Specious Sites: Tracking the Spread and Sway of Spurious News Stories at Scale

Hans W. A. Hanley
Stanford University

Deepak Kumar
Stanford University

Zakir Durumeric
Stanford University

***Abstract*—Misinformation, propaganda, and outright lies proliferate on the web, with some narratives having dangerous real-world consequences on public health, elections, and individual safety. However, despite the impact of misinformation, the research community largely lacks automated and programmatic approaches for tracking news narratives across online platforms. In this work, utilizing daily scrapes of 1,404 unreliable news websites, the large-language model MPNet, and DP-Means clustering, we introduce a system to automatically identify and track the narratives spread within online ecosystems. Identifying 55,301 narratives on these 1,404 websites, we describe the most prevalent narratives spread in 2022 and identify the most influential websites that originate and amplify narratives. Finally, we show how our system can be utilized to detect new narratives originating from unreliable news websites and to aid fact-checkers like Politifact, Reuters, and AP News in more quickly addressing misinformation.**

1. Introduction

Over the last decade, spurious, misleading, and outright false information has spread throughout online ecosystems [1]. Digital misinformation has influenced elections [2], promoted bogus health cures leading to unnecessary deaths [3], and incited mob violence throughout the world [4], [5]. Worsening the problem, misleading stories have been shown to spread at over ten times the rate of true information [6].

The security community, likening disinformation to an attack similar to spam, phishing, and censorship [7], [8], has applied a range of methodologies to ameliorate its spread [9], [10], [11], [12], [13], [14]. For example, by examining features similar to those used to identify spam accounts, researchers have identified networks of state-propagandists throughout Reddit and Twitter [11], [15]. However, despite these advances, most investigations into false narratives remain limited in scope and retroactive, primarily conducted through time-consuming, qualitative approaches [16], [17]. To make fundamental progress in combating the threat posed by disinformation, we argue that the security community needs to build technical approaches for tracking the online spread of false narratives globally and in real time.

In this work, we present an NLP-based approach for programmatically identifying and tracking the spread of narratives and stories across *unreliable* news websites and social media platforms. Between January 1 and November 1, 2022, we crawl a set of 1,404 known misinformation, state-propaganda, biased, and otherwise unreliable news websites as well as two fringe forums, 8kun and 4chan. We extract passages from these news articles, which we embed using an MPNet model [18] fine-tuned with contrastive learning on the semantic textual similarity task. Then, employing a modified version of nonparametric DP-Means, we cluster these embeddings to identify specific narratives/stories.

Our approach enables us to isolate and track 55,301 narrative threads that spread on unreliable news websites during 2022. We do not attempt to determine whether individual stories are factual, which is a qualitative task that ML-based approaches have failed to reliably achieve. Rather, we track all narratives from these specious sites across online ecosystems and quantify these websites' influence. We find, unsurprisingly, that many of the most prominent narratives in 2022 concerned the Russo-Ukrainian War and inflation, with websites like southfront.org, rt.com, and zerohedge.com dominating these topics. Next, identifying which websites play outsized roles in *originating* and *amplifying* narratives across our set of unreliable news websites, we find that a website's popularity has a small correlation with its ability to propagate narratives with seemingly minor websites like infostormer.com or barenakedislam.com playing massive roles in popularizing stories.

Next, we investigate how our method can be used to focus limited investigative resources on the most pernicious narratives. We show that, like an early detection alarm, our approach can identify when new narratives emerge. Comparing when popular false narratives appeared to when three major organizations (AP News, Reuters, and Politifact) fact-checked them, we show that our system can prioritize checking misleading narratives months before their peak when they first start to gain traction. We hope that this type of real-time visibility can enable fact-checkers and journalists to more efficiently track and respond to new, potentially problematic narratives as soon as they emerge.

Similar to past large-scale empirical analysis in the security community (*e.g.*, [19], [20], [21], [22], [23], [24], [25], [26], [27]), our work shows that a programmatic approach to tracking narratives at scale uncovers a set of online propagation patterns that would have been difficult to uncover through manual, small-scale investigations. We also discuss how having a continuous tracking process can help analysts uncover and track the most worrisome influence operations. We stress that our approach does not make

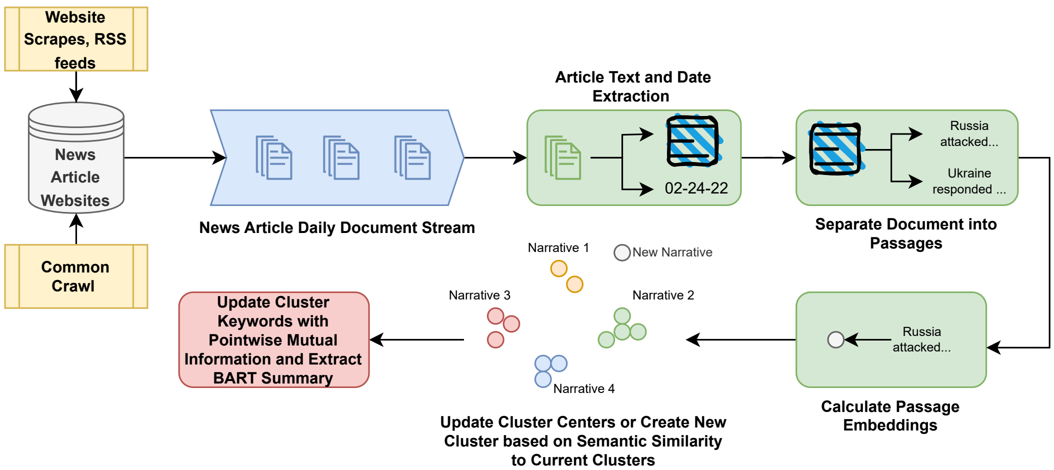

Figure 1: Our pipeline for identifying and labeling narrative clusters from the daily publications of unreliable news websites.

factual judgments on individual stories or on the reliability of websites. Indeed, our system takes *as input* websites that human experts previously labeled as unreliable. Rather, our approach provides the critical, real-time visibility into the spread of news narratives online that human experts need to effectively identify and respond to misinformation.

2. Background

Unreliable information often spreads through multiple avenues as individual users, state-supported actors, and even entire platforms participate in the dissemination of falsities. Unreliable information can take the form of *misinformation*, *disinformation*, *fake-news*, *propaganda*, among others [28]. *Misinformation* is any information that is false or inaccurate regardless of the author's intent [28], [29], [30], [31]. The term "fake news" is often used interchangeably with misinformation. Disinformation, in contrast, is inaccurate information spread with the deliberate purpose to mislead [28], [32]. Similar to disinformation, *propaganda* refers to "deliberate, systematic information campaigns, usually conducted through mass media forms" regardless of whether the information is true or false [28]. A single narrative can be considered *disinformation* when spread by a state actor and as *misinformation* when spread by users. We refer readers to Jack et al. [28] for a more detailed taxonomy.

3. Methodology

The goal of our work is to programmatically track how narratives spread amongst "*unreliable*" news websites. In this section, we define a narrative and then describe how we collect news articles and extract narratives. We emphasize that while we focus on websites known to publish misleading, false, or state-controlled narratives, we do not assume that *all* narratives from these sites are "misinformation". Indeed, many stories are not. We do not seek to label new narratives as "misinformation", which is a qualitative, investigative task. As such, we refer to only individual stories that have been previously identified by experts as *misinformation* or *disinformation*, as false. We label the websites that spread these verified false narratives as *unreliable*.

3.1. Narrative Definition

Tracking misinformation narratives requires a high degree of specificity. Unlike traditional topic modeling, which seeks to identify *themes* and/or statistical word correlations [33], [34], [35], misinformation tracking requires distinguishing between specific narratives and stories. Within this work, we define a narrative/story using the same definition as the Event Registry [36], Hanley et al. [37], and Miranda et al. [38]: collections of documents that seek to address the same *event* or *issue*. For example, two example events in the Event Registry are "Felix Baumgartner's jump from a helium balloon on October 14, 2012" and "bombings during the Boston Marathon on April 15, 2013." Within our dataset, events constitute ideas like "election fraud in the 2020 US election" and "the COVID-19 vaccine leading to mass death." An example of two ideas—while related—that we do not consider to be the same narrative are "US funds Ukrainian War" and "Russia attacks Ukraine."

3.2. System Architecture

As shown in Figure 1, our system (1) collects news articles from unreliable news sites on a daily basis through our own web scraping and from Common Crawl data [39]; (2) parses out news articles, extracting article text, which it then segments into constituent passages; and (3) embeds text passages into a shared subspace utilizing the large language model MPNet [18]. Then, to track the spread of narratives, our system (4) clusters semantically similar content using DP-means, and (5) extracts keywords and auto-generated summaries to provide human-interpretable labels for the narrative clusters.

We opt for this LLM-based approach because our system needs to track specific narratives. Prior approaches that utilize simpler, more generic keyword-based topic modeling tools like LDA fall short in identifying specific narratives across different news websites [40]. Furthermore, keyword-based approaches often rely on pre-existing expert knowledge of disinformation campaigns and largely cannot adapt to the rapid pace of the news ecosystem [41]. By utilizing this new approach, our system can update and track news stories without *a priori* or domain-specific knowledge in an efficient and fine-grained manner.

For this paper, we use data from January 1 to November 1, 2022, but we emphasize that our system runs continuously, enabling us to identify new narratives in near real time. In the remainder of this section, we detail each step and validate that our system captures specific and coherent narratives.

3.3. Data Collection

Our study is based on scraping and parsing articles from websites known to spread unreliable information.

Unreliable News Websites. We collect articles from 2,514 candidate websites that have been labeled as "politically biased", "misinformation", "disinformation", "conspiracy", "fake news", or "state-based propaganda" by past studies (Iffy Index [42], OpenSources [43], Politifact [44], Snopes [45], Melissa Zimdars [46], and Hanley et al. [47]). This list includes politically-biased websites like dailywire.com, conspiracy-oriented websites like x22report.com, and state-propaganda outlets like rt.com.

Scraping Articles. We crawl websites using Colly[1] and Headless Chrome orchestrated with Python Selenium. For each website, we collect the homepage and linked pages daily from January 1 to November 1, 2022. To ensure full coverage of each site's published articles, we additionally gather the HTML pages indexed by Common Crawl [48] for each site during this same period. We emphasize that under 1% of articles were only found in the Common Crawl dataset, indicating that our scrapes found the vast majority of published content on each site. We then parse each HTML page to collect the published *articles* using the Python libraries `newspaper3k` and `htmldate`.

Of our 2,514 candidate websites, 1,404 were operational and published articles during our 2022 measurement period (many sites that spread unreliable information are short-lived [12]). Altogether, we collect 2,126,738 articles from these sites. We provide this data to researchers upon request.

3.4. Preprocessing and Embedding

To prepare data for embedding, we first remove any non-English articles using the Python `langdetect` library and then remove URLs, emojis, and HTML tags. We then segment each article into its constituent paragraphs by splitting article text based on newline and tab characters. Then, in line with prior work, we subsequently divide these paragraphs into article passages with 10–100 words [37], [49], [50]. There were several instances (*e.g.*, sputniknews.com) where this approach failed; in these cases, we manually identified and built custom parsers for each site based on site-specific HTML elements.

After preprocessing, we embed the constituent passages that make up each article. Embedding passages rather than entire articles is in line with prior work [49] for topic analysis as articles often address multiple narratives but embeddings should represent only a single narrative or idea [37], [49], [50]. We thus embed passages to capture context while also obtaining an embedding for the (often) one narrative/idea present within the passage.

GloVe	BERT	USE	All-MPNet	Our Model
0.580	0.464	0.749	0.840	**0.856**

TABLE 1: Evaluation–based on Pearson Correlation—of our MPNet-contrastive model and other models on the SemEval STS-benchmark [51]. Data for GloVe [52], BERT [53], the Universal Sentence Encoder (USE) [54] is from Reimers et al. [55].

We specifically embed passages using a version of MPNet that we fine-tune on the semantic text similarity (STS) task [37], [56] using unsupervised contrastive learning for sentence embeddings as specified in Gao et al. [57] on a random assortment of passages from January 2022 from our websites. We perform this fine-tuning with the default hyperparameters (learning rate 3×10^{-5}, batch size=128, and 1M examples) specified in Gao et al. and by freezing all but the last two layers of a public version of MPNet.[2] See Appendix A for details. This ensures that our model is attuned to the language present on our set of websites. As seen in Table 1, despite not being trained on the SemEval STS Benchmark [51], a benchmark for measuring the quality of text embeddings, our model outperforms the fine-tuned publicly released version of MPNet. After fine-tuning our model, from the 2.1M articles, we embed 27,850,016 passages (11.00 hours on an Nvidia RTX A6000).

3.5. Comparing Semantic Content

We compare the semantic content of our embeddings utilizing cosine similarity [18]. Prior work [37], [59], [60] has found that a cosine similarity threshold of 0.60–0.80 can be utilized to determine whether two pieces of text are about the same narrative. However, to ensure that we select a minimum cosine similarity threshold that accurately models whether passages are about the same *narrative* as defined in Section 2, we: (1) benchmark our model on the English portion of the Multilingual SemEval2022 dataset [58], and (2) manually validate the coherency of a random sample of passage pairs.

1. https://github.com/gocolly/colly

2. https://huggingface.co/sentence-transformers/all-mpnet-base-v2

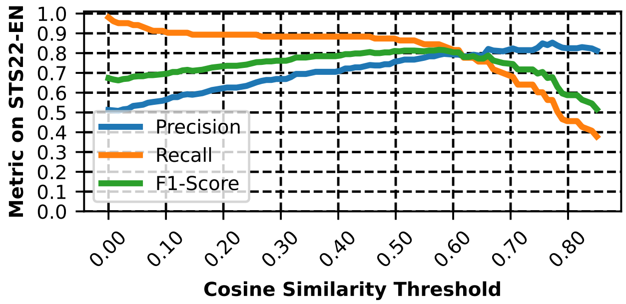

Figure 2: Evaluation of our model's precision, recall, and F_1 scores on the English portion of the SemEval22 dataset [58] (using 3.0 as the cut-off for the two articles being about the same event [50]).

0.45 Thresh.	0.50 Thresh.	0.55 Thresh.	0.60 Thresh.
$80.6\% \pm 6.33\%$	$85.7\% \pm 8.29\%$	$94.7\% \pm 6.94\%$	$96.29\% \pm 7.12\%$

TABLE 2: Evaluation of the precision of embedded messages having the same narrative at various thresholds utilizing 2000 random passages. We provide 95% Normal confidence intervals.

As seen in Figure 2, on the SemEval22 [58] dataset, as the similarity threshold increases, our model's precision in determining whether two passages are about the same narrative increases while the corresponding recall decreases, reaching a peak F_1 score near 0.60 cosine similarity.

To confirm this result, we perform a manual evaluation by selecting 2,000 random passage pairs from our dataset with similarities at varying thresholds and have two experts determine whether the passage pairs are about the same narrative (per our definition) and then determine the corresponding precision at various thresholds. We calculate a Kappa's Cohen of 0.80 between our two raters, indicating a high degree of agreement. As seen in Table 2, as the threshold used to determine similarity increases, we see an increase in precision. We choose a threshold of 0.60 (as a lower bound) as it has acceptable manually calculated precision (96.29%) and F_1 score on the SemEval22 [58] dataset (0.809). We present an example passage pair at our selected threshold of 0.60 in Figure 3 and other thresholds in Appendix B.

3.6. Identifying Narratives

To identify *narrative/stories* in our dataset, we cluster our passage embeddings using cosine similarity and DP-Means, a non-parametric version of K-means (Appendix C). Prior work has identified narratives using a similar high-level approach [17], [37], but our methodology differs in several ways based on several unique requirements. First, our approach must be *highly scalable*. While Hanley et al. [37] utilize a BERTopic-based method [61], we find that this does not scale to the approximately 100K embeddings per day in our dataset. Second, we need to update our clusters on a daily basis as new news articles are published, which past approaches like BERTopic do not allow. Third, since the number of narratives is unknown *a priori*, the methodology must automatically infer the number of clusters, which precludes parametric algorithms like incremental K-means clustering.

PASSAGE 1: The MaraLago search warrant served Monday was part of an ongoing Justice Department investigation into the discovery of classified White House records recovered from Trump's home earlier this year. The Archives had asked the department to investigate after saying 15 boxes of records it retrieved from the estate included classified records.
PASSAGE 2: The FBI raided Donald Trump's estate in MaraLago using the pretext of Trump supposedly violating the Presidential Records Act by keeping documents after he left office to initiate a siege to terrorize Joe Biden's chief political rival.

Figure 3: Example passage pair at our selected minimum similarity threshold (0.60).

We specifically adapt Dinari et al.'s efficient and parallelizable version of DP-Means [62] (Appendix C), making four main alterations. First, we cluster embeddings based on their cosine similarity rather than their Euclidean distance. We set $\lambda = 0.60$ (the minimum cosine distance an embedding can be from a cluster before a new cluster is created) to ensure that clusters have high semantic similarity, informed by our prior manual investigation (Table 2). Second, we perform partial fits over each day's worth of news article embeddings. Specifically, on each day throughout 2022, we embed that day's passages and update the previous day's cluster centers (*i.e.,* we update our given clusters on a daily basis with the DP-Means algorithm until convergence utilizing that day's article embeddings). Third, we remove the random reinitialization of clusters added by Dinari et al. [62] from the algorithm; we find that this step often led to over-clustering given that many website passages are slight variations of each other. Lastly, we note that rather than relying on Dinari et al.'s released code, we re-implement their algorithm to take advantage of the matrix multiplication speedups that come from utilizing a GPU (3 times speedup with an Nvidia RTX A6000).

For this work, we utilize the clusters from November 1, 2022. From January 1 to November 1, clustering all embeddings required the equivalent of 1.5 days. We filter out clusters where 50% or more of the passages are from only one website (*e.g.*, author bios) or there were fewer than 25 passages to remove spam, similar to the methodology specified by Leskovec et al. [17]. After this removal, we identify 55,301 narrative clusters. Each article's passages are part of an average of 5.02 narrative clusters (4.0 median). On average, each embedding has an average similarity of 0.686 to its cluster center, which shows that our embeddings are assigned to clusters with high semantic similarity. Each narrative cluster has an average cosine similarity of 0.0169 with other identified narrative clusters, which indicates that our approach identified distinct narratives.

3.7. Interpretability and Narrative Specificity

We create human-interpretable identities for our narrative clusters using two approaches. First, we extract the most distinctive and representative keywords of the cluster using pointwise mutual information [63], [64] (detailed in Appendix D). As in Kessler et al. [65], rather than finding the pointwise mutual information between different words,

Narr.	Keywords	Passages Checked	Prec.	Narr.	Keywords	Passages Checked	Prec.
1	kiril, patriarch, orthodox, church, putin	427	99.53%	16	peterson, suspend, rubin, jordan, elliot	252	100.00%
2	abbott, texas, border, greg, lone	500	99.40%	17	norwegian, ellingsen, feminist, lesbian, norway	108	100.00%
3	sinema, manchin, filibuster, kyrsten, senate	500	100.00%	18	lantsman, trudeau, melissa, mp, swastika	127	100.00%
4	hurricane, atlantic, storm, tropic, season	500	99.80%	19	polio, 1979, eradicate, virus, disease	109	99.08%
5	balloon, leaflet, korea, korean, north	100	100.00%	20	humanitarian, aid, shelter, refuge, relief	500	100.00%
6	monkeypox, york, nyc, outbreak, city	385	96.62%	21	kiev, coup, nationalist, neonazi, nazi	143	98.60%
7	johnson, resign, poll, boris, tory	500	95.20%	22	noah, flood, ark, wives, genesis	155	90.97%
8	antibody, monoclon, regeneron, omicron, variant	302	100.00%	23	fda, prescribe, offlabel, drug, treatment	195	100.00%
9	fauci, anthoni, kennedy, pharma, gate	384	100.00%	24	refugee, asylum, persecute, seeker, migrant	192	100.00%
10	nucleic, acid, test, shanghai, province	500	99.40%	25	swift, sanction, bank, sberbank, vtb	500	100.00%
11	finland, sweden, nato, deploy, nuclear	133	100.00%	26	orban, fidesz, viktor, hungary,victory	469	99.60%
12	windfall, profit, tax, oil, barrell	379	99.21%	27	unvaccined, infection, recipe, covid19, covid	167	100.00%
13	energy, europe, crisis, price, electric	500	100.0%	28	nuclear, closer, brink, cuban, war	364	96.43%
14	pen, macron, le, french, marin	500	99.20%	29	alexandra, pelosi, footage, nancy, daughter	100	100.00%
15	protein, spike, mrna, inject, cell	173	100.00%	30	civilian, kabul, afghan, drone, strike	500	94.00%
						Prec.	**98.90%**

TABLE 3: Evaluation of the precision of our narrative analysis model on a random set of 30 stories/narratives derived from the articles in our dataset. Keywords were extracted utilizing pointwise mutual information. We checked all available passages in cases where there are fewer than 500 passages in the story/narrative cluster.

we utilize the measure to understand individual words' association with narrative clusters. In this manner, we find the set of words most *distinctive* to each cluster. Second, after identifying the top five passages closest (*i.e.*, with the largest cosine similarity) to the center of the cluster, we subsequently use an off-the-shelf state-of-the-art BART [66] summarization tool from Huggingface fine-tuned on news data, to summarize the contents of the cluster. We utilize this approach because while keywords provide an identifiable "handle" for each cluster, keywords typically do not fully capture the full semantic meaning or the specificity of our narrative clusters. For example, the auto-generated summary for the cluster with keywords *Age, Pfizer, Booster, Children, Vaccine* is:

> *The U.S. Food and Drug Administration FDA in October 2021 authorized the PfizerBioNTech COVID vaccine for children 5 through 11. Children under 5 remain the only segment of the US population that isn't eligible for one of the COVID vaccines.*

where a random passage from the cluster states:

> *As of now, U.S. children aged five and older are eligible for the COVID19 vaccine, though only Pfizer's shot has received authorization. The Pfizer jab is also available as a booster for children 12 and older.*

However, the auto-generated summary for a similar cluster with keywords *Children, Risk, Adult, Covid, Immunity* is:

> *Children have a minuscule risk of COVID mortality. There is very limited safety data for vaccines from the trials on children. If the risk of adverse reactions is the same as for adults, the harms outweigh the risks.*

where a random passage from the cluster states:

> *COVID poses no danger to children. They have a statistically zero chance of dying from that disease. The COVID shots, however, are already linked to innumerable adverse reactions, and their longterm side effects are unstudied.*

This illustrates the need for further specificity using summarization to understand the narratives being spread. We provide several additional examples in Appendix F.

3.8. Validating Narrative Clusters

We evaluate our narrative clustering technique by validating whether a random sample of 500 passages (or maximum present) for a random set of 30 narrative clusters are about the same narrative using the methodology outlined in Section 3.5. Our methodology identifies coherent story/narrative clusters with an overall 98.9% precision and a minimum precision of 90.97% for Topic 22 (Table 3).

3.9. Ethical Considerations

Our analysis is based on analyzing publicly posted news articles. We limit the load that each news site experiences by checking for new articles daily at a maximum rate of one request every 10 seconds. We further follow the guidelines as outlined by others for scraping data [5]. During our study period, we received no requests from websites to opt-out.

4. Narratives on Unreliable News Sites

In the last section, we presented and validated our methodology for programmatically extracting the narratives promoted by unreliable news websites. Here, we describe the most prolific narratives, trace three misinformation/propaganda stories, and derive communities of topically-related websites.

4.1. The Largest Narratives

We start by analyzing the narratives most prolifically covered by our set of unreliable news sites in 2022. As can be seen in Table 4, the most popular narratives concerned the Russo-Ukrainian War, inflation, and Elon Musk's criticism and later acquisition of the social media platform Twitter. As can be seen in Figure 4, we further observe peaks in coverage of specific stories, as well as narratives that maintained consistent coverage throughout our study. For example, stories about abortion peak both before the US Supreme

Narr.	Keywords	Articles	Websites	Most Profilic Domains	Auto-Generated Summary
1	ukraine, troop, kyiv, russian, donbas	10,225	416	express.co.uk (626), southfront.org (523), dailymail.co.uk (464)	The Russian military has not been able to fully encircle and neutralize the grouping of Kyiv s forces in the Donbass so far. At the same time, the Russians managed to liberate a number of important territories and towns.
2	zelensky, volodymyr, ukraine, kyiv, president	9,285	426	dailymail.co.uk (548), nypost.com (466), express.co.uk (411)	Ukrainian President Volodymyr Zelensky has accused Russian forces of committing genocide in his country. He also slammed the West.
3	index, consumer, inflation, cpi, price	7,496	361	shorenewsnetwork.com (963), theepochtimes.com (514), dailymail.co.uk (336)	The consumer price index climbed 0.6 percent from a month before. Compared with January of last year, consumer prices are up 7.5 percent. The Consumer Price Index increased 9.1 percent in the year through June.
4	musk, elon, twitter, platform, tesla	7,416	472	nypost.com (364), dailymail.co.uk (281), theepochtimes.com (222)	Tech Mogul and Tesla Boss Elon Musk is wellknown for his wisecracks and witty posts he shares on Twitter. Musk has been critical of social media, particularly Twitter, over its enforcement of rules that critics say targets conservative voices.
5	germany, europe, oil, sanction, energy	7,152	393	express.co.uk (355), zerohedge.com (323), rt.com (283)	Russia has been hit by sweeping sanctions on its economy and trade since the start of Putin's war in Ukraine. But measures by EU governments have not targeted oil and gas contracts with Moscow. Europe is heavily reliant on Russia for its energy needs.

TABLE 4: Top 5 narratives—by number of articles—in our 2022 dataset.

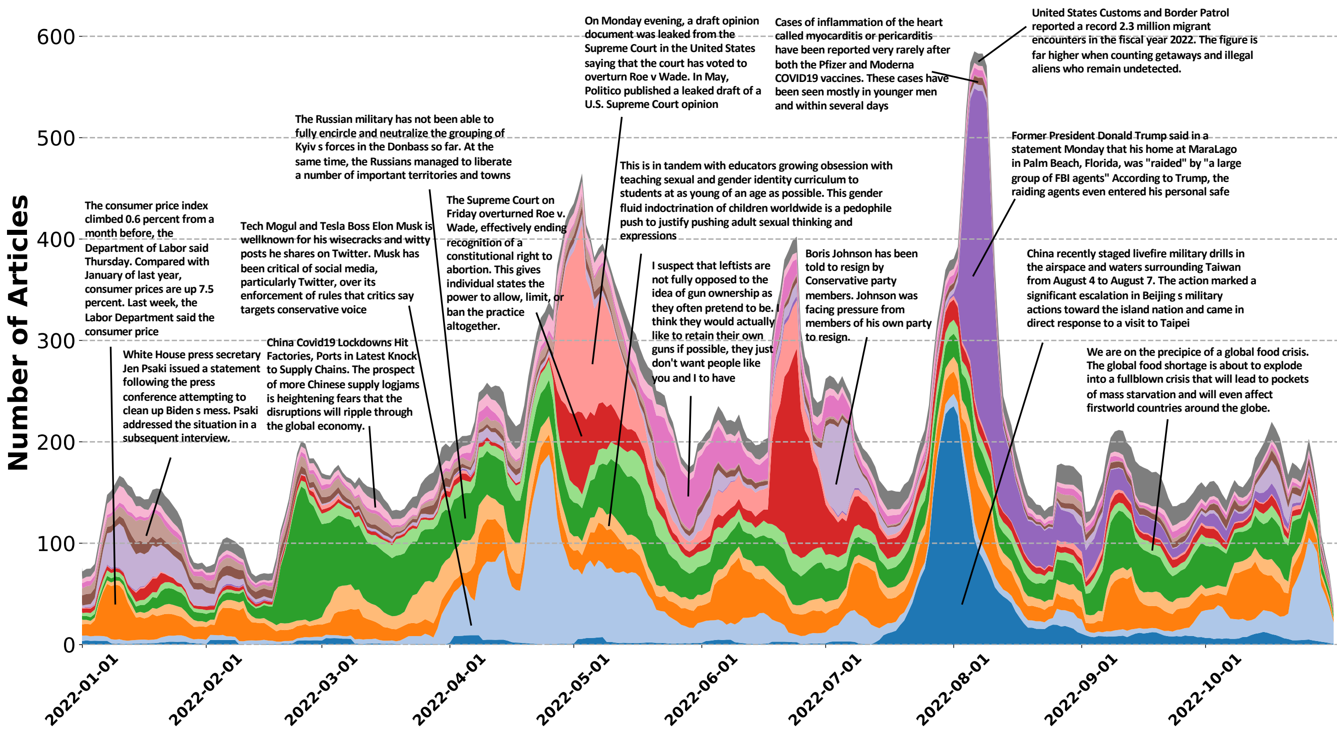

Figure 4: Article volume of the most popular narratives from January 1, 2022, to November 1, 2022.

Court decision (Dobbs v. Jackson) about federal abortion rights was leaked and following the official decision [67]. In contrast, a narrative about the EU's role in NATO has seen a steady stream of articles throughout the year, with a slight uptick following the Russian invasion of Ukraine. Analyzing the specific news sites that post about each narrative, we find that many Russian-backed and controlled websites [68] such as rt.com and southfront.org, in addition to several UK-based tabloids express.co.uk and dailymail.co.uk were the most prolific in writing about the Russian invasion of Ukraine (Narratives 1, 2, 5 in Table 4). This largely matches previous studies of the Russian-controlled media in influencing discussions on the war [37].

4.2. Misinformation/Propaganda Case Studies

As seen in the last section, many of the most common narratives are mainstream news topics. However, one of our goals is to track the spread of misinformation narratives. In this section, we show that our technique is capable of tracking known unreliable narratives by investigating the evolution of one confirmed *propaganda* and two confirmed *misinformation* stories.

Ukrainian Nazis (Keywords: Azov, Battalion, Regiment, Far-right, Ukraine): One of the most prominent propaganda narratives utilized by Russian media in justifying the Russian Federation's invasion of Ukraine was that the Ukrainian government was controlled by "neo-nazis" [63]. This is despite Ukraine's relatively low level of antisemitism [69]. Our method is able to find that even before the Russian invasion of Ukraine on February 24, 2022, there were heavy references to Nazism in Ukraine by Russian-controlled or influenced outlets. For example, on January 27, gloablresearch.ca penned:[3]

> *If we are to draw parallels between the current crisis on the Ukraine border and WW2 we should compare the Neo-Nazi ideology which dominates Ukrainian nationalism with that of Nazi Germany.*

However, as seen in Figure 5, the major increase in the number of articles promoting this narrative occurred in the weeks prior to the Russo-Ukrainian War (specifically jumping in volume on February 8, 2022). The most prominent, websites that pushed this narrative were unsurprisingly known Russian propaganda outlets including globalresearch.ca (68 articles), sputniknews.com (55), and rt.com (46). Beyond these known pro-Russian websites, we find US-based websites like veteranstoday.com (64 articles), sott.net (62), and thegatewaypundit.com (21) repeating this narrative.

Killer Covid-19 Vaccines (Keywords: Vaccine, Safe, Adverse, MRNA, Effect): One prominent misinformation narrative about COVID-19 that we identify is that COVID-19 vaccines are "killer vaccines" and a major cause of death around the world. For example on lewrockwell.com, an author wrote:[4]

> *Whatever they may be, these vaccines are most definitely not safe. We can very clearly see this from the explosion of reports of death to the Vaccine Adverse Event Reporting System (VAERS), which coincided with the introduction of the Covid injections in late 2020.*

As seen in Figure 5, stories about "killer vaccines" have remained prominent throughout 2022, increasing in popularity several times throughout the year. The sites that most prominently echoed this narrative were theepochtimes.com (53 articles), pandemic.news (36), and vaccines.news (31). This is consistent with prior studies [70], [71].

2020 Election Denialism (Keywords: Fraud, Election, 2020, Irregular, Voter): The narrative that the presidential election was stolen and that current President Biden is illegitimate [72] spread throughout social media and was a key aspect of the January 6, 2021 attack on the US Capitol [5]. We see in our dataset that this false narrative maintained

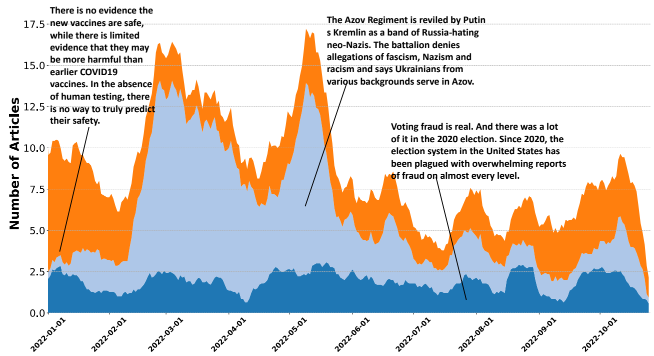

Figure 5: Volume over time for case-study narratives of Ukrainian Nazis, Killer COVID-19 vaccines, and 2020 Election Denialism.

a substantial presence amongst unreliable news sites (Figure 5). For example, the fringe website thetrumpet.com wrote on January 6, 2022:[5]

> *The insurrection hoax is a cover-up for the stolen election*

The websites that most prominently repeated this narrative were welovetrump.com (143 articles), thegatewaypundit.com (55), and votefraud.news (14).

4.3. Communities of Unreliable News Websites

To begin to capture the semantic communities that exist within the unreliable news ecosystem, we utilize each website's distinct distribution of articles among our discovered set of 55,301 stories/narratives. To compare each website's reporting choices and semantic content, we represent each website's narratives as a multinomial distribution. For example, if we had three narratives (rather than 55.3K) and a website that wrote five articles about Narrative 1, four articles about Narrative 2, and one article about Narrative 3, the website's distribution would be [0.5, 0.4, 0.1]. We do this for all 55,301 narratives and 1,404 websites, thus representing each website as a 55,301-dimensional vector of probabilities. We then use Jensen-Shannon Divergence [73] (detailed in Appendix E) to compare websites' probability vectors. For example, the JS-Divergence of rt.com and sputniknews.com.com, two Russian state-sponsored websites that often address similar topics [37], [68] is 0.413 while the JS-Divergence of rt.com and nypost.com, a US-based website, is 0.600.

After calculating each website's narrative similarities using JS-Divergence with every other website in our dataset, we build an undirected graph with edge weights based on these values (*i.e.,* an edge between a website P and Q is given a weight of $1 - JS(P||Q)$) where $JS(P||Q)$ is the JS-Digerence between websites P and Q. We determine communities of websites using the Louvain clustering algorithm [74]. Louvain clustering identified 3 communities,

3. https://web.archive.org/web/20220127162858/https://www.globalresearch.ca/war-fever-air-west-confuses-russia-nazi-germany/5768335

4. https://web.archive.org/web/20220120061509/https://www.lewrockwell.com/2022/01/vasko-kohlmayer/dangerous-and-deadly-over-1000-scientific-studies-referencing-injuries-and-deaths-from-covid-vaccines/

5. https://web.archive.org/web/20220107091002/https://www.thetrumpet.com/stephen-flurry/25070-the-insurrection-hoax-is-a-cover-up-for-the-stolen-election

and from these communities, we qualitatively identified the corresponding three semantic communities: US-focused, International, and Conspiratorial. We label these clusters based on the top topics found within each cluster, with the US-focused cluster writing on Abortion and the Biden Administration, the International cluster writing on Russo-Ukrainian War, and the Conspiratorial cluster heavily writing about COV1D-19 vaccines.

US-focused Community. 364 websites fall into our US-focused community including sites like dailywire.com, breitbart.com, and welovetrump.com. The most common narrative in the community concerned the US Supreme Court Dobbs v. Jackson decision to overturn the 1973 Roe v. Wade decision that provided the federal right to abortion (Keywords: Roe, Abortion, Wade, Overturn, 1973). To further examine the role of this website community, particularly in regard to its most prominent narrative, we collect a larger set of narratives that more broadly relate to the topic of abortion by aggregating all 284 narrative clusters whose centers have a 0.50 similarity to the Abortion/Roe cluster.

We consider a website to *originate* a narrative if they published an article about the narrative on the first day that the narrative appeared in our dataset (more than one website can originate a narrative). Altogether, we find that 61.27% of Roe/Abortion narratives originated from this community, with the website rawstory.com originating the most Roe/Abortion narratives (36). In addition to originating most of the narratives about abortion, these websites contribute 76.93% of the articles on the 121 narratives about abortion; theepochtimes.com (13,30 articles across 146 Roe/Abortion narratives) and breitbart.com (1,264 articles across 136 Roe/Abortion narratives) have the most. Largely expected, many International websites such as dailymail.co.uk (1,104 articles across 116 narratives) and Conspiracy websites like evil.news (65 articles across 25 narratives) also picked up on these US-centered political narratives, evidencing the spread of stories from this community.

International Community. 408 websites fall into our International community including rt.com and dailymail.co.uk. The top story was one of our top overall narratives and Russian invasion of Ukraine (Keywords: Ukraine, Kyiv, Troop, Russian, Donbas). We gather a larger set of 504 narrative clusters that discuss the Russo-Ukrainian War using the same methodology outlined in the prior section.

We find that 38.2% of Russo-Ukrainian War narratives started from the International community of websites with 5.4% of these narratives specifically starting on nine pro-Russian propaganda websites [68]. Southfront.org.com (7 Ukraine narratives) and tass.com (10 Ukraine narratives) originate the most narratives among these Russian websites. We again find that this cluster of websites is responsible for a large portion (56.12%) of articles about the war. Again, largely expected, other websites such as nypost.com (2,934 articles across 160 Ukraine narratives) or treason.news (32 articles across 23 Ukraine narratives) write extensively about the conflict as well.

Conspiratorial Community. 632 websites belong to our Conspiratorial community, including popular sites known for spreading conspiracy theories about QAnon and COVID-19 [5], [47] like unz.com, qresear.ch, and radiopatriot.net. Unsurprisingly, the top narrative within this community concerns COVID-19 (Keywords: Children, Risk, Adult, Covid, Immunity). We gather a more extensive set of narrative clusters that discuss the COVID-19 and/or COVID-19 vaccines using the same methodology outlined before; altogether gathering 162 narratives. Most prominently, the website childrenshealthdefense.org, the nonprofit run by Robert F. Kennedy Jr., wrote about nearly every COVID-19 story in our dataset (1250 articles across 94 COVID narratives).

We find that our set of Conspiratorial websites originate 37.3% of narratives about COVID-19, the most prominent of these being nvic.org (18 COVID narratives) and healthimpactnews.com (7 COVID narratives). We note that COVID-19 narratives originated not only from these websites but from our International (23.8%) and US-focused cluster (38.9%) as well, with the Conspiratorial cluster being responsible for only 20.3% of COVID-19 articles. This evidences that COVID-19 narratives spread and originated among a wide variety of different websites in 2022.

5. Originating and Amplifying Narratives

As seen throughout the last section, several websites play dominant roles in perpetuating and promoting certain types of stories. In this section, we quantify and identify which websites have the most pivotal roles in originating and amplifying narratives throughout the ecosystem of unreliable websites. As before, we consider a website to *originate* a narrative if they published an article about the narrative on the first day that the narrative appeared in our dataset (more than one website can originate a narrative). We consider a website to have *amplified* (*i.e.*, increased the popularity of) a narrative if it (1) posted an article about the narrative before the narrative peaked in popularity, (2) did not originate the narrative, and (3) if the posted article appeared in the first 15% of the total volume of that given narrative. We utilize the 15% cutoff as it ensures that the vast majority of a narrative's articles have not been published yet (*i.e.*, the story has not dramatically increased in popularity already), allowing us to observe how amplification affects the narrative's popularity. This is consistent with prior work [17].

Using this approach, we further investigate how the popularity of a website influences its effectiveness in originating and amplifying narratives. To do so, we utilize the website rank data provided by the Google Chrome User Report (CrUX) from October 2022, which Ruth et al. showed to be the most reliable website popularity metric [75].

5.1. Originating Narratives

To measure the efficacy of websites in originating and amplifying narratives, we perform a correlational comparison of the number of external non-origin articles that are written about a given narrative in the week after a given website originated a narrative vs. the number of non-origin

Domain	CrUX Rank	Wtd. Ext. Art.Δ	Cohen's D	To Peak Δ (Days)	Cohen's D
dailymail.co.uk	< 1K	0.334	0.754	-24.31	-0.426
express.co.uk	< 1K	0.186	0.007*	6.43	0.056*
breitbart.com	1K–5K	0.417	1.205	-56.50	-0.923
newsmax.com	1K–5K	0.345	1.004	-48.30	-0.787
zerohedge.com	1K–5K	0.196	0.678	-35.69	-0.639
thegatewaypundit.com	5K–10K	0.316	1.055	-63.80	-1.016
newsmax.com	5K–10K	0.345	0.844	-35.58	0.731
dailystar.co.uk	5K–10K	0.216	0.252*	-40.60	-0.502
redstate.com	10K–50K	0.526	1.492	-71.09	-1.181
twitchy.com	10K–50K	0.476	1.407	-72.93	-1.377
dailywire.com	10K–50K	0.472	1.222	-74.94	-1.388
theconservativetreehouse.com	50K–100K	0.643	1.916	-82.49	-1.707
halturnerradioshow.com	50K–100K	0.514	1.635	-50.76	-1.012
justthenews.com	50K–100K	0.472	1.372	-72.86	-1.268
–	–	–	–	–	–

Domain	CrUX Rank	Wtd. Ext. Art.Δ	Cohen's D	To Peak Δ (Days)	Cohen's D
therightscoop.com	100K–500K	0.694	1.653	-82.59	-1.805*
weaselzippers.us	100K–500K	0.694	1.613	-81.45	-1.829
toddstarnes.com	100K–500K	0.607	1.456	-72.43	-1.499
yournews.com	500K–1M	0.350	1.363	-128.56	-2.338
nationalfile.com	500K–1M	0.432	1.321	-72.62	-1.410
ussanews.com	500K–1M	0.306	1.320	-122.24	-2.723
infostormer.com	1M–5M	**0.764**	**2.451**	**-155.62**	**-4.332**
projectveritas.com	1M–5M	0.626	1.837	-64.70	-1.162
defconnews.com	1M–5M	0.597	1.700	-81.09	-1.404
libertyunyielding.com	5M–10M	0.412	1.415	-73.19	-1.225
redstatenation.com	5M–10M	0.641	1.395	-92.20	-1.433
thejeffreylord.com	5M–10M	0.581	1.351	-65.01	-1.182
thefreedomtimes.com	10M–50M	0.548	1.586	-61.65	-1.237
presscorp.org	10M–50M	0.555	1.531	-84.19	-1.469
trueviralnews.com	10M–50M	0.276	0.787	-56.26	-0.780

TABLE 5: We present the weighted average change (and effect-sizes) in the number of external articles that are published by a random subset of 100 external domains in the week after the website publishes the narrative (*i.e.,* articles not written by the origin domain) and the average change in time (and effect-sizes) for a story to peak in popularity when a website originates a narrative. We utilize the U-Mann Whitney test for significant differences in the means. After applying the Bonferroni correction, we conclude that a value is significant if the p-value is < 0.0017 (*i.e.*, 0.05/29). We star values that are *not significant*.

Domain	CrUX Rank	Wtd. Ext. Art.Δ	Cohen's D	To Peak Δ (Days)	Cohen's D
dailymail.co.uk	< 1K	0.652	1.741	-13.19	-0.101
express.co.uk	< 1K	0.493	0.740	-14.71	-0.174
nypost.com	1K–5K	0.840	1.684	-15.20	-0.072
breitbart.com	1K–5K	0.841	1.682	-20.61	-0.172
zerohedge.com	1K–5K	0.564	1.195	-23.59	-0.254
thegatewaypundit.com	5K–10K	0.788	1.630	-20.03	-0.168
newsmax.com	5K–10K	0.752	1.241	-24.90	-0.276
rawstory.com	5K–10K	0.540	1.018	-15.04	-0.184
redstate.com	10K–50K	1.105	1.878	-20.12	-0.166
twitchy.com	10K–50K	1.124	1.735	-17.25	-0.088
democraticunderground.com	10K–50K	1.067	1.699	-19.19	-0.176*
rumormillnews.com	50K–100K	0.160	1.995	-10.83	0.132*
beforeitsnews.com	50K–100K	0.859	1.980	-28.75	-0.286
justthenews.com	50K–100K	0.865	1.624	-19.58	-0.166
–	–				

Domain	CrUX Rank	Wtd. Ext. Art.Δ	Cohen's D	To Peak Δ (Days)	Cohen's D
populistpress.com	100K–500K	1.511	1.894	**-44.28**	-0.428*
henrymakow.com	100K–500K	1.100	1.885	-19.80	-0.208*
sgtreport.com	100K–500K	0.663	1.743	-28.27	-0.326
ussanews.com	500K–1M	1.106	2.078	-8.30	-0.111*
yournews.com	500K–1M	1.161	1.746	-40.09	**-0.448**
barenakedislam.com	1M–5M	**1.972**	**2.189**	-23.84	0.121*
americafirstreport.com	1M–5M	1.030	2.135	-15.57	-0.142*
conservativeangle.com	1M–5M	1.353	1.995	-20.99	-0.285*
patriotjournal.org	5M–10M	1.590	2.035	-17.83	-0.083*
legitgov.org	5M–10M	1.521	1.866	-21.89	-0.220*
gopdailybrief.com	5M–10M	1.090	1.706	-38.67	-0.411
trueviralnews.com	10M–50M	0.748	1.375	-23.43	-0.280
roguereview.net	10M–50M	0.868	1.290	-19.95	-0.217*
thefreedomtimes.com	10M–50M	0.745	1.127	-12.58	-0.023*

TABLE 6: We present the weighted average change (and effect-sizes) in the external articles that are published by a random subset of 100 external domains for a given domain's amplified narratives (*i.e.,* articles not written by the origin domain) and the average change in time (and effect-sizes) for a story to peak in popularity when a website amplifies a narrative. We utilize the U-Mann Whitney test for significant differences in the means. After applying the Bonferroni correction, we conclude that a value is significant if the p-value is < 0.0017 (*i.e.*, 0.05/29). We star values that are *not significant*.

external articles that are written in the week after origination if the website did not originate the narrative but still eventually wrote about that narrative. We note that for this analysis, we weight the number of articles by the log inverse of its CrUX popularity ranking [76], [77] to ensure that we do not consider an article from a highly popular website such as breitbart.com the same as from a relatively obscure website such as welovetrump.com.

To ensure each website has a marked effect on the full unreliable news ecosystem and improve the robustness of our approach, we further utilize a bootstrapping procedure ($B = 250$) to measure the influence of each website by taking a random subset of 100 websites in each bootstrap and then measuring the weighted increase in the number of articles across this set of a random set of 100 websites [78]. We provide the average effect size (Cohen's D) and the p-value for the change in the number of external articles in Table 5. For this section, we limit our analysis to websites that consistently originate articles throughout our study period by only considering websites with at least 25 instances of *originating* an article.

We observe a small correlation (Pearson correlation $\rho = 0.213$) between a website's popularity and its ability to originate and perpetuate narratives amongst other unreliable news websites. For example, considering express.co.uk and dailymail.co.uk, two tabloids known to engage in sensationalism and biased reporting with the highest CrUX popularities, while dailymail.co.uk is fairly effective at originating narratives (Cohen's D = 0.334), express.co.uk is one of the worst at originating new narratives (Cohen's D = 0.007). Furthermore, a seemingly unpopular website, infostormer.com, is one of the best at propagating narratives it originates to other sites. Infostormer.com, with a header labeled the *"Jewish Problem"*, writes heavily sensationalist and antisemitic perspectives on the news that is taken up by other websites. For example, after writing an article on how the CNBC news host Jim Cramer was promoting Meta stock,[6] this news story was later covered by more popular websites like activistpost.com[7] and hannity.com.[8]

In addition to quantifying each website's efficacy in originating narratives, we determine how quickly after a website

6. http://web.archive.org/web/20221028232854/https://infostormer.com/jew-jim-cramer-cries-and-apologizes-for-hyping-metas-stock/

7. https://web.archive.org/web/20221028014725/https://www.activistpost.com/2022/10/the-big-tech-companies-are-telling-us-exactly-where-the-economy-is-headed-in-2023.html

8. http://web.archive.org/web/20220901000000*/https://hannity.com/media-room/sad-money-jim-cramer-in-tears-after-meta-stock-nosedives-i-made-a-mistake/

originates a narrative that the story peaks in popularity. Here, a negative Cohen's D indicates that the "time for a narrative to peak" occurs faster. This metric, combined with the previous metric, describes how effective a website is at reorienting online conversations to its own narratives. We see only a slight correlation ($\rho = 0.125$) with a website's CrUX-defined popularity. Rather, we again see that many smaller websites are highly effective in originating narratives that peak quickly (*i.e.*, writing about narratives that become of immediate interest). For example, we see that when the small website infostormer.com originates narratives, those narratives peak in popularity 155 days earlier than when infostormer.com does not originate narratives (Table 5). We similarly observe that the right-wing and conspiratorial website ussanews.com, which had often written about QAnon [5], is also highly effective at quickly landing its narratives on other websites, with one of the lowest "time to peak" in our dataset.

5.2. Amplifying Narratives

To understand how effective websites are at amplifying narratives, we correlationally compare the number of external non-origin articles that are written about a given narrative when a given website amplifies the narrative versus when the website does not amplify the narrative but still eventually wrote about that narrative. We utilize the same weighting and bootstrapping procedure as in the previous section, again limiting our analysis to websites that amplify at least 25 narratives across our period of study. We again observe only a slight correlation between a website's popularity and its ability to amplify narratives (Pearson correlation $\rho = 0.279$). As before, we see in Table 6 that websites across different CrUX popularities excel at amplifying narratives. One of the most effective websites is barenakedislam.com, an anti-Islam website with the slogan *"It isn't Islamaphobia when they really ARE trying to kill you."* For example, after echoing a narrative about how Muslim men were targeting Ukrainian refugees,[9] this news story traveled to an additional 12 other unreliable news websites including more popular websites like americanthinker.com[10] and breitbart.com.[11]

Finally, we determine the effect of narrative amplification by each website on how quickly the narrative peaks. There is again a small correlation between website popularity and amplification (Pearson correlation $\rho = 0.169$). As seen in Table 6, the most effective website at quickly amplifying narratives to their peak popularity is populistpress.com, a drudge-style news website that hosts hyperlinks to different news articles. Despite not hosting many articles itself, we see that when it does mention a narrative, this news story is more likely to peak in popularity sooner.

9. https://web.archive.org/web/20220422091130/https://barenakedislam.com/2022/03/21/what-a-surprise-not-ukrainian-female-refugees-say-they-dont-feel-safe-in-multicultural-sweden/

10. http://web.archive.org/web/20220619083901/https://www.americanthinker.com/articles/2022/06/the_only_rape_where_the_left_says_victimblaming_is_okay.html

11. https://web.archive.org/web/20220527122734/https://www.breitbart.com/europe/2022/05/27/sweden-asylum-home-tells-ukrainian-women-dress-modestly-to-not-provoke-migrant-men/

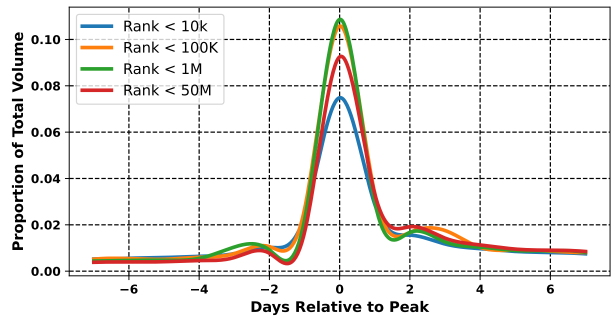

Figure 6: Time lag for differently ranked websites. The most popular websites write more of their articles prior to the peak of a narrative's popularity. In contrast, less popular websites tend to respond to narratives and write most of their articles after the peak.

Social Media Platform	Posts	Posts With Corresponding News Article Narrative
8kun.top	632,091	35,227 (5.57%)
4chan.org	4,690,669	455,327 (9.71%)

TABLE 7: Our dataset of social media posts and their relationship to the narratives published by unreliable news websites.

Trend Setting. Despite not seeing clear discernible patterns in how website popularity correlates with a website's ability to originate and amplify narratives, we do observe differences in *when* these websites write about given narratives. As seen in Figure 6, across all narratives, more popular websites tend to write fewer articles on the day that a given narrative peaks. A slightly higher percentage (33.0%) of articles from websites with a CrUX rank <100K come before the peak versus the 28.76% of articles with a rank above 1M. This indicates, as also found by Leskovec et al. [17], that across all narratives, popular websites have *some* ability to set the agenda for what smaller websites write.

5.3. The Role of Fringe Forums/Social Media

We now analyze the relationship between our set of unreliable news websites and the fringe social media sites 8kun and 4chan. As with our set of unreliable news websites, we scrape 8kun and 4chan `/pol` posts published between January 1 and November 1, 2022 (Section 3.4). 8kun data is readily available on their website 8kun.top; 4chan `/pol` posts are archived through the website archive.4plebs.org. Altogether, we gather 632,091 posts from 8kun.top and 4.69 million posts from 4chan (Table 7).

To find the correspondence of 8kun and 4chan comments between news narratives, we preprocess, embed, and assign each 8kun post to its most similar narrative cluster. As before, we utilize a threshold of 0.60 for matching a comment to its corresponding news article narrative. Altogether, we find that 35,227 (5.57%) of comments on 8kun.top and

Top Narratives on 8kun	Comments
ukraine, nato, putin, russia, war	625
hillary, collusion, clinton, mueller, lie	454
antisemit, jew, israel, zionist, israel	373
trucker, trudeau, canadian, ottawa, convoy	365
ukraine, kyiv, troop, russia, donbas	286
Top Narratives on 4chan	**Comments**
ukraine, kyiv, troop, russia, donbas	4,361
volodymyr, zelenskyy, ukrainie, kyiv, president	3,003
ukraine, conflict, war, escalation, tension	2,792
race, white, theory, black, crt	2,134
jew, white, supremacy, goyim, zionist	2,079

TABLE 8: The top topics from our set of unreliable news websites present on 8kun and 4chan /pol.

Platform	Narratives Originated	Wtd. Ext. Art.Δ	Cohen's D	To Peak Δ (Days)	Cohen's D
8kun	392	0.416	**1.781**	**-72.94**	**-1.149**
4chan	**2,432**	**0.438**	0.961	-28.60	-0.445
Platform	**Narratives Amplified**	**Wtd. Ext. Art.Δ**	**Cohen's D**	**To Peak Δ (Days)**	**Cohen's D**
8kun	2,695	**0.903**	0.267	**-20.28**	**-0.372**
4chan	**12,184**	0.370	**0.938**	0.646	0.006*
Platform	**JS-Sim. to News**	**Most Similar News Sites**			
8kun	0.242	lucianne.com, thebulwark.com, radiopatriot.net			
4chan	0.250	unz.com, beforeitsnews.com, thetruthseeker.co.uk			

TABLE 9: The influence of 8kun and 4chan on the ecosystem of unreliable news websites and their similarity (by JS-Divergence) to the unreliable news dataset. We star values that are *not significant*, according to the U-Mann-Whitney test.

455,327 (9.71%) of comments from 4chan.org correspond to a narrative on our set of unreliable news websites (Table 7).

5.3.1. 8kun. Examining the top narratives posted on 8kun that corresponded with a news narrative (Table 8), we see that the most commonly shared narratives on 8kun.top concern the Russo-Ukrainian war, the investigation of Donald Trump by Special Counsel Robert Mueller [79], antisemitic beliefs, and the 2022 Trucker Convoy in Ottawa Canada [80]. This largely corresponds with 8kun being known as the home of hard-right, conspiratorial, and antisemitic posts [81]. As in Section 4.3, we determine the distribution of narratives from our set of unreliable websites that are present on 8kun to understand the similarity between 8kun and the collective narratives on our set of unreliable news websites. Altogether, 8kun has a JS-Divergence of 0.242 with the collective narrative distribution of unreliable news websites (Table 9). Performing this on an individual site level, we observe several of the websites with the most similar narrative distributions prominently discuss conspiratorial ideas (*e.g.*, lucianne.com and radiotpatriot.net) [5].

Having examined similarities between the narratives discussed on 8kun and those on particular websites in our dataset, we next determine the influence of 8kun on our unreliable news website ecosystem. We utilize the same definitions of *originate* and *amplify* as well as the same methodology as in Sections 5.1 and 5.2. As seen in Table 9, 8kun originating or amplifying a particular narrative has a modest effect on the number of articles written about that narrative. In the week after 8kun originates a given narrative, we see an average Cohen's D of 1.781. In contrast, in the week after 8kun users amplify a narrative, we observe a Cohen's D of 0.267, illustrating that 8kun is somewhat better at originating narratives than amplifying narratives. Thus, while not as effective as some of the websites in our dataset (Tables 5 and 6), when 8kun users comment on narratives, this correlates with a slight increase in the narrative's popularity. We see this further mirrored in the effect that 8kun has in expediting narratives to peak earlier. On average, if 8kun originates a narrative, it peaks in popularity 72.4 days earlier than if 8kun did not originate the narrative. Similarly, if 8kun amplifies a narrative it peaks in popularity 20.8 days earlier on average.

5.3.2. 4chan /pol. Looking at the top corresponding shared narratives on 4chan /pol, we see several that target Judaism and the Jewish people (Table 8). As with 8kun, 4chan has a reputation for antisemitism and racist language. Besides the narratives that center on the Russo-Ukrainian war, we see this racism and antisemitism reflected in the top shared narratives on the website [82]. Determining the distribution of narratives from our set of unreliable news websites that are present on 4chan, altogether, 4chan has a JS-Divergence of 0.250 with the collective narrative distribution of unreliable news websites (Table 9). Examining the most similar websites to 4chan, we observe several with known conspiratorial reputations. As documented by Medias-Bias/FactCheck, the most similar website to 4chan, unz.com is a conspiratorial and hate-oriented website that often cites white nationalist groups in its articles [83]. Similarly, beforeitsnews.com [84] and thetruthseeker.co.uk [85] are known to "promote conspiracy theories and pseudoscience."

Finally, we determine the role 4chan has in promoting and amplifying narratives within our ecosystem of unreliable news websites. We observe a similar effect to 8kun, in terms of the weighted increase of articles when 4chan originates a narrative compared to when it does not (Cohen's D of 0.438). However, unlike 8kun, we do observe that 4chan is better at amplifying narratives, with an effect size of Cohen's D of 0.938 (Table 9). However, in contrast to 8kun, 4chan is relatively less effective at getting the narrative to peak earlier. If 4chan users originate a narrative, it peaks in popularity 28.6 days earlier compared to 72.9 days earlier when 8kun originates a narrative. When 4chan amplifies a narrative, it has little effect on when that narrative peaks.

6. Detecting Narratives and Fact-Checking

In the last two sections, we analyzed the narratives and behavior of unreliable news websites during 2022. In this section, we present two case studies that highlight how our programmatic approach can also identify new narratives and assist in focusing fact-checking efforts.

6.1. Identifying New Trending Narratives

By examining the week-over-week percentage increases in story volumes, we programmatically determine which narratives are receiving new or renewed focus on unreliable news websites, which is imperative for ameliorating the spread of specious information [9], [10], [11]. The narratives that increased most in volume during the last week of our experiment (October 26 to November 1, 2022) were:

The Attack of Paul Pelosi, Keywords: Pelosi, Depap, hammer, Nancy, Paul. On October 28, 2022, the husband of Congresswoman Nancy Pelosi was attacked in his home [86]. Largely due to the proximity of the time of the attack to the 2022 US midterm elections, the attack became a source of conspiracy theories and wild speculation. For example, one user wrote on thegatewaypundit.com:[12]

> *Whenever bad things happen to Paulie P, his wife always manages to have an alibi.*

364 articles (compared to zero the week before) were written about the event within our dataset across 175 websites. The thegatewaypundit.com had 44 articles, dailymail.co.uk.com had 40, and nypost.com had 37.

The Seoul Halloween Stampede, Keywords: Halloween, Seoul, Itaewon, Festivity, Stampede. On October 29, 2022, a crowd rush in the Seoul neighborhood of Itaewon resulted in the death of 158 people. Across our dataset, we see 132 articles across 46 websites written about this event, with 18 articles from dailymail.co.uk, 11 from republicworld.com, and 8 from mirror.co.uk.

Elon Musk's First Visit to Twitter Headquarters, Keywords: Sink, Headquarters, Musk, Carry, Twitter. After officially purchasing the social media company Twitter, on his first visit to the company on October 26, Elon Musk carried a sink into the headquarters with him. This prop humor by Musk was supposed to be a play on "let that sink in" but with a real sink. We see 176 articles from 84 domains about the story, with 10 articles from the dailymail.co.uk, 9 from westernjournal.com, and 8 from conservativeangle.com.

6.2. Fact-Checking

One approach to combating the spread of new misinformation stories that many organizations have adopted is fact-checking. Fact-checking a story requires hours to deeply understand its context and nuance [87]. Unfortunately, this means that propaganda and misinformation often spread significantly in a rapidly evolving media landscape before journalists can respond. Our approach can serve as a way to programmatically identify new misinformation narratives as they appear and begin to gain traction, ideally curbing the amount of time from when a story is published to when a fact-checker can respond.

To show how our system might be useful to fact-checking organizations, we utilize our approach to analyze the behaviors of particular narratives before being fact-checked by three organizations: Politifact, Reuters, and AP-News [88], [89], [90]. For the three agencies, we gathered the set of fact-checking articles that each published in 2022. Altogether we scraped 1,524, 3,090, and 140 articles from Politifact [91], Reuters [92], and APNews [93], respectively. To augment our system to perform fact-checking (i.e., determine whether a fact-check article *refutes* a given narrative), we additionally train a DeBERTa-based [94] classifier on the FEVER [95] dataset that takes a claim (*i.e.*, an article) and a query (*i.e.*, a fact-check) and labels the query as either *supporting* the claim, *refuting* the claim, or *not having enough information* to say anything about the claim. Using 10% of the FEVER dataset as a held-out test set, our DeBERTa-based model achieves an overall 90.7% accuracy on this test set (90.5% precision in labeling refutations).

	Narr. Fact-Checked	Med. Art. Prior Fact-Check	Med. Days to Fact-Check	Med. Days from Narr. Peak	0-Day Fact-Checks
Politifact	6,231	6	55.0	4.0	110
Reuters	9,604	3	49.0	0.0	647
AP News	230	15	83.0	3.0	8

TABLE 10: Efficacy of fact-checking websites. All three websites most commonly fact-check—by the number of articles with the same narrative as the fact-check—the articles of resear.ch (433 articles), gatesofvienna.net (414), dailymail.co.uk (406).

For each fact-checking article, as with articles from unreliable websites, we divide the article into its constituent passages and embed them utilizing our MPNet model. We consider a narrative to have been addressed by a fact-checker if the fact-checker writes about the narrative. We note that articles frequently "fact-check" or "add context" to multiple narratives. To provide fact checkers with the greatest number of "opportunities" to fact-check a narrative, we map each fact-check passage to *all* articles above our cosine similarity threshold of 0.60 rather than map the fact-check passage to only the single closest narrative. After mapping these fact-checking passages to our set of articles, we utilize our DeBERTa-based fact-checking classifier to ensure that the corresponding "fact-check" *refutes* the information of the corresponding unreliable news article passage. We provide an example of an identified fact-check below. To ensure that our model is able to properly identify "fact-checks", we manually validate 100 random fact-check-article refutation pairs, finding that 94% of them are indeed refutations.

Article Passage: This is a shining example and a small part of why it is so vitally important to find the underlying cause of the fraud that took place both in November 2020 and the lead-up to that election.

Fact-Check: THE FACTS: To be clear, no widespread corruption was found and no election was stolen from Trump.

As seen in Table 10, on average, narratives can spread one to two months before being fact-checked by these reputable websites. Furthermore, on average both Politifact and AP News write about stories after they have peaked in popularity; AP News, with the fewest fact-checks, writes about narratives right as they peak in popularity. We further see a heavy overlap between the narratives that each

12. https://web.archive.org/web/20221028130434/https://www.thegatewaypundit.com/2022/10/breaking-pelosis-home-broken-early-morning-san-francisco-paul-pelosi-violently-beaten-taken-hospital/

website fact-checks. Reuters and Politifact have an overlap of 2,646 stories/narratives; AP News and Politifact, have an overlap of 137 narratives; and Reuters and AP News have an overlap of 141 narratives. Furthermore, the unreliable websites that have articles most commonly fact-checked by the three fact-checking organizations are the same: qresear.ch (433 articles), gatesofvienna.net (414), and dailymail.co.uk (406). This underscores that these fact-checking websites are duplicating effort, often fact-checking the same narratives [96]. We note, however, that while narratives often spread for long periods before being fact-checks, the number of articles, on average, is often low (3–15 articles). We thus see that many fact-checkers *are* effective at fact-checking narratives when they peak in popularity, but often understandably do not fact-check narratives that have just begun to spread among different unreliable news websites.

We thus see that many narratives spread for long periods on unreliable news websites before they are fact-checked near their narrative peak. However, our system can surface these narratives to fact-checkers long before they peak in popularity, aiding in the fact-checking organizations' typical workflow. This can enable fact-checkers to identify and address misleading narratives concurrent to when they first rise in popularity.

7. Related Work

Our study builds on considerable prior work on both the spread of misinformation online and language models. There have been several past quantitative studies of the spread of information online. Leskovec et al. identify the trends in the propagation of "memes" [17]. They find that while the majority of memes originate from mainstream websites, key phrases that start on smaller blogs are often adopted by larger platforms. Gomez-Rodriguez et al. adopt a cascade transmission model and identify how best to estimate the relative influence of different news outlets in spreading stories [97]. Similar to our use of DP-Means, prior works have utilized CluStream among other clustering techniques to track information or news over time [98], [99], [100]. For example, Curiskis et al. [101] utilize document clustering based on dictionaries to track topics.

Analyzing the Spread of Misinformation. Several works have tracked the spread and impact of misinformation. Shu et al. [102] present the largest overall overview of misinformation detection issues, presenting various paradigms for tracking and labeling misinformation. These include tracking news content features and social content features. For example, Cao et al. [103] and Meel et al. [104] explore utilizing image and text-based features to label misinformation. Abdali et al., in contrast, use screenshots of websites to identify the trustworthiness of websites and label misinformation [105]. Extensive work has studied individual campaigns that spread unreliable information, on topics like QAnon [5], [47], Syrian White Helmets [106], the Russo-Ukrainian War [37], [63], and COVID-19 [107].

Recent work from the security community has focused on identifying and curbing misinformation. Kaiser et al. [13] studied how borrowing techniques from the security warning landscape might help to inform mis/disinformation warnings. Paudel et al. [14] recently demonstrated how techniques like Learning To Rank (LTR) can be used to soft-moderate misinformation on Twitter. On the human-level, Sharevski et al. identified folk models of misinformation on social media that could inform potential defenses [108].

Language Models, Semantic Search, and Topic Analysis. Many previous topic analysis methods have been built on Latent Dirichlet Allocation (LDA). Albalawi et al. show that LDA is one of the most effective methodologies for extracting topics from short text data compared to other computationally light alternatives proposed within the last decade (*e.g.*, LSA, LDA, NMF, PCA, RP) [109]. Meng et al. [110], Angelov [111], and Grootendorst [61] have enabled users to perform topic modeling utilizing large language models. Utilizing these techniques and online document clustering [112], [113], others have performed robust, but smaller scoped semantic analysis (e.g., on Russian disinformation campaigns [37], [50]).

8. Discussion and Conclusion

In this work, we introduced and validated a new, scalable methodology for tracking news narratives online. Applying the methodology to study the stories published on 1,404 unreliable news websites during 2022, our work shows how a large-scale, quantitative analysis can identify propagation patterns and significant players that may otherwise have been difficult to uncover through qualitative investigations of individual disinformation campaigns. Specifically, we showed that less frequented websites and fringe social media platforms can have marked effects on amplifying the narratives discussed on unreliable news websites.

Our study also highlights the need to programmatically detect the rise of false narratives in real time. Prior work has shown that misinformation can spread ten times faster than legitimate news [6] and our analysis finds that false narratives can often start on small, seemingly unpopular websites. In many cases, these false narratives spread for months online before being fact-checked.

While our study illustrates the potential for programmatically tracking news narratives, it also simultaneously surfaces areas for further research. For example, as found in past works [37], [114], though our approach can identify precise stories/narratives within our dataset, medical information poses challenges for large language models like MPNet. For example, one of the misclassifications (Narrative 6 in Table 3) concerned COVID-19 rather than Monkeypox. Given the high knowledge level needed in understanding medical misinformation, past works have recommended utilizing models specifically trained for medical misinformation for topic analysis of this stories [114]. We hope that the potential of and demonstrated need for programmatic approaches for tracking news narratives and misinformation online motivates further work on these topics.

References

[1] Z. Stanton, "You're living in the golden age of conspiracy theories - politico," 6 2020. [Online]. Available: https://www.politico.com/news/magazine/2020/06/17/conspiracy-theories-pandemic-trump-2020-election-coronavirus-326530

[2] A. Bovet and H. A. Makse, "Influence of fake news in twitter during the 2016 us presidential election," *Nature Communications*, 2019.

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

[4] 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.

[5] H. W. Hanley, D. Kumar, and Z. Durumeric, "No calm in the storm: investigating qanon website relationships," in *International AAAI conference on Web and social media*, 2022.

[6] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.

[7] K. Thomas, D. Akhawe, M. Bailey, D. Boneh, E. Bursztein, S. Consolvo, N. Dell, Z. Durumeric, P. G. Kelley, D. Kumar *et al.*, "Sok: Hate, harassment, and the changing landscape of online abuse," in *IEEE Symposium on Security and Privacy*, 2021.

[8] M. E. Zurko, "Disinformation and reflections from usable security," *IEEE Security & Privacy*, vol. 20, no. 3, pp. 4–7, 2022.

[9] 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.

[10] L. Wu, F. Morstatter, K. M. Carley, and H. Liu, "Misinformation in social media: definition, manipulation, and detection," *ACM SIGKDD Explorations Newsletter*, vol. 21, no. 2, pp. 80–90, 2019.

[11] 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 and Privacy*.

[12] 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.

[13] 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.

[14] 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.

[15] S. Zannettou, T. Caulfield, E. De Cristofaro, M. Sirivianos, G. Stringhini, and J. Blackburn, "Disinformation warfare: Understanding state-sponsored trolls on twitter and their influence on the web," in *World wide web conference*, 2019.

[16] F. Plasser, "From hard to soft news standards? how political journalists in different media systems evaluate the shifting quality of news," *Harvard International Journal of Press/Politics*, 2005.

[17] 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.

[18] 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.

[19] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver, "Inside the slammer worm," *IEEE Security & Privacy*, vol. 1, no. 4, pp. 33–39, 2003.

[20] 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.

[21] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxson, and S. Savage, "Spamalytics: An empirical analysis of spam marketing conversion," in *15th ACM conference on Computer and communications security*, 2008.

[22] M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. M. Voelker, "An analysis of underground forums," in *ACM Internet measurement conference*, 2011.

[23] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis *et al.*, "Understanding the mirai botnet," in *26th USENIX security symposium*, 2017.

[24] 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*.

[25] 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.

[26] 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.

[27] 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.

[28] C. Jack, "Lexicon of lies: Terms for problematic information," *Data & Society*, vol. 3, no. 22, pp. 1094–1096, 2017.

[29] S. Jiang and C. Wilson, "Linguistic signals under misinformation and fact-checking: Evidence from user comments on social media," *ACM CSCW*, 2018.

[30] S. Lewandowsky, U. K. Ecker, C. M. Seifert, N. Schwarz, and J. Cook, "Misinformation and its correction: Continued influence and successful debiasing," *Psychological science in the public interest*, vol. 13, no. 3, pp. 106–131, 2012.

[31] H. Allcott, M. Gentzkow, and C. Yu, "Trends in the diffusion of misinformation on social media," *Research & Politics*, 2019.

[32] S. Z. Akbar, A. Panda, D. Kukreti, A. Meena, and J. Pal, "Misinformation as a window into prejudice: Covid-19 and the information environment in india," *CSCW*, 2021.

[33] J. Allan, "Detection as multi-topic tracking," *Information Retrieval*, vol. 5, no. 2-3, pp. 139–157, 2002.

[34] 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.

[35] 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.

[36] G. Leban, B. Fortuna, J. Brank, and M. Grobelnik, "Event registry: learning about world events from news," in *Proceedings of the 23rd International Conference on World Wide Web*, 2014, pp. 107–110.

[37] 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.

[38] 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.

[39] C. Crawl, "Common crawl," 10 2022. [Online]. Available: https://commoncrawl.org/

[40] W. Min, B.-K. Bao, C. Xu, and M. S. Hossain, "Cross-platform multi-modal topic modeling for personalized inter-platform recommendation," *IEEE Transactions on Multimedia*, 2015.

[41] R. Bal, S. Sinha, S. Dutta, R. Joshi, S. Ghosh, and R. Dutt, "Analysing the extent of misinformation in cancer related tweets," in *International AAAI Conference on Web and Social Media*, 2020.

[42] Barret Golding, "Iffy index of unreliable sources," https://iffy.news/index/, 2022.

[43] M. Szpakowski, "Fake news corpus," https://github.com/several27/FakeNewsCorpus/, 2020.

[44] P. Staff, "Politifact's guide to fake news websites and what they peddle," https://www.politifact.com/article/2017/apr/20/politifacts-guide-fake-news-websites-and-what-they/, 2017.

[45] Aloisius Regen, "fake-news," https://github.com/Aloisius/fake-news, 2017.

[46] Melissa Zidmars, "Misinformation and news literacy: Home," https://library.athenstech.edu/fake, 2017.

[47] 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.

[48] 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 *51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2013.

[49] 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.

[50] H. W. Hanley and Z. Durumeric, "Partial mobilization: Tracking multilingual information flows amongst russian media outlets and telegram," *arXiv preprint arXiv:2301.10856*, 2023.

[51] 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.

[52] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in *empirical methods in natural language processing (EMNLP)*, 2014.

[53] 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.

[54] 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," *arXiv preprint arXiv:1803.11175*, 2018.

[55] N. Reimers and I. Gurevych, "Sentence-bert: Sentence embeddings using siamese bert-networks," in *Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, 2019.

[56] A. Rondinelli, L. Bongiovanni, and V. Basile, "Zero-shot topic labeling for hazard classification," *Information*, 2022.

[57] T. Gao, X. Yao, and D. Chen, "SimCSE: Simple contrastive learning of sentence embeddings," in *Empirical Methods in Natural Language Processing (EMNLP)*, 2021.

[58] N. Goel and R. Reddy, "Semeval-2022 task 8: Multi-lingual news article similarity," *arXiv preprint arXiv:2208.09715*, 2022.

[59] D. Vetter, J. J. Tithi, M. Westerlund, R. V. Zicari, and G. Roig, "Using sentence embeddings and semantic similarity for seeking consensus when assessing trustworthy ai," *arXiv:2208.04608*, 2022.

[60] G. Bernard, C. Suire, C. Faucher, A. Doucet, and P. Rosso, "Tracking news stories in short messages in the era of infodemic," in *International Conference of the Cross-Language Evaluation Forum for European Languages*, 2022.

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

[62] O. Dinari and O. Freifeld, "Revisiting dp-means: fast scalable algorithms via parallelism and delayed cluster creation," in *Uncertainty in Artificial Intelligence*. PMLR, 2022, pp. 579–588.

[63] 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.

[64] G. Bouma, "Normalized (pointwise) mutual information in collocation extraction," *Proceedings of GSCL*, 2009.

[65] J. Kessler, "Scattertext: a browser-based tool for visualizing how corpora differ," in *ACL, System Demonstrations*, 2017.

[66] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, "Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," in *58th Annual Meeting of the Association for Computational Linguistics*, 2020.

[67] W. P. Staff, "Read the full opinion in dobbs v. jackson women's health o - washington post," 6 2022. [Online]. Available: https://www.washingtonpost.com/politics/interactive/2022/roe-wade-decision-pdf/

[68] G. E. Center, "Pillars of russia's disinformation and propaganda ecosystem," 2020.

[69] D. Masci, "Most poles accept jews as fellow citizens and neighbors — pew research center," 3 2018. [Online]. Available: https://www.pewresearch.org/fact-tank/2018/03/28/most-poles-accept-jews-as-fellow-citizens-and-neighbors-but-a-minority-do-not/

[70] A. Perrone and D. Loucaides, "A key source for covid-skeptic movements, the epoch times yearns for a global audience - coda story," 3 2022. [Online]. Available: https://www.codastory.com/disinformation/epoch-times/

[71] T. Daigle, "Canadian professor's website helps russia spread disinformation, says u.s. state department — cbc news," 4 2021. [Online]. Available: https://www.cbc.ca/news/science/russian-disinformation-global-research-website-1.5767208

[72] P. Baker and M. Haberman, "In torrent of falsehoods, trump claims election is rigged - the new york times," 11 2020. [Online]. Available: https://www.nytimes.com/2020/11/05/us/politics/trump-presidency.html

[73] F. Nielsen, "On a generalization of the jensen–shannon divergence and the jensen–shannon centroid," *Entropy*, 2020.

[74] X. Que, F. Checconi, F. Petrini, and J. A. Gunnels, "Scalable community detection with the louvain algorithm," in *IEEE International Parallel and Distributed Processing Symposium*, 2015.

[75] 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.

[76] 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 *22nd ACM Internet Measurement Conference*, 2022.

[77] K. Järvelin and J. Kekäläinen, "Cumulated gain-based evaluation of ir techniques," *ACM Transactions on Information Systems*, 2002.

[78] M. A. Little and R. Badawy, "Causal bootstrapping," *arXiv preprint arXiv:1910.09648*, 2019.

[79] M. Zapotosky and S. S. Hsu, "Mueller prosecutor says special counsel 'could have done more' to hold trump accountable," 09 2020. [Online]. Available: https://www.washingtonpost.com/national-security/andrew-weissmann-book-mueller-trump/2020/09/21/6a7967e8-fc10-11ea-b555-4d71a9254f4b_story.html

[80] J. Murphy, "Emergencies act: Us was 'worried' over canada freedom convoy protests," 11 2022. [Online]. Available: https://www.bbc.com/news/world-us-canada-63736482

[81] A. Glaser, "8chan is back as 8kun, but its racist users found other places to go." https://slate.com/technology/2019/11/8chan-8kun-white-supremacists-telegram-discord-facebook.html, 2019, accessed: 2021-04-15.

[82] A. Zelenkauskaite, P. Toivanen, J. Huhtamäki, and K. Valaskivi, "Shades of hatred online: 4chan duplicate circulation surge during hybrid media events," *First Monday*, 2021.

[83] D. V. Zandt, "Media Bias Fact Check," https://mediabiasfactcheck.com/the-unz-report/, 2022.

[84] ——, "Media Bias Fact Check," https://mediabiasfactcheck.com/before-its-news/, 2022.

[85] ——, "Media Bias Fact Check," https://mediabiasfactcheck.com/the-truth-seeker/, 2022.

[86] T. N. Y. Times, "The facts about the attack on paul pelosi, according to prosecutors - the new york times," 11 2022. [Online]. Available: https://www.nytimes.com/article/pelosi-paul-nancy-attack-facts.html

[87] G. Kessler, "Fact-checking movement grapples with a world awash in false claims," https://www.washingtonpost.com/politics/2022/06/29/fact-checking-movement-grapples-with-world-awash-false-claims/, 2022.

[88] D. Zlatkova, P. Nakov, and I. Koychev, "Fact-checking meets fauxtography: Verifying claims about images," in *Conf. on Empirical Methods in NLP and the 9th Intl. Joint Conf. on NLP*, 2019.

[89] B. Nyhan and J. Reifler, "When corrections fail: The persistence of political misperceptions," *Political Behavior*, 2010.

[90] H. Rashkin, E. Choi, J. Y. Jang, S. Volkova, and Y. Choi, "Truth of varying shades: Analyzing language in fake news and political fact-checking," in *Proceedings of the 2017 conference on empirical methods in natural language processing*, 2017, pp. 2931–2937.

[91] Politifact Staff, "Politifact latest fact checks," https://www.politifact.com/factchecks/list/, 2022.

[92] Retuers Staff, "Reuters fact check," https://www.reuters.com/fact-check, 2022.

[93] "AP Fact Check," https://apnews.com/hub/ap-fact-check, 2022.

[94] 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.

[95] J. Thorne, A. Vlachos, C. Christodoulopoulos, and A. Mittal, "Fever: a large-scale dataset for fact extraction and verification," in *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2018.

[96] D. Graves, "Understanding the promise and limits of automated fact-checking," 2018.

[97] M. Gomez-Rodriguez, J. Leskovec, and A. Krause, "Inferring networks of diffusion and influence," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 5, no. 4, pp. 1–37, 2012.

[98] H. Tajalizadeh and R. Boostani, "A novel stream clustering framework for spam detection in twitter," *IEEE Transactions on Computational Social Systems*, 2019.

[99] A. Alsayat and H. El-Sayed, "Social media analysis using optimized k-means clustering," in *IEEE 14th International Conference on Software Engineering Research, Management and Applications*, 2016.

[100] W. Fan, Z. Guo, N. Bouguila, and W. Hou, "Clustering-based online news topic detection and tracking through hierarchical bayesian nonparametric models," in *44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2021.

[101] S. A. Curiskis, B. Drake, T. R. Osborn, and P. J. Kennedy, "An evaluation of document clustering and topic modelling in two online social networks: Twitter and reddit," *Information Processing & Management*, 2020.

[102] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD explorations newsletter*, vol. 19, no. 1, pp. 22–36, 2017.

[103] J. Cao, P. Qi, Q. Sheng, T. Yang, J. Guo, and J. Li, "Exploring the role of visual content in fake news detection," *Disinformation, Misinformation, and Fake News in Social Media*, pp. 141–161, 2020.

[104] P. Meel, H. Agrawal, M. Agrawal, and A. Goyal, "Analysing tweets for text and image features to detect fake news using ensemble learning," in *Intl. Conf. on Intelligent Computing and Smart Communication*, 2020.

[105] S. Abdali, R. Gurav, S. Menon, D. Fonseca, N. Entezari, N. Shah, and E. E. Papalexakis, "Identifying misinformation from website screenshots," *arXiv preprint arXiv:2102.07849*, 2021.

[106] 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.

[107] 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.

[108] F. Sharevski, A. Devine, E. Pieroni, and P. Jachim, "Folk models of misinformation on social media," 2023.

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

[110] 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.

[111] D. Angelov, "Top2vec: Distributed representations of topics," 2020.

[112] 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.

[113] 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 (JACM)*, 2010.

[114] S. L. Isaac, G. Irving, and I. Gabriel, "Ethical and social risks of harm from language models," 2021.

[115] 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.

[116] B. Kulis and M. I. Jordan, "Revisiting k-means: New algorithms via bayesian nonparametrics," *arXiv:1111.0352*, 2011.

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

[118] A. Aizawa, "An information-theoretic perspective of tf–idf measures," *Information Processing & Management*, 2003.

Appendix A. Training with Unsupervised Contrastive Loss

To train our MPNet model, we utilize unsupervised contrastive learning to better the quality of our embeddings [57]. For training, this is such that we embed each example $x_i = (text_i) \in D_{News}$ (where $text_i$ is the text and t_i is whether the text is toxic or not) using a contextual word model by inputting $[CLS]text_i[SEP]$ and averaging the contextual word vectors of the resulting output as a hidden vector $\mathbf{h}_i$ for $text_i$ as its representation. Then, given a set of hidden vectors $\{\mathbf{h}_i\}_{i=0}^{N_b}$, where N_b is the size of the batch, we perform a contrastive learning step on that 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 $\mathbf{h_i}, \mathbf{h_j} \in \mathcal{B}$, 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 such that:

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

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

where, as in prior work [115], we utilize a temperature $\tau = 0.07$.

Appendix B. Passage Pairs

0.45 Similarity

PASSAGE 1: The growing possibility that nuclear weapons might be used, as hostilities in Ukraine continue to escalate, merits your full attention.
PASSAGE 2: Raising the alert level of Russian nuclear forces is a bonechilling development, Guterres declared. The prospect of nuclear conflict, once unthinkable, is now back within the realm of possibility.

0.50 Similarity

PASSAGE 1: When you actually look at the bill and it says no sexual instruction to kids preK through three, how many parents want their kids to have transgenderism or something injected into classroom instruction? DeSantis said earlier this month.
PASSAGE 2: Parents watchdog group Parents Defending Education PDE has warned that a school district in Minnesota is pushing transgender and pride books and materials on to children as young as three years old.

0.55 Similarity

PASSAGE 1: Protests in the Netherlands became violent with police cars being set ablaze as the public grows angry with their enforcement of COVID edicts to restrict their civil liberties:
PASSAGE 2: Thousands of Dutch citizens lined up in the streets defiantly even after government officials banned protest, using the neverending pandemic as an excuse to brutally crackdown on civil liberties.

0.60 Similarity

PASSAGE 1: The raid by over 30 plain clothes agents from the Southern District of Florida and the FBI s Washington Field Office extended through the Trump family s entire 3,000squarefoot private quarters, as well as to a separate office and safe, and a locked basement storage room in which 15 cardboard boxes of material from the White House were stored.
PASSAGE 2: Donald Trump lamented Wednesday that the FBI blocked his lawyers from the property during the raid at his Palm Beach, Florida residence and suggested that agents may have 'planted' evidence.

Figure 7: Example of passage pairs at different levels of cosine similarities.

Appendix C. DP-Means Algorithm

DP-Means [116] 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. [62] 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) [62].

Appendix D. Pointwise Mutual Information

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

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

where P is the probability of occurrence and a scaling parameter α is added to the counts of each word. This scaling parameter α prevents single-count or one-off words in each cluster from having the highest PMI values. Given the scale of our dataset and the number of clusters within our dataset, we determine that a baseline count of 1 (α =1) for each word in the full dictionary in each cluster led to the best results [117].

Appendix E. JS-Divergence

Formally JS-Divergence between two distributions P and Q is calculated as follows:

$$JS(P||Q) = \frac{1}{2} KL(P|| \frac{(P+Q)}{2}) + \frac{1}{2} KL(Q|| \frac{(P+Q)}{2})$$

$$KL(P||Q) = \sum_x P(x) \log(\frac{P(x)}{Q(x)})$$

For our purposes, given that every website does not address every topic, as recommended in other works, we add a small value $\epsilon = 0.1$ to the counts of every website's topics before calculating each website's probability distribution. We utilize this approach rather than cluster-normalized TF-IDF as in other works [37], [61] because class TF-IDF is dependent on document classes being similar in length [61], [118] and the number of articles within each of our clusters varies widely. PMI finds the distinct characteristics of individual clusters and is not dependent on how often words appear in other individual clusters, avoiding this issue.

Appendix F.
Auto-Generated Summaries and Cluster Specificity

Narr.	Keywords	Auto-Generated Summary	Random Sample Passage
1	trudeau, motion, 151, 185, emergency,	The Canadian Parliament voted Monday night to approve Prime Minister Justin Trudeau's motion to invoke the Emergencies Act by a vote of 185 for and 151 against.	On Monday night, Canada's parliament voted to confirm Prime Minister Trudeau's declaration of the Emergencies Act in response to the freedom protests that have swept across the nation for three weeks.
2	manchin, filibuster, schumer, sinema, senate	Republicans and other critics immediately started to wonder: If Democrats extract what they want out of Manchin, couldn't even a small number of them promptly refuse to go along with the secondary assurances he's been promised?	[The pipeline Manchin was promised] would require passage of legislation that would overhaul the permitting process for energy infrastructure, according to The American Prospect, a liberal website. Apparently, progressives in the House are not keen on supporting a measure that could undercut the IRA s down payment on clean energy by accelerating approval for energy projects that could ramp up U.S. fossil fuel production and exports of natural gas, The American Prospect reported.
3	agrawal, musk, parag, ceo, twitter	Twitter CEO Parag Agrawal tweeted he was "excited" that Musk would join Twitter's board after it was revealed that Musk bought a 9.2 percent stake in the company, and in doing so became its largest shareholder.	When Musk's takeover of Twitter became official, Agrawal and Bret gave comments alongside the Tesla CEO.
4	capitol, committe, hearing, select, january	The House of Representatives committee investigating the Jan. 6, 2021, attack on the U.S. Capitol is planning to hold its next hearing on Sept. 28.	The U.S. House of Representatives select committee investigating the deadly Jan. 6, 2021, attack on the Capitol will conduct its next hearing on Oct. 13, the panel said in a statement on Thursday.
5	smith, slap, black, oscar, rock	Apparently, whites can't be outraged by Will Smith s slap without being racist. And never mind that plenty of blacks including Kareem Abdul-Jabbar were also outraged	I elaborated that Will Smith proved he believes violence is the way to handle disagreement. He makes blacks look bad his slap reinforces the widely held stereotype that blacks are violent. He shamed AMPAS before the world.
6	extremist, maryland, walkby, virginia, protest	The group that calls itself Ruth Sent Us announced its plans on a website to harass the justices. It said: Announcing Walkby Wednesday, May 11, 2022!At the homes of the six extremist justices, three in Virginia and three in Maryland.	RSU subsequently announced a WalkBy Wednesday protest on May 11, to be held in front of the homes of the six extremist justices
7	cornyn, boo, convent, texas, gop	U.S. Senator John Cornyn RTexas was loudly booed at the Republican Party of Texas Convention in Houston, where the state GOP adopted a resolution condemning the bipartisan gun control framework he has negotiated in the Senate.	Very loud boos for John Cornyn as he takes the stage at the TexasGOP convention. Cornyn has faced opposition within the party for working with Democrats on a gun package after the shooting in Uvalde.
8	threat, truss, china, uk, britain	K Prime Minister Liz Truss is for the first time due to officially declare China a threat to the UK within days. The designation would be a formal update to former PM Boris Johnson s Integrated Review of Defense and Foreign Policy published in March 2021.	On October 11, The Guardian reported that the Liz Truss government is going to formally designate China a national "threat" to Britain in its upcoming strategic defense review. Under former Prime Minister Boris Johnson, China was named just a "systemic competitor.
9	lithuania, vilniu, baltic, beijing, export	China has called for a corporate boycott of the small Baltic nation. The move is in retaliation for Lithuania's decision to open a Taiwanese representative office in its capital of Vilnius in November 2021.	In a letter last month, the GermanBaltic Chamber of Commerce demanded that Lithuania come to a constructive solution with the communist nation, saying per Reuters: The basic business model of the companies is in question and some will have no other choice than to shut down production in Lithuania.
10	curfew, quebec, province, legault, montreal	Quebec first imposed a COVID curfew on January 9, 2021, which was lifted on May 28. Quebec is the only province in Canada to have imposed a curfew during the pandemic	Quebec is the first province to impose such a system on its citizens. Itś also the only province in Canada that has a curfew in place.

Example of the auto-generated summaries and passages from a set of 10 random narrative clusters to illustrate the specificity and the precision of our approach.