Harvard Kennedy School Misinformation Review[1]
June 2026, Volume 7, Issue 2

DOI: https://doi.org/10.37016/mr-2020-200
Website: https://misinforeview.hks.harvard.edu/

Research Article

Quantifying the "misinformation beat": 38 years of coverage in major U.S. daily newspapers

Media have made misinformation conversations part of daily life. We looked at nearly four decades' worth of news stories about misinformation to see exactly what this coverage looked like. We searched five major U.S. daily newspapers for articles containing the misinformation-related terms—disinformation, misinformation, conspiracy theory, fake news, and propaganda—then extracted words in proximity to these key terms to identify associative patterns. Propaganda *was the dominant term used by major newspapers prior to 2016, when term frequency and variety increased, peaking in 2020. Since 2016, newspaper usage of these terms has focused primarily on Donald Trump, Russia, social media, and U.S. elections.*

Authors: Bryce Greene (1), Brian P. Harper (1), Christena E. Nippert-Eng (1)
Affiliations: (1) School of Informatics, Computing, and Engineering, Indiana University Bloomington, USA
How to cite: Greene, B., Harper, B. P., & Nippert-Eng, C. E. (2026). Quantifying the "misinformation beat": 38 years of coverage in major U.S. daily newspapers. *Harvard Kennedy School (HKS) Misinformation Review*, *7*(2).
Received: August 29th, 2025. Accepted: May 22nd, 2026. Published: June 23rd, 2026.

Research questions

- How has the frequency of misinformation coverage in U.S. newspapers changed over time?
- What topics or entities have been associated with misinformation coverage, and how have these changed over time?
- How are different misinformation-related terms being used?

Essay summary

- We examine how frequently and in what context newspapers use misinformation-related terms.
- Our ProQuest query yielded 196,989 articles written between 1987 and 2024 containing 292,266 instances of misinformation-related terms.
- After we extracted the 50 words surrounding each instance of a term, we ranked and compared them to determine the most common words in proximity to the search term.
- Misinformation-related term use rose dramatically in late 2016; thus, much of our analysis focuses on 2016–2024.
- *Propaganda* and *conspiracy theory* were relatively consistently used from 1987 to 2015. They were then overshadowed by the explosive rise of *misinformation* and *disinformation*.

[1] *A publication of the Shorenstein Center on Media, Politics and Public Policy at Harvard University, John F. Kennedy School of Government.*

- Donald Trump, Russia, and U.S. elections were by far the top focus of misinformation-related coverage throughout the time series.
- Newspapers used misinformation-related terms in distinct yet overlapping contexts: *disinformation* with state actors, predominantly Russia; *misinformation* with reference to social media and COVID-19; *conspiracy theory* with Trump and various right-wing interests (QAnon, January 6th, etc.); *fake news* with Trump and the media in general; *propaganda* with U.S. adversaries (Russia, China, North Korea, Iran).
- This work shows the value of using large datasets that span decades, computational methods of analysis, and a range of misinformation-related terms.

Implications

Journalists and scholars share a sustained interest in the "misinformation beat" (McClure Haughey et al., 2020). To help bring misinformation research in line with current newspaper reporting, we designed a study that uses computational methods to analyze 38 years' worth of stories in five of the six top-circulated U.S. newspapers, focusing on a set of five misinformation-related terms (disinformation, misinformation, conspiracy theory, fake news, and propaganda). Large-scale temporal analyses have proven useful for researchers interested in U.S. newspaper coverage (e.g., Kwak et al., 2020; Wardle & Derakhshan, 2017). Researchers have a well-established interest in the news media's use of misinformation-related terms, too, focusing largely on the single term *fake news* (Egelhofer et al., 2020; Farhall et al., 2019; Farkas, 2023a; Farkas, 2023b; Riebling & von der Wense, 2019); less has been written about the specific uses of increasingly popular terms like *misinformation* and *disinformation*. Of the studies that do examine specific term use, Thorson (2024) includes content analysis of 800 newspaper articles on *mis-* and *disinformation* as part of research into news media effects on perceptions and trust. Farkas and Schousboe (2024) included a diversity of misinformation terms in their analysis of the Danish press, but their work was a self-assessment by journalists rather than an investigation of the text they produced. More expansive analyses that include more terms like *propaganda* and *conspiracy theory* have received fairly little treatment, despite both terms being highly salient concepts in the U.S. media ecosystem (Anderson, 2021).

Our quantitative, multi-term, longitudinal study helps fill an important research gap. We show that, over time, newspapers reflect and reinforce associations between specific misinformation-related terms and an assortment of people, places, and social media platforms. The result is an emergent pattern of distinct yet overlapping contexts in how newspaper reporters have used each of these misinformation-related terms.

While single-term research is valuable, future research that includes multiple misinformation-related terms is also important for capturing the breadth of public discussion. In addition, using natural language processing methods like sentiment analysis, named entity recognition, and grammatical analysis on this dataset would allow for a deeper study of specific aspects of misinformation discourse in U.S. newspapers. Our method of concordance analysis could also be usefully applied to corpora beyond newspapers and the context of the U.S. news media.

Findings

Finding 1: Misinformation-related term usage began sharply increasing in 2016, peaking in 2020.

Figures 1 and 2 are, respectively, a 100% stacked area chart and a compound line graph of our full dataset (1987–2024), allowing us to see both the relative percentage and total number of occurrences for all of our search terms in a given month. In Figure 1, each color corresponds to a given term and the extent of that color along the vertical axis indicates the percentage of that term's usage in that month as a proportion of all five misinformation terms we examined. In Figure 2, the height of each spike corresponds to the number of times that any of these terms are used in a given month. To ensure the accuracy of our interpretation of the raw counts, we also investigated these terms as a proportion of the total words per month in our dataset, contained in our Appendix. Figure A2 shows the same content as Figure 2, but with a normalized line representing all of our terms as a proportion of the total word count per month.

From 1987 until about 2016, the misinformation beat was largely one of reporting on *propaganda*, as seen by the orange portion of both graphs; it remained relatively consistent during that time. In contrast, *fake news* (dark blue) appeared in force in October 2016 and decreased after 2019. The term "misinformation" (light blue) had been used in small amounts prior to 2016 and saw a large and sustained increase afterwards. All terms saw a major spike in usage in 2020. Misinformation-related coverage for the entire 38-year period peaked at 4,753 instances in October 2020.

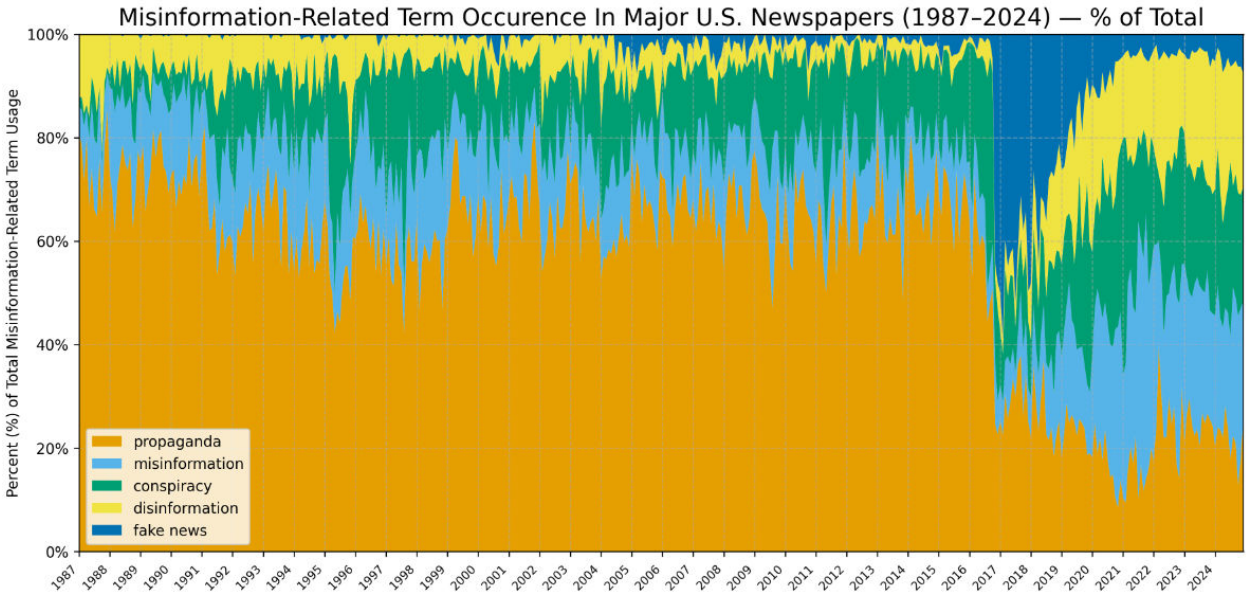

Figure 1. All misinformation-related terms by month from 1987 to 2024 as a percentage of the total number of misinformation-related terms.

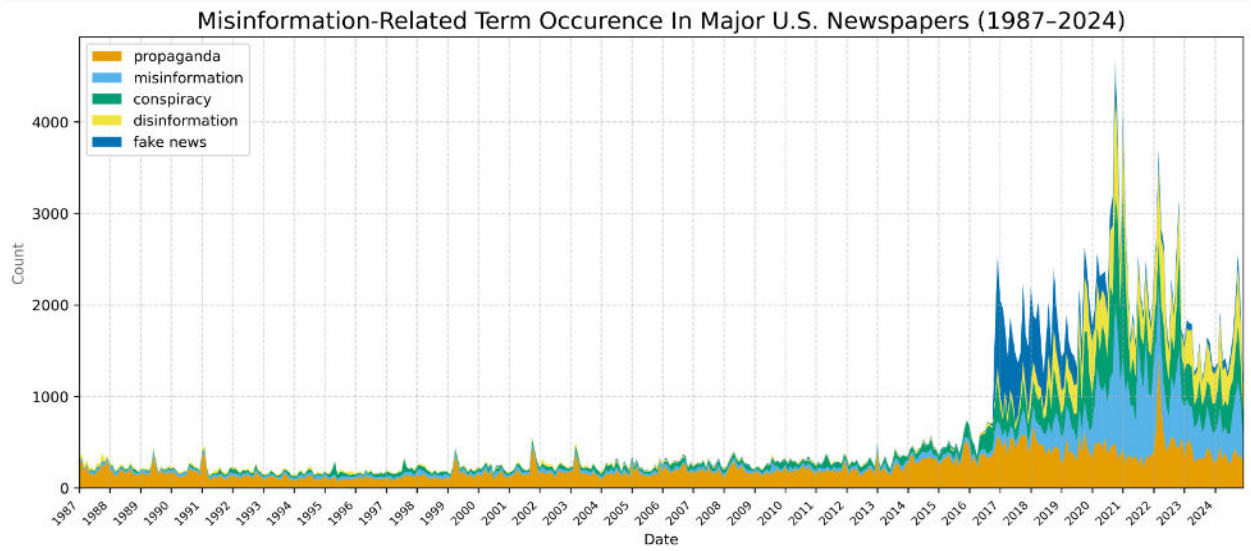

Figure 2. Raw counts of misinformation-related terms by month from 1987 to 2024.

Table 1 provides statistical summaries of our raw term frequencies separated into the periods of 1987–2015 and 2016–2024. From this, we can see that the term *propaganda* was considerably more common in the 1987–2015 period relative to the other terms with a mean of 170.64, more than all the other terms' means combined, a finding also seen in Figure 1. In contrast, the 2016–2024 period sees more balance between the term frequencies. This provides additional support to our decision to focus on the 2016–2024 period, as this period sees more interesting and dramatic differences in the relative use of these terms compared to the previous period.

Table 1. *Descriptive statistics for the occurrences by month for each of our searched terms, separated into the periods of 1987–2015 and 2016–2024.*

	1987–2015			2016–2024		
Term	Mean	Max	*SD*	Mean	Max	*SD*
Propaganda	173.11	512	70.76	428.18	1471	147.18
Misinformation	34.62	114	12.28	4267.56	1535	313.50
Disinformation	14.23	88	9.36	329.66	1219	209.65
Conspiracy theory	36.18	170	23.34	426.56	1535	313.50
Fake news	3.20	28	5.00	261.98	1214	281.24

In Figures 3 and 4, we focus on the period from 2016 to 2024 (i.e., the right side of the graph in Figures 1 and 2). In Figure 3, we see a greater diversity of terms compared to Figure 1, with *fake news* appearing in late 2016 and maintaining prominence until fading in relative use by 2020. The period 2019 onward sees the growing diversity and equity in misinformation term usage, as seen in Table 1. Figure 4 shows that all our searched terms showed a sizable increase in frequency during late 2016 and early 2017, corresponding with the first election of Donald Trump as U.S. President. Several notable spikes follow the peak coverage for all terms in October 2020: the COVID-19 pandemic (beginning March 2020) and vaccine rollout (Summer 2021), the 2020 U.S. presidential election, the Russian invasion of Ukraine (February 2022), the U.S. midterm elections that same year (October), and the 2024 U.S. presidential election, with peak coverage in October.

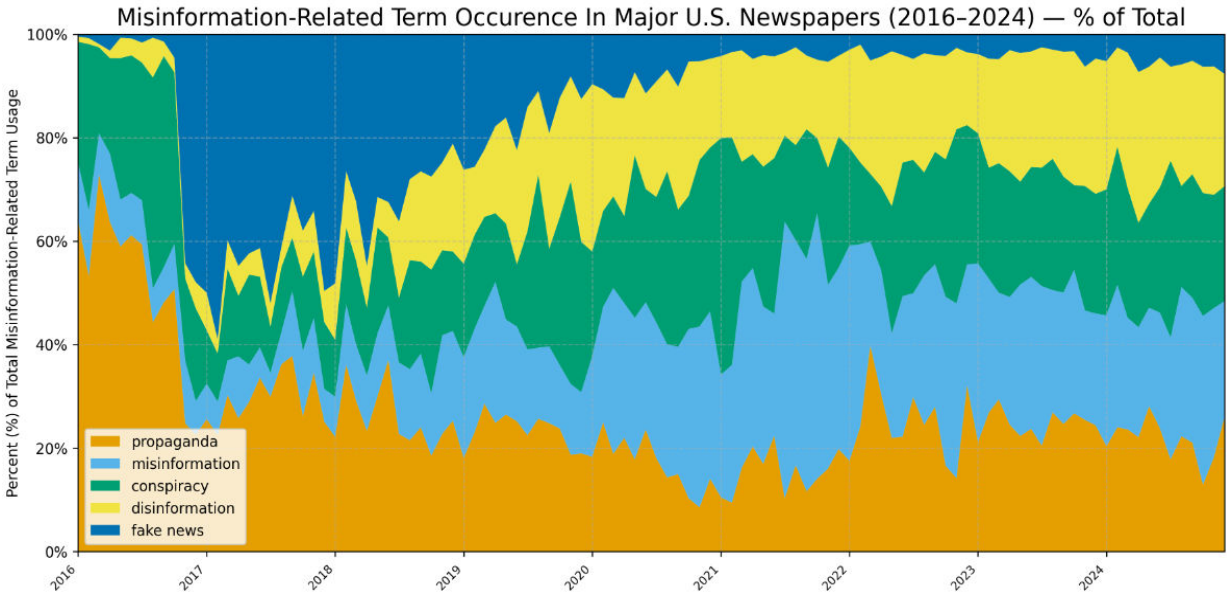

Figure 3. All misinformation-related terms by month from 2016 to 2024 as a percentage of the total number of misinformation-related terms.

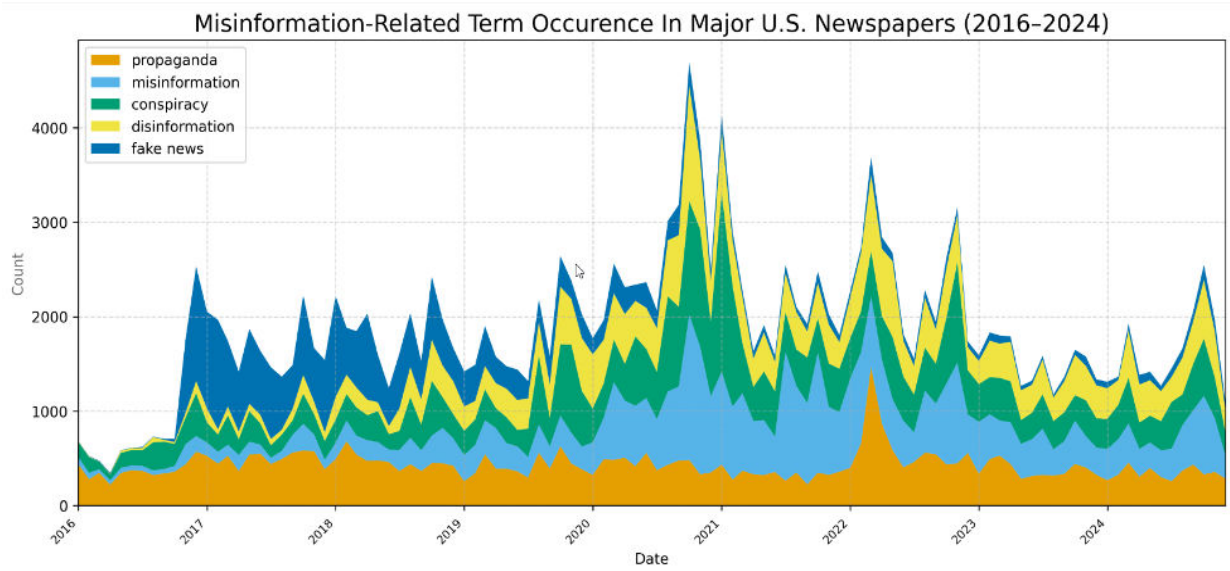

Figure 4. Raw counts of search terms by month from 2016 to 2024.

Finding 2: Trump, Russia, and U.S. elections were the top subjects of misinformation coverage since 2016.

Researchers rapidly identified 2016 as a key inflection point in the emergence and academic coverage of what was then referred to as *fake news* (Allcott & Gentzkow, 2017; Del Vicario et al., 2016; Lazer et al., 2018); our findings confirm this timing. Given this markedly increased volume of reporting on our search terms in 2016, we chose to focus on the period from 2016 to 2024 (the last full year for which we have data) for much of our remaining analysis. During this time, *Trump*, *Russia*, *election*, and *campaign* were among the most frequent terms to co-occur with (i.e., appear within 25 words of) our misinformation-related terms (see Table 2). The relatively high occurrence of these four terms in proximity to our misinformation-related terms was driven in large part by reporting on alleged Russian interference in the 2016 election, further allegations and investigations that Trump colluded with Russia in the election

interference, and Russia's invasion of Ukraine in 2022. *Facebook* was also among the most frequent terms, as the conversation about misinformation also focused on the public effects of consuming misinformation on social media. Thus, during the last ten years, the "misinformation beat" appears to have been focused largely on Trump, Russia, and U.S. presidential elections, including the role of social media in spreading misinformation.

Table 2. *Top words that co-occur with misinformation-related terms, 1987–2024.*

Rank	Word	# Of Occurrences
1	Trump	46,366
2	Russia	37,133
3	election	30,357
4	media	27,759
5	spread	27,043
6	president	26,853
7	news	23,996
8	Facebook	22,808
9	campaign	21,012
10	America	19,338

Table 3. *Annual top 20 words co-occurring with misinformation-related terms, 2016–2024.*

	2016	2017	2018	2019	2020	2021	2022	2023	2024
1	trump	trump	trump	trump	trump	vaccine	russia	russia	trump
2	news	russia	russia	russia	election	trump	election	spread	election
3	facebook	news	news	president	president	facebook	ukraine	media	russia
4	russia	media	facebook	facebook	spread	spread	spread	election	spread
5	president	president	media	campaign	facebook	election	media	government	media
6	media	twitter	president	media	covid19	covid19	trump	trump	president
7	clinton	election	spread	election	media	company	twitter	company	campaign
8	america	campaign	campaign	spread	russia	media	president	news	america
9	campaign	facebook	election	news	campaign	president	company	year	year
10	election	america	twitter	company	twitter	america	covid19	china	government
11	story	story	america	political	china	republican	war	new	official
12	year	fake	company	twitter	america	misinformation	new	use	post
13	state	new	year	2016	company	year	america	president	news
14	new	spread	use	ukraine	news	post	year	campaign	political
15	group	year	new	america	new	group	platform	social	social
16	spread	use	political	new	political	new	official	ukraine	use
17	government	political	fake	government	official	news	state	america	online
18	use	government	government	year	post	social	news	include	company
19	even	states	social	use	social	platform	misinformation	war	new
20	site	information	2016	social	government	campaign	government	platform	claim

Finding 3: Trump overwhelms other U.S. and world leaders.

Figure 5 shows that Trump was associated with misinformation concepts to a greater degree and more consistently than any other domestic leaders and their families (Biden, Obama, Clinton) or international rivals. Notable surges include that of U.S. President Joe Biden in 2019 and Russian President Vladimir Putin in 2022. Chinese President Xi Jinping and Israeli Prime Minister Benjamin Netanyahu hardly registered.

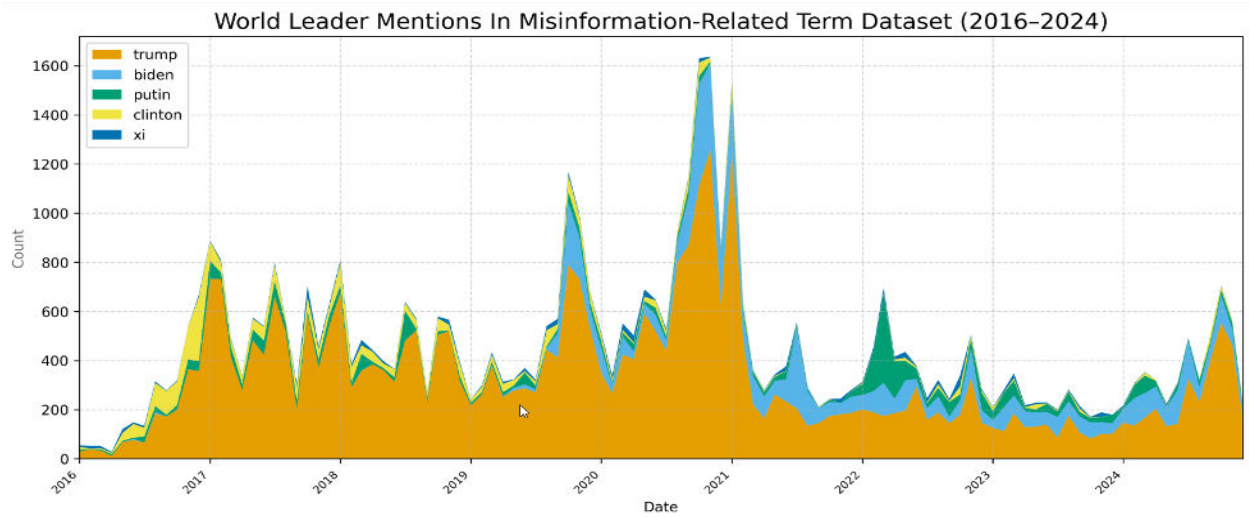

Figure 5. The co-occurrence of the names of world leaders with our search terms by month from 2016 to 2024.

Finding 4: Russia is the consistent frontrunner, followed by China and Ukraine.

Table 3 shows that for seven out of nine years from 2016 to 2024, *Russia* was either the first or second word most frequently associated with our misinformation-related terms. Figure 6 focuses just on non-U.S. countries in relation to our search terms. Throughout the period, *Russia* is a consistent presence, while *China* and *Ukraine* appear with more sporadic coverage associated with specific events, that is, for *China,* COVID-19 in 2020, and for *Ukraine*, the first Trump impeachment in 2019, and the Russian invasion in 2022.

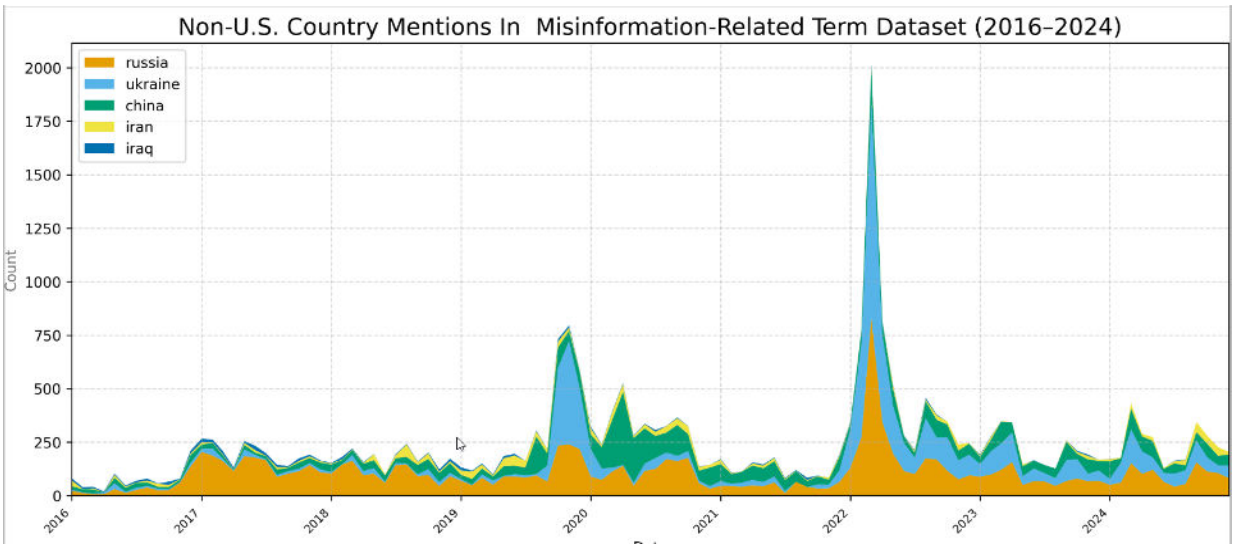

Figure 6. The co-occurrence of the names of non-US countries with our search terms by month from 2016 to 2024.

Finding 5: Facebook dominates other social media platforms, followed by Twitter.

The most common social media platform associated with misinformation-related terms was Facebook (mentions of Meta were negligible), followed by Twitter (see Table 4); YouTube and Instagram were mentioned much less. Figure 7 shows that mentions of Facebook peaked in 2021, then sharply declined as TikTok mentions occurred more frequently, with TikTok often becoming the most commonly mentioned platform.

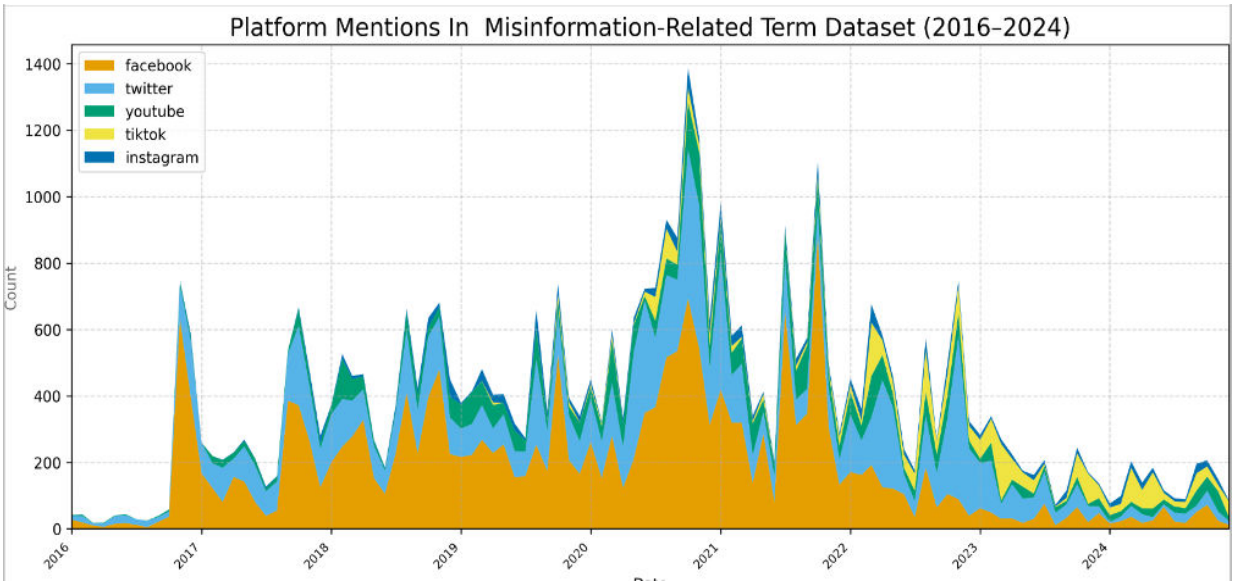

Figure 7. The co-occurrence of social media platform labels with our search terms by month from 2016 to 2024.

Table 4. *The most frequent social media platform occurrences in our combined searches.*

Rank	**Word**	**# Of Occurrences**
1	Facebook	22,808
2	Twitter	18,243
3	YouTube	4,443
4	Tiktok	2,453
5	Instagram	1,664
6	Whatsapp	1,219
7	Telegram	858
8	Reddit	369
9	Snapchat	100

Finding 6: Newspapers had distinct use cases for each misinformation-related term.

In Table 1, we showed that if we look at each search term separately, we see differences in how often the term *propaganda* was used compared to other terms. We can get a more nuanced sense of how newspapers disambiguate each of these terms by individually investigating their associations.

Disinformation: Associated with elections and the Russian foe

Table 5 shows that *Russia* was the most frequent term associated with *disinformation* during 2016–2020. Other top words indicate that this was driven by stories related to allegations of Russian interference in the U.S. election and the first Trump impeachment. Even with the pandemic dominating the discourse of 2020–2021, *Russia* was still the second most frequently associated word in 2021, moving back to the top position in 2022. During peak usage of *disinformation* in 2022 (see Figure 2), top words *Ukraine* and *war* show that Russia's invasion of Ukraine became the dominant driver of coverage. We also see *disinformation* as a term associated with *disinformation*, meaning that newspapers often used the word repeatedly in close proximity to itself. One other notable pattern is how *Facebook* was commonly associated with *disinformation* and *vaccine*. *Twitter* was more commonly associated with *disinformation* and *Ukraine*.

Table 5. *Disinformation time series. The most frequent words for each year associated with disinformation, 2016–2024.*

	2016	2017	2018	2019	2020	2021	2022	2023	2024
1	russia	russia	russia	russia	russia	spread	russia	russia	russia
2	campaign	campaign	campaign	campaign	election	election	ukraine	spread	election
3	media	trump	facebook	facebook	campaign	vaccine	election	campaign	campaign
4	europe	spread	election	election	trump	russia	campaign	media	spread
5	public	election	company	company	spread	campaign	spread	government	trump
6	trump	president	spread	spread	facebook	facebook	disinformation	election	official
7	president	facebook	effort	trump	president	media	media	company	media
8	government	news	2016	president	disinformation	trump	official	use	america
9	use	media	media	disinformation	china	company	misinformation	china	government
10	news	propaganda	twitter	2016	media	disinformation	company	social	online
11	propaganda	america	america	media	america	america	president	disinformation	president
12	spread	intelligence	trump	social	official	president	twitter	platform	effort
13	states	use	account	political	company	group	new	new	disinformation
14	election	company	social	effort	effort	online	government	official	company
15	political	twitter	president	twitter	twitter	effort	america	year	year
16	year	political	new	use	intelligence	social	war	online	china
17	foreign	information	platform	america	covid19	covid19	online	ukraine	intelligence
18	new	government	use	new	social	official	information	news	use
19	military	states	news	government	information	platform	platform	misinformation	social
20	even	official	year	platform	new	year	security	include	platform

Misinformation: Associated with social media, elections, and public health

Table 6 shows that the most frequent words associated with *misinformation* are related to concerns about its impact on the public, particularly regarding social media. Compared to *disinformation*, newspapers used more words related to social media when discussing *misinformation*. The word *Facebook* was the most frequently occurring word from 2016 through 2020 and was second in 2021. In 2021 and 2022, terms like *vaccine* and *COVID-19* rose to high positions. Social media terms remained in high positions in 2023 and 2024 but without a specific named platform.

Table 6. *Misinformation time series. The most frequent words for each year associated with misinformation, 2016–2024.*

	2016	2017	2018	2019	2020	2021	2022	2023	2024
1	facebook	facebook	facebook	facebook	facebook	vaccine	spread	spread	election
2	spread	spread	spread	spread	spread	facebook	election	company	spread
3	news	news	company	vaccine	election	covid19	covid19	media	trump
4	media	company	election	company	covid19	spread	twitter	election	media
5	campaign	election	news	platform	trump	company	company	platform	company
6	election	russia	platform	campaign	company	misinformation	vaccine	covid19	social
7	trump	president	russia	media	twitter	platform	platform	social	post
8	president	campaign	social	social	president	post	media	news	platform
9	even	trump	campaign	political	misinformation	media	misinformation	new	political
10	fake	twitter	media	election	media	election	new	content	online
11	company	information	twitter	trump	post	social	content	twitter	year
12	social	media	year	content	social	health	social	misinformation	president
13	site	platform	2016	news	platform	trump	year	year	vaccine
14	false	fake	online	new	information	twitter	facebook	include	government
15	public	america	new	year	health	information	health	online	news
16	information	social	user	president	new	content	post	vaccine	claim
17	year	political	political	public	campaign	new	public	ai	use
18	new	use	president	information	news	group	information	public	campaign
19	lot	account	use	group	online	year	disinformation	government	misinformation
20	story	post	information	online	vaccine	online	online	speech	health

Conspiracy theory: Associated with Trump and right-wing politics

In Table 7, we see that the phrase *conspiracy theory* was used largely in relation to narratives coming from the right wing of the U.S. political/media spectrum. Most years in this time series include the words *Trump*, *Republican*, and *Fox* (a reference to the Republican-oriented Fox News network). *Trump* was the most frequently co-occurring word with *conspiracy theory* from 2016 to 2024, except for 2022. Coverage of Trump's allegations that the 2020 election was stolen regularly included the phrase *conspiracy theory*, remaining a focus of coverage through 2024. The QAnon movement and the January 6, 2021, attack on the U.S. Capitol by Trump supporters were also a significant focus of reporting in 2021 and 2022.

Table 7. *The most frequent words for each year associated* with conspiracy theory/theories, *2016–2024.*

	2016	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**	**2024**
1	trump	trump	trump	trump	trump	trump	election	trump	trump
2	clinton	president	president	president	qanon	election	trump	election	election
3	president	news	news	ukraine	president	qanon	republican	president	president
4	america	russia	twitter	2016	election	republican	2020	republican	republican
5	campaign	media	political	russia	covid19	president	president	fox	year
6	obama	america	media	election	republican	vaccine	qanon	news	claim
7	republican	new	russia	former	twitter	covid19	america	claim	america
8	year	year	year	political	spread	spread	spread	year	spread
9	news	political	america	biden	america	america	claim	2020	post
10	political	twitter	republican	democratic	claim	claim	state	new	political
11	new	election	conspiracy	republican	media	media	year	spread	media
12	even	clinton	spread	campaign	news	year	media	media	former
13	believe	story	campaign	news	facebook	party	new	america	biden
14	conspiracy	claim	soros	conspiracy	political	capitol	political	former	new
15	media	even	post	america	campaign	new	voting	include	2020
16	election	conspiracy	last	twitter	group	group	official	post	false
17	story	government	even	debunked	new	greene	former	dominion	even
18	twitter	obama	new	investigation	post	political	twitter	political	include
19	government	campaign	facebook	spread	conspiracy	news	group	government	kennedy
20	time	national	claim	white	year	include	include	even	news

Fake news: Associated with Trump and the media

Table 8 shows the shift in use of the phrase *fake news* from a reference to fraudulent news sites deliberately posting false stories to an attack adopted by President Trump and waged against journalists, news outlets, and political leaders he deemed hostile. The latter interpretation is what drove the explosion in coverage after 2016. Interestingly, in 2022, *fake news* started to co-occur with *misinformation* and, in 2023 and 2024, *disinformation*.

Table 8. *The most frequent words for each year associated with* fake news*, 2016–2024.*

	2016	2017	2018	2019	2020	2021	2022	2023	2024
1	facebook	trump	trump	trump	trump	trump	russia	trump	trump
2	news	news	news	media	media	media	trump	news	media
3	story	media	media	president	president	news	media	media	news
4	election	president	president	news	news	president	news	president	russia
5	trump	russia	twitter	twitter	covid19	new	president	spread	post
6	site	twitter	fake	fake	twitter	government	year	post	president
7	fake	story	facebook	facebook	america	covid19	spread	year	america
8	media	fake	story	new	election	post	law	new	year
9	spread	election	russia	story	fake	year	ukraine	government	spread
10	president	cnn	election	times	new	vaccine	war	social	dont
11	real	facebook	new	election	story	election	new	russia	social
12	social	america	year	america	dont	spread	twitter	fake	campaign
13	fake news	new	use	spread	campaign	called	election	called	story
14	company	campaign	cnn	post	spread	america	fake	political	site
15	new	fake news	america	political	china	fake	government	yoon	political
16	false	political	spread	year	virus	political	post	include	use
17	clinton	use	journalist	government	russia	journalist	information	disinformation	election
18	zuckerberg	post	press	russia	use	use	america	use	new
19	post	even	fake news	use	even	first	misinformation	story	government
20	twitter	called	political	even	amp	times	military	information	fake

Propaganda: Associated with state actors and the Russia-Ukraine war

Table 9 shows that *propaganda* was most frequently used in the context of U.S. geopolitical adversaries like the Islamic State, China, or Russia. From 2016–2024, the most frequent term for six out of the nine years was *Russia*. The Islamic state was the dominant focus in 2016, and the pandemic in 2020 and 2021 saw *China* rise to the most frequently co-occurring term. In 2022, the Russian invasion of Ukraine brought *Russia* back to the top.

Table 9. *The most frequent words for each year associated with* propaganda*, 2016–2024.*

	2016	2017	2018	2019	2020	2021	2022	2023	2024
1	russia	russia	russia	china	china	china	russia	russia	russia
2	state	trump	trump	russia	trump	america	ukraine	china	china
3	islamic	america	america	trump	russia	government	war	ukraine	ukraine
4	group	government	media	media	government	russia	china	war	government
5	government	news	china	government	america	media	putin	government	war
6	china	year	news	america	media	year	media	use	media
7	america	media	korea	state	news	use	government	america	america
8	year	state	north	year	president	trump	state	media	use
9	trump	use	government	news	state	new	use	year	state
10	use	group	facebook	president	official	group	news	state	year
11	official	president	use	group	political	official	official	putin	spread
12	media	war	president	official	use	news	america	group	campaign
13	states	campaign	campaign	use	covid19	state	propaganda	news	israel
14	news	states	year	new	campaign	political	president	spread	official
15	new	new	state	war	new	president	new	new	trump
16	war	official	war	campaign	spread	vaccine	year	official	group
17	terrorist	korea	spread	political	states	campaign	spread	military	news
18	islamic state	north	states	spread	war	spread	even	campaign	political
19	campaign	facebook	official	states	year	public	world	israel	president
20	president	islamic	new	even	united	last	campaign	president	election

Methods

We used ProQuest TDM Studio to access the ProQuest U.S. Major Dailies dataset, which includes historical archives of The Chicago Tribune, Los Angeles Times, New York Times, Wall Street Journal, and Washington Post. We searched all articles that used the phrases *misinformation*, *disinformation*, *propaganda*, *conspiracy theory*, *conspiracy theories*, and *fake news* from January 1, 1987, to December 31, 2024.

The data received from ProQuest came in the form of XML files for each of the 196,989 articles that met the criteria. To accurately parse the data, we removed any markup tags to ensure that only the readable text was captured. To clean the dataset and correct for any issues from text identification during scanning, we removed all non-alphanumeric and non-ASCII characters other than hyphens and ending punctuations; question marks and periods were replaced with spaces. Since hyphens are often attached to the keywords listed above, we removed hyphens that preceded or trailed a given keyword.

We then extracted every use of each of these search terms throughout all of the articles from each dataset. We included the 25 words that preceded and followed each instance of each search term to capture its context. Within a single article, multiple occurrences of each search term were counted

separately. We then removed any duplicate title/context combinations. The results of this process are displayed in Table 10.

Table 10. *Total number of term occurrences based on our query for our search terms.*

Term	# of Occurrences
disinformation	40,556
misinformation	58,116
fake news	29,409
propaganda	106,486
conspiracy theory/ conspiracy theories	57,699

In order to answer our research questions, we analyzed the dataset for word frequencies during certain periods. We removed all multi-letter stopwords (e.g., "the," "and," "of") using the Natural Language Toolkit package, augmented with our own list of stopwords developed during data exploration**.** We further cleaned the text using a set of stem mappings developed during exploration (i.e., turning *ukrainian* to *ukraine*, *trumps* to *trump*, or *coronavirus* to *covid19*). We also included a custom list of n-grams based on data exploration to better capture core concepts (i.e., *fake news*, *white house*, *u k*, or *u s s r*). We then removed any remaining single-letter words that were not a part of an n-gram. All data cleaning choices can be seen in our code in the Data Availability section.

Bibliography

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236. http://doi.org/10.1257/jep.31.2.211

Anderson, C. W. (2021). Propaganda, misinformation, and histories of media techniques. *Harvard Kennedy School (HKS) Misinformation Review*, *2*(2). https://doi.org/10.37016/mr-2020-64

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E. & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554–559. http://doi.org/10.1073/pnas.1517441113

Farkas, J. (2023a). Fake news in metajournalistic discourse. *Journalism Studies, 24*(4), 423–441. http://doi.org/10.1080/1461670X.2023.2167106

Farkas, J. (2023b). News on fake news: Logics of media discourses on disinformation. *Journal of Language and Politics, 22*(1), 1–21. http://doi.org/10.1075/jlp.22020.far

Farkas, J., & Schousboe, S. (2024). Facts, values, and the epistemic authority of journalism: How journalists use and define the terms fake news, junk news, misinformation, and disinformation. *Nordicom Review*, *45*(1).http://doi.org/10.2478/nor-2024-0016

Kwak, H., An, J., & Ahn, Y.-Y. (2020). A systematic media frame analysis of 1.5 million New York Times articles from 2000 to 2017. In *WebSci '20: Proceedings of the 12th ACM Conference on Web Science* (pp. 305–14). Association for Computing Machinery. https://doi.org/10.1145/3394231.3397921

McClure Haughey, M., Muralikumar, M. D., Wood, C. A., & Starbird, K. (2020). On the misinformation beat: Understanding the work of investigative journalists reporting on problematic information online. Proceedings of the ACM on Human-Computer Interaction, *4*(CSCW2), 1–22. https://doi.org/10.1145/3415204

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. http://doi.org/10.1126/science.aao2998

Riebling, J. R., & von der Wense, I. (2019). Framing the mass media: Exploring 'fake news' as a frame embedded in political discourse. *Journal of Alternative & Community Media, 4*(1), 57–76. http://doi.org/10.1386/joacm_00043_1

Thorson, E. (2024). How news coverage of misinformation shapes perceptions and trust. Cambridge University Press. https://www.cambridge.org/core/elements/how-news-coverage-of-misinformation-shapes-perceptions-and-trust/6DBC0ED06F52900F0FE207C7E9056AAD

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking* (Vol. 27, pp. 1-107). Council of Europe.

Acknowledgements
We thank Colby Shumway and Moira Kehoe for their assistance with data collection.

Funding
No funding has been received to conduct this research.

Competing interests
The authors declare no competing interests.

Data availability
All materials needed to replicate this study are available via the Harvard Dataverse:
https://doi.org/10.7910/DVN/AMJTZY

Appendix: Changes in total corpus word count over time

This appendix addresses the occurrences of misinformation terms relative to the total output of newspapers as it appears in the ProQuest archive for the months of our archived dataset. The ProQuest dataset may or may not be the same as the actual production and publication rates of these papers, depending on what material each paper submits to the ProQuest archive. This is typically limited by the contract between the publisher and the archive.

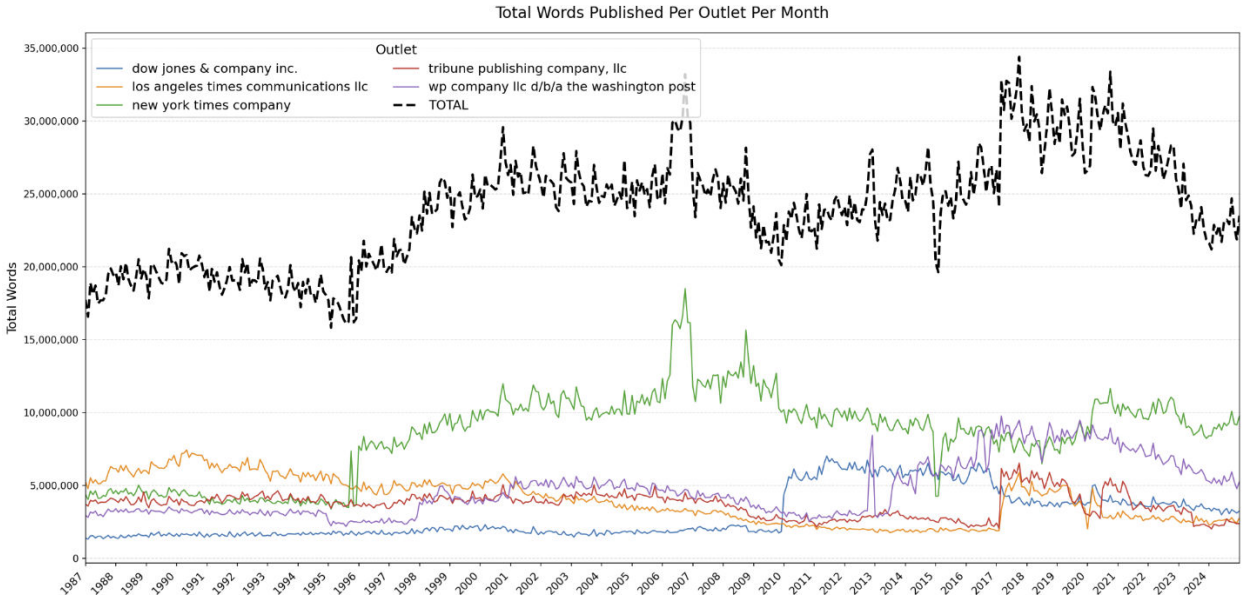

Figure A1. The total number of words published per month in total and also separated by newspaper outlet from 1987* to *2024.

Figure 1 shows the words published per month, both in total and separated by outlet over our timeline. There are plenty of granular changes, with word counts per month fluctuating between 1987 and 2024. The most dramatic changes appear to occur when an outlet changes the amount of data sent to ProQuest. We suspect these large discontinuities are indicative of either contract or technical changes in what is submitted to the archive rather than changes in journalistic output. For instance, there is a spike in total words published on January 1, 2017, near our time period of interest, stemming from a considerable increase in the number of words received from the Chicago Tribune and the L.A. Times. Such fluctuations are decoupled from the observed increases in misinformation terms, such as in November 2016 or July 2020.

As the charts below demonstrate, these alterations in the total word count do not affect the conclusions we reached in our investigation. The changes in the proportion of total words track very closely with the raw count of the key words, indicating that raw counts of key words and their changes over time are largely accurate indicators of the relative prominence of these terms in the press.

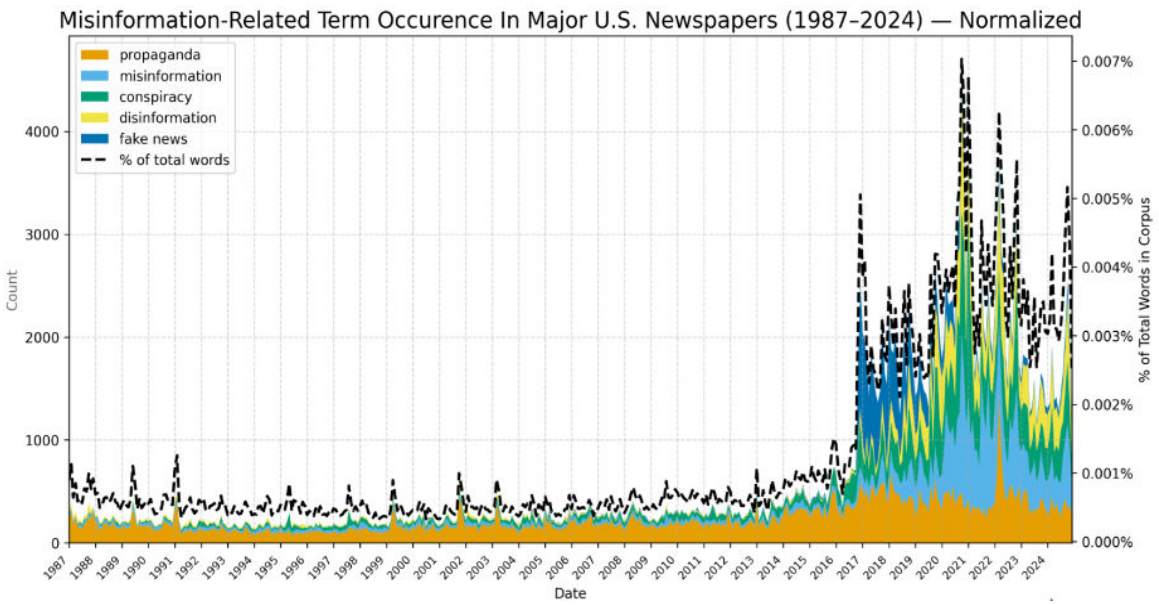

Figure A2. Our misinformation term counts alongside the percentage of the total words per month that were misinformation terms from 1987 to 2024.

The dotted line plotted in Figure 2 represents the word count of misinformation terms over the total number of words for that month, as can be seen in the secondary y-axis. This ratio has been placed atop our existing word counts at a scale to show their generally consistent parallel movement.

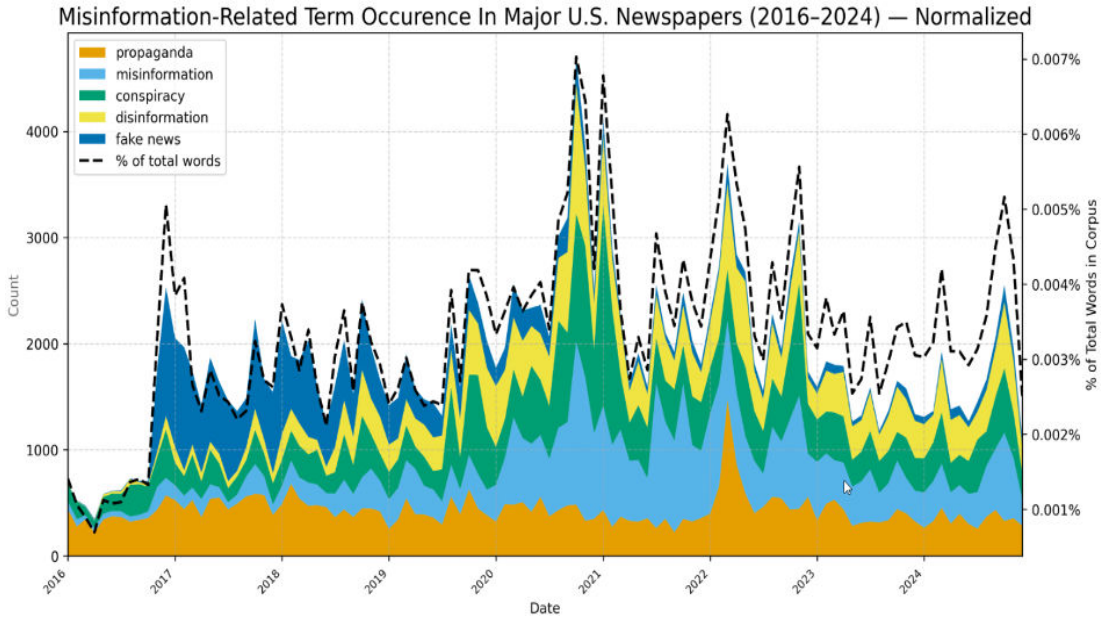

Figure A3. Our misinformation term counts alongside the percentage of the total words per month that were misinformation terms from 2016 to 2024.

Figure 3 focuses on our time period of acute interest (2016–2024). Overall, we see that the parallel pattern of total word count and our terms of interest continues in this period. Our analysis supports the conclusion that the patterns we see in the relative occurrence of misinformation terms over time are a reflection of their changing prominence in newspaper reporting, not in the changing amount of reporting that occurred during this same time.