Tracking the Takes and Trajectories of English-Language News Narratives across Trustworthy and Worrisome Websites

Hans W. A. Hanley
Stanford University

Emily Okabe
Stanford University

Zakir Durumeric
Stanford University

Abstract

Understanding how misleading and outright false information enters news ecosystems remains a difficult challenge that requires tracking how narratives spread across thousands of fringe and mainstream news websites. To do this, we introduce a system that utilizes encoder-based large language models and zero-shot stance detection to scalably identify and track news narratives and their attitudes across over 4,000 factually unreliable, mixed-reliability, and factually reliable English-language news websites. Running our system over an 18 month period, we track the spread of 146K news stories. Using network-based interference via the NETINF algorithm, we show that the paths of news narratives and the stances of websites toward particular entities can be used to uncover slanted propaganda networks (*e.g.*, anti-vaccine and anti-Ukraine) and to identify the most influential websites in spreading these attitudes in the broader news ecosystem. We hope that increased visibility into our distributed news ecosystem can help with the reporting and fact-checking of propaganda and disinformation.

1 Introduction

Misinformation has promoted dangerous fake health cures [16], promoted jingoism and propaganda during wars [84, 109, 121], and incited violence [6, 17]. While there has been significant investigation into how misleading information spreads across social media platforms and fringe websites [80, 81, 127], recent work has emphasized that the vast majority of people do not visit fringe websites or regularly encounter misinformation on social media [10, 101]. Rather, most people consume news through more mainstream platforms like television news [10]. However, systematically tracking how misleading, propagandistic, and outright false information spreads from untrustworthy websites into mainstream media and how fringe sites influence the broader news ecosystem remains a significant technical challenge due to the magnitude and distributed nature of the global news ecosystem [9, 19, 61, 127].

In this work, we introduce and validate a system for scalably identifying and tracking potentially unreliable news narratives across different English-language media ecosystems. Building on past work [3, 62, 92, 145], our proposed approach: (1) collects articles by continually crawling news sites from across media ecosystems; (2) extracts semantic narratives and sites' stances towards different topics using a fine-tuned version of the e5-base-v2 large language model [138], DP-Means clustering [38], and zero-shot stance detection [8]; and (3) identifies the relationships between news sites and broader ecosystems using the NETINF algorithm [49]. We emphasize that our approach does not make factual assessments of individual stories, which is a deeply nuanced task. Rather, our system allows us to shed light on how stories travel across the distributed news ecosystem.

We analyze the results from our deployed system for an 18 month period during which we collected articles from pre-curated lists of 1,003 factually unreliable news websites (*e.g.*, twisted.news), 1,012 mixed factuality reliability websites (*e.g.*, foxnews.com), and 2,061 factually reliable news websites (*e.g.*, washingtonpost.com) maintained by Media-Bias/Fact-Check [33]. Analyzing 146K stories that our system extracted from 29M articles on these news websites, we observe significant crossover in the stories covered by different news ecosystems [136]. We show that reliable and mixed-reliability news websites play the largest role in setting the stories and narratives addressed by other websites. However, despite covering similar topics, our stance analysis reveals that each type of website adopts distinctive stances towards shared topics, with reliable news websites generally being left-leaning and pro-Ukraine and unreliable websites being the most right-leaning and anti-Ukraine.

Framing our story clusters as cascades, our system uses NETINF [49] to uncover relationships between news sites and to detect potential networks of coordinating websites that spread particular slanted content and narratives. For example, using this approach we identify a network of right-leaning news websites that ostensibly act as local-news websites all operated by Metric Media, LLC. From the out-

putted results, we further identify the sites most influential in spreading stories to unreliable sites (*e.g.*, thegatewaypundit.com) and the sites from which both reliable and unreliable news websites most commonly adopt stories (*e.g.*, dailymail.co.uk and ussanews.com). Additionally, we identify the sites that most effectively promote specific types of information across ecosystems like anti-vaccine misinformation (naturalnews.com, theepochtimes.com, and vaccines.news) and anti-Ukrainian propaganda (rt.com, sputniknews.com, and news-front.info).

Ultimately, our work introduces an end-to-end system for building a near-global perspective of the English-language news ecosystem and explores how tracking how narratives travel within it can help us to understand how misleading information enters mainstream news and to uncover previously unknown relationships between news sites. We hope that our approach can serve as the foundation for further study of how information spreads online. Our code and URL data is available at https://github.com/hanshanley/tracking-takes.

2 Related Work

Significant prior work has studied news ecosystems and analyzed how misinformation spreads online. Here, we summarize the prior work that our study builds on:

Tracking Narratives on News Websites. Several studies have utilized online document clustering [22, 142] for tracking news stories. For example, Zhang et al. [146] identify potential events by monitoring for the appearance of specific phrases or keywords, cluster identified phrases that may indicate news events, and train a series of classifiers to assign news articles to identified clusters. Similarly, by clustering a collection of short phrases or "memes," Leskovec et al. find that smaller blogs often play a definitive role in encouraging the adoption of particular language onto mainstream websites [85]. Rodriguez et al. [49, 50] examine the changing relationships between websites during the discussion of news events, finding that connections between websites increase during periods of high activity.

While many studies have analyzed topics using statistical word-association approaches like Latent Dirichlet Allocation (LDA) and Dynamic Topic Models [5, 100, 147], recent works such as those by Meng et al. [96], Hanley et al. [56, 61], and Grootendorst [52] have used large language models (LLMs) for more granular topic modeling. In line with our work, Nakshatri et al. [104] utilize peak detection and HDB-SCAN [93] on news article embeddings to identify the most prominent news events in a stream of news articles. Saravanakumar et al. [119] similarly utilize an external named entity recognition system to embed entity knowledge into a BERT language model to differentiate between news articles about different events. Beyond these quantitative approaches, many prior works have qualitatively investigated the spread

of individual news stories (*e.g.*, [112, 120, 127]).

Most similar to our work, Hanley et al. [62], using MP-Net and DP-Means clustering, track news narratives across a smaller number of fringe websites to determine the role that individual unreliable news websites play in originating and amplifying news narratives. Their work finds that less-popular websites oftentimes play an outsized role in promoting narratives that reverberate across the unreliable news ecosystem.

In contrast to these prior works, our study accounts for the *stance* towards each topic in order to better differentiate between articles that cover the same topic. Tracking stance enables our work to understand the widespread understanding of individual websites' ideologically skew, changes in coverage of individual topics, and the detection of websites that coordinate in spreading particular types of propaganda.

Analyzing the Spread of Misinformation. While our approach is one of the first to track both topic and valence/stance towards that topic in a programmatic manner, several prior works have focused on the peculiarities, detection, and the spread of misinformation. For example, Ma et al. [88] and Jin et al. [71] utilize recurrent neural networks to analyze and detect the spread of unreliable rumors on social media. Abdali [1] et al., taking a domain-based approach, use website screenshots to assess the credibility of news websites. In addition to analyzing the spread of general misinformation on particular social platforms, other works have further investigated the spread of specific narratives, including those concerning the Syrian White Helmets [127], QAnon [11, 58, 107], the Russo-Ukrainian War [59, 61, 109], and COVID-19 [4, 31, 89]. We note that because work utilizes topic analysis followed by stance detection, our system can be used to quickly identify websites and topics that deserve in-depth investigation of particular types of coverage of individual events.

Building off these studies, several works have analyzed the characteristics of misinformation. Juul and Ugander find that often false information on Twitter spreads faster and wider than factual information [73]. Indeed, Kwon et al. [81], utilizing the distinct temporal differences between reliable information and unreliable rumors, are able to classify these rumors with an F_1-score as high as 0.878. In a different work [80], Kwon et al. analyze the semantic and structural characteristics of rumors on Twitter. In a different vein, Using a learning-to-rank-based approach and ClaimBuster API, Paudel et al. [108] identify potential claims that should be fact-checked on Twitter [64]. Beyond studying the dynamics of misinformation, Bak et al. [15] have proposed concrete steps to ameliorate the spread of misinformation, including removal and nudges. Finally, Kaiser et al. [74] have studied how borrowing techniques from the security warning landscape might help inform users of potential misinformation.

Unlike the past approaches outlined above, by utilizing fine-tuned encoder-based large language models, our work scalably tracks and identifies unique news stories across thousands of news websites without depending on particular key-
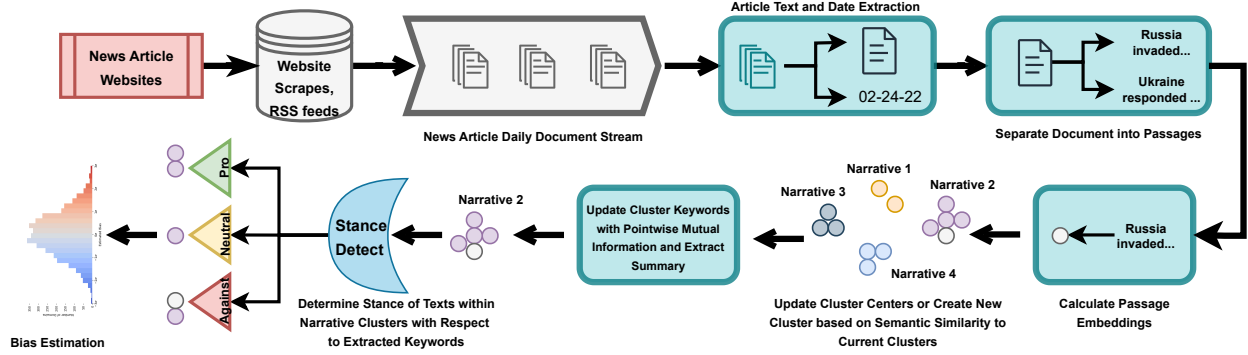
2

Figure 1: Our pipeline for identifying, labeling, and extracting the stance of story clusters from the daily publications of news websites.

words or by limiting analysis to a subset of unreliable websites previously fact-checked or curated by experts [62, 127]. By utilizing network analysis combined with stance detection, our work further provides a highly interpretable means of understanding the spread and dynamics of propagandistic, biased, or factually unreliable narratives across multiple types of media ecosystems.

3 Methodology

In this section, we provide an overview of our data collection and approach for extracting and tracking narratives across different types of news websites.

3.1 News Websites

Our study analyzes articles collected from three sets of English-language news websites of varying factual reliability. We specifically track narratives on websites rated by Media-Bias/Fact-Check [33], a media monitoring website founded by Dave M. Van Zandt to assess the factual reliability of individual websites given its widespread use in prior work [14, 62, 103, 139] and its ratings' high agreement with other organizations like NewsGuard.

Unreliable News Websites. We collect news articles from 1,003 websites labeled as having "low" or "very low" factual reporting by Media-Bias/Fact-Check [33]. We extend this list with conspiracy theory-promoting websites identified by Hanley et al. [60]. Our list of *unreliable* news websites includes pseudo-science sites like vaccine.news, state-propaganda outlets such as rt.com, and partisan websites with low-factuality ratings like the liberal-leaning occupydemocrats.com.

Mixed-Reliability News Websites. We collect articles from 1,012 *mixed-reliability* news websites labeled as having "mixed" factual reporting by Media-Bias/Fact-Check [33]. This list includes websites across the political spectrum, such as foxnews.com, nypost.com, and theguardian.com.

Reliable News Websites. We collect articles from 2,061 *reliable* news websites labeled as having "high", "very high", or "mostly factual" reporting by Media-Bias/Fact-Check [33]. The category "mostly factual" is included to capture sources with strong reputations like The Washington Post. This list features websites such as reuters.com and apnews.com.

We lastly note that we utilize the full set of English-language news websites from the lists of Media-Bias/Fact-Check [33] and Hanley et al. [60] that were accessible to us from the beginning of our study.

3.2 Definition of a News Story

Our approach tracks specific news stories and their propagation across websites rather than analyzing broader themes as captured by methods like LDA [7, 36, 70]. Following previous research [61, 62], we adopt Event Registry's definition of a *news story* as "collections of documents that seek to address the same *event* or *issue*" [83, 98]. It is important to note that even if two ideas are related, they may not constitute the same news story. For example, while "Florida Governor Ron De-Santis declares for President" and "Nikki Haley surpasses Ron DeSantis in the polls" are related, they are considered separate news stories in our work.

3.3 System Architecture

Our approach for capturing and tracking news stories builds on the LLM-based narrative tracking methodology introduced by Hanley et al. [62]. However, while Hanley et al.'s method scalably tracks individual topics, their work does not incorporate articles' attitudes towards a topic. While this was not problematic for their work, which focused on the spread of stories amongst *unreliable* news websites, the approach cannot track news stories across a broader set of news websites that present stories in dramatically different ways. We expand their method to additionally account for the *stance*/valence of news articles towards a topic (*i.e.*, we distinguish between

		all-mpnet	all-mpnet	e5-base-v2
BERT	USE	specious	peft+lora	peft+lora
0.464	0.749	0.856	0.860	**0.866**

Table 1: Model Performance on SemEval STS Benchmark. Our PEFT+LoRA models fine-tuned using unsupervised contrastive loss perform better than prior work [26, 27, 37, 62, 114].

articles that cover vaccines positively vs. negatively).

As shown in Figure 1, our system identifies stories by: (1) scraping articles from news sites, (2) splitting articles into passages of 100 words [62, 110], (3) embedding passages with a fine-tuned LLM [138], and (4) clustering news articles using an optimized version of the DP-Means algorithm [38, 72]. To describe clusters that each represent a story, we extract keywords from the resulting cluster using Pointwise Mutual Information (PMI) and performing multi-document summarization with an open source LLM. Building on the clusters, we utilize network inference techniques to identify website relationships and *zero-shot* stance detection [8, 55] to determine the stance/position of individual passages within each cluster. Finally, based on individual websites' stances toward their given topics, we perform bias estimation to quantify websites' biases along various political and non-political axes. We detail each stage below:

Collecting and Preparing News Articles. We crawl our set of 4,076 websites daily using the Go Colly library [125] from January 1, 2022 to July 1, 2023. Each day, we collect every site's homepage, RSS feeds, and linked articles. We collected a total 29.0M articles: 17.9M articles from reliable news sites (median 2,467 articles/site), 8.7M articles from mixed-reliability news sites (median 964 articles/site), and 2.5M articles from unreliable news websites (median 219 articles/site). We will provide to URLs researchers on request.

To prepare our news article data for embedding, we first remove any URLs, emojis, and HTML tags from the text. Then, inline with prior work, after first separating articles into paragraphs by splitting text on (\n) or tab (\t) characters [57], we subsequently divide paragraph into constituent *passages* with at most 100 words [57, 61, 110]. This enables us to fit passages into the context window of our LLM embedding model. Further, given that articles often address multiple ideas, embedding passages allows us to track the often single idea present within the passage [57, 110]. Our dataset consists of 428M passages. For additional details, see Appendix A.

Embedding Passages. Before embedding our articles' passages, to ensure that our embedding model is attuned to the language of news articles, we tailor our model to our domain of our collected articles using *Parameter Efficient Fine-Tuning*/PEFT [86] through *Low-Rank Adaption*/LoRA [69] with an unsupervised contrastive learning step based on SimCSE [47]. Rather than directly fine-tuning the original model's weights as in Hanley et al. [62], this approach freezes the originally trained large language model and introduces an additional set of parameters of reduced dimensionality that are then fine-tuned for purposes, allowing for better generalizability [69]. We utilize default LoRA hyperparameters of rank=8 and α=16.[1] See Appendix B and C for additional details. We utilize cosine similarity of embeddings to determine passages' estimated semantic similarity [28, 47, 52, 110].

We specifically fine-tune and evaluate two public open-source large language models, e5-base-v2 [138] and MPNet [126] using this approach. We benchmark these two fine-tuned models on the SemEval STS-benchmark (Table 1) and find that the approach outperforms prior work as general models for semantic similarity. We use the fine-tuned e5-base-v2 model in this work given its top performance.

Story Identification. We base our story-identification algorithm on Dinari et al.'s optimized and parallelizable version of the DP-Means algorithm, a non-parametric version of K-means [38]. We utilize this approach as it is highly scalable (able to cluster our 428M embeddings) unlike other LLM-based approaches [52] while also allowing us to identify stories without *a priori* knowledge. To further scale the approach, we re-implement DP-Means [38] to use the GPU-enhanced FAISS library [72] to perform the embedding-to-cluster assignments and similarity calculations required by DP-Means. To determine a suitable threshold for clustering two news passages together, after fine-tuning e5-base-v2, we benchmark our model on the English portion SemEval 2022 Task 8 dataset [29]. The SemEval 2022 Task 8 dataset consists of two parallel lists of news articles where each pair is graded on whether they are about the same news story. Our model achieves a max F_1-score of 0.793 on this dataset near a cosine similarity threshold of 0.50, which we use in this work. We provide examples of passage pairs in Appendix F.

From January 1, 2022 to July 1, 2023, clustering all our embeddings required the equivalent of 12 days using an NVIDIA A100 GPU. After clustering, like in other works [62, 85], we filter out clusters where 50% or more of the passages are from only one website (*e.g.*, website-specific headers or author bios). After this pruning, we identified 146,212 story clusters. We provide 30 cluster examples in Appendix I and evaluate these 30 clusters to ensure that they contain coherent stories using the method outlined by Hanley et al. [62]. We achieve an estimated precision of 99.3% of assigning passages to appropriate story clusters where each passage matches the summary, keywords, and other passages in the cluster.

Story Summarization and Labeling. To build human-understandable representations of our clusters, we extract keywords using pointwise mutual information (PMI), an information-theoretic for uncovering associations [23], to uncover the words most associated with each story cluster [59]. To make these words more uniform, we lemmatize each word to each cluster before calculating PMI. For details on PMI,

---

[1]https://huggingface.co/docs/peft/task_guides/semantic-similarity-lora

see Appendix D. In addition, we perform multi-document summarization utilizing an instruction fine-tuned version of Llama 3 [39].[2] This enables us to summarize the different perspectives of the passages within a given cluster, while also allowing humans to easily understand a story cluster's contents. We utilize the following prompt to summarize the contents of each of our clusters: *You work for a news researcher and your job is to summarize articles. Write a single concise collective abstractive summary of the texts, where individual texts are separated by |||||, and return your response as a single summary that covers the key points of the text.*

Website Relationship Inference. To further understand the relationships between news sites, we analyze how stories spread across websites over time. We consider the set of articles in a cluster as a time cascade based on the date that each article was published, and we use an open source version of NETINF [49] to infer the underlying structure and relationship amongst our set of news websites.[3] Given a set of time cascades (*e.g.,* the time steps for when a particular website posts an article within a given story cluster), while assuming that each node in a particular cascade is influenced by exactly one other node, the NETINF algorithm attempts to infer the optimal network to explain the observed posting behavior [49]. Based on each website posting behavior across the different cascades, NETINF estimates the number of times that each website copied information from another as well as the time delay between copies. We provide additional details in Appendix G.

Stance Detection. While passages may cover the same story, they often adopt different *stances* [61,78,99] in addressing the same event. After identifying the stories on our set of news websites, we employ stance detection to understand how different websites address each story. Stance detection methods determine the attitude of an author toward a specific topic or target [20]. Typically, stance detection involves taking a passage p_i and a topic or target t_i, and outputting the stance $s_i \in \{Pro, Against, Neutral\}$ of the passage towards the target, where the target is a *noun* or a *noun phrase*. Given that most stance detection methods heavily rely on the topic or target, with many models struggling to generalize to topics or targets outside their domain, various models have been developed to perform stance detection in *zero-shot* (where the tested topics or targets are not in the training data) and *few-shot* (where very few examples of the tested topics or targets are in the training data) settings [8,87].

To perform this stance detection, we utilize the current state-of-the-art zero-shot TATA model [55], which was trained on the VAST dataset [8]. We note that the size of our dataset of stance pairs precluded us from using popular large language model services like GPT-4 or Claude Sonnet. To enhance this model, we retrained it on both the VAST dataset and news-
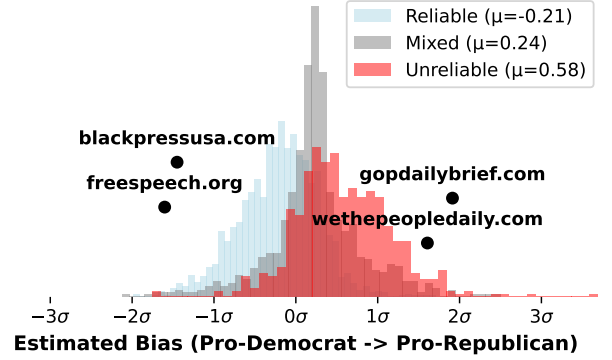

Figure 2: Estimated Partisanship via Bayesian regression of our websites based on their stances to articles' topics.

specific stance detection NewsMTSC dataset [53]. By training the TATA model using this extended dataset, we achieved state-of-the-art F_1 scores of 0.781 in the zero-shot setting and 0.741 in the few-shot setting on the VAST test dataset, and a macro F_1 score of 0.849 on the NewsMTSC test dataset.

Specifically, rather than performing stance detection on a pre-determined set of topics [51,78,82], we leverage our topic and story modeling to conduct stance detection across each story cluster. Once we extract story keywords using PMI, we utilize the Python `NLTK` library's Part-of-Speech (POS) tagging function to identify the most distinctive noun keywords [21], capturing the topic addressed in each passage. We further use the `NLTK` library to filter out common first names (*e.g.*, Michael, Jessica) from our stance detection algorithm and employ the Python `spaCy` library [135] to exclude nouns that fall into the following categories: *FAC, LOC, WORK_OF_ART, DATE, TIME, PERCENT, MONEY, QUANTITY, ORDINAL, CARDINAL*. This approach ensures that passages are not erroneously categorized as *Pro* or *Against* particular dates or monetary amounts. To ensure robust measurements of the collective ecosystems' and websites' stances toward specific entities, we gather the top 5,000 noun entities from our data and perform stance detection on each passage within each cluster where it appears among the top 10 PMI keywords. Altogether, this process involves running stance detection on 96.3M passage and keyword pairs, requiring the equivalent of 15 days of computation on a NVIDIA A100.

Interpretable Mapping of Websites' Biases. A simplistic approach to understanding a website's overall bias (*i.e* how anti or pro) toward an entity such as "Ukraine" would involve aggregating the percentage of their articles that had pro-"Ukraine" and anti-"Ukraine" stances (*i.e.*, % pro-Ukraine articles − %anti-Ukraine articles). However, this approach could potentially fail given that some websites may not have an abundance of articles focused on Ukraine or may only discuss Ukraine-related entities to obfuscate their bias. As such, taking inspiration from Waller et al. [137] who train Word2Vec models to predict subreddit's bias, we instead take

---

[2] https://huggingface.co/meta-llama/Meta-Llama-3.1-8B-Instruct
[3] https://snap.stanford.edu/netinf/

Reliable News	Mixed News	Unreliable News
Pro CDC	Against Kardashian	Against Pfizer
Pro Quantum	Pro Gunnar	Against Vaccine
Pro Senate	Pro Alnassar	Against Wuhan

Table 2: Keywords most associated with each news ecosystem estimated using PMI.

Republican Stances	Coeff.	Std.
Pro Souza	0.311	0.083
Pro America	0.245	0.122
Against Communist	0.215	0.102
Democratic Stances	Coeff.	Std.
Against Santos	-0.323	0.105
Against Texas	-0.315	0.075
Against Conservative	-0.282	0.098

Table 3: The stances most associated with U.S. partisanship factions, estimated using Bayseian regression.

a holistic approach by aggregating each website's respective stances to their written-about entities and predicting bias via Bayesian regression models.

To estimate websites' biases toward a subject along a given axis, we first gather a seed set of websites with at least 250 articles[4] discussing the entity and compute their simplistic bias score (*i.e.*, % pro-entity articles − % anti-entity articles). To make these values more interpretable, we normalize these scores as z-scores (*i.e.*, mean 0 and variance 1), such that a score of 1.0 can be interpreted as bias in favor of entity one standard deviation above the mean [137]. Following this calculation, we subsequently train a linear Bayesian regression model with L_2 regularization to predict this bias score by utilizing our seed set of websites' stances to other entities (besides the one in question). Finally, once trained, using the model, we estimate the rest of our websites' bias scores to the given entity. We adopt a Bayesian model approach as this directly enables us to quantify how individual stances contribute to our prediction of a given website's bias.

To validate this approach, we mapped our websites to partisanship scores along the US left–right political spectrum (Figure 2) using the keywords "democrat" and "republican," and a seed set of 105 websites. The partisanship scores from the resulting model had a $\rho = 0.51$ Spearman correlation with the partisanship labels (Far-Right, Right, Right-Center, Center, Left-Center, etc.) provided by Media-Bias/Fact-Check. As seen in Table 3, some of the most right-leaning partisan keywords included positive stances towards Dinesh D'Souza, a right-leaning commentator [140], and America, while having a negative stance toward communism. On the Democratic side, the associated stances include being against Texas, conservatives, and the former Republican Congressman George Santos. Similarly, as seen in Figure 3 and matching the partisan labels from Media-Bias/Fact-Check, we broadly observe that our set of reliable websites is left-leaning and the unreliable websites are right-leaning.

4 Characterizing News Ecosystems

Having detailed our methodology, we now characterize the ecosystem of reliable, mixed reliability, and unreliable news websites. Visualized in Figure 3, the most heavily discussed stories among our set of reliable news websites included

the U.S. Republican primary (62,911 articles), business news quarterly revenue (57,453 articles), the U.S. Supreme Court's decision to overturn federal abortion rights (Roe v. Wade) (38,358 articles), and the Russian invasion of Ukraine (35,135 articles). Looking at the top stories spread by unreliable news websites, we observe many of the same topics, most notably one concerning the U.S. Republican primary (13,393 articles). Indeed, across all shared story clusters (91,390 stories, 62.5%), we observe an average Pearson correlation of 0.501 between the volume of articles from our unreliable and reliable news websites.

Beyond these shared stories, we observe a focus on corruption and government failures (9,094 articles), the U.S. Federal Bureau of Investigation's (FBI) search of President Donald Trump's Mar-a-Lago estate (7,678 articles), and the investigation into Hunter Biden's (U.S. President Joe Biden's son) laptop (7,509 articles) [105] on unreliable websites. Finally, for our set of mixed-reliability news websites, we observe a heavy focus on sports and pop culture; two of the top five topics focus on the celebrity Kardashian family and one on the footballer Cristiano Ronaldo. Mixed-reliability news volume isalso highly correlated with the volume of stories on reliable (127,106/86.9% shared stories with a $\rho = 0.689$ Pearson correlation for the story volumes) and unreliable new websites (91,205/62.4% shared stories with a $\rho = 0.646$). We detail each ecosystem's stories in Appendix K.

Using the stance of each website toward the top 5,000 entities in our dataset, as output by our augmented TATA model, we further characterize the attitudes of our reliable, mixed-reliability, and unreliable news websites. To do this, we utilize PMI to determine the non-neutral stances most associated with each ecosystem (we limit this analysis to stances represented in at least 500 articles within each ecosystem to avoid spurious values; see Section D for details). As seen in Table 2, reliable news websites are more pro-CDC (Centers for Disease Control), pro-Quantum, and pro-Senate (than mixed-reliability and unreliable websites). The most distinctive stances of mixed-reliability websites concern pop culture and football (Gunnar is a Norwegian football manager and Al Nassr Football Club is a Saudi-Arabian football team). In contrast, the most distinctive stances among the unreliable news

---

[4]This ensures that the margin of error for probabilities is below 0.10 with a 95% confidence interval based on the normal distribution.
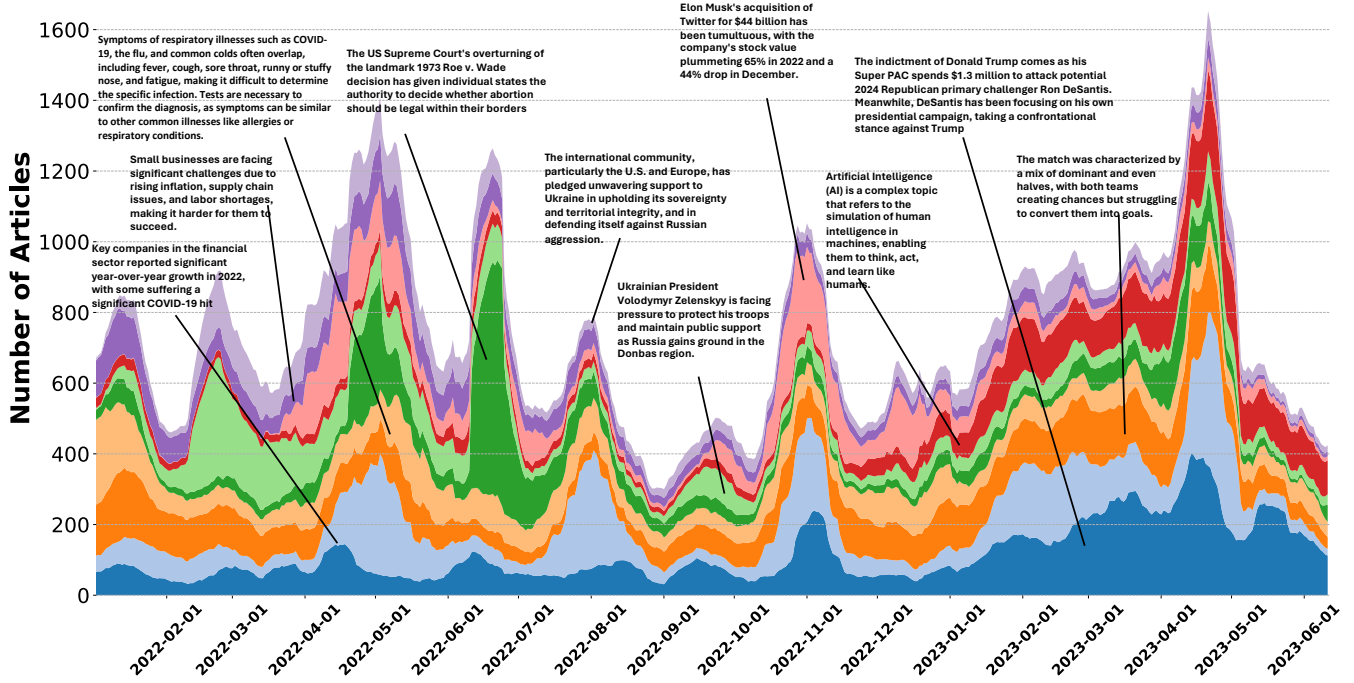
Figure 3: The most commonly discussed stories on reliable news websites labeled with their LLM-generated summaries.

websites primarily concern the COVID-19 pandemic, with these websites distinctly opposing vaccines, Pfizer (one of the leading companies that developed a COVID-19 vaccine), and Wuhan, China (the origin of COVID-19) [102].

The stances between different news ecosystems are fairly distinctive. Indeed, by fitting a random forest classifier to 80% (3,260 websites) of the websites' stance data based on their percentage for and against different entities (using 10% of the websites as validation (408 websites) and 10% as test data), we achieve an accuracy of 85.9% and an AUC of 0.889 in differentiating unreliable news websites from reliable and mixed-reliability websites. This illustrates the ease of differentiating between types of websites by their stances and the ability to predict a potentially unlabeled website's reliability based on its stance towards popular news stories.

Bias Case Study: Ukraine and Vaccines. Beyond the most distinctive stances that each website has, to further understand the underlying attitudes within each ecosystem, we perform a case study on each website ecosystem's attitudes towards *Ukraine* and *Vaccines*—two of the most commonly covered topics in our dataset—using the methodology outlined in Section 3.3. While this analysis specifically addresses Ukraine and vaccines, similar to how we analyzed U.S.-based political partisanship in Section 3.3, this approach can be applied to any popular entity within our dataset. We additionally present analyses for America, China, and Iran in Appendix J.

Fitting our Bayesian regression models with the stances for the remaining 4,999 entities for both Ukraine and vaccines, we map all of our news articles to a bias latent for both

Pro-Ukraine Stances	Coeff.	Std.
Pro Zelenskyy	0.378	0.114
Pro Zelensky	0.368	0.110
Against Syria	0.225	0.114
Anti-Ukraine Stances		
Against Zelenskiy	-0.500	0.111
Against Biden	-0.370	0.103
Against DHS	-0.345	0.119

Table 4: Stances associated with Ukraine estimated by a Bayesian regression model.

entities in Figure 4. We observe that reliable news websites express higher support for Ukraine and vaccines ($\mu_{vaccine}$=0.32, $\mu_{ukraine}$=0.36), while unreliable news websites oppose both ($\mu_{vaccine}$=-0.82, $\mu_{ukraine}$=-0.87), and mixed-reliability websites in the middle ($\mu_{vaccine} = -0.11, \mu_{ukraine} = 0.04$). This matches the average distribution where 31.9% of unreliable news articles were anti-Ukraine and 23.8% were anti-vaccine; for mixed-reliability websites, 17.8% were anti-Ukraine and 8.5% were anti-vaccine; and for reliable news websites 12.9% of articles were anti-Ukraine and 7.1% were anti-vaccine.

Among our dataset, the news sites most anti-Ukraine include rt.com ($z_{ukraine}$ = -2.36), strategic-culture.org ($z_{ukraine}$ = -2.54), and southfront.org ($z_{ukraine}$ = -2.41)—three websites known for spreading Russian propaganda [106]. Some of the most pro-Ukraine websites include nationaljournal.com ($z_{ukraine}$ = +2.53), a U.S. political policy-oriented website, kyivpost.com ($z_{ukraine}$ = +0.80), a Ukrainian website, as well
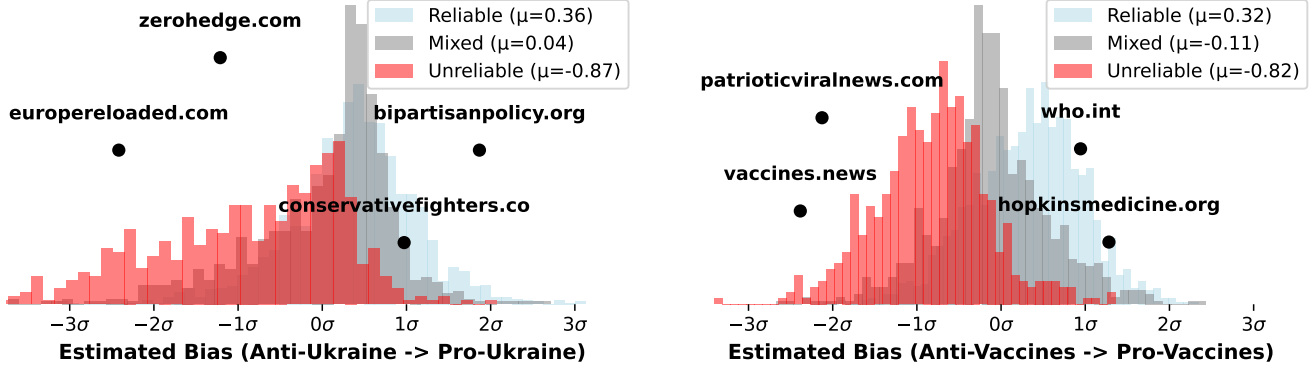
Figure 4: Distribution of Ukraine and Vaccine bias across unreliable, mixed-reliability, and reliable news websites estimated by Bayesian regression models.

Table 5: Stances associated with vaccines estimated by a Bayesian regression model.

Pro-Vaccine Stances	Coeff.	Std.
Pro Ukraine	0.343	0.111
Pro Trans	0.259	0.110
Pro Healthcare	0.233	0.093
Anti-Vaccine Stances		
Against COVID	-0.360	0.144
Against FDA	-0.334	0.122
Against Pfizer-BioNTech	-0.333	0.143

as a selection of NBC and ABC affiliate websites including wbaltv.com ($z_{ukraine} = +2.04$), wvtm13.com ($z_{ukraine} = +2.14$), and ketv.com ($z_{ukraine} = +2.55$) [33]. The most anti-vaccine websites are vaccineimpact.com ($z_{vaccine} = -3.41$) and pantsonfirenews.com ($z_{vaccine}$ = -2.56), both known for spreading misinformation [33]. Conversely, the most pro-vaccine websites include Johns Hopkins ($z_{vaccine} = +1.28$) and the World Health Organization ($z_{vaccine} = +0.94$).

Examining the stances most associated with each topic latent (Tables 4 and 5), we observe that for Ukraine, this includes being pro the current president of Ukraine, Volodymyr Zelensky [131]. Beyond this entity, we further observe the entities associated with attitudes towards toward Ukraine include other Ukrainian allies (*e.g.*, Biden and DHS) and countries in the Global South that have battled for attention and aid following the Russian invasion of Ukraine [25]. For the vaccine latent, we observe that website stances most associated with being pro vaccines have to do with being pro-health interventions like healthcare, as well as left-leaning causes like transgender rights and Ukraine [75, 77]. In contrast, we observe that being against vaccines is associated with being against COVID (the cause of the polarization of vaccination [75]), the US Food and Drug Administration (FDA), and Pfizer, one of the companies that developed COVID-19 vaccines [75, 122].

5 Underlying Website Relationships

As observed in Section 4, unreliable, mixed-reliability, and reliable news websites often cover the same stories simultaneously, suggesting an interdependence [127]. To further understand these relationships, we utilize an open source version of the NETINF [49] algorithm to infer the underlying structure and relationships amongst our sets of news websites [85]. Specifically, we first run NETINF using all of the extracted stories within our dataset as time cascades. To determine the appropriate number of iterations to run NETINF algorithm, as in Gomez et al. [49], we utilize the point at which the marginal gain of adding new edges plateaus (90% of the total marginal gain). We find that margin gain reaches a plateau at 37,670 iterations in our dataset.

Ecosystem Relationships Across All News Stories. Using the estimated number of copies between websites and the time delay between copies as found by NETINF, we first examine the overall relationships between ecosystems. We find that reliable and mixed-reliability news websites have a large role in introducing stories adopted by the rest of the news ecosystem. As seen in Figure 5a, 60% of the news articles on reliable news websites that were copied/influenced from elsewhere came from other reliable news websites, 43% on mixed-reliability websites came from reliable sites, and 32% on unreliable sites came from reliable sites. Unreliable news websites had significantly less influence, with only 10% of the stories on reliable news websites originating from unreliable news websites (13% for mixed-reliability, 27% for unreliable).

Looking at the set of reliable websites that are the most common sources of copied stories throughout the entire news ecosystem, we see several popular websites including ya-hoo.com (1.19%), apnews.com (0.73%), abcnews.go.com (0.60%), and cnn.com (0.60%). Despite popular, reliable news websites being common sources, website popularity had only a slight Pearson correlation with their percentage of copies. Using data from the Google Chrome User Report (CrUX)

(a) Copies for All Stories   (b) Δ-Copy Times for All Stories   (c) Copies for Unrel. Stories   (d) Δ-Copy Times for Unrel. Stories
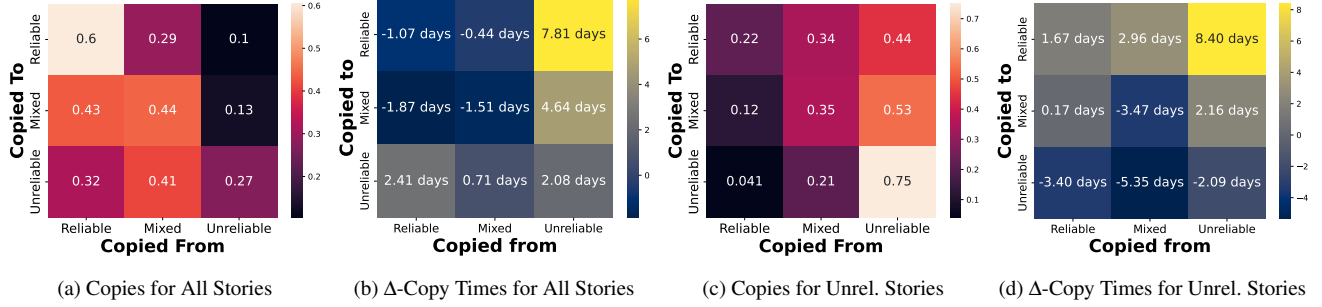
Figure 5: The percentage of each ecosystems copied stories that came from each different ecosystem as well as the change in the average time delay between website copy/reposting on the same narrative depending on the combination of news ecosystems.

from October 2022 (which Ruth et al. [116,117] showed to be the most reliable website popularity metric), we find that for unreliable news websites copying from reliable websites, the corresponding reliable websites' popularity had a correlation of ρ=0.225 with the tendency of unreliable website's to copy from them (ρ=0.189 for reliable websites copying from reliable websites, ρ=0.311 for mixed-reliability news websites copying from reliable websites).

As seen in Figure 5b, reliable news websites adopt the stories of other reliable news websites more quickly (-1.07 days) compared to the average copy delay (38.4 days). Mann-Whitney U-tests indicate that these differences are all significant. This compares to a nearly +7.81 day additional delay of reliable news websites picking up the stories from unreliable news websites and -0.44 days from mixed-reliability websites. We find a similar pattern amongst mixed-reliability websites, who adopt stories from reliable news sites (-1.87 days) more quickly than from unreliable news sites (+4.64 days).

Influence on the Full News Ecosystem. Having examined the website copies and rates of adoption between the different ecosystems, we next consider which websites are the most influential using the graph of the edge connections between individual news sites. Eigenvector centralities are utilized to determine the relative influence of nodes within graphs [115] and, as such, we utilize this metric to understand websites' influence. We further compute hub centralities as a metric for websites' influence in originating stories that spread to other websites (given the directionality of the arrows in our graph, this metric determines the most important websites for supplying content [76]). We show the most influential sites in Table 6 and Figure 6.

All stories	Hub	Eign.
yahoo.com	0.149	0.111
apnews.com	0.101	0.092
dailymail.co.uk	0.097	0.099
nypost.com	0.052	0.076
independent.co.uk	0.049	0.097

Table 6: Websites with the largest influence in the underlying influence graph determined by NETINF with all stories considered.

We find that website popularity is correlated with the relative influence of websites within the news ecosystems (when looking at all news sites compared to only reliable sites in the last section). Again using website popularity data from the Google Chrome User Report (CrUX), we find that a website's eigenvector centrality/influence has a Spearman correlation of ρ = 0.571 (0.396 for hub centrality) with that website's popularity rank. Despite making up 24.6% of the news websites in our dataset, unreliable news websites do not make up a proportional percentage amongst the most influential news websites. Directly comparing the eigenvector centralities of the reliable news websites to those of the unreliable news websites, we find that reliable news websites are significantly more influential in this ecosystem than unreliable news websites (Cohen's D = 0.64, p-value ≈ 0),[5] with mixed reliability websites having comparable influence to reliable ones (no significant difference through Mann Whitney U-test). In terms of origination (hub centralities), we observe a slightly different trend with mixed-reliability websites having slightly more influence in originating stories compared to reliable news websites (Cohen's D = 0.04, p-value ≈ 0) and unreliable news websites (Cohen's D = 0.05, p-value ≈ 0).

Stories Spread by Unreliable News Websites. To identify the websites most influential in spreading potentially unreliable stories, we run the NETINF algorithm on the set of 6,762 news stories where unreliable news websites posted the plurality of articles about those stories. Again applying the same methodology for identifying the appropriate number of edges to add, the marginal gain plateaus at 16,196 edges. The most popular story amongst these clusters was about government censorship and control (9,094 articles) summarized as: *There is censorship, propaganda, and government control in the US. Cancel culture is a form of censorship, and that government-funded media outlets can exercise control over editorial content. The text also warns about the influence of the "Deep State" and far-left communists in US institutions, including the government, media, education, and Big Business.*

As expected, given how we narrow our set of stories, as seen in Figure 5c, relative to all news stories, unreliable news

---

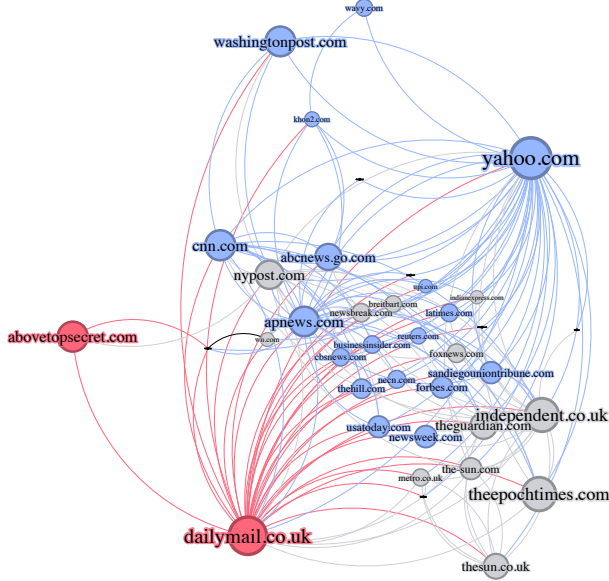[5]The p-value is computed using the Mann-Whitney U-test.

Figure 6: The most influential websites and their interactions. The size of nodes is proportional to their hub centrality. Reliable news websites are colored blue, mixed-reliability websites are colored grey, and unreliable news websites are colored red.

Reliable	Propor.
dailymail.co.uk	0.140
abovetopsecret.com	0.039
ussanews.com	0.035
Mixed	
dailymail.co.uk	0.062
thegatewaypundit.com	0.030
ussanews.com	0.027
Unreliable	
naturalnews.com	0.025
ussanews.com	0.021
theburningplatform.com	0.020

Table 7: Websites that are the most common source of unreliable news stories for each news ecosystem.

Predom. Unreliable News Stories	Hub	Eign.
thegatewaypundit.com	0.129	0.109
dailymail.co.uk	0.075	0.125
theburningplatform.com	0.060	0.103

Table 8: Websites with the largest influence in the underlying influence graph for stories predominated spread by unreliable websites.

websites had significantly more influence in originating potentially unreliable content. For example, while for all stories, reliable news websites sourced less than 10% of all of their stories from unreliable news websites, within this specific set of news stories, the figure was 44%. Similarly, for mixed-reliability websites, this percentage increased from 13% to 53%. Furthermore, we find that unreliable websites source the majority of their influenced or copied stories from other unreliable news websites, at a rate of 75%. Looking at the set of websites that are the common source for other sites to copy from (Table 7), we find a heavy reliance on dailymail.co.uk, a United Kingdom-based tabloid that Media-Bias/Fact-Check describes as having "low" factual reporting due to "numerous failed fact checks and poor information sourcing." We also find that ussanews.com, described by Media-Bias/Fact-Check as promoting "entirely false, so-called facts," was a common source of unreliable news stories.

For this selection of news stories predominately published by unreliable news websites, comparing the copy times of these stories in Figure 5d to those in Figure 5b, we find that reliable news websites are slower to adopt the stories, regardless of from which news ecosystem the story originated. We thus observe a reticence amongst our reliable news websites to report on the news stories primarily spread by unreliable news outlets. However, we find that for mixed-reliability websites, if the news story began amongst other mixed-reliability news outlets, these news outlets are faster to adopt the story (-3.47 days). We further observe that unreliable news websites are the fastest at picking up these news stories compared, picking them up quicker if they initially came from a mixed-reliability (-5.35 days) or reliable news website (-3.40 days).

Influence in the Unreliable News Ecosystem. To understand which websites are the most influential in the unreliable news ecosystem, we utilize the eigenvector centrality of each website in the resultant graph created by running NETINF on our set of predominantly unreliable news stories (Table 8). For this ecosystem, we find the popularity of websites is only slightly correlated with eigenvector centrality/influence ($\rho = 0.158$) and hub centrality ($\rho = 0.175$). Examining the set of websites that are most prominent within the unreliable news ecosystem (Table 8 and Figure 7), we find that many well-documented websites known for spreading unreliable information are among the most prominent, including theepochtimes.com, dailymail.co.uk, and thegatewaypundit.com [127].

Comparing the eigenvector centralities of the unreliable news websites to those of the authentic news websites, we find that unreliable news websites are more influential within this ecosystem (Cohen's D = 0.219, p-value < 0.001), but that unreliable and mixed-reliability websites had comparable influence (no significant difference via the Mann-Whitney U-test). However, most notably, we observe that among the top influencers within this ecosystem are the reliable news website, Yahoo News, and the mixed-reliability Fox News (not shown in the table). Yahoo News primarily serves as a news aggregator, gathering reports from various sources including Fox News, the BBC, and Reuters [143]. Given its role as an aggregator, Yahoo News appears to have a prominent role in disseminating current events that are reported by other outlets.
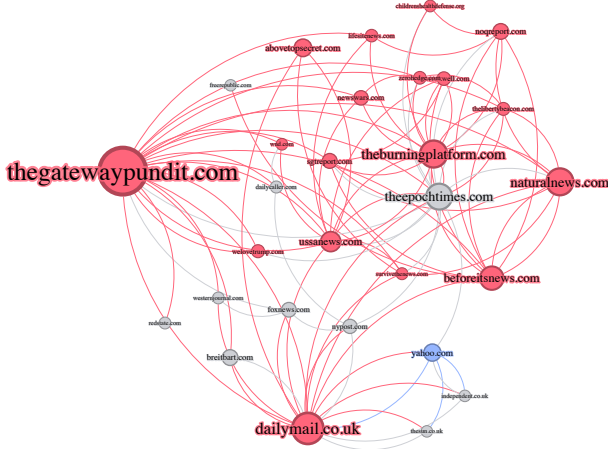
Figure 7: Most influential websites and their interaction for stories that are predominantly spread by unreliable news websites. Nodes sizes are proportional to their hub centralities.

Figure 8: Anti-Ukraine Copy Matrix.

Classified as a mixed-reliability news source, Fox News has been widely commented upon for its role in disseminating hyperpartisan news and misinformation [18, 34, 67].

Case Study: News Website Coordination. To identify potential coordination among our websites, we utilize NETINF to discern the relationships between websites involved in stories predominantly published articles spread by unreliable and mixed-reliability news websites (encompassing 40,325 news stories). After running the NETINF algorithm, we further clustered the resulting graph using the Louvain clustering algorithm [35]. Qualitatively, the largest of these clusters comprised 885 relatively mainstream and tabloid websites that report on general news (*e.g.*, wpxi.com, nbc29.com, wvva.com), with the top stories concerning the Kardashians (*Keywords: Kourtney, Kardashian, Travis, Khloe, Barker*). The second largest cluster consisted of 492 locally-oriented news websites (*e.g.*, cbs4local.com, idahostatejournal.com), where the top stories focused on immigration (*Migrant, Border, Patrol, Customs, Smuggling*) and the US Constitution (*Constitution, Oath, Amendment, Constitutional*). Finally, the third largest cluster (*Crore, Yoy, Profit, FY23, Quarter*) included 334 international websites (*e.g.*, sputniknews.com, alarabiya.net), where the top story involved international companies' profits.

Most notably among our clusters was a set of 338 websites, all with seemingly innocuous names such as southindynews.com and northalaskanews.com, which appeared to be dedicated to local news. Upon further investigation through querying WHOIS, we discovered that each of these websites was registered by the domain registrar Epik, Inc., a popular provider for misinformation and online hate sites [54]. We find that this set of 338 ostensibly local websites is owned and operated by the same entity, Metric Media LLC, which produces algorithmically generated content and promotes rightwing views [144]. Indeed, using our mapping of websites to their respective political partisanship, we found that despite
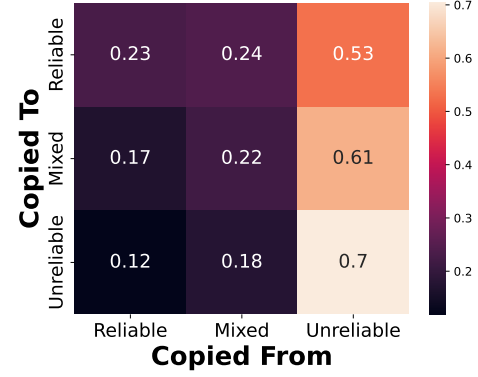
these websites rarely writing articles about Republicans or Democrats, they have an average partisanship $\mu_{politics} = 0.22$, indicating a slight right-leaning bias, with 88.2% of these websites classified as right-leaning. These websites largely repeat the same text including articles promoting herd immunity from COVID-19 in the United States: *More than 50 percent of US citizens are considered fully vaccinated against COVID-19, nearing the target for "herd immunity" Herd immunity happens when enough of the population has become immune to the virus from the previous infection that it effectively protects those who are not immune.*

6 Propaganda and Slanted Influence Networks

As seen in the last section, news sites, regardless of their factual reliability, often report on the same stories, with unreliable news websites in select cases influencing both reliable and mixed-reliability news platforms. Furthermore, while reliable and mixed-reliability news websites predominantly adopt stories from other reliable and mixed-reliability sources (Figure 5a), for topics primarily spread by unreliable news websites, these specious sources often act as the originators of the content (Figure 5c). Within this vein, tracking the spread of unreliable news and propaganda and determining which sources are most effective at seeding these stories into the mainstream media is critical for fact-checkers, journalists, and researchers [62, 127]. To this effect, in this section, we utilize our system to map out and understand the sites originating and spreading specific propaganda and influence campaigns.

To map the influence networks targeting specific entities (either positively or negatively), we gather news articles and the associated sites that exhibit a particular valence towards a given entity (*e.g.*, anti-vaccine articles). Upon gathering this subset of news articles, we run the NETINF algorithm over these cascades of news article clusters with specific stances. We subsequently perform network analysis using eigenvector centrality and hub centrality, as discussed in the previous section, to identify the most prominent and influential sites promoting a given stance towards a particular subject. By

further examining day-to-day increases in news stories with specific stances and comparing their spread in reliable and unreliable news ecosystems, we further document *new* individual stories meant to spread particular views or stances.

This programmatic approach can help identify stories that are receiving renewed focus from unreliable news websites and which websites are influential in propagating stances towards entities of interest in a particularly damaging manner, thereby facilitating the identification and mitigation of misinformation [62, 113, 118, 141]. To illustrate, we perform this analysis for anti-vaccine and anti-Ukraine news stories.

Anti-Ukraine	Hub	Eign.
rt.com	0.155	0.210
sputniknews.com	0.073	0.129
news-front.info	0.054	0.163
Anti-Vaccine		
naturalnews.com	0.141	0.179
theepochtimes.com	0.086	0.196
vaccines.news	0.049	0.170

Table 9: Websites with the largest influence in the underlying influence graph of anti-vaccine and anti-Ukraine news.

Anti-Ukraine Messaging. As seen in Figure 9 and Table 9, the most prominent anti-Ukrainian news websites during our study included well-known Russian propaganda websites such as Russia Today (RT), Sputnik News, and News-Front [106]. Beyond known Russian propaganda sites, we also find that antiwar.com, described as a "libertarian non-interventionist website" [33], is one of the most prominent websites in spreading anti-Ukrainian content. Altogether, as seen in Figure 8, unreliable news websites largely supply the majority of the stories used across the entire news ecosystem, with reliable websites copying 53% of anti-Ukrainian stories from unreliable outlets, mixed-reliability websites copying 61%, and unreliable sites 70%.

The most common story pushed in this ecosystem of websites concerned justifications for Russia's invasion of Ukraine, with one Russia Today article writing [132]: *Moscow attacked the neighboring state in late February, following Ukraine's failure to implement the terms of the Minsk agreements signed in 2014, and Russia's eventual recognition of the Donbass republics of Donetsk and Lugansk.* Further, the anti-Ukrainian news story that received the largest increase in relative popularity among our unreliable news websites in the last week of our study (June 25 to July 1, 2023) featured a series of articles with the keywords *Ukraine, MacGregor, Douglas, Colonel, Zelensky*, showing a ratio of 9 articles in the unreliable news ecosystem for every 1 article in the reliable news ecosystem. This story, which predominantly spread within the unreliable news ecosystem, concerned an interview with retired US Colonel Douglas MacGregor suggesting that the war between Ukraine and Russia was unwinnable and that Ukrainian
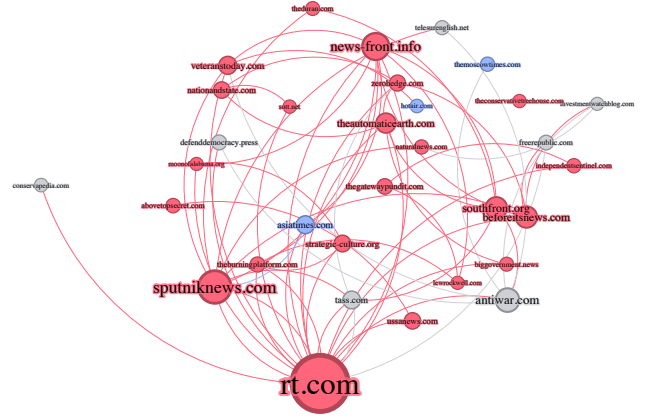

Figure 9: Anti-Ukraine Influence Network determined by the NET-INF algorithm. The nodes' sizes are proportional to their hub centralities.

President Zelensky was a puppet of Western powers: *"The war is really over for the Ukrainians. I don't see anything heroic about the man. And I think the most heroic thing he can do right now is to come to terms with reality," retired Army Colonel Douglas MacGregor told Fox Business News. "I think Zelensky is a puppet, and he is putting huge numbers of his own population in unnecessary risk," he said.* The website that spread this story the most was paulcraigroberts.org ($z_{ukraine}$=-3.59, 3 articles).

Beyond the set of unreliable news websites spreading anti-Ukrainian messaging, we further observe several international news websites including asiatimes.com ($z_{ukraine}$ =-1.13), themoscowtimes.com ($z_{ukraine}$ =-0.67), and the right-leaning website hotair.com ($z_{ukraine}$ =-1.21) as purveyors of influential anti-Ukrainian content in this ecosystem. Based on the inward-weighted edges, the most influenced mainstream news website in this ecosystem was haaretz.com, an Israeli outlet ($z_{ukraine}$ =+0.002 for Ukraine bias), and the most influenced mixed-reliability news website was salon.com ($z_{ukraine}$ = -0.056), a US-based left-leaning news outlet. We thus observe that even relatively neutral and pro-Ukrainian websites can be potentially influenced by anti-Ukrainian news articles.

Anti-Vaccine Messaging. As seen in Figure 11 and Table 9, the largest source of anti-vaccine stories was naturalnews.com, while the most influential anti-vaccine website was theepochtimes.com, both known for spreading anti-vaccine misinformation [33, 111]. As with anti-Ukraine stories, we observe that each website category predominantly sourced their content from unreliable news websites: 51% for reliable news websites, 66% for mixed-reliability news websites, and 81% for unreliable news websites. In addition to the theepochtimes.com and naturalnews.com, we find that childrenshealthdefense.org, a website associated with former presidential candidate Robert F. Kennedy Jr., had a major influence on spreading anti-vaccine content, including one article sug-
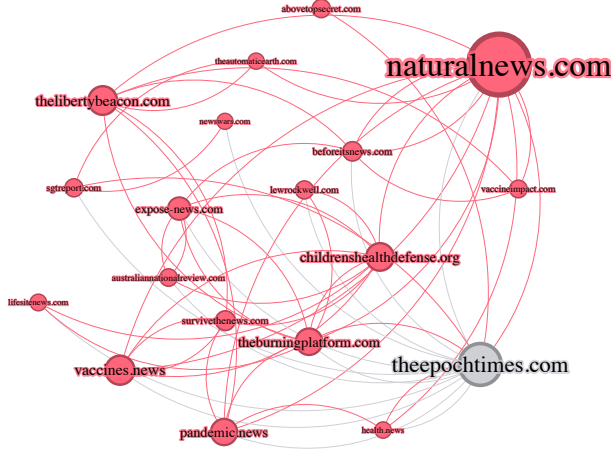
Figure 10: Anti-Vaccine Influence Network determined by NETINF. The nodes' sizes are proportional to their hub centralities.
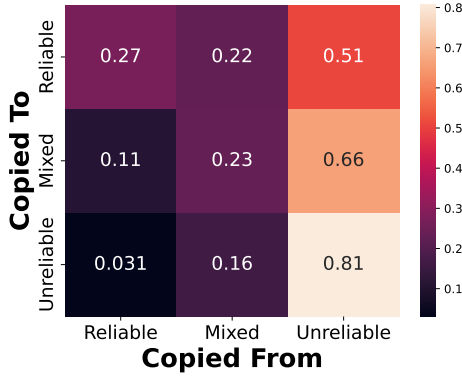

Figure 11: Anti-Vaccine Copy Matrix

gesting that a vaccine was not as safe as the US Food and Drug Administration claimed [30].

The most prominent-anti-vaccine story in terms of article volume raised concerns about children receiving COVID-19 vaccines, as highlighted by childrenshealthdefense.org [97]: *Pfizer, at the urging of federal health officials, is hustling to get infants and toddlers injected with experimental COVID vaccines*. The story that saw the largest relative increase in news articles (14 articles in the unreliable news ecosystem for every 1 in the reliable news ecosystem) was one with the keywords *Pfizer, Batch, Danish, Bnt162b2, Adverse*. This story concerned Danish scientists ostensibly discovering that batches of Pfizer vaccines were actually placebos: *The Danish scientists uncovered "compelling evidence" that a significant percentage of the batches distributed in the EU likely consisted of "placebos and non-placebos," prompting the researchers to call for further investigation.*. The top sites that spread this narrative were sgtreport.com ($z_{vaccine}$=-1.31 for vaccine-bias), theautomaticearth.com ($z_{vaccine}$ =-1.08), and theburningplatform.com ($z_{vaccine}$=-1.86) with two articles each.

We find that the reliable news website most influenced (by the weighted in-degree within the resulting NETINF graph) was sciencebasedmedicine.org ($z_{vaccine}$= +0.06), which frequently reports on and quotes anti-vaccine information [94], detected by our system. Additionally, the most influenced mixed-reliability website (besides theepochtimes.com) was thelibertyloft.com ($z_{vaccine}$=-0.81), a right-leaning website that Media-Bias/Fact-Check has identified as spreading COVID-19 related misinformation [33].

7 Limitations and Future Work

Our work shows the promise of mapping the global trajectories of news stories and the takes of news sites towards specific entities. However, we the emphasize the complexity of the global news ecosystem and the considerable future work that remains to understand how information travels online. Below, we discuss the limitations of our work and potential future research directions.

English-Language Websites. Our work is limited to English-language news articles and focuses predominantly on US, UK, and Australian websites. As a result, our analysis of the spread of particular stories is limited largely to the English speaking world and could miss other sources of news (*i.e.*, a Russian-language website for example may be more influential in spreading pro-Russian propaganda than the websites in our dataset). This restriction is largely due to our use of PMI for identifying keywords for stance detection amongst our story clusters, which does not directly work in a multilingual setting. Similarly, we currently lack highly accurate multilingual topic-agnostic stance-detection models [55, 63]. We encourage future work to consider how to semantically map both news topics and stances towards them in multilingual settings, as well as to consider how to source news content from websites in additional languages.

Automated Fact-Checking of Narratives. As previously noted, we do not fact check individual news stories, which we argue is a journalistic task beyond the scope of our automated approach. While our system can be utilized to uncover networks of websites pushing potentially unreliable news narratives allowing journalists to prioritize which stories need to be fact-checked by their relative spread, these stories still require human investigation to determine their veracity. However, we note that for stories that have already fact-checked on reputable websites, it may be possible to incorporate the approaches of Hanley et al. [62], Zhou et al. [148], and others to automatically label particular stories. Hanley et al.'s approach involves gathering fact-checks from reputable sources and using a DeBERTa-based model [65] to identify unreliable news stories that directly contradict these fact-checks [62]. In a similar fashion, Zhou et al.'s [148] approach involves using a LLM agent and Google Search to identify which unreliable news stories contradict fact-checks.

Ephemeral Unreliable News Websites. Factually unreliable news tend to be ephemeral [32, 58, 68, 101], often only being active long enough to spread misinformation to other platforms before shutting down themselves. As such, finding news sites as soon as they come online is critical long term. We note that while our current system relies on previously curated lists of websites, it can easily incorporate new websites as they appear (*e.g.*, using the methods outlined by Hounsel et al. [68] for identifying new unreliable news websites based on their domain registration and network infrastructure characteristics). This inclusion would enable our system to surface potentially unreliable news stories that have not spread onto more popular websites. It would also potentially enable uncovering malicious Doppelgänger sites that masquerade as ordinary local websites but that actually spread propaganda as soon as they come online [46, 90, 134] similar to past work that has detected phishing and malware domains [12, 13].

8 Discussion and Conclusion

In this work, we investigated the spread and stance of news stories across 4,076 news websites from January 1, 2022, to July 1, 2023. Our approach, which advances previous methodologies for understanding news flows by incorporating stance into how we track semantic narratives, allows us to track stories across a mix of reliable, mixed-reliability, and unreliable news websites. (Neglecting stance in understanding the spread of narratives, while helpful for examining a singular ecosystem [62, 127], would likely led to misrepresentations of the interactions between different types of websites.)

Our work demonstrates the key role that reliable news platforms play in dictating the stories covered by the entire news ecosystem. These popular and largely factual websites maintain the largest degree of influence on the broader news ecosystem (Figure 6) and are the source of much content on mixed-reliability and unreliable websites (Figure 5). To understand which stories unreliable websites will spin or contort, researchers should consider reliable outlets as agenda-setters [24, 45, 91]. However, we simultaneously highlight that while a minimum of 62.4% of stories are shared between different types of news websites (Section 4), different ecosystems often have distinctive attitudes towards stories. For example, using our analysis, inline with prior work [44, 48], we show that current lists of unreliable websites, among other biases, tend to be more conservative and have distinctive biases against COVID-19 vaccines and Pfizer (Table 2), informing which larger topics may particularly need additional fact-checking. We also find that biased coverage of particular entities (*e.g.*, Ukraine, or vaccines) that otherwise reliable news websites produce are often sourced from unreliable news sites.

Finally, our work demonstrates how, by analyzing the stance of articles towards specific topics, we can uncover and understand influence networks directed at specific entities, facilitating the tracking of propaganda (*e.g.*, anti-Ukraine) or misinformation (*e.g.*, anti-vaccine) within the news ecosystem. This method also aids in identifying which otherwise reliable news sources may be influenced by disinformation and propaganda campaigns. Our approach, which considers the context of authentic and mainstream websites, provides a valuable tool for identifying dubious networks of websites spreading particular types of slanted information, which we argue can assist fact-checkers, journalists, and researchers in better understanding potential online misinformation.

We hope that our work encourages further quantitative analysis of the distributed news ecosystem, particular as social media platforms become more opaque to researchers. Prior security research has uncovered weaknesses and attacks through large-scale analysis (*e.g.*, [2, 40, 42, 66, 95, 128–130]), and we argue that there is significant potential for future work within the security community on understanding attacks against and strengthening the resilience of the news ecosystem.

Ethical Considerations

Trustworthy news media is fundamental to a democratic society. Previously, false information has incited real-world violence and had major consequences on public health and elections. Disinformation and propaganda are *attacks*, and it behooves the security community to understand how these attacks are conducted and how to build better defenses against them. Advances in this space help both citizens and news outlets themselves, who regularly fact-check articles. At the same time, like all active measurements, web crawling and programmatic analysis of online content have potential ethical ramifications that we must carefully consider.

Our work collects only publicly available news content in line with prior work (*e.g.*, [60, 123, 124]). We follow best practices when scraping websites by slowly collecting content over time to reduce load. Our scraping also includes built-in safety mechanisms to prevent making requests more often than once every 10 seconds. We never attempt to access any privileged or private data but rather focus on public stories that are linked from news platforms' public homepages.

We also adhere to the best practices set forth for conducting active Internet measurements [2, 41, 43]. The servers we use for collecting content are identified as part of a research study through WHOIS, reverse DNS, and informational websites

that indicate how to reach the researchers. Our IT and security teams are also informed about how to route any questions, requests, or complaints to our team. We received no requests to opt out of our data collection during our study.

Our study does not generate any new content or redistribute existing content. Instead, we analyze how context spreads. We emphasize that while we utilize labels of individual websites as *unreliable* or *mixed-reliability* from Media-Bias/Fact-Check [33] and on existing previously-curated lists, this does not necessarily mean that every news story spread by these websites is misinformation. Many unreliable news websites report factual information [127], and at times, otherwise reliable websites may mistakenly report incorrect information. We only label stories that have been previously and individually expertly labeled as *misinformation*.

Open Science

We are committed to sharing our data with other researchers at academic or non-profit institutions seeking to conduct future work or re-implement our approach. We will publicly release the weights and the code for the models used in this study. Additionally, we will supply the URLs of crawled news stories used in this study upon request.

References

[1] S. Abdali, R. Gurav, S. Menon, D. Fonseca, N. Entezari, N. Shah, and E. E. Papalexakis. Identifying misinformation from website screenshots. In *Intl. AAAI Conf. on Web & Social Media*, 2021.

[2] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz. The web never forgets: Persistent tracking mechanisms in the wild. In *ACM SIGSAC Conference on Computer and Communications Security*, 2014.

[3] S. Afroz, A. C. Islam, A. Stolerman, R. Greenstadt, and D. McCoy. Doppelgänger finder: Taking stylometry to the underground. In *IEEE Symposium on Security and Privacy*, 2014.

[4] H. Aghababaeian, L. Hamdanieh, and A. Ostadtaghizadeh. Alcohol intake in an attempt to fight covid-19: A medical myth in iran. *Alcohol*, 88:29–32, 2020.

[5] R. Albalawi, T. H. Yeap, and M. Benyoucef. Using topic modeling methods for short-text data: A comparative analysis. *Frontiers in artificial intelligence*, 3:42, 2020.

[6] M. Aliapoulios, A. Papasavva, C. Ballard, E. De Cristofaro, G. Stringhini, S. Zannettou, and J. Blackburn. The gospel according to Q: Understanding the QAnon conspiracy from the perspective of canonical information. In *Intl. AAAI Conf. on Web & Social Media*, 2022.

[7] J. Allan. Detection as multi-topic tracking. *Information Retrieval*, 5(2-3):139–157, 2002.

[8] E. Allaway and K. McKeown. Zero-Shot Stance Detection: A Dataset and Model using Generalized Topic Representations. In *Empirical Methods in Natural Language Processing*, 2020.

[9] H. Allcott and M. Gentzkow. Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2), 2017.

[10] J. Allen, B. Howland, M. Mobius, D. Rothschild, and D. J. Watts. Evaluating the fake news problem at the scale of the information ecosystem. *Science advances*, 6(14), 2020.

[11] A. Amarasingam and M.-A. Argentino. The qanon conspiracy theory: A security threat in the making. *CTC Sentinel*, 13(7):37–44, 2020.

[12] M. Antonakakis, R. Perdisci, D. Dagon, W. Lee, and N. Feamster. Building a dynamic reputation system for DNS. In *USENIX Security Symposium*, 2010.

[13] M. Antonakakis, R. Perdisci, W. Lee, N. Vasiloglou II, and D. Dagon. Detecting malware domains at the upper DNS hierarchy. In *USENIX Security Symposium*, 2011.

[14] M. Babaei, J. Kulshrestha, A. Chakraborty, E. M. Redmiles, M. Cha, and K. P. Gummadi. Analyzing biases in perception of truth in news stories and their implications for fact checking. *IEEE Transactions on Computational Social Systems*, 9(3):839–850, 2021.

[15] J. B. Bak-Coleman, I. Kennedy, M. Wack, A. Beers, J. S. Schafer, E. S. Spiro, K. Starbird, and J. D. West. Combining interventions to reduce the spread of viral misinformation. *Nature Human Behaviour*, 6(10):1372–1380, 2022.

[16] P. Ball and A. Maxmen. The epic battle against coronavirus misinformation and conspiracy theories. *Nature*, 2020.

[17] S. Banaji, R. Bhat, A. Agarwal, N. Passanha, and M. Sadhana Pravin. Whatsapp vigilantes: An exploration of citizen reception and circulation of whatsapp misinformation linked to mob violence in India. 2019.

[18] A. Bauer, A. Nadler, and J. L. Nelson. What is fox news? partisan journalism, misinformation, and the problem of classification. *Electronic News*, 16(1):18–29, 2022.

[19] P. Biancovilli, L. Makszin, and C. Jurberg. Misinformation on social networks during the novel coronavirus pandemic: a quali-quantitative case study of Brazil. *BMC Public Health*, 21(1):1–10, 2021.

[20] D. Biber and E. Finegan. Adverbial stance types in english. *Discourse processes*, 11(1), 1988.

[21] S. Bird, E. Klein, and E. Loper. *Natural language processing with Python: analyzing text with the natural language toolkit*. 2009.

[22] D. M. Blei, T. L. Griffiths, and M. I. Jordan. The nested chinese restaurant process and bayesian nonparametric inference of topic hierarchies. *Journal of the ACM*, 2010.

[23] G. Bouma. Normalized (pointwise) mutual information in collocation extraction. *GSCL*, 2009.

[24] G. R. Boynton and G. W. Richardson Jr. Agenda setting in the twenty-first century. *New Media & Society*, 18(9):1916–1934, 2016.

[25] M. Brosig and R. Verma. The war in ukraine, the global south and the evolving global order. *Global Policy*, 2024.

[26] D. Cer, M. Diab, E. Agirre, I. Lopez-Gazpio, and L. Specia. Semeval-2017 task 1: Semantic textual similarity multilingual and crosslingual focused evaluation. In *11th International Workshop on Semantic Evaluation*, 2017.

[27] D. Cer, Y. Yang, S.-y. Kong, N. Hua, N. Limtiaco, R. S. John, N. Constant, M. Guajardo-Cespedes, S. Yuan, C. Tar, et al. Universal sentence encoder for English. In *Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 2018.

[28] D. Chandrasekaran and V. Mago. Evolution of semantic similarity—a survey. *ACM Computing Surveys (CSUR)*, 54(2):1–37, 2021.

[29] X. Chen, A. Zeynali, C. Camargo, F. Flöck, D. Gaffney, P. Grabowicz, S. Hale, D. Jurgens, and M. Samory. Semeval-2022 task 8: Multilingual news article similarity. In *16th International Workshop on Semantic Evaluation*, 2022.

[30] J. Comber. Fda authorizes 'traditional' novavax covid vaccine, but critics question safety claims. https://web.archive.org/web/20220713202551/https://childrenshealthdefense.org/defender/fda-authorize-traditional-novavax-covid-vaccine-safety-claims/, 7 2022.

[31] J. Y. Cuan-Baltazar, M. J. Muñoz-Perez, C. Robledo-Vega, M. F. Pérez-Zepeda, and E. Soto-Vega. Misinformation of covid-19 on the internet: infodemiology study. *JMIR public health & surveillance*, 2020.

[32] R. Dahlke, D. Kumar, Z. Durumeric, and J. T. Hancock. Quantifying the systematic bias in the accessibility and inaccessibility of web scraping content from url-logged web-browsing digital trace data. *Social Science Computer Review*.

[33] Dave Van Zandt. Media bias/fact check. https://mediabiasfactcheck.com/, 2023.

[34] R. C. David Bauder and G. Mulvihill. Fox, dominion reach 787m settlement over election claims. https://web.archive.org/web/20230418103714/https://apnews.com/article/fox-news-dominion-lawsuit-trial-trump-2020-0ac71f75acfacc52ea80b3e747fb0afe, 4 2023.

[35] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti. Generalized louvain method for community detection in large networks. In *Intl. Conf. on intelligent systems design and applications*, 2011.

[36] P. Devine and K. Blincoe. Unsupervised extreme multi label classification of stack overflow posts. In *1st International Workshop on Natural Language-based Software Engineering*, 2022.

[37] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1*, 2019.

[38] O. Dinari and O. Freifeld. Revisiting dp-means: fast scalable algorithms via parallelism and delayed cluster creation. In *Uncertainty in Artificial Intelligence*, 2022.

[39] A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Yang, A. Fan, et al. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*, 2024.

[40] Z. Durumeric, D. Adrian, A. Mirian, J. Kasten, E. Bursztein, N. Lidzborski, K. Thomas, V. Eranti, M. Bailey, and J. A. Halderman. Neither snow nor rain nor mitm... an empirical analysis of email delivery security. In *ACM Internet Measurement Conference*, 2015.

[41] Z. Durumeric, D. Adrian, P. Stephens, E. Wustrow, and J. A. Halderman. Ten years of zmap. In *ACM Internet Measurement Conference*, 2024.

[42] Z. Durumeric, J. Kasten, M. Bailey, and J. A. Halderman. Analysis of the https certificate ecosystem. In *ACM Internet measurement conference*, 2013.

[43] Z. Durumeric, E. Wustrow, and J. A. Halderman. ZMap: fast internet-wide scanning and its security applications. In *22nd USENIX Security Symposium*, 2013.

[44] U. K. Ecker and L. C. Ang. Political attitudes and the processing of misinformation corrections. *Political Psychology*, 2019.

[45] L. Erbring, E. N. Goldenberg, and A. H. Miller. Front-page news and real-world cues: A new look at agenda-setting by the media. *American journal of political science*, pages 16–49, 1980.

[46] EU Disinfo Lab. What is the doppelganger operation? https://www.disinfo.eu/doppelganger-operation/.

[47] T. Gao, X. Yao, and D. Chen. SimCSE: Simple contrastive learning of sentence embeddings. In *EMNLP*, 2021.

[48] R. K. Garrett and R. M. Bond. Conservatives' susceptibility to political misperceptions. *Science Advances*, 2021.

[49] M. Gomez-Rodriguez, J. Leskovec, and A. Krause. Inferring networks of diffusion and influence. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 5(4):1–37, 2012.

[50] M. Gomez Rodriguez, J. Leskovec, and B. Schölkopf. Structure and dynamics of information pathways in online media. In *ACM international conference on Web search and data mining*, 2013.

[51] L. Grimminger and R. Klinger. Hate towards the political opponent: A twitter corpus study of the 2020 us elections on the basis of offensive speech and stance detection. In *Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 2021.

[52] M. Grootendorst. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*, 2022.

[53] F. Hamborg and K. Donnay. Newsmtsc: (multi-)target-dependent sentiment classification in news articles. In *European Chapter of the Association for Computational Linguistics*, 2021.

[54] C. Han, D. Kumar, and Z. Durumeric. On the infrastructure providers that support misinformation websites. In *Intl. AAAI Conf. on Web & Social Media*, 2022.

[55] H. Hanley and Z. Durumeric. Tata: Stance detection via topic-agnostic and topic-aware embeddings. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 11280–11294, 2023.

[56] H. W. Hanley and Z. Durumeric. Machine-made media: Monitoring the mobilization of machine-generated articles on misinformation and mainstream news websites. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18, pages 542–556, 2024.

[57] H. W. Hanley and Z. Durumeric. Partial mobilization: Tracking multilingual information flows amongst russian media outlets and telegram. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18, pages 528–541, 2024.

[58] H. W. Hanley, D. Kumar, and Z. Durumeric. No calm in the storm: investigating QAnon website relationships. In *Intl. AAAI Conf. on Web & Social Media*, 2022.

[59] H. W. Hanley, D. Kumar, and Z. Durumeric. "A Special Operation": A quantitative approach to dissecting and comparing different media ecosystems' coverage of the Russo-Ukrainian war. In *International AAAI Conference on Web and Social Media*, 2023.

[60] H. W. Hanley, D. Kumar, and Z. Durumeric. A golden age: Conspiracy theories' relationship with misinformation outlets, news media, and the wider internet. *CSCW*, 2023.

[61] H. W. Hanley, D. Kumar, and Z. Durumeric. Happenstance: Utilizing semantic search to track russian state media narratives about the russo-ukrainian war on reddit. In *International AAAI conference on web and social media*, 2023.

[62] H. W. Hanley, D. Kumar, and Z. Durumeric. Specious sites: Tracking the spread and sway of spurious news stories at scale. In *IEEE Symposium on Security and Privacy*, 2024.

[63] M. Hardalov, A. Arora, P. Nakov, and I. Augenstein. Few-shot cross-lingual stance detection with sentiment-based pre-training. In *AAAI Conference on Artificial Intelligence*, 2022.

[64] N. Hassan, G. Zhang, F. Arslan, J. Caraballo, D. Jimenez, S. Gawsane, S. Hasan, M. Joseph, A. Kulkarni, A. K. Nayak, et al. Claimbuster: The first-ever end-to-end fact-checking system. *VLDB*, 2017.

[65] P. He, J. Gao, and W. Chen. Debertav3: Improving deberta using electra-style pre-training with gradient-disentangled embedding sharing. In *11th Intl. Conf. on Learning Representations*, 2022.

[66] N. Heninger, Z. Durumeric, E. Wustrow, and J. A. Halderman. Mining your Ps and Qs: Detection of widespread weak keys in network devices. In *21st USENIX Security Symposium*, 2012.

[67] J. Hoewe, K. C. Brownell, and E. C. Wiemer. The role and impact of fox news. In *The forum*, 2020.

[68] A. Hounsel, J. Holland, B. Kaiser, K. Borgolte, N. Feamster, and J. Mayer. Identifying disinformation websites using infrastructure features. In *USENIX Workshop on Free and Open Communications on the Internet*, 2020.

[69] E. J. Hu, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen, et al. Lora: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2021.

[70] H. Jelodar, Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li, and L. Zhao. Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 2019.

[71] Z. Jin, J. Cao, H. Guo, Y. Zhang, and J. Luo. Multimodal fusion with recurrent neural networks for rumor detection on microblogs. In *25th ACM international conference on Multimedia*, 2017.

[72] J. Johnson, M. Douze, and H. Jégou. Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, 7(3), 2019.

[73] J. L. Juul and J. Ugander. Comparing information diffusion mechanisms by matching on cascade size. *Proceedings of the National Academy of Sciences*, 118(46), 2021.

[74] B. Kaiser, J. Wei, E. Lucherini, K. Lee, J. N. Matias, and J. Mayer. Adapting security warnings to counter online disinformation. In *30th USENIX Security Symposium*, 2021.

[75] J. Kerr, C. Panagopoulos, and S. Van Der Linden. Political polarization on covid-19 pandemic response in the united states. *Personality and individual differences*, 179:110892, 2021.

[76] J. M. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan, and A. S. Tomkins. The web as a graph: Measurements, models, and methods. In *Computing and Combinatorics Conference*, 1999.

[77] P. Kreko. Political tribalism, polarization, and the motivated rejection of science. In *The Tribal Mind and the Psychology of Collectivism*. 2024.

[78] D. Küçük and F. Can. Stance detection: A survey. *ACM Computing Surveys (CSUR)*, 2020.

[79] B. Kulis and M. I. Jordan. Revisiting k-means: new algorithms via bayesian nonparametrics. In *International Conference on Machine Learning*, 2012.

[80] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang. Aspects of rumor spreading on a microblog network. In *Social Informatics: 5th International Conference*, 2013.

[81] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang. Prominent features of rumor propagation in online social media. In *IEEE international conference on data mining*, 2013.

[82] M. Lai, V. Patti, G. Ruffo, and P. Rosso. Stance evolution and twitter interactions in an italian political debate. In *Conference on Applications of Natural Language to Information Systems*, 2018.

[83] G. Leban, B. Fortuna, J. Brank, and M. Grobelnik. Event registry: learning about world events from news. In *23rd International Conference on World Wide Web*, 2014.

[84] D. Leonhardt. Revisiting the gaza hospital explosion. https://www.nytimes.com/2023/11/03/briefing/gaza-hospital-explosion.html, 11 2023.

[85] J. Leskovec, L. Backstrom, and J. Kleinberg. Meme-tracking and the dynamics of the news cycle. In *15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009.

[86] B. Lester, R. Al-Rfou, and N. Constant. The power of scale for parameter-efficient prompt tuning. In *Conference on Empirical Methods in Natural Language Processing*, 2021.

[87] B. Liang, Q. Zhu, X. Li, M. Yang, L. Gui, Y. He, and R. Xu. Jointcl: A joint contrastive learning framework for zero-shot stance detection. In *60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2022.

[88] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K.-F. Wong, and M. Cha. Detecting rumors from microblogs with recurrent neural networks. 2016.

[89] G. Madraki, I. Grasso, J. M. Otala, Y. Liu, and J. Matthews. Characterizing and comparing covid-19 misinformation across languages, countries and platforms. In *The web conference*, 2021.

[90] A. Martin. Russians impersonate washington post and fox news with anti-ukraine stories. https://therecord.media/russians-fake-news-anti-ukraine.

[91] M. McCombs and D. Shaw. The agenda-setting function of the press. *Oxford University Press*, 2005.

[92] D. McCoy, A. Pitsillidis, J. Grant, N. Weaver, C. Kreibich, B. Krebs, G. Voelker, S. Savage, and K. Levchenko. PharmaLeaks: Understanding the business of online pharmaceutical affiliate programs. In *21st USENIX Security Symposium*, 2012.

[93] L. McInnes, J. Healy, S. Astels, et al. hdbscan: Hierarchical density based clustering. *J. Open Source Softw.*, 2(11):205, 2017.

[94] S. B. Medicine. Vaccines. https://sciencebasedmedicine.org/category/vaccines/, 2024.

[95] S. Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. M. Voelker, and S. Savage. A fistful of bitcoins: characterizing payments among men with no names. In *ACM Internet measurement conference*, 2013.

[96] Y. Meng, Y. Zhang, J. Huang, Y. Zhang, and J. Han. Topic discovery via latent space clustering of pretrained language model representations. In *ACM Web Conference*, 2022.

[97] J. Mercola. Fda anxious for pfizer to rush covid shots for babies and toddlers. but why? https://web.archive.org/web/20220208225347/https://childrenshealthdefense.org/defender/fda-pfizer-rush-covid-shots-babies-toddlers/, 2 2022.

[98] S. Miranda, A. Znotiņš, S. B. Cohen, and G. Barzdins. Multilingual clustering of streaming news. In *Conference on Empirical Methods in Natural Language Processing*, 2018.

[99] S. Mohammad, S. Kiritchenko, P. Sobhani, X. Zhu, and C. Cherry. Semeval-2016 task 6: Detecting stance in tweets. In *International workshop on semantic evaluation*, 2016.

[100] E. Momeni, S. Karunasekera, P. Goyal, and K. Lerman. Modeling evolution of topics in large-scale temporal text corpora. In *Intl. AAAI Conf. on Web & Social Media*, 2018.

[101] R. C. Moore, R. Dahlke, and J. T. Hancock. Exposure to untrustworthy websites in the 2020 us election. *Nature Human Behaviour*, 2023.

[102] B. Mueller and S. G. Stolberg. Fauci grilled by lawmakers on masks, vaccine mandates and lab leak theory. https://web.archive.org/save/https://www.nytimes.com/2024/06/03/science/fauci-hearing-covid-origins.html, 6 2024.

[103] P. Nakov and G. Da San Martino. Fake news, disinformation, propaganda, and media bias. In *ACM International Conference on Information & Knowledge Management*, 2021.

[104] N. Nakshatri, S. Liu, S. Chen, D. Roth, D. Goldwasser, and D. Hopkins. Using llm for improving key event discovery: Temporal-guided news stream clustering with event summaries. In *Empirical Methods in Natural Language Processing*, 2023.

[105] D. Ng. Left-wing guardian triggered by my son hunter: Stop trying to make hunter biden conspiracy theories happen. https://web.archive.org/web/20220903140730/https://www.breitbart.com/entertainment/2022/09/03/leftwing-guardian-triggered-by-my-son-hunter-stop-trying-to-make-hunter-biden-conspiracy-theories-happen/, 9 2022.

[106] U. D. of State Global Engagement Center. Gec special report: Russia's pillars of disinformation and propaganda - united states department of state. https://www.state.gov/russias-pillars-of-disinformation-and-propaganda-report/, 8 2020.

[107] A. Papasavva, J. Blackburn, G. Stringhini, S. Zannettou, and E. D. Cristofaro. "is it a qoincidence?": An exploratory study of qanon on voat. In *Web Conference*, 2021.

[108] P. Paudel, J. Blackburn, E. De Cristofaro, S. Zannettou, and G. Stringhini. Lambretta: learning to rank for twitter soft moderation. In *IEEE Symposium on Security and Privacy*, 2023.

[109] F. Pierri, L. Luceri, N. Jindal, and E. Ferrara. Propaganda and misinformation on facebook and twitter during the russian invasion of ukraine. In *Proceedings of the 15th ACM web science conference 2023*, pages 65–74, 2023.

[110] A. Piktus, F. Petroni, V. Karpukhin, D. Okhonko, S. Broscheit, G. Izacard, P. Lewis, B. Oğuz, E. Grave, W.-t. Yih, et al. The web is your oyster–knowledge-intensive nlp against a very large web corpus. *arXiv preprint arXiv:2112.09924*, 2021.

[111] T. A. Press. Not real news: A look at what didn't happen this week. https://web.archive.org/web/20231124154018/https://apnews.com/article/fact-check-misinformation-f3c1d54f2d059de0532360d638335e99, 11 2023.

[112] S. Prochaska, K. Duskin, Z. Kharazian, C. Minow, S. Blucker, S. Venuto, J. D. West, and K. Starbird. Mobilizing manufactured reality: How participatory disinformation shaped deep stories to catalyze action during the 2020 us presidential election. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW1):1–39, 2023.

[113] M. Rajdev and K. Lee. Fake and spam messages: Detecting misinformation during natural disasters on social media. In *Intl. Conf. on Web Intelligence and Intelligent Agent Technology*, 2015.

[114] N. Reimers and I. Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *EMNLP*, 2019.

[115] B. Ruhnau. Eigenvector-centrality—a node-centrality? *Social networks*, 22(4):357–365, 2000.

[116] K. Ruth, A. Fass, J. Azose, M. Pearson, E. Thomas, C. Sadowski, and Z. Durumeric. A world wide view of browsing the world wide web. In *ACM Internet Measurement Conference*, 2022.

[117] K. Ruth, D. Kumar, B. Wang, L. Valenta, and Z. Durumeric. Toppling top lists: Evaluating the accuracy of popular website lists. In *22nd ACM Internet Measurement Conference*, 2022.

[118] M. H. Saeed, S. Ali, J. Blackburn, E. De Cristofaro, S. Zannettou, and G. Stringhini. Trollmagnifier: Detecting state-sponsored troll accounts on reddit. In *IEEE Symposium on Security & Privacy*, 2022.

[119] K. K. Saravanakumar, M. Ballesteros, M. K. Chandrasekaran, and K. Mckeown. Event-driven news stream clustering using entity-aware contextual embeddings. In *ACL*, 2021.

[120] M. S. Schäfer and J. Painter. Climate journalism in a changing media ecosystem: Assessing the production of climate change-related news around the world. *Wiley Interdisciplinary Reviews: Climate Change*, 12(1):e675, 2021.

[121] G. K. Shahi. Warclaim: A dataset for fake news on 2023 Israel–Hamas war. In *ACM Web Science Conference*, 2024.

[122] L. Silver and A. Connaughton. Partisanship colors views of covid-19 handling across advanced economies. https://www.pewresearch.org/global/2022/08/11/partisanship-colors-views-of-covid-19-handling-across-advanced-economies/, 8 2022.

[123] V. Singrodia, A. Mitra, and S. Paul. A review on web scrapping and its applications. In *2019 international conference on computer communication and informatics (ICCCI)*, pages 1–6. IEEE, 2019.

[124] J. Smith, H. Saint-Amand, M. Plamadă, P. Koehn, C. Callison-Burch, and A. Lopez. Dirt cheap web-scale parallel text from the common crawl. In *ACL*, 2013.

[125] V. Smith. *Go Web Scraping Quick Start Guide: Implement the power of Go to scrape and crawl data from the web*. 2019.

[126] K. Song, X. Tan, T. Qin, J. Lu, and T.-Y. Liu. Mpnet: Masked and permuted pre-training for language understanding. *Adv. in Neural Information Processing Systems*, 2020.

[127] K. Starbird, A. Arif, T. Wilson, K. Van Koevering, K. Yefimova, and D. Scarnecchia. Ecosystem or echo-system? exploring content sharing across alternative media domains. In *International AAAI Conference on Web and Social Media*, 2018.

[128] R. Sundara Raman, L.-H. Merino, K. Bock, M. Fayed, D. Levin, N. Sullivan, and L. Valenta. Global, passive detection of connection tampering. In *ACM SIGCOMM*, 2023.

[129] R. Sundara Raman, P. Shenoy, K. Kohls, and R. Ensafi. Censored planet: An internet-wide, longitudinal censorship observatory. In *ACM Conference on computer and communications security*, 2020.

[130] K. Thomas, E. Bursztein, C. Grier, G. Ho, N. Jagpal, A. Kapravelos, D. McCoy, A. Nappa, V. Paxson, P. Pearce, et al. Ad injection at scale: Assessing deceptive advertisement modifications. In *IEEE Symposium on Security and Privacy*, 2015.

[131] F. Timotija. Ukraine defense minister presses us to allow use of long-range weapons on russia. https://web.archive.org/web/20240831192927/https://thehill.com/policy/international/4857208-ukraine-defense-minister-rustem-umerov-long-range-missiles-restriction/, 8 2024.

[132] R. Today. Finland rules on fate of seized russian artwork. https://web.archive.org/web/20220524233445/https://www.rt.com/russia/553529-finland-russia-art-decision/, 4 2022.

[133] P. D. Turney. Mining the web for synonyms: Pmi-ir versus lsa on toefl. In *European conference on machine learning*, 2001.

[134] U.S. Department of Justice. Justice department disrupts covert russian government-sponsored foreign malign influence operation targeting audiences in the united states and elsewhere. https://www.justice.gov/opa/pr/justice-department-disrupts-covert-russian-government-sponsored-foreign-malign-influence.

[135] Y. Vasiliev. *Natural language processing with Python and spaCy: A practical introduction*. No Starch Press, 2020.

[136] S. Vosoughi, D. Roy, and S. Aral. The spread of true and false news online. *Science*, 2018.

[137] I. Waller and A. Anderson. Quantifying social organization and political polarization in online platforms. *Nature*, 2021.

[138] L. Wang, N. Yang, X. Huang, B. Jiao, L. Yang, D. Jiang, R. Majumder, and F. Wei. Text embeddings by weakly-supervised contrastive pre-training. *arXiv preprint arXiv:2212.03533*, 2022.

[139] G. Weld, M. Glenski, and T. Althoff. Political bias and factualness in news sharing across more than 100,000 online communities. In *Intl. AAAI Conf. on Web & Social Media*, 2021.

[140] C. Wilkie. Dinesh dśouza election fraud film, book ́2000 mules pulled after defamation suit. https://web.archive.org/web/20240601000808/https://www.cnbc.com/2024/05/31/dinesh-dsouza-election-film-2000-mules-pulled.html, 5 2024.

[141] L. Wu, F. Morstatter, K. M. Carley, and H. Liu. Misinformation in social media: definition, manipulation, and detection. *ACM SIGKDD Explorations Newsletter*, 21(2):80–90, 2019.

[142] J. Yin, D. Chao, Z. Liu, W. Zhang, X. Yu, and J. Wang. Model-based clustering of short text streams. In *24th ACM SIGKDD international conference on knowledge discovery & data mining*, 2018.

[143] D. V. Zandt. Media-Bias/Fact-Check. https://mediabiasfactcheck.com/yahoo-news/, 2022.

[144] D. V. Zandt. Media-Bias/Fact-Check. https://mediabiasfactcheck.com/north-alaska-news/, 2022.

[145] E. Zeng, T. Kohno, and F. Roesner. Bad news: Clickbait and deceptive ads on news and misinformation websites. In *Workshop on Technology and Consumer Protection*, 2020.

[146] Y. Zhang, F. Guo, J. Shen, and J. Han. Unsupervised key event detection from massive text corpora. In *ACM SIGKDD conference on knowledge discovery and data mining*, 2022.

[147] D. Zhou, H. Xu, and Y. He. An unsupervised Bayesian modelling approach for storyline detection on news articles. In *Conference on Empirical Methods in Natural Language Processing*, 2015.

[148] X. Zhou, A. Sharma, A. X. Zhang, and T. Althoff. Correcting misinformation on social media with a large language model. *arXiv preprint arXiv:2403.11169*, 2024.

A Article Preprocessing

After collecting each page's HMTL, we then parse the content to extract the news article text and publication date using the Python libraries `newspaper3k` and `htmldate`. We subsequently remove any leftover boilerplate language (*i.e.*, navigation links, headers, and footers) from the text using the `justext` Python library and remove any non-English articles based on labels provided by the Python `langdetect` library.

To prepare our news article data for embedding, we first remove any URLs, emojis, and HTML tags from the text. Then, in line with prior work, we subsequently divide these our articles in constituent *passages* with at most 100 words [57, 61, 110]. Specifically, after first separating articles into different paragraphs by splitting their text on (`\n`) or tab (`\t`) characters and breaking apart each paragraph into its constituent sentences [21], we add sentences from a single paragraph to create a *passage* until its length is at most 100 words. We embed the constituent *passages*, rather than full articles given the context window size limitations of the large language that we use in this work. Furthermore, as argued by Hanley et al. [62] and shown by Pikbus et al. [110], given that articles often address multiple ideas, embedding *passages* allows us to track the often single idea present within the passage.[6] Our dataset consists of 428,051,085 passages.

B PEFT through LoRA

We utilize *Parameter Efficient Fine-Tuning*/PEFT [86] through *Low-Rank Adaption*/LoRA [69] to fine-tune and adapt pre-trained models to our datasets or to better their performance. LoRA, specifically, after freezing the weights of the original pre-trained model learns pairs of low-rank-decomposition matrices, reducing the amount of parameters that need to be learned. LoRA has been shown to often outperform other types of adaptions including full-tuning [69]. Once learned, these matrices are merged with the original frozen weights. LoRA requires the specification of the rank of the matrices learned and an α value that scales the learned parameters. Within this work, we learn LoRA matrices for the attention and the dense/linear layers of our models and utilize the commonly used defaults of rank=8 and α=16 [69].

C Training with Unsupervised Contrastive Loss

To adapt our embedding models to our news dataset, we utilize unsupervised contrastive learning [47]. For training, this is such that we embed each example $x_i = (passage_i) \in D_{News}$

---

[6]For example, a New York Times article addressing a novel virus spreading in China may also address the Chinese government's previous approaches to dealing with the COVID-19 pandemic.https://web.archive.org/web/20231125001740/https://www.nytimes.com/2023/11/23/world/asia/who-china-respiratory-children.html

(where $passage_i$ is the passage text) twice (with dropout both times) with a given model by inputting $[CLS]text_i[SEP]$ and averaging the contextual word vectors of the resulting output as a hidden vector $\mathbf{h}_i$ and $\tilde{\mathbf{h}}_i$ for $passage_i$ as its representations. Then, given a set of hidden vectors $\{\mathbf{h}_i\}_{i=0}^{N_b}$ and $\{\tilde{\mathbf{h}}_j\}_{j=0}^{N_b}$ (different dropout), where N_b is the size of the batch, we perform a contrastive learning step for each batch. This is such that for each Batch $\mathcal{B}$, for an *anchor* hidden embedding $\mathbf{h_i}$ within the batch, the set of hidden vectors $\mathbf{h_i}, \tilde{\mathbf{h}}_j \in \mathcal{B}$, vectors where $i = j$ are positive pairs. Other pairs where $i \neq j$ are considered negative pairs. Within each batch $\mathcal{B}$, the contrastive loss is computed across all positive pairs in the batch:

$$L_{sim} = -\frac{1}{N_b} \sum_{\mathbf{h}_i \in \mathcal{B}} l^c(\mathbf{h}_i)$$

$$l^c(\mathbf{h}_i) = \log \frac{\sum_{j \in \mathcal{B}} \mathbb{1}_{[i=j]} \exp(\frac{\mathbf{h}_i^\top \tilde{\mathbf{h}}_j}{\tau ||\mathbf{h}_i|| ||\tilde{\mathbf{h}}_j||})}{\sum_{j \in \mathcal{B}} \exp(\frac{\mathbf{h}_i^\top \tilde{\mathbf{h}}_j}{\tau ||\mathbf{h}_i|| ||\tilde{\mathbf{h}}_j||})}$$

where, as in prior work [87], we utilize a temperature $\tau = 0.07$. When performing fine-tuning, we utilize default hyperparameters (learning rate 3×10^{-5}, batch size=128, and 1M examples) specified in Gao et al. [47].

D Pointwise Mutual Information

The PMI of a word $word_i$ in a cluster C_j is calculated:

$$PMI(word_i, C_j) = \log_2 \frac{P(word_i, C_j)}{P(word_i)P(c_i)}$$

where P is the probability of occurrence and a scaling parameter $\alpha = 1$ is added to the counts of each word per cluster. This scaling parameter α prevents low-frequency words in each cluster from having the highest PMI value [133].
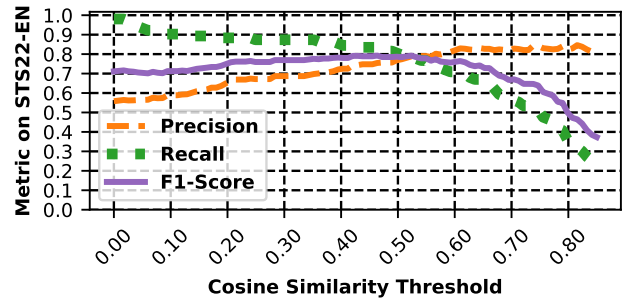
E Evaluation on SemEval22

Figure 12: Evaluation of our model's precision, recall, and F_1 scores on the English portion of the SemEval22 test dataset [29] (using 3.0 as the cut-off for the two articles being about the same event [57]).

F Passage Pairs at Various Thresholds

0.35 Similarity

PASSAGE 1: Russias war in Ukraine created a full-on energy crisis as Moscow reduced or cut off natural gas flows to European countries that rely on the fuel to power industry, generate electricity and heat and cool homes. Shrinking supplies, higher demand and fears of a complete Russian cutoff have driven natural gas prices to record highs, further fueling inflation that has squeezed peoples ability to spend and raised the risk of a recession in Europe and the U.K.

PASSAGE 2: Fingrid had been expecting that Olkiluoto 3 alone would more than compensate for the loss of Russian power imports this winter, he added. Imports of Russian power stopped in May after Russian utility Inter RAO said it had not been paid for the power it sold via pan European exchange Nord Pool since May 6. Without Olkiluoto 3 the situation is quite tight because that would have been more than 10 of the peak demand alone, Ruusunen said.

0.40 Similarity

PASSAGE 1: HONG KONG, Jan 13 (Reuters Breakingviews) - WM Motor, a Chinese electric-car star that has fallen on hard times, is reversing into an ambitious Plan C. After two earlier attempts at an initial public offering floundered, the group founded by Zhejiang Geely veteran Freeman Shen will at last go public by selling its auto business to Hong Kong-listed Apollo Future Mobility (0860.HK) for $2 billion. They make an odd couple. Although both are car companies, their cars have little in common.

PASSAGE 2: Global and Chinese automakers plan to unveil more than a dozen new electric SUVs, sedans and muscle cars this week at the Shanghai auto show, their first full-scale sales event in four years in a market that has become a workshop for developing electrics, self-driving cars and other technology.

0.45 Similarity

PASSAGE 1: Diseases like Ebola and other hemorrhagic fevers were responsible for 70% of those outbreaks, in addition to illnesses like monkeypox, dengue, anthrax and plague. We must act now to contain zoonotic diseases before they can cause widespread infections and stop Africa from becoming a hotspot for emerging infectious diseases, WHOs Africa director, Dr. Matshidiso Moeti said in a statement.

PASSAGE 2: Equatorial Guinea has confirmed 13 cases of Marburg disease since the beginning of the epidemic, its health officials said on Wednesday after the head of the World Health Organization (WHO) urged the Central African country's government to report new cases officially. Marburg virus disease is a viral haemorrhagic fever that can have a fatality rate of up to 88%, according to the WHO.

0.50 Similarity

PASSAGE 1: The fatal accident occurred around 4 a.m., and it took several hours to clear the freeway. The firetruck had to be towed away. The Model S was among the nearly 363,000 vehicles Tesla recalled on Thursday because of potential flaws in its Full Self-Driving system. While the recall is aimed at correcting possible problems at intersections and with speed limits, it comes amid a broader investigation by U.S. safety regulators into Teslas automated driving systems.'

PASSAGE 2: The National Highway Traffic Safety Administration (NHTSA) is opening a special investigation into the crash of a 2021 Tesla Model Y vehicle that killed a motorcyclist in California, it said on Monday. Since 2016, NHTSA has opened 37 special investigations of crashes involving Tesla vehicles and where advanced driver assistance systems such as Autopilot were suspected of being used. A total of 18 crash deaths were reported in those Tesla-related investigations, including the most recent fatal California crash.

Figure 13: Example of passage pairs at different levels of cosine similarity.

G NETINF Algorithm

NETINF is a greedy algorithm that iteratively computes the marginal gain (*i.e.*, the explanatory power of adding the edge given the set of cascades) of adding a particular weighted edge between two entities/nodes and by only considering the most probable transmission tree (*i.e.*, the steps taken for a given narrative to reach a particular website), NETINF efficiently infers the relationships within the underlying network. The NETINF algorithm returns both the marginal gain as well the rate of information flow between two nodes/websites for each edge. Thus given a set of cascades C, NETINF aims to find the the graph $\hat{G}$ that solves the optimization problem [49]:

$$\hat{G} = \arg\max_{|G| \leq k} P(C|G)$$

where

$$P(C|G) = \prod_{c \in C} P(c|G)$$

$$P(c|G) = \sum_{T \in T(G)} P(c|T)P(T|G) \propto \sum_{T \in T(G)} \prod_{(i,j) \in T} P_c(i,j)$$

and where c is a cascade, $T(G)$ is the set of all directed spanning trees on G, andd $P_c(i,j)$ (or that the node i influences node j in a cascade) is proportional to the time difference between when the two nodes are infected (*i.e.*, when the two websites post a given story about a particular narrative), given an exponential waiting time for infection.

To optimize this formulation, NETINF only considers the most likely propagation tree T for a given cascade c

$$P(C|G) = \prod_{c \in C} \max_{T \in TG} P(c|T) = \prod_{c \in C} \max_{T \in TG} \prod_{(i,j) \in T} P_c(i,j)$$

The improvement of the log-likelihood of a given cascade c for a graph G over the empty graph K is then:

$$F_c(G) = \max_{T \in T(G)} \log P(c|T) - \max_{T \in T(K)} \log P(c|T)$$

and the NETINF algorithm optimizes the following objective function by iteratively and greedily adding edges with the highest marginal gain to the objective function:

$$F_C(G) = \sum_{c \in C} F_c(G)$$

H Optimized DP-Means

DP-Means [79] is a non-parametric extension of the K-means algorithm that does not require the specification of the number of clusters *a priori*. Within DP-Means, when a given datapoint is a chosen parameter λ away from the closest cluster, a new cluster is formed. Dinari et al. [38] parallelize this algorithm by *delaying cluster creation* until the end of the assignment step. Namely, instead of creating a new cluster each time a new datapoint is discovered, the algorithm instead determines which datapoint is furthest from the current set of clusters and then creates a new cluster with that datapoint. By delaying cluster creation, the DP-means algorithm can be trivially parallelized. Furthermore, by delaying cluster creation, this version of DP-Means avoids over-clustering the data (*i.e., only the most disparate datapoints create new clusters*) [38].

I Evaluation of Clusters

Narr.	Keywords	Passages Checked	Prec.
1	laissez-faire, progressivism, liberalism, laissez, corpus	193	96.89%
2	quake, earthquake, aftershock, turkey, rubble	500	100.00%
3	sinema, manchin, filibuster, kyrsten, senate	500	100.00%
4	williamson, marianne, self-help, williamsons, sander	500	97.40%
5	dysphoria, puberty, blocker, crosssex, hormone	500	100.00%
6	sudan, anand, evacuation, sudanese, khartoum	500	100.00%
7	rioter, slogan, bearing, capitol, drum	500	99.20%
8	teixeira, dighton, guardsman, teixeiras, massachusetts	500	99.40%
9	fdny, firefighter, firehouse, klein, kavanagh	500	97.00%
10	bragg, alvin, rouser, nypd, rabble	500	95.60%
11	taliban, afghan, afghanistan, hunger, malnutritio	500	100.00%
12	eyesight, blindness, blind, eye, sight	500	99.40%
13	maralago, classified, ballroom, fundraiser, document	500	100.0%
14	carolina, vetoproof, map, raleigh, cooper	500	99.80%
15	tarantino, quentin, pulp, cinema, filmmaker	500	99.20%
16	seoul,korea, posco, compensate, keb	500	100.00%
17	miscarriage, pregnant, pregnancy, csection, motherhood	500	100.00%
18	faucis, niaid, anthony, gain-of-function, allergy	500	100.00%
19	crump, arbery, breonna, ahmaud, trayvon	500	99.08%
20	portuguese, slave, plantation, colony, dutch	500	100.00%
21	cadet, guard, harassment, assault, adjutant	500	100.00%
22	spam, bot, musk, twitter, elon	500	98.80%
23	ufo, roswell, sighting, saucer, alien	500	100.00%
24	cpu, intel, x86, processor, amd	500	96.40%
25	chappelle, comedian, isaiah, onstage, attacker	500	100.00%
26	burisma, pozharskyi, vadym, hunter, zlochevsky	500	100.00%
27	ubridgerton, penelope, featherington, daphne, coughland	500	100.00%
28	naloxone, narcan, over-the-counter, emergent, nasal	500	100.00%
29	schmitt, greitens, hartzler, missouri, trudy	100	99.20%
30	currency, dollar, yuan, reserve, de-dollarization	500	99.80%
		Prec.	**99.26%**

J Additional Stances

Pro-China Stances	Coeff.	Std.
Pro self-reliance	0.286	0.092
Pro communist	0.180	0.093
Pro africa	0.126	0.103
Anti-China Stances		
Against communist	-0.288	0.087
Against covid	-0.228	0.091
Against russia	-0.215	0.072

Table 10: The stances associated with China according to the Bayesian model.

Pro-Iran Stances	Coeff.	Std.
Against yemeni	0.055	0.061
Pro islamic	0.052	0.042
Pro armenia	0.049	0.036
Anti-Iran Stances		
Against tehran	-0.073	0.060
Against islamic	-0.058	0.043
Against communist	-0.047	0.038

Table 11: The stances associated with Iran according to the Bayesian model.

Pro-America Stances	Coeff.	Std.
Pro harvard	0.111	0.064
Pro allstar	0.085	0.055
Pro mvp	0.070	0.045
Anti-America Stances		
Against graham	-0.143	0.070
Against cia	-0.126	0.069
Against canada	-0.126	0.060

Table 12: The stances associated with America according to the Bayesian model.

(a) Dist. of Pro America

(b) Dist. of Against America

(c) Dist. of Pro China

(d) Dist. of Against China

(e) Dist. of Pro Iran
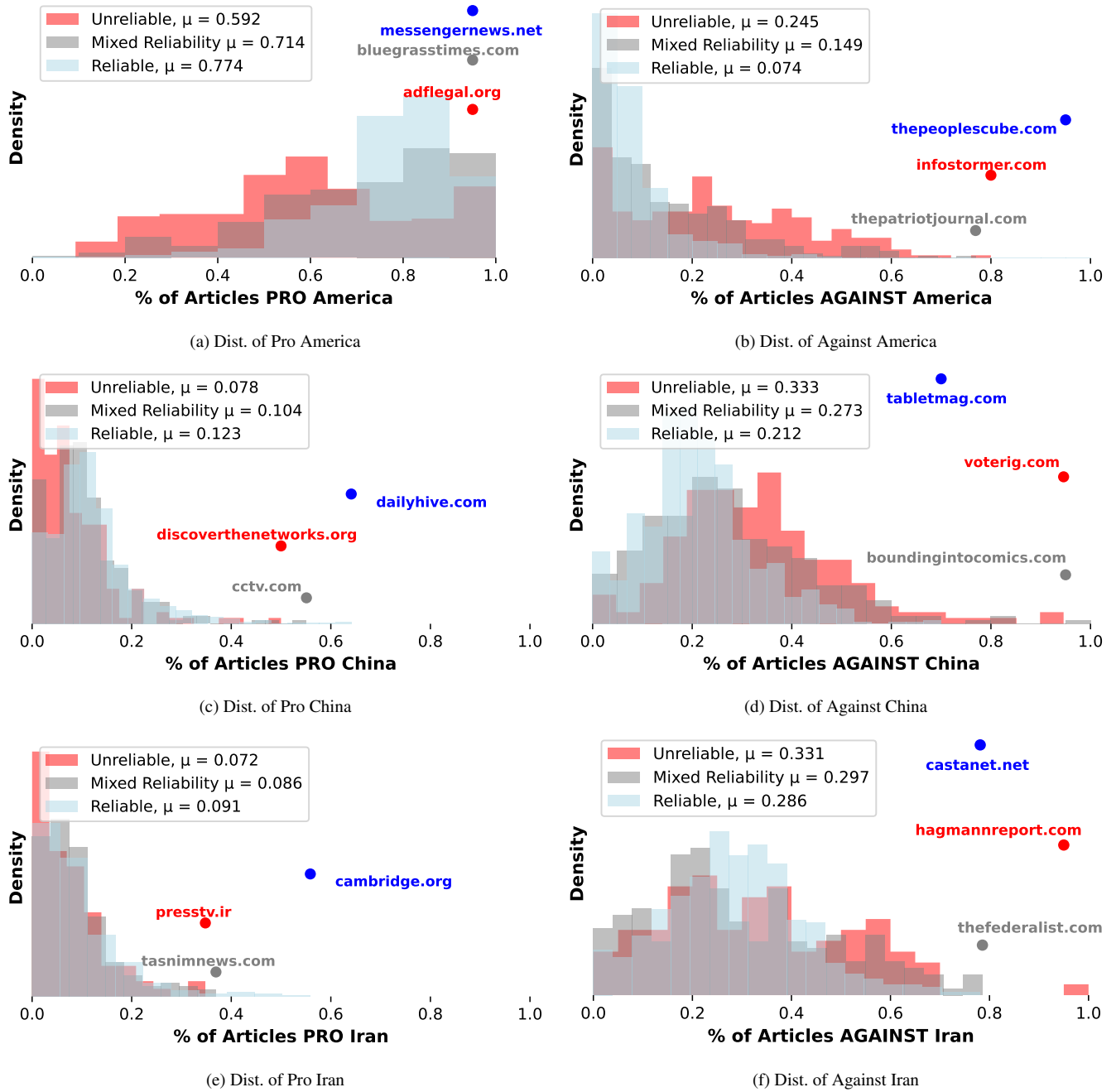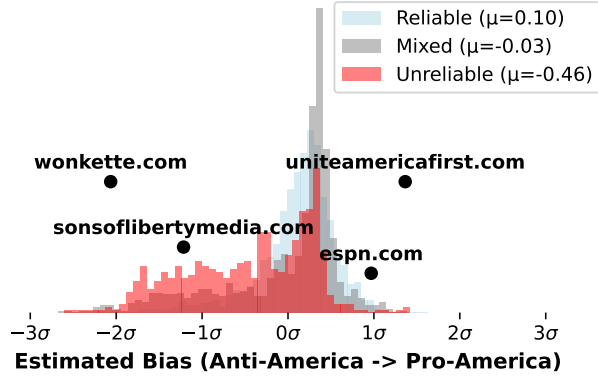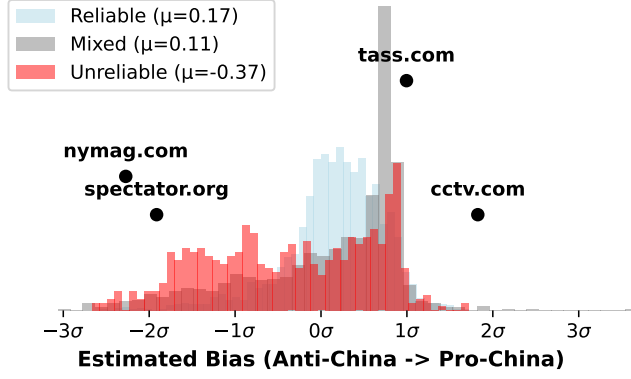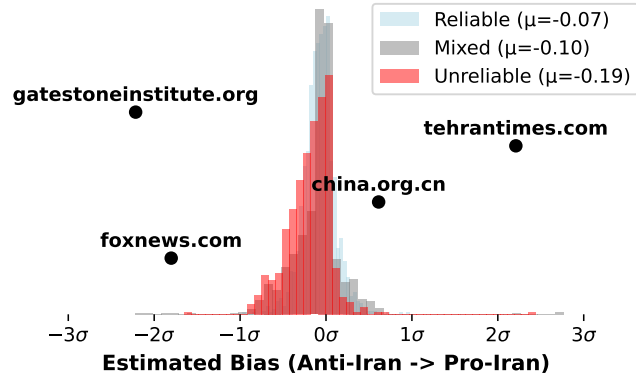
(f) Dist. of Against Iran

Figure 14: Distribution of stances to various entities.

(a) America Latent

(b) China Latent

(c) Iran Latent

Figure 15: Latents for Biases to entities.

Pro America

Reliable	Mixed	Unreliable
messengernews.net (100.0%)	bluegrasstimes.com (100.0%)	adflegal.org (100.0%)
nationalgeographic.com (100.0%)	tribunnews.com (100.0%)	worldcouncilforhealth.org (100.0%)
fee.org (100.0%)	yaf.org (100.0%)	nvic.org (100.0%)

Against America

Reliable	Mixed	Unreliable
thepeoplescube.com (100.0%)	thepatriotjournal.com (76.9%)	infostormer.com (80.0%)
cjr.org (50.0%)	greanvillepost.com (68.9%)	paulcraigroberts.org (71.4%)
techdirt.com (46.2%)	newscorpse.com (58.8%)	gellerreport.com (64.9%)

Pro China

Reliable	Mixed	Unreliable
dailyhive.com (64.1%)	cctv.com (55.1%)	discoverthenetworks.org (50.0%)
emerging-europe.com (64.0%)	egypttoday.com (47.1%)	occupydemocrats.com (41.7%)
tdn.com (58.3%)	thecountersignal.com (45.5%)	medicalkidnap.com (37.8%)

Against China

Reliable	Mixed	Unreliable
tabletmag.com (70.0%)	boundingintocomics.com (100.0%)	voterig.com (94.6%)
outsidethebeltway.com (69.8%)	faithwire.com (81.8%)	govtslaves.com (93.3%)
icij.org (63.6%)	fff.org (81.8%)	conservativeplaylist.com (82.4%)

Pro Iran

Reliable	Mixed	Unreliable
cambridge.org (55.9%)	tasnimnews.com (36.9%)	presstv.ir (34.8%)
lamag.com (50.0%)	almanar.com.lb (34.5%)	theconservativetreehouse.com (33.3%)
msmagazine.com (46.4%)	mehrnews.com (32.6%)	healthimpactnews.com (25.0%)

Against Iran

Reliable	Mixed	Unreliable
castanet.net (78.1%)	thefederalist.com (78.6%)	hagmannreport.com (100.0%)
dailysignal.com (72.2%)	smirkingchimp.com (78.6%)	clarionproject.org (69.2%)
wearethemighty.com (70.0%)	patriotnewsalerts.com (76.9%)	thewashingtonstandard.com (68.8%)

Table 13: The set of news websites with the highest percentage of *Pro*-articles about various entities/topics.
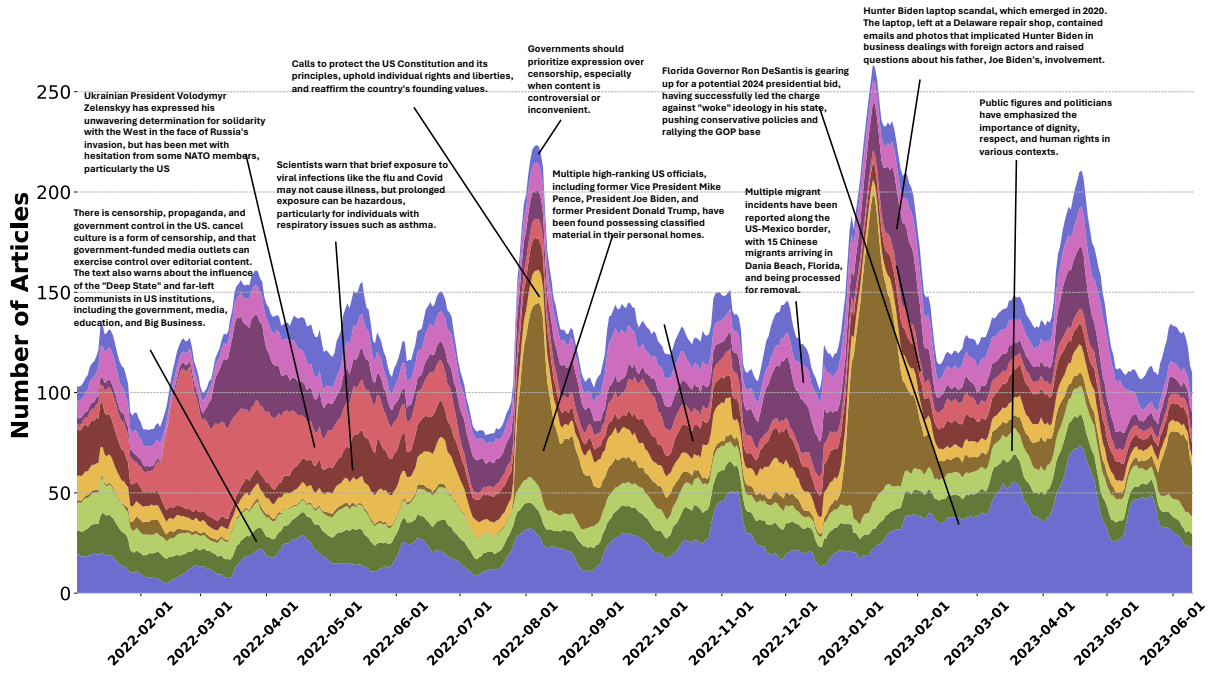
(a) Story Spread on Unreliable News Websites
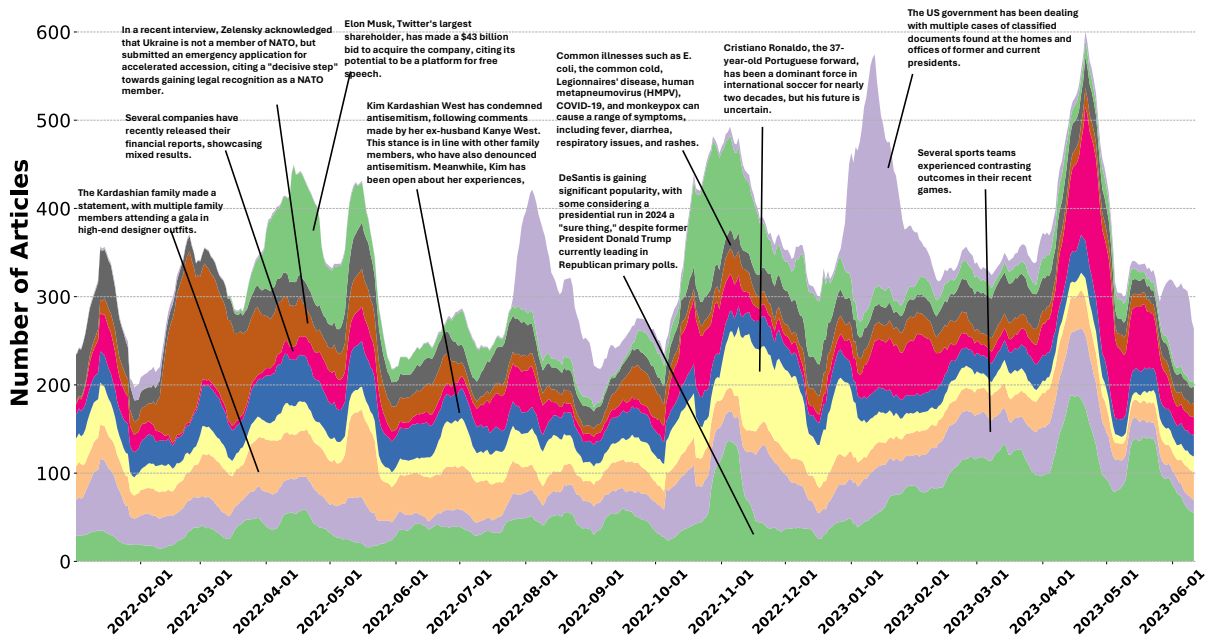
(a) Story Spread on Mixed-Reliability News Websites

Reliable News			
Keywords	Articles	Websites	Most Prolific Domains
desantis, ron, florida, gov, governor	62,911	1,330	yahoo.com (5,454), floridapolitics.com (3,098), thehill.com (1,900)
quarter, net, revenue, expense, loss	57,453	824	benzinga.com (19,266), yahoo.com (3,356), marketwatch.com (3,106)
half, halftime, minute, goal, possession	52,362	809	yahoo.com (2,785), espn.com (2,527), rte.ie (2,124)
symptom, fever, cough, throat, sore	42,581	1,241	yahoo.com (1,685), webmd.com (1,023), news-medical.net (869)
roe, wade, abortion, dobbs, supreme	38,358	1,317	news-yahoo.com (2,118), thehill.com (798), cnn.com (615)

Mixed-Reliability News			
Keywords	Articles	Websites	Most Prolific Domains
desantis, ron, florida, gov, governor	32,723	386	tampafp.com (1,762), theepochtimes.com (1,441), nypost.com (1,220)
half, halftime, minute, goal, possession	21,113	228	the-sun.com (2,402), thesun.co.uk (2,354), independent.co.uk (1,968)
kourtney, kardashian, travis, khloe, barker	19,226	117	the-sun.com (7,046), thesun.co.uk (4,951), metro.co.uk (1,028)
ronaldo, cristiano, portugal, football, manchester	17,791	126	the-sun.com (3,307), thesun.co.uk (3,024), dailystar.co.uk (1,732)
kardashian, khloe, kim, kourtney, jenner	16,578	145	the-sun.com (5,603), thesun.co.uk (3,781), metro.co.uk (987)

Unreliable News			
Keywords	Articles	Websites	Most Prolific Domains
desantis, ron, florida, gov, governor	13,963	328	dailymail.co.uk (2,831), ussanews.com (902), occupydemocrats.com (496)
dictatorship, democracy, censorship, propaganda, authoritarian	9,094	358	abovetopsecret.com (2,549), noqreport.com (384), theburningplatform.com (293)
respect, dignity, equal, equality, freedom	7,950	333	dailymail.co.uk (3,090), abovetopsecret.com (2,183), lifesitenews.com (160)
classified, document, maralago, archive, wilmington	7,678	249	dailymail.co.uk (1,393), thegatewaypundit.com (658), survivethenews.com (253)
constitution, oath, amendment, constitutional, defend	7,655	353	abovetopsecret.com (2,044), dailymail.co.uk (452), theqtree.com (307)

Table 14: Top Narratives from Different News Sources by Number of Articles in Our Dataset.

L Keywords and Auto-Generated Summaries

Story.	Keywords	
		Auto-Generated Summary
1	transgender, trans, minor, banning, restrict	Across the United States, anti-trans legislation is on the rise, with multiple states imposing restrictions on transgender individuals, including bans on gender-affirming care for minors, drag shows, and participation in female school sports. At least 20 states have imposed similar limits on trans athletes, and nearly half of all states have laws that make it difficult for transgender youth to receive necessary care. Companies and lawmakers are being urged to take a stance against these laws, which are seen as a personal attack on the LGBTQ+ community.
		Random Sample Passage
		Contino described how anti-trans legislation has grown across the country, including two bills in Tennessee that ban gender-affirming care for minors as well as drag shows. She says the overarching impact this culture can have on transgender people is huge, and companies need to take a stance.
		Auto-Generated Summary
2	buckingham, mall, palace, procession, royal	Large crowds have gathered in London to pay respects to the Queen, with many camping out for days to secure a prime viewing spot for the procession and potentially catch a glimpse of the newly crowned monarch and Queen Camilla.
		Random Sample Passage
		Some 100 heads of state, representatives from 200 countries and hundreds of thousands of visitors are expected to descend on London for the historic event. Many die-hard royal fans are already camped out near Buckingham Palace to secure the best viewing spot.
		Auto-Generated Summary
3	card, debit, atm, credit, fraudulent	To protect against credit card theft and unauthorized transactions, it's essential to keep important documents secure and be aware of potential scams, such as card skimming. Credit card data can also indicate changes in consumer spending habits. However, the use of credit and debit cards has become increasingly popular, with many retailers now accepting these forms of payment. Despite the convenience of these cards, there is still a risk of theft and unauthorized transactions, with some cases involving large-scale data breaches and individuals using sleight-of-hand techniques to steal gift cards.
		Random Sample Passage
		Card skimming involves the use of a device that looks like a normal part of a point-of-sale machine or PIN pad but instead copies EBT card information for the thief. These card skimmer devices are difficult to spot, to the point that retailers may not be aware of their installation.