projects, rural america, infrastructure investment and jobs act, bipartisan infrastructure law, rural clean energy, credits, investment, projects	Clean energy advocacy	Clean energy advocacy	Yes	No	Yes	Yes
Utah families are being charged for lawmakers' fossil fuel favoritism. Contact your legislator today before it's too late.	tax, energy, bill, credit, cost	utah, investinginamerica, iija, act utah, legislation, legislature, lawmakers, legislature's decision, act, bill	Energy cost alarm	Energy cost alarm	Yes	No	Yes	Yes
Already struggling with your electrical bill? It might get worse if lawmakers roll back clean energy incentives. Call your legislator today to protect American energy innovation.	tax, energy, bill, credit, cost	credits, rural energy, energy tax, legislation, credit, act, taxes, utah, bill, change	Solar energy advocacy	Energy cost alarm	Yes	No	No	Yes

Table 8: Evaluation of topic-model keywords and LLM theme assignments using human judge (Meta).

cally favor these methods based on the presence of salient keywords, even when the overall semantic framing of the post is only weakly captured. In contrast, human annotators place greater emphasis on interpretability and semantic coherence. The descriptive, high-level theme labels generated by our approach more directly convey the underlying intent and meaning of the posts, making them easier for humans to evaluate accurately. This divergence in evaluation criteria helps explain why the LLM judge exhibits a relative bias toward LDA. The effect is less pronounced for BERTopic, likely due to its high proportion of outlier assignments and loosely formed clusters, which reduce the consistency of its keyword sets. We further observe that both Text-to-Summary and Text-to-Theme achieve lower accuracy under LLM-based evaluation than under human judgment. This gap likely occurs because LLM judges rely more on direct word overlap between the label and the text than on deeper semantic interpretation, which makes abstract theme labels harder to assess than explicit keyword-based representations.

We see that, although our Text-to-Summary and Text-to-Theme assignment variants achieve higher accuracy than LDA and BERTopic, TopicGPT, much like for Meta, achieves the highest accuracy overall by far. However, much of this accuracy can be attributed to the highly generic themes derived from the TopicGPT assignments. For example, "Energy" applies to a very wide range of the climate-related dataset. We provide the full lists of TopicGPT generated topics for Bluesky and Meta such that the generality of the topic set can be noticed in Tables 9, 10 11, 12.

Level 1 Topics
Agriculture, Climate Science, Climate Change Impacts and Adaptation, Emergency Response, Energy, Energy Policy and Transition, Environment, Climate Change, Health, Social Justice, Climate Justice, Sustainable Development

Table 9: TopicGPT Level 1 topic labels for Bluesky

To validate our annotation results for the Bluesky dataset, we look at Inter-annotator agreement scores for our methods as well as TopicGPT. We

Level 2 Topics
Agricultural Technology, Clean Energy and Technological Responses, Climate Change, Climate Change Communication and Public Perception, Climate Change Impacts and Adaptation, Climate Change Policy and Regulation, Climate Justice, Climate Justice and Socioeconomic Issues, Climate and Environment, Disaster Management, Disease Prevention and Health Behavior, Economic and Social Equity, Energy, Energy Policy and Transition, Environment, Environmental Impact of Agriculture, Environmental Impact on Society and Health, Environmental Policy and Governance, Fossil Fuels, Health, Health Policy and Social Issues, Pollution, Public Health and Environment, Renewable Energy, Social Justice, Social and Economic Sustainability, Sustainable Agriculture, Sustainable Development, Sustainable Infrastructure and Mobility

Table 10: TopicGPT Level 2 topic labels for Bluesky

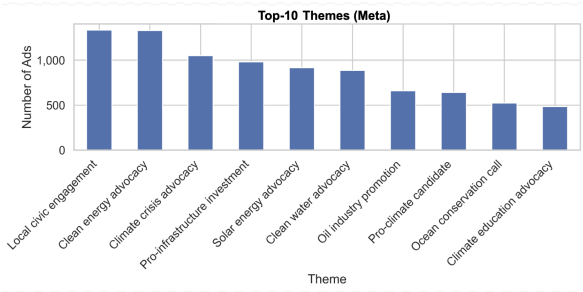


Figure 16: Top 10 themes by number of ads for Meta data.

see that for our method, there is 71.4% raw agreement between the human and LLM judge ($\kappa = 0.44$) for the summary-based judgments, and furthermore, a 70.4% raw agreement ($\kappa = 0.40$) for theme-based judgments. Both Cohen's Kappa scores fall in the moderate agreement range. Further analysis suggests that the disagreements are not random: the LLM judge is systematically more conservative, achieving higher precision on positive labels (0.86–0.87) but lower recall (0.64). This means the LLM judge tends to miss true positives flagged by human annotators, rather than hallucinating false positives. This asymmetry points to a somewhat consistent difference in labeling threshold between the two judges, rather than unreliable or noisy behavior. For TopicGPT, we see 90.2% agreement between the human and LLM judge with

Level 1 Topics
Agriculture, Sustainable Farming, Economics, Energy, Energy Policy and Regulation, Energy Technology and Innovation, Renewable Energy, Environment, Conservation and Natural Resources, Sustainable Energy and Technology, Health and Well-being, Human Experience, Social Issues and Community Well-being, Infrastructure, Transportation Infrastructure, Politics, Political Campaigns and Elections, Political Discourse and Public Sentiment, Public Policy and Governance, Science

Table 11: TopicGPT Level 1 topic labels for Meta

a Cohen's Kappa score of 0.43.

C.3 Human Validation

Seven NLP and CSS researchers, age range 20-40, 3 males and 4 females, did three $\approx$ 1.5-hour sessions to validate whether the generated and refined themes, as well as text $\rightarrow$ theme mapping, are correct or not for each dataset. In total, it took $\approx$ 11 hours for both datasets. Annotators were provided with the same instructions that were provided to the LLM prompt. The annotators included advanced undergraduate students, graduate students, and faculty.

D Joint-Clustering Details

To support direct cross-platform comparison in §6.1, we re-run a modified version of our clustering pipeline on the joint combination of all Meta and Bluesky texts (36,206 total). The pipeline follows the same stages as the per-platform pipeline (App. B) but operates on the merged embedding space. We retain hierarchical structure by extracting both top-level clusters and sub-clusters within each top-level group, producing 140 top-level clusters and 471 sub-clusters. Coherence filtering, summarization, and theme labeling are then applied to sub-clusters as in the per-platform pipeline, yielding 354 coherent themes.

We define three theme categories based on within-theme platform composition. A theme is *Meta-dominant* if at least 70% of its texts come from Meta, *Bluesky-dominant* if at least 70% come from Bluesky, and *shared* otherwise. The 70% threshold ensures that dominant themes reflect platform-specific discourse rather than minority

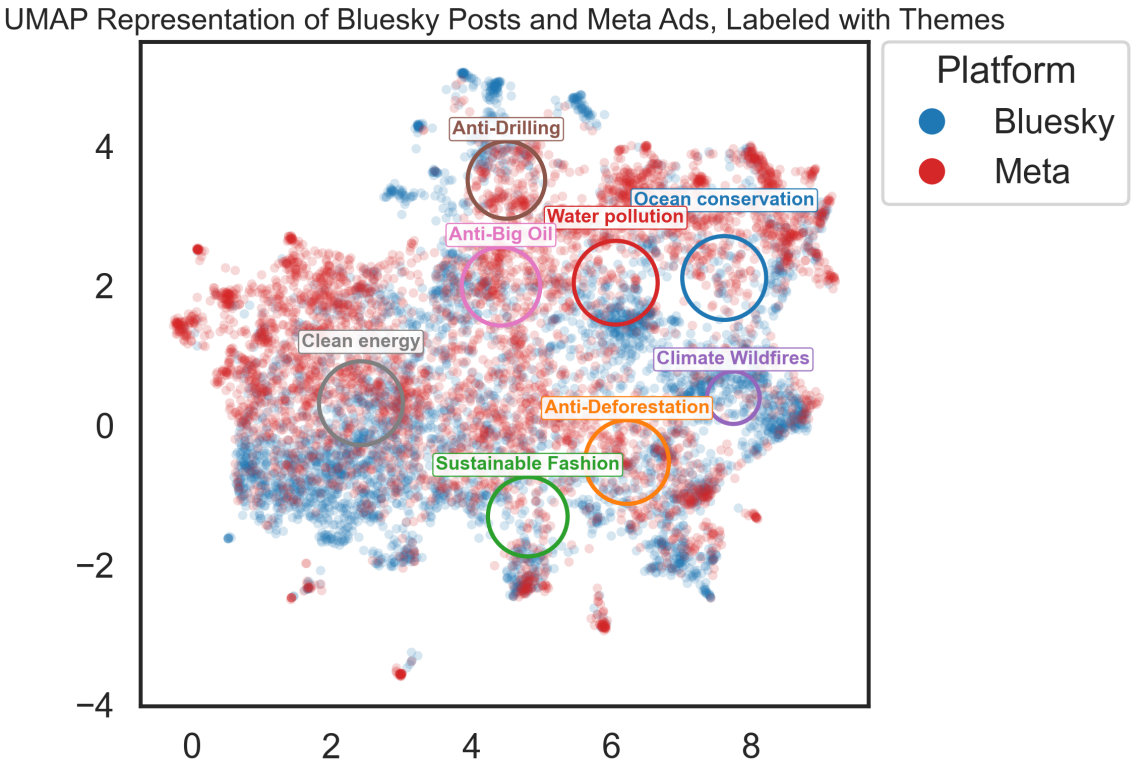


Figure 17: UMAP representation of Meta ads and Bluesky posts, annotated with themes.

representation, while still admitting a meaningful number of shared themes. The resulting partition contains 145 Meta-dominant, 112 Bluesky-dominant, and 97 shared themes.

D.1 Shared Themes

Fig. 18 shows the top 20 shared themes ranked by total cluster size. Shared themes span a range of climate-relevant topics (e.g., *Sacred land defense*, *Climate defense mobilization*, *Plastic crisis resistance*) and provide a contrast to the strongly platform-dominant themes discussed in §6.1. Furthermore, Fig. 19 shows the distribution of joint themes for the Meta and Bluesky datasets. We see that Bluesky discourse concentrates heavily on a small number of top themes, whereas Meta discourse does not see as sharp of a long-tailed distribution.

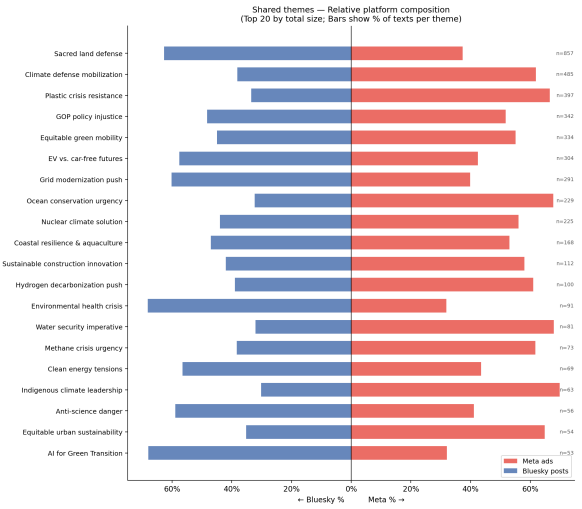


Figure 18: Top 20 shared themes with similar numbers of Meta and Bluesky texts.

E Named Entity Recognition

E.1 Pipeline

We apply a Named Entity Recognition pipeline combining SpaCy's pretrained `en_core_web_lg` model with a custom climate-domain entity dictionary and alias normalization map. The pipeline runs in two stages per text: a domain-dictionary pass that matches multi-word climate-relevant phrases that SpaCy might miss (e.g., *climate change*, *Inflation Reduction Act*, *Paris Agreement*), followed by SpaCy's NER for general-purpose en-

Level 2 Topics
Agriculture and Land Use, Climate Change and Emissions, Conservation and Natural Resources, Cultural and Personal Expression, Economic Development, Economic Education, Economics, Education and Leadership Development, Energy, Energy Access and Affordability, Energy Policy and Regulation, Energy Technology and Innovation, Energy and Community Impact, Environment, Environmental Justice and Community Action, Environmental Science and Climate Research, Environmental and Climate Concerns, Health and Well-being, Housing and Urban Development Infrastructure, Infrastructure, Political Campaigns and Elections, Political Discourse and Public Sentiment, Politics, Public Policy and Governance, Pollution and Waste Management, Renewable Energy, Science, Scientific Research and Education, Social Issues and Community Well-being, Social and Environmental Justice, Sustainable Energy and Technology, Sustainable Farming, Transportation Infrastructure, Water Infrastructure

Table 12: Level 2 topic labels for Meta

tity extraction.

Extracted entities are normalized via an alias map that consolidates phrases which refer to the same entity: people (e.g., *Trump*, *President Trump*, *Donald Trump* → *Donald Trump*), places (*U.S.*, *USA*, *America*, *United States of America* → *United States*), organizations (*GOP*, *Republicans*, *Republican Party* → *Republican Party*), and policy/issue concepts (*climate crisis*, *global warming* → *climate change*; *IRA* → *Inflation Reduction Act*).

We aggregate recognized entities into four high-level analytic types by mapping SpaCy's labels and our custom dictionary entries into a common taxonomy: PERSON (SpaCy's PERSON); ORG (SpaCy's ORG, plus domain-specific organizational entries such as *fossil fuel industry* and *MAGA*); GPE/LOC/GROUP (SpaCy's GPE, LOC, NORP, and FAC); and POLICY/EVENT/ISSUE (SpaCy's EVENT and LAW, plus all custom climate-policy and climate-issue entries). SpaCY labels for dates, quantities, money, products, etc. are excluded as not informative for our analysis. Within each text, entity mentions are deduplicated such that repeated references to the same entity count once per text. Reported entity-salience percentages in App. E.2 are computed as the proportion of texts mentioning each entity at least once.

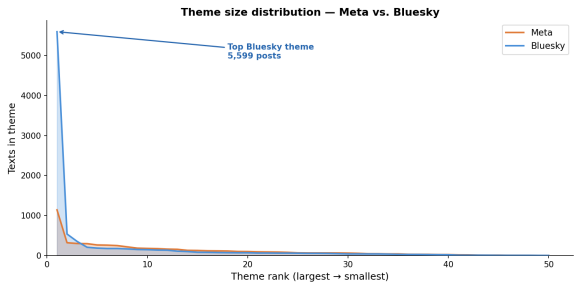


Figure 19: Distribution of joint themes for Meta vs. Bluesky

E.2 Results

Entity Type Distribution. Entity-type prevalence differs sharply by platform (Fig. 20). Bluesky contains substantially more POLICY/EVENT/ISSUE entities than Meta, while Meta dominates the GPE/LOC/GROUP and PERSON categories. This suggests that Meta ads more frequently invoke specific groups, locations, and named individuals, while Bluesky discourse more frequently invokes general policy concepts. ORG entities are common on both platforms.

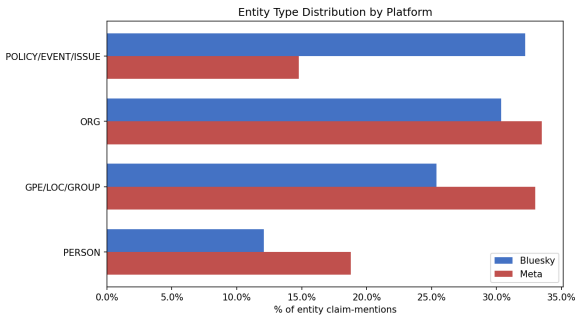


Figure 20: Named Entity Type Distribution for the Bluesky and Meta datasets.

Top Entity Salience. The most frequent entities on both platforms are general (*clean energy*, *climate change*), with long-tailed distributions thereafter (Fig. 21). Bluesky features prominent references to countries and continents (*United States*, *United Kingdom*, *Europe*, *Australia*), while Meta features specific U.S. states (*Wisconsin*, *Iowa*, *Michigan*, *California*), consistent with state-targeted advertising delivery.

Entity Salience by Cluster Category. Disaggregating by joint-clustering category (Fig. 22), Meta-dominant clusters contain more U.S. state names

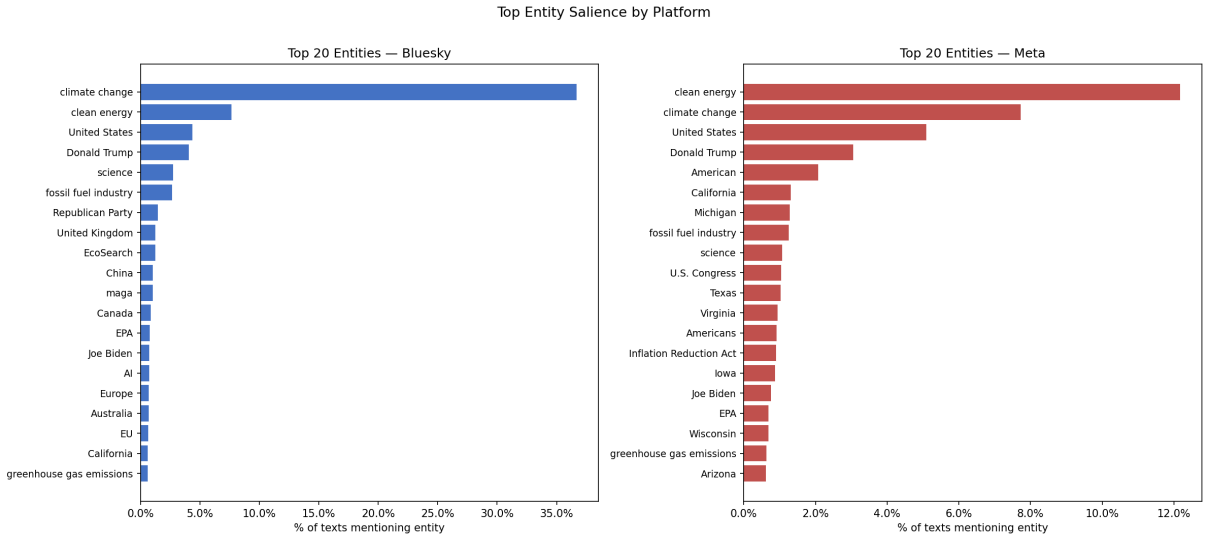


Figure 21: Top 20 entities (% of total texts) by platform.

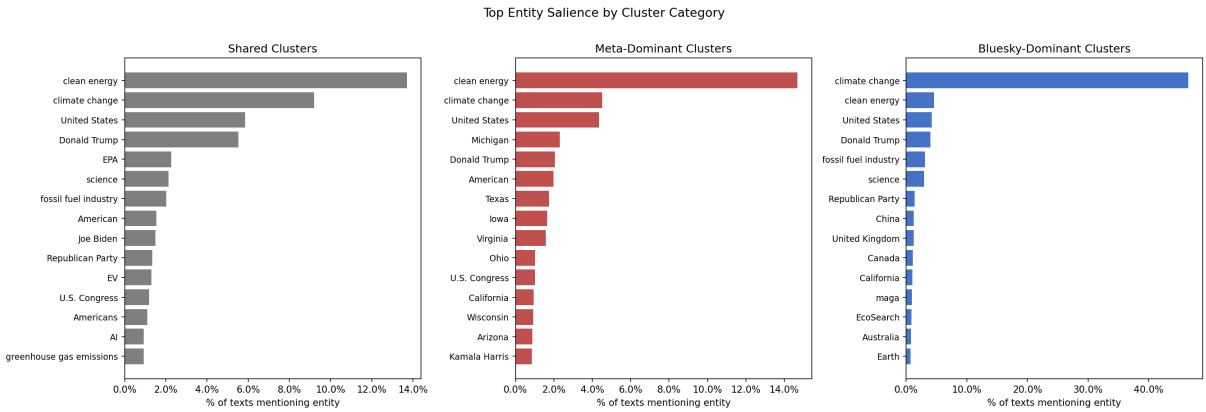


Figure 22: Top 10 entity salience for the three cluster categories: Shared, Meta-Dominant, and Bluesky-Dominant.

while Bluesky-dominant clusters center on countries and large-scale collective groups (including *Earth*). This reinforces the thematic-structure finding in §6.1: Meta-dominant discourse is found to be geographically targeted, while Bluesky-dominant discourse operates at a larger systemic scale.

F Rhetorical Register: Additional Analyses

This appendix provides the lexicon construction details, cluster- and demographic-level framing breakdowns, and theme-level LIWC profiles that support the main-text analysis in §6.2.

F.1 Entman Framing Lexicon Construction

We construct a framing lexicon seeded with five hand-defined terms per subcategory, each accompanied by a category description. We then prompt Mistral-Large-Instruct-2407 to generate 40 additional terms conditioned on each category's description and seed examples, yielding a lexicon of approximately 45 terms per subcategory. Framing scores are computed as the proportion of lexicon-matched tokens to total tokens in each text.

F.2 Cluster-Level Framing

Fig. 23 presents mean framing scores across nine dimensions, aggregated by joint-clustering category (Bluesky-dominant, Shared, Meta-dominant). Bluesky-dominant clusters score highest on emotional urgency framing (0.79), political problem framing (0.72), and institutional actor framing (0.71), suggesting a discourse centered on systemic accountability and urgent risk. Meta-dominant clusters score lowest on scientific framing (0.33); this is notable given that scientific authority framing grounds claims in empirical evidence, and suggests that paid climate advertising favors emotive and community-oriented persuasion over evidence-

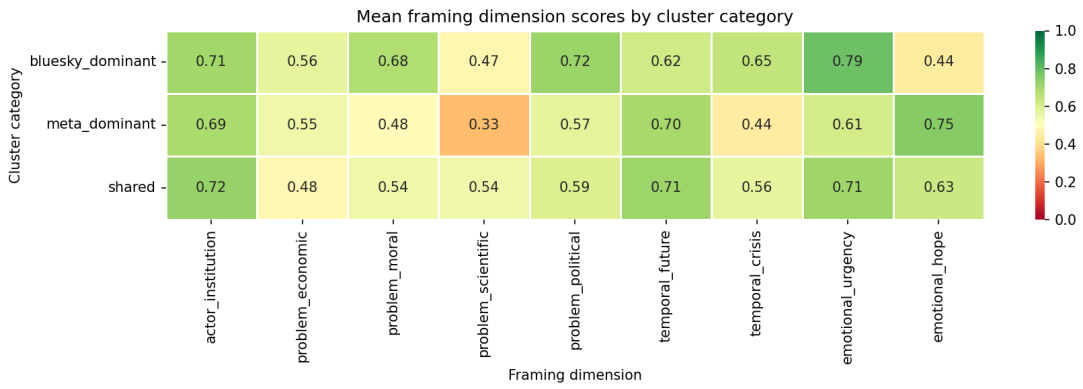


Figure 23: Mean Entman framing dimension scores per cluster category.

heavy rhetoric. Shared clusters exhibit the most balanced profile, with elevated future framing (0.71), institutional actor framing (0.72), and emotional urgency framing (0.71).

F.3 Demographic Targeting Patterns (Meta)

The Meta Ad Library exposes certain delivery and demographic metadata (age, gender, state, spend) that allows us to examine how framing varies across audience segments. We compute delivery-weighted framing scores for each segment. Bluesky does not expose comparable metadata, so this analysis is necessarily Meta-only.

Age and gender. Fig. 24(a, b) shows several monotone trends. Solution framing decreases steadily with age, while individual actor framing and populist skepticism increase with age. Institutional actor framing forms a U-shape: highest in the 18–24 group, lowest in middle-aged groups, and rising again in 65+. Female-targeted ads exhibit substantially higher institutional actor framing and moral framing than male-targeted ads; male-targeted ads exhibit higher economic and solution framing.

Political lean and ad spend. Fig. 24(c) shows that democratic-leaning audiences receive the most solution framing and scientific authority framing, while republican-leaning audiences receive more institutional actor and crisis framing. Fig. 24(d) shows that low-spend ads carry more solution framing than high-spend ads, while both institutional and individual actor framing are concentrated in high-spend ads.

Geography. Fig. 25 shows delivery-weighted framing concentration by state. Scientific authority framing concentrates in left-leaning states, consistent with the political-lean pattern above. Other framing dimensions show state-level concentrations that are not cleanly explained by political lean alone.

F.4 LIWC Content Categories

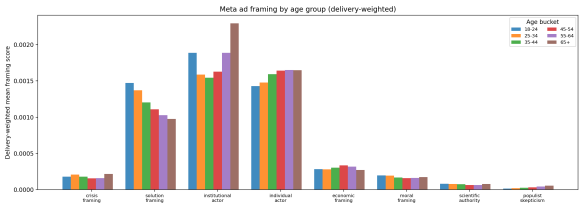
In addition to the four summary variables reported in §6.2, we analyze eleven LIWC content categories mapped to the four Entman framing functions: *tone_neg*, *emo_anx*, and *risk* for problem definition; *cause* and *power* for causal attribution; *moral*, *certitude*, and *tentat* for moral evaluation; and *achieve* and *focusfuture* for treatment recommendation.

F.5 Theme-Level LIWC Profiles

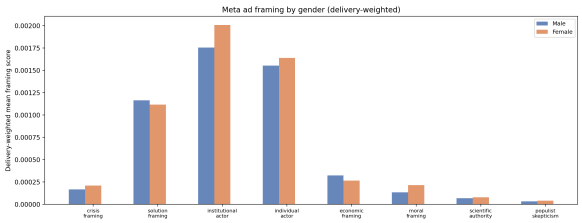
Fig. 26 shows per-theme LIWC profiles for Bluesky and Meta, z-scored within each LIWC dimension and color-coded by Entman framing function.

Bluesky. Themes organized around systemic critique exhibit distinctive signatures. *Corporate climate sabotage* is the strongest outlier on *cause*, reflecting concentrated attribution to corporate actors. *Neoliberal authoritarianism* peaks on *conflict* and *moral*. *Plastic crisis resistance* registers the highest *emo_anx* score. *Anti-capitalist resistance* combines low *Analytic* and *Clout* with elevated *certitude*, *achieve*, and *tentat*, suggesting an informal, non-authoritative voice that mixes strong ideological messages with uncertainty about concrete paths forward. *Grid modernization push* is the most analytically formal Bluesky theme, diverging from the platform's overall register due to the technical nature of the topic.

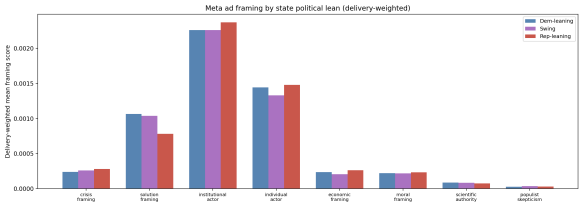
Meta. *Clean Energy Defense* exhibits the highest *power* and *conflict* scores on the platform. *Ocean*



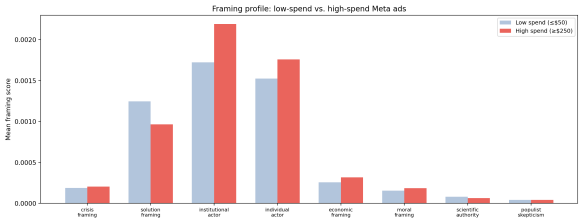
(a) By age group.



(b) By gender.



(c) By state political lean.



(d) By ad spend.

Figure 24: Delivery-weighted Entman framing for Meta ads across audience segments.

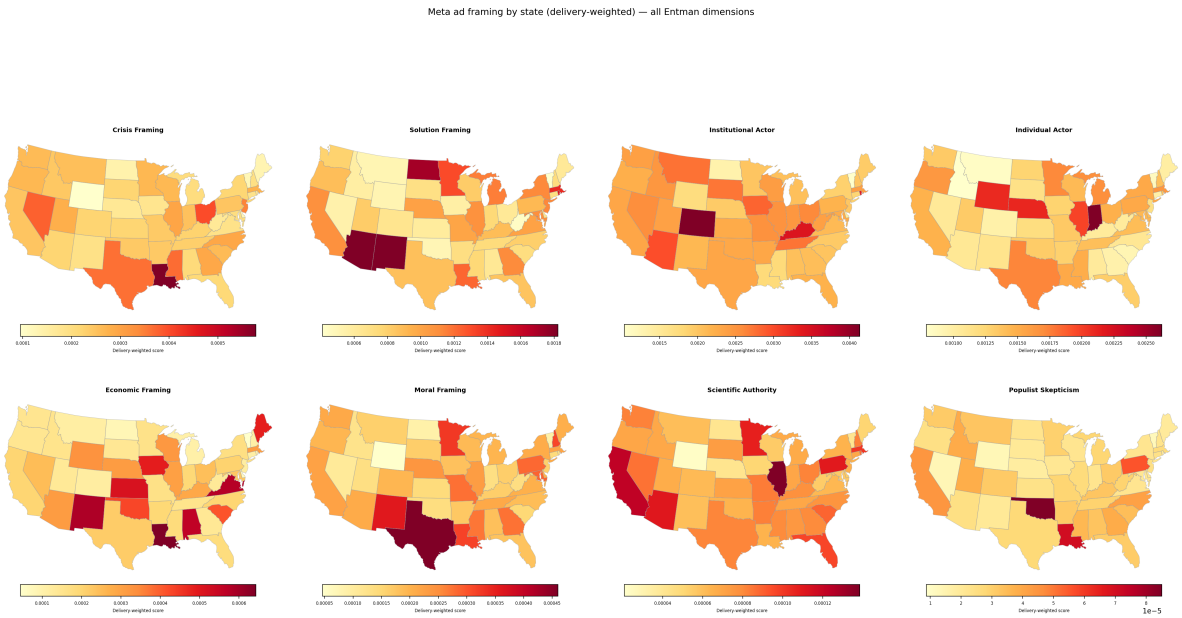


Figure 25: Heatmap of the eight Entman framing dimensions, delivery-weighted by state.

Theme	# Ads	Example ad text
Local Civic Engagement	1365	Tired of the status quo? Community-led change is THE game-changer. From global crises to climate action, we're your go-to for the latest updates and ways to support local heroes.
Clean Energy Advocacy	1356	We're committed to Environmental Justice & Social Equity so no community is left behind in our clean energy future.
Climate Crisis Advocacy	1087	Mongolia's herders face rising climate risks. Watch how we are working with partners and the government of Mongolia in solving complex challenges in the country.
Pro-infrastructure Investment	1010	These upgrades will make the Bells Mill Road and Valley Green Road bridges more resilient to extreme weather, ensuring safer, stronger connections for generations to come.
Solar Energy Advocacy	929	One of the most significant benefits of community solar projects is the potential to reduce energy costs for low-income households.

Table 13: Top themes for Meta ads.

conservation urgency registers the highest *risk* score. *Corporate climate sabotage* combines high *cause*, *emo_anx*, and *Authentic*, suggesting that paid ads invoke corporate blame through an anxious, personally engaged voice rather than formal argumentation. *Sacred land defense* scores very low on *achieve* and carries negative *Tone*, operating less as a prescriptive call to action than as a moral grievance framing.

Shared themes. Themes that appear on both platforms reveal that identical conceptual content can be presented through diverging linguistic strategies. *Corporate climate sabotage* is a high-*cause* theme on both platforms, but Bluesky's rendering carries a stronger causal attribution signal while Meta's is additionally marked by high *emo_anx*. *Grid modernization push* is analytically formal on Bluesky but *Clout*-suppressed on Meta. *Sacred land defense* on Meta is *achieve*-suppressed relative to its Bluesky counterpart.

G Qualitative Analysis Details

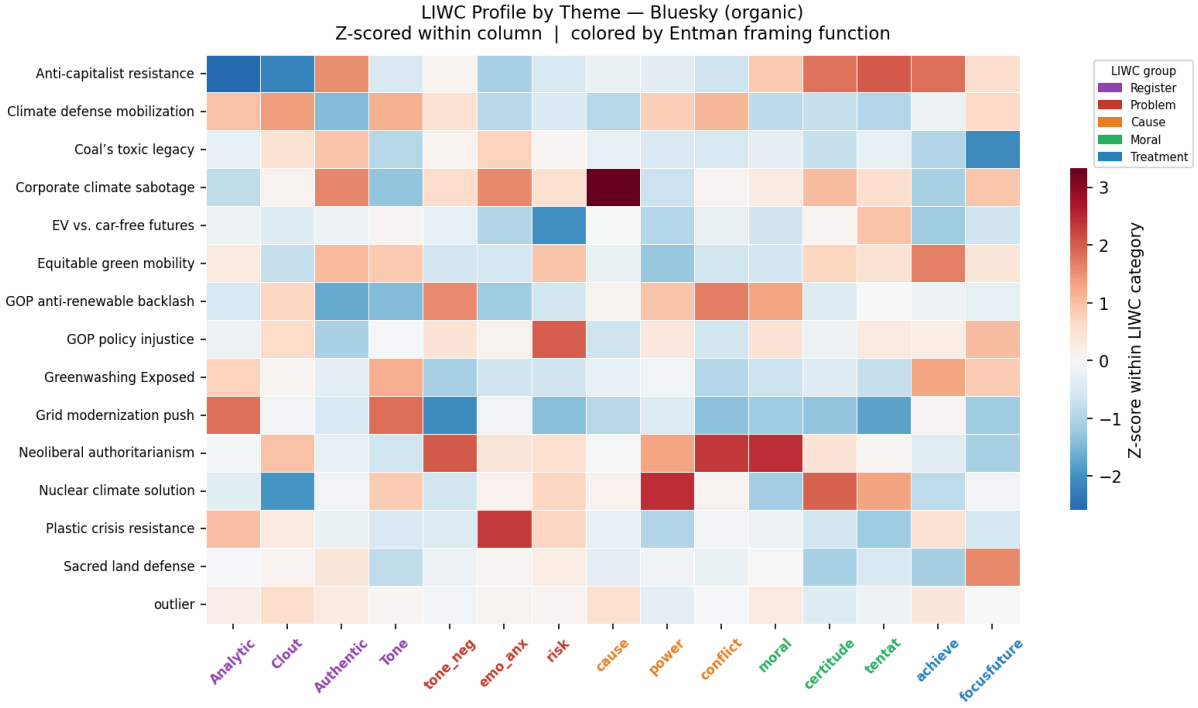
G.1 Per-Platform Theme Distributions

Table 13 and Fig. 16 report the top themes by ad count for Meta; Table 14 and Fig. 28 report the corresponding distributions for Bluesky.

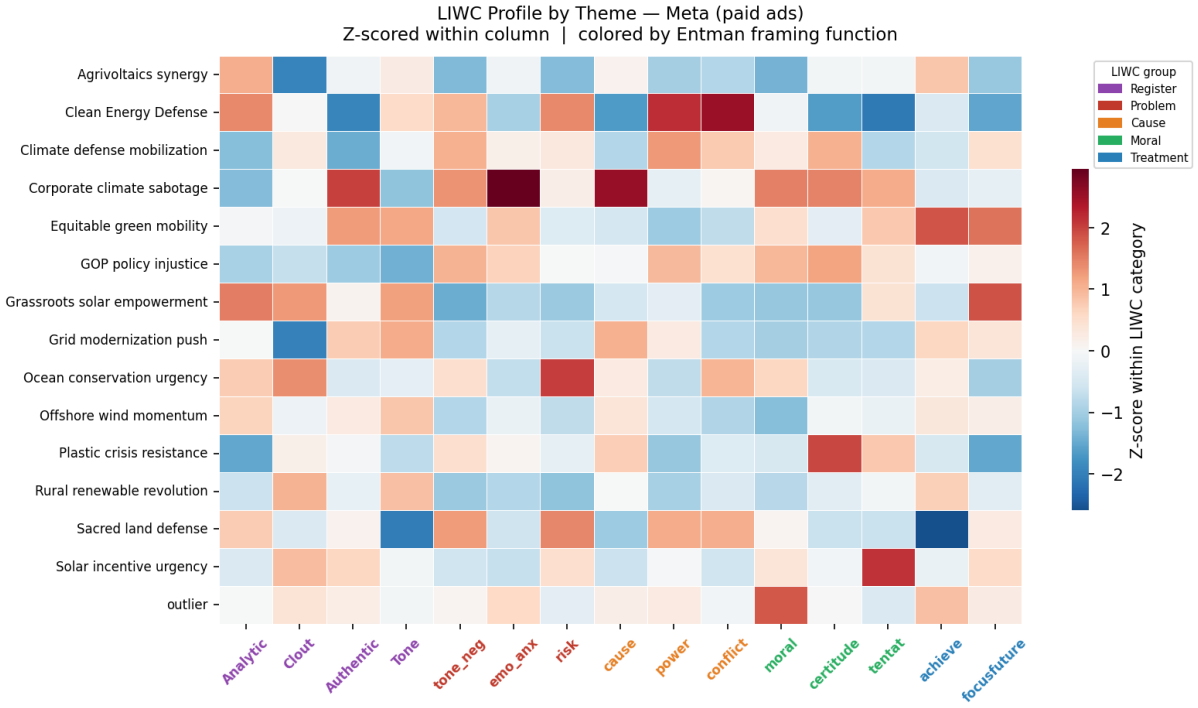
G.2 Full Theme–Stance Correlation

Figs. 29 and 30 report theme–stance correlations for all five methods on Meta and Bluesky, extending the LDA and Text-to-Theme panels shown in the main text. Note that we manually label 500 Meta texts and 500 Bluesky texts as Pro-Climate, Pro-Energy, or Neutral. Although both platforms lean heavily *Pro-Climate*, we see a more balanced structure for Meta than for Bluesky. The breakdown is as follows: for Meta, we have 46 Pro-Energy, 101 Neutral, and 353 Pro-Climate texts. For Bluesky, we have 4 Pro-Energy, 55 Neutral, and 441 Pro-Climate texts. Organic climate discourse on Bluesky is essentially a Pro-Climate-dominated space, consistent with the systemic-critique orientation observed in §6.1.

Meta. The two probabilistic baselines yield diffuse, low-magnitude correlations: LDA peaks near 0.22 and BERTopic near 0.21, with no clear separation between *Pro-Climate* and *Pro-Energy* stances. TopicGPT is the strongest baseline: its substantive climate topics align with *Pro-Climate* (*Climate Change and Emissions* 0.30, *Environmental Justice* 0.28, *Pollution and Waste Management* 0.26) and its supply-side topics with *Pro-Energy* (*Energy Access and Affordability* 0.30, *Energy Policy and Regulation* 0.20). However, much of its signal is carried by broad categories that load onto *Neutral* (*Political Campaigns and Elections*, *Politics*, *Human Experience*), reflecting a generic high-level taxonomy rather than dataset-specific structure. Our variants produce sharper alignment on the *Pro-Energy* pole, where pro-fossil themes map strongly to *Pro-Energy* (*Oil industry promotion* 0.48 for Text-to-Theme, 0.31 for Text-to-Summary; *Energy cost alarm* 0.27–0.30). We also see moderate alignment of pro-climate themes with *Pro-Climate* (*Clean energy advocacy*, *Big Oil accountability*), and of civic or non-energy themes with *Neutral* (*Local civic engagement*, *Pro-childcare candidates*). Text-



(a) Bluesky (organic).



(b) Meta (paid ads).

Figure 26: LIWC profiles by theme, z-scored within column and colored by Entman framing function.

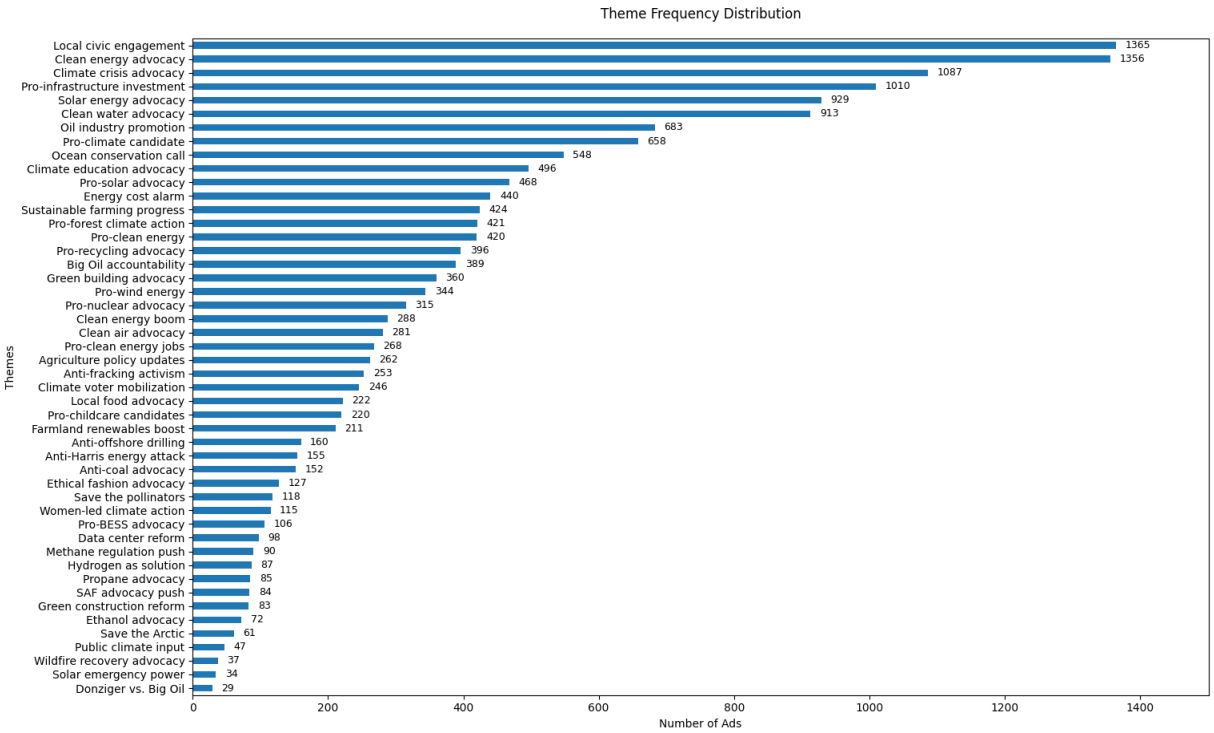


Figure 27: Distribution of themes across Meta dataset.

Theme	# Posts	Example text
Climate rollbacks	2051	Trump's EPA wants to eliminate regulation for greenhouse gases.
Science under attack	1399	Science Trump Administration Decommissions Sea Ice Data That Sounded an Alarm on Arctic Climate Change.
Climate unease	1256	That our kids will come to fear summer rather than look forward to it due to climate change just crushes me every time.
Water pollution crisis	928	tudy finds climate change limits River Thames cleanup efforts by accelerating algal growth.
Renewable Energy Momentum	848	Democrats invested $400,000,000,000.00 in renewable energy during Biden's term.

Table 14: Top themes for Bluesky posts.

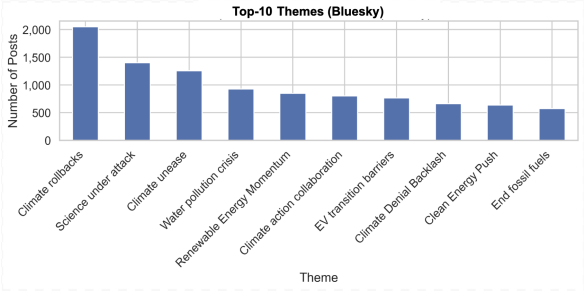


Figure 28: Top 10 themes by number of posts for Bluesky data.

to-Theme is marginally more concentrated than Text-to-Summary. Our derived themes carry a comparable number of *Neutral* labels to TopicGPT, but these are more diffuse, without TopicGPT's reliance on broad catch-all categories.

Bluesky. Bluesky exhibits extreme stance imbalance (4 *Pro-Energy*, 55 *Neutral*, 441 *Pro-Climate*), which limits the stability of theme–stance correlations relative to Meta. LDA and BERTopic exhibit more variable correlations on Bluesky than on Meta, but these are more likely driven by noise than by coherent topic–stance alignment. Our LLM-based methods show reduced absolute correlation magnitudes relative to Meta but remain directionally consistent. TopicGPT patterns more like our methods than like LDA and BERTopic: its climate- and energy-relevant topics are directionally coherent (*Climate Change*, *Renewable Energy*, and *Energy Policy* correlate positively with *Pro-Climate*, 0.11–0.14; *Fossil Fuels* is the sole topic with a positive *Pro-Energy* correlation, 0.12). However, its strongest correlations are carried by broad, climate-peripheral topics (*Health*, *Sustainable Development*) that load onto *Neutral*, reflecting its reliance on a generic high-level taxonomy rather than the dataset-specific themes our methods recover.

G.3 Bluesky: LIWC Event Dynamics

Fig. 31 shows the LIWC shift around both 2024 U.S. election and 2025 Southern California wildfire events. We can see that the only significant shift in the U.S. election event is *Cause*, which is reduced post-election. However, surrounding the wildfire event, *Cause* increases significantly and *Tone* decreases, suggesting that wildfire discourse is characterized by more negative affect and increased causal reasoning, consistent with crisis-driven language.

G.4 Meta: U.S. Election Event Analysis

Volume, spend, and impressions. Fig. 32 reports total number of ads, total estimated spend, and total impressions for the five themes with the largest before/after shifts in a $\pm$3-day window around the election. All three metrics decline, suggestion a reduction in overall advertising activity post-election.

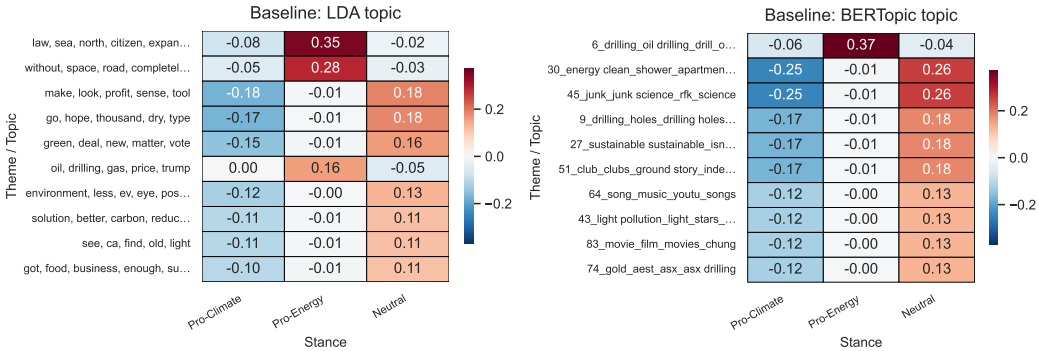
Longer-horizon shifts. Within a 30-day before/after window, the theme *Wildfire recovery advocacy* disappears entirely following the election. This change may loosely coincide with the electoral transition but is also potentially driven by the seasonal timing of wildfire-related advertising.

LIWC dynamics. Fig. 33 shows the LIWC shift around the U.S. election. We see that *moral* ($p < .01$), *risk* ($p < .05$), *power* ($p < .05$), and *focusfuture* ($p < .01$) all decline, whereas *Cause* and *tentat* increase significantly. In terms of register variables, there is increase in *Authentic* ($p < .01$) and *Analytic* ($p < .01$). Overall, this linguistic pattern suggests increased tentativeness, authenticity, and causal reasoning post-election. The decline in normative, authoritative, and future-focused language may reflect the rhetorical demands of pre-election campaigning.

G.5 Meta: California Wildfires Event Analysis

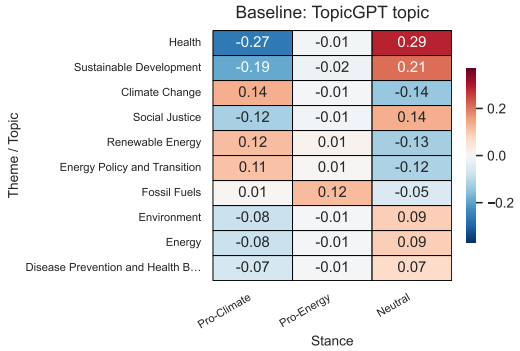
For the January 2025 Southern California wildfires, we utilize an analysis window of 30 days before and after the event onset to cover the month-long duration of the event. The five themes with the largest pre/post shifts in ad volume are *Solar energy advocacy*, *Oil industry promotion*, *Pro-clean energy jobs*, *Pro-solar advocacy*, and *Pro-climate candidate*. As shown in Fig. 34, the directionality across metrics is less uniform than around the election. Ad volume increases for *Solar energy advocacy*, *Oil industry promotion*, and *Pro-climate candidate*, while *Pro-solar advocacy* and *Pro-clean energy jobs* decline. Over a short 3-day horizon, top themes differ from the 30-day window—*Pro-infrastructure investment* and *Climate crisis advocacy* are most active.

LIWC dynamics. Fig. 35 shows the LIWC shift around the wildfires. *Tone* declines noticeably post-event, suggesting a less positive emotional register. *Power* language increases significantly post-wildfire ($p < .01$), pointing to a surge in language

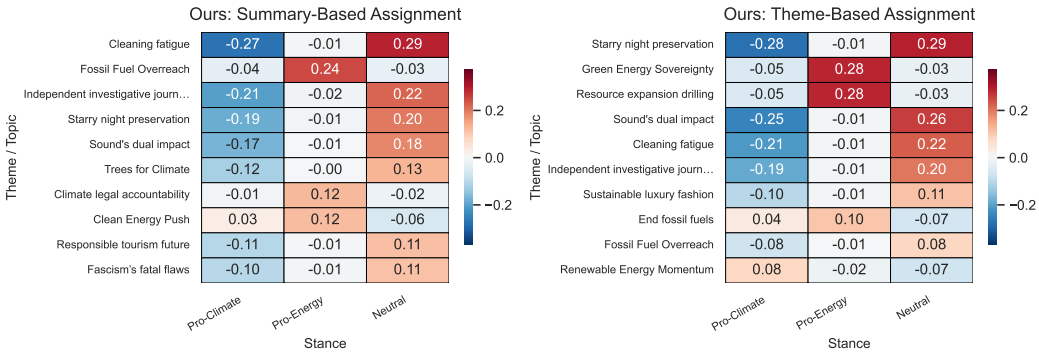


(a) **Baseline**: LDA.

(b) **Baseline**: BERTopic.



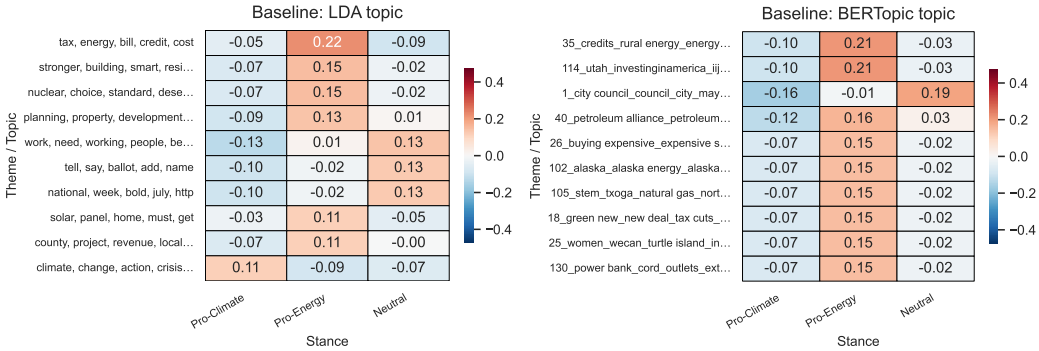
(c) **Baseline**: TopicGPT.



(d) **Ours**: LLM Summary.

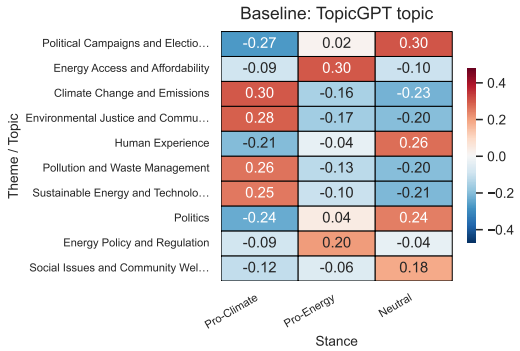
(e) **Ours**: LLM Theme.

Figure 29: Correlations between themes and stances for Bluesky.

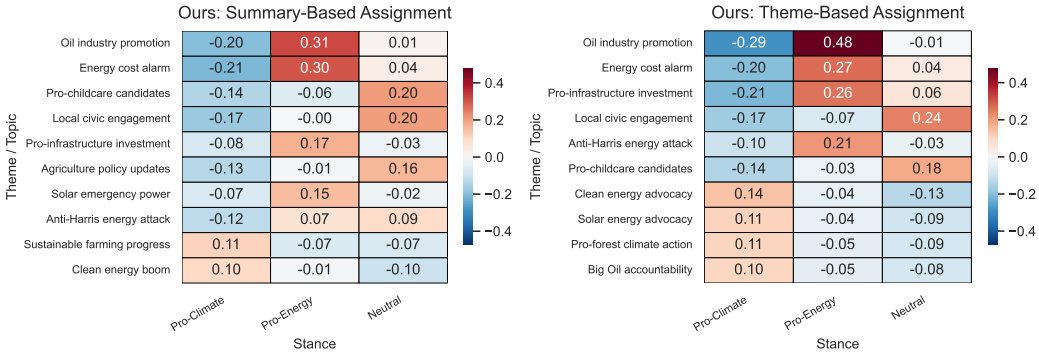


(a) **Baseline**: LDA.

(b) **Baseline**: BERTopic.



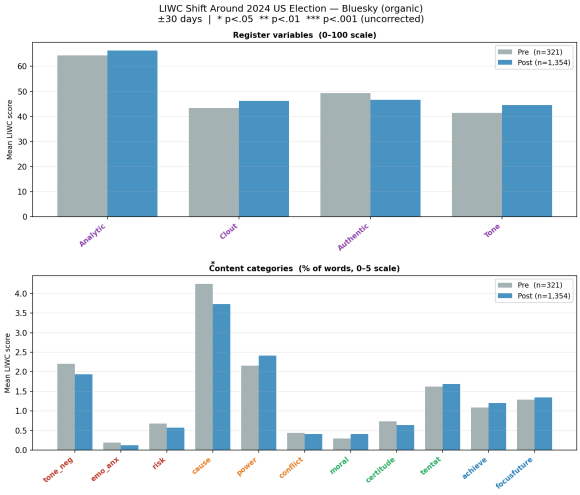
(c) **Baseline**: TopicGPT.



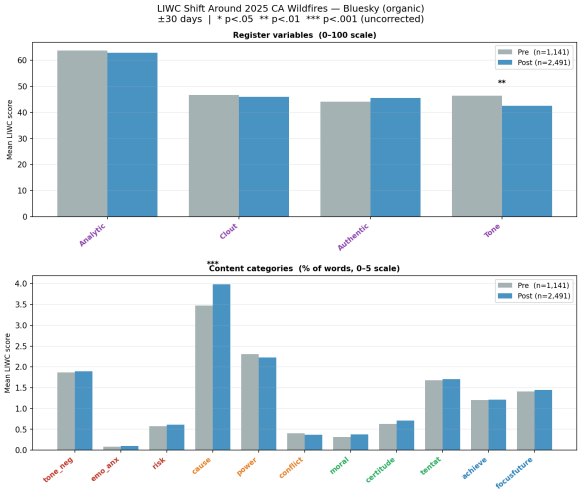
(d) **Ours**: LLM Summary.

(e) **Ours**: LLM Theme.

Figure 30: Correlations between themes and stances for Meta (all baselines and methods).

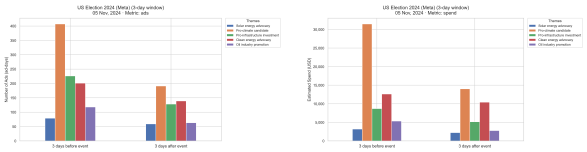


(a) LIWC shift ±30 days around U.S. election (Bluesky).

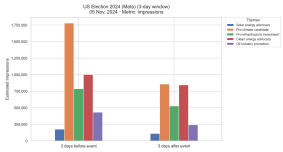


(b) LIWC shift ±30 days around CA wildfires (Bluesky).

Figure 31: LIWC content categories and register variables around the two events for Bluesky. Election: *cause* declines despite post-volume expansion. Wildfires: *Tone* declines and *cause* increases.



(a) Total number of ads. (b) Total spend.



(c) Total impressions.

Figure 32: Meta advertising activity ±3 days around the November 2024 U.S. election.

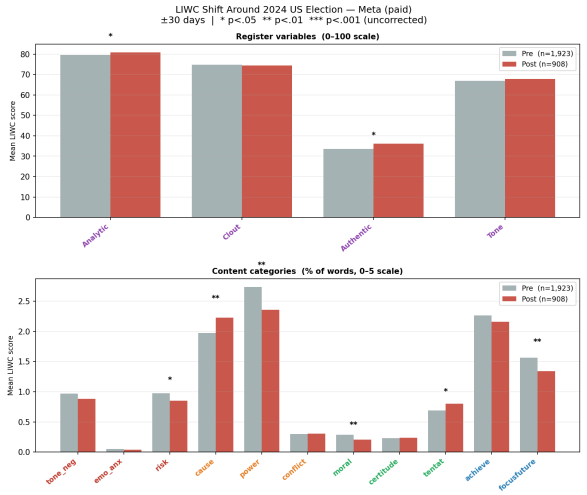


Figure 33: LIWC content categories and register variables ±30 days around the U.S. election (Meta).

centered on authority, control, and systemic agency. Unlike the post-election pattern, *focusfuture* and *moral* remain stable, and *risk* does not significantly increase despite the salience of the event.

H Details of Downstream Tasks

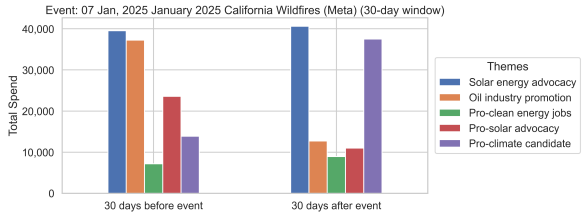
H.1 Downstream Task: Stance Prediction

In our stance prediction downstream task, the model is provided with three input variants: text only (`txt`), theme only (`thm`), and their concatenation (`txt+thm`), and is asked to predict the stance of the text (pro-climate, pro-energy, or neutral). We split the data using **stratified** sampling to preserve stance label distributions across splits. Specifically, we first create an 80/20 stratified split into training–validation and test sets. The training–validation portion is then further split using stratified sampling into training and validation sets (80/20). All splits are performed with a fixed random seed for reproducibility.

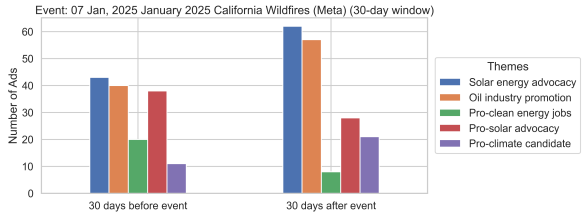
We evaluate open-source LLMs including Llama 3.1 (llama-3.1-8b-instant) (Touvron et al., 2023) and kimi k2 (moonshotai/kimi-k2-instruct-0905) (Team et al., 2025). For both models, we use default inference settings with zero-shot prompting and run them via the Groq API[3]. Prompts (all three input variants) for the stance prediction downstream task are given in Fig. 36.

In addition, we include a pre-trained language model, RoBERTa (roberta-base) (Liu et al., 2019), as well as a simple logistic regression (LR) classifier (Cox, 1958) trained on term frequency–inverse document frequency (tf-idf) features.

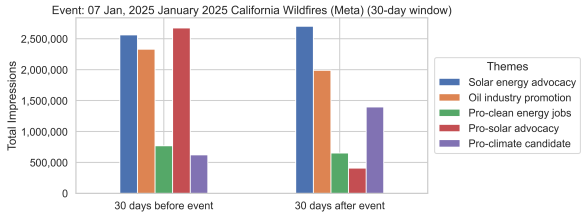
[3] https://wow.groq.com/



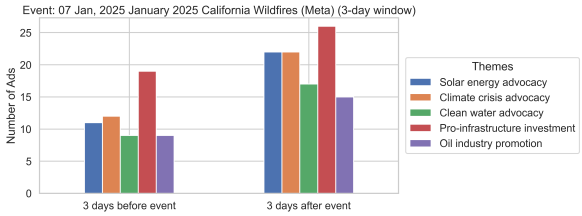
(a) Total spend



(b) Total number of ads



(c) Total number of impressions



(d) Total number of ads (Short-Horizon)

Figure 34: Advertising activity before and after the Southern California wildfires: (a) total spend, (b) total number of ads, and (c) total number of impressions.

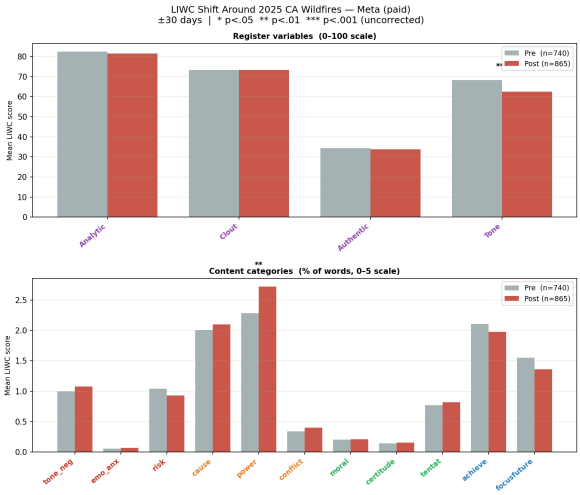


Figure 35: LIWC content categories and register variables 30 days before/after 2025 CA Wildfires for Meta dataset.

For RoBERTa, we use the 'roberta-base' model, batch size = 8, epoch = 3, learning rate = $2e-5$, weight decay = 0.01. Our early stopping criterion is the **lowest validation loss**. To run the stance classifier, we use a single GPU GeForce GTX 1080 Ti GPU, with 4 Intel Core $i5$-7400 CPU @ 3.00 GHz processors, and it takes around 2 minutes to run the model.

As a lightweight baseline, we train a multinomial logistic regression classifier using TF-IDF features. For the combined setting, TF-IDF features are independently computed for text and theme and then concatenated. We tune the n-gram range and the regularization strength C. We consider unigram and short higher-order n-grams to capture both lexical cues and limited phrasal patterns, and select a compact search space ($\{(1,1),(1,2),(2,3)\}$ for Meta and $\{(1,1),(1,2),(2,3),(3,4)\}$ for Bluesky) to avoid overfitting under class imbalance. The regularization strength C is chosen from a logarithmic grid ($C \in \{0.1, 1.0, 10.0\}$ for Meta and $C \in \{0.01, 0.1, 1.0, 10.0, 100\}$ for Bluesky) to balance model flexibility and generalization. All hyperparameters are selected based on the performance of **validation** set.

H.2 Downstream Task: Theme Guided Retrieval

Fig. 37 shows the t-SNE projection of text and theme embeddings for the same 500 Meta ads and 500 Bluesky posts. Meta's themes and ads exhibit a denser, more centralized semantic struc-

You are a stance classification system.

Task:
Given the following text, predict the stance expressed toward the text.

Text:
{text}

Possible stances:
{Pro-Climate, Pro-Energy, Neutral}

Respond with exactly one stance label from the list above without using any other words.

(a) Text only.

You are a stance classification system.

Task:
Given the following theme, predict the stance expressed toward the theme.

Theme:
{theme}

Possible stances:
{Pro-Climate, Pro-Energy, Neutral}

Respond with exactly one stance label from the list above without using any other words.

(b) Theme only.

You are a stance classification system.

Task:
Given the following text and theme, predict the stance expressed toward the text.

Text:
{text}

Theme:
{theme}

Possible stances:
{Pro-Climate, Pro-Energy, Neutral}

Respond with exactly one stance label from the list above without using any other words.

(c) Text+Theme.

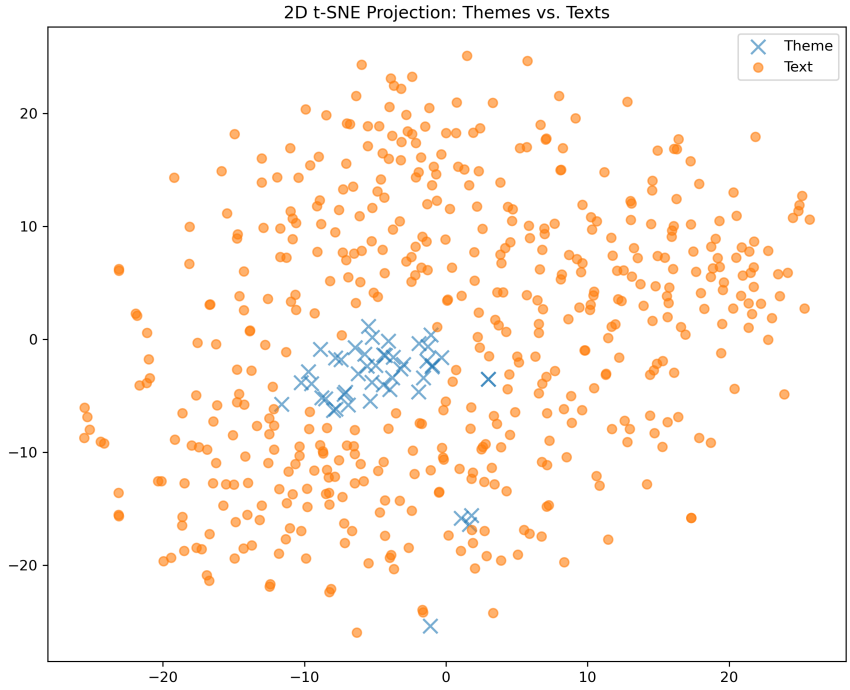
Figure 36: Prompts for downstream task: stance prediction.

Theme	Top Retrieved Ad
Climate crisis advocacy	Time & time again we see how most of our elected officials continue to fail at protecting our health & safety, which is why we need YOUR help to create the world that we want to live in. For more information on how to get involved, click the link in our bio to our landing page to sign up to be involved in upcoming rulemakings and engagement opportunities #environment #climate #legislation #policy #environment #climatecrisis #community #love #joy #resilence #pollution #solutions #action #hope #climatecrisis #denver #denvercolorado
Ocean conservation call	Our ocean is everything. From providing most of the oxygen we breathe to shielding us from the worst impacts of climate change, it's our greatest ally. This Giving Season, your support helps us protect vulnerable marine ecosystems. Give now to multiply your impact for our oceans!
Methane regulation push	An emergency brake for our climate: cutting methane pollution is the best thing we can do to stop our planet from boiling. Climate activist Sophia Kianni explains.
Energy cost alarm	Already struggling with your electrical bill? It might get worse — contact your senator today to tell them we need to keep American energy competitive.
Oil industry promotion	For nearly five decades, the oil & gas industry has been the backbone of Alaska's economy. Its impact includes: Helping create 69,250 additional jobs, statewide Our state and local governments collecting nearly $4.4 in oil & gas revenues Over $1 billion added being added to the Permanent Fund Learn more about the impacts the oil & gas industry has on Alaska in this video.

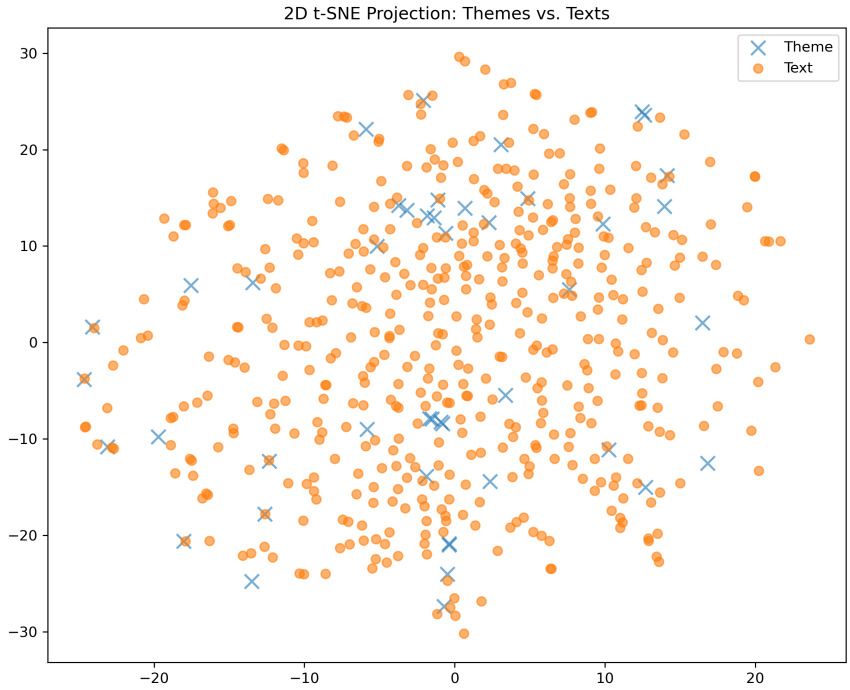
Table 15: Theme-guided retrieval examples from Meta. Ads tend to reflect coordinated or campaign-style messaging.

ture, likely reflecting a narrower and coordinated message space (e.g., advertisements or campaigns). In contrast, Bluesky exhibits a broader semantic spread, indicating more diverse themes and discourse patterns characteristic of open social platforms.

Qualitative examples (Tables 15 and 16) show that theme-guided retrieval surfaces semantically relevant posts across platforms. Meta examples tend to cluster around coordinated messages, while Bluesky samples are more linguistically and topically diverse. Retrieved examples reflect stylistic and semantic differences across platforms, with Meta ads being more coordinated and campaign-like (Table 15), and Bluesky showcasing open-ended user discourse (Table 16).



(a) Meta



(b) Bluesky

Figure 37: t-SNE projection of post and theme embeddings. Themes (blue ×) tend to cluster near semantically aligned texts (orange •), validating theme-text embedding consistency in (a) Meta and (b) Bluesky.

Theme	Top Retrieved Post
End fossil fuels	The fossil fuel industry won't go down without a fight. It's too lucrative to keep trying to convince people they're essential to our survival.
Greenwashing deception	Authenticity sells. Aligning with nature = no greenwashing. It shows you walk the talk on sustainability and values.
Science under attack	We have no time to waste in opposing the worst of President Trump's Cabinet nominees with respect to science and the role it should play in keeping us safe. Urge your senators to vote NO on the 'Junk Science Six'.
Climate action collaboration	Climate action is health action. From reducing waste to supporting renewable energy and green spaces, every step we take protects our health and our planet. Learn how we can work together to champion climate resilience and safeguard our future.
Sustainable luxury fashion	Sustainable fashion can be considered an antidote to fast fashion. It recognizes the economic and sociocultural contributions of the fashion industry while also advocating for the need to promote sustainability across all aspects of production and consumption.

Table 16: Theme-guided retrieval examples from Bluesky. Posts show more diverse, user-driven expression and framing.