Global Disinformation Index

Disinformation Risk Assessment: The Online News Market in the United States

www.disinformationindex.org

The Global Disinformation Index is a not-for-profit organisation that operates on the three principles of neutrality, independence and transparency. Our vision is a world free from disinformation and its harms. Our mission is to catalyse industry and government to defund disinformation. We provide disinformation risk ratings of the world's news media sites. For more information, visit <u>www.disinformationindex.org</u>.

The Global DisInformation Lab (GDIL) at the University of Texas at Austin encourages collaborative interdisciplinary academic research on the global circulation of a broad spectrum of information, misinformation, and disinformation via digital media. Our faculty and student researchers are committed to promoting a deeper understanding of national and regional contexts in which such information is generated with a hope that such research will generate policy recommendations and solutions.

Design: www.designbythink.co.za





December 2022. Published under a Creative Commons License (CC BY-NC-SA 4.0)

Table of contents

| Executive summary | 4 | |
|---|----|--|
| Introduction: GDI's approach | 6 | |
| Ensuring neutrality | 8 | |
| The United States media market: Key features and scope | 9 | |
| Disinformation risk ratings | 10 | |
| Market overview | 10 | |
| Pillar overview | 13 | |
| Operations pillar | | |
| Content pillar | | |
| Adversarial narratives | | |
| Ten lowest risk online news outlets | 19 | |
| Ten riskiest online news outlets | 20 | |
| Appendix: Methodology | 22 | |
| Site selection | 22 | |
| Data collection | 22 | |
| Data analysis and indicator construction | 23 | |
| Risk ratings | 25 | |
| Endnotes | 26 | |

3

Executive summary

Since the news business has expanded to the online world, transformations in news production and distribution have exposed the industry to new disinformation risks.

News websites have financial incentives to spread disinformation in order to increase their online traffic and, ultimately, their advertising revenue. Meanwhile, the dissemination of disinformation has disruptive and impactful consequences. Disinforming narratives surrounding the COVID-19 pandemic are a recent example. By disrupting society's shared acceptance of scientific advances, these narratives undermine public health, safety and government responses.

To combat ad-funded disinformation, the Global Disinformation Index (GDI) deploys its assessment framework to rate news domains' risk of disinforming their readers. These independent, trusted and neutral ratings are used by advertisers, ad tech companies, and platforms to redirect their online ad spending, in line with their brand safety and disinformation risk mitigation strategies.

GDI defines disinformation as "adversarial narratives, which are intentionally misleading; financially or ideologically motivated; and/or, aimed at fostering long-term social, political or economic conflict; and which create a risk of harm by undermining trust in science or targeting institutions or at-risk individuals."

The GDI risk rating provides information about a range of indicators related to the likelihood that a given news website will disinform its readers by spreading these adversarial narratives. The GDI risk rating methodology is not an attempt to identify and label disinformation sites or trustworthy news sites. Rather, GDI's approach is based on the idea that a combined set of indicators can reflect a site's overall risk of carrying disinformation. These ratings should be seen as offering initial insights into the U.S. media market and its overall levels of disinformation risk, along with the strengths and challenges the sites face in mitigating disinformation risks.

GDI's research looked at 69 U.S. news sites, selected on the basis of online traffic and social media followers, as well as geographical coverage and racial, ethnic and religious community representation. The index scores sites across 16 indicators – indicators which themselves contain many, many more individual data points – and generates a score for the degree to which a site is at risk of disinforming its readers. These indicators are grouped under the index's **Content** and **Operations pillars**, which respectively measure the quality and reliability of a site's content and its operational and editorial practices and policies. A domain's overall risk rating is based on that site's aggregated score across all the indicators, and ranges from zero (maximum risk level) to 100 (minimum risk level).

This report highlights the ten sites with the lowest levels of risk out of the 69 that were reviewed, as well as the ten riskiest domains. The ten lowest-risk sites published content that was generally free of Bias, Negative targeting and Sensational language. These outlets also disclosed and adhered to standard journalistic policies and practices much more frequently than other sites in the study, with an average **Operations pillar** score that was 25 points above the sample-wide average. On the other hand, the ten riskiest online news outlets struggled in both pillars. These sites published content with far fewer fact-based *Ledes* and far more *Bias*, *Negative* targeting and Sensational language than the rest of the sample. For each of these indicators, their average score was at least 20 points below the market average. The average **Operations pillar** score for this group of sites was 35% lower than the market average, largely due to their lack of policies to ensure the accuracy and attribution of their content.

Overall, the data from the study corroborate today's general impression that hyperbolic, emotional, and alarmist language is a feature of the U.S. news media landscape. Every site displayed some degree of cherrypicking facts, omitting relevant information, making unsubstantiated claims, and/or using logical fallacies. Many of the sites that regularly posted this kind of misleading, biased content also used sensational language to elicit an emotional response from the reader. Moreover, every site in the study used, to some degree, what GDI terms "targeting language," which demeans or belittles people or organisations rather than simply presenting the news. Importantly, this type of language is distinct from criticism or satire; it is sarcastic, derogatory or hateful language that promotes division and distrust. Furthermore, many of the sites that relied on this type of content specifically covered politics. The data showed that this divisive language appeared on both sides of the aisle with similar prevalence. Of all the articles containing targeting language, 38% targeted Democrats, and another 38% targeted Republicans.

Taken together, bias, sensationalism and targeting distract and divide. Readers may be drawn in by the sensationalism and clickbait, to the benefit of advertisers, but when the institutions of science and democracy are undermined by the Fourth Estate itself, everyone loses.

Key findings

In reviewing the media landscape for the United States, GDI's assessment found that:

- A third of the domains in the study (26 out of 69) were assessed as low- or minimum- risk.
- Twenty sites fell into the medium-risk category, while another 20 sites were high-risk.
- Three sites were rated as maximum-risk.
- Weaknesses in the Operations pillar were the predominant risk factor across the market, especially for sites in the medium-, high- and maximum-risk groups. More than 60% of sites scored below 50 on the Operations pillar. Only one site — NPR — scored better on the Operations pillar than the Content pillar.
- Among high-risk sites, average indicator scores for the **Operations pillar** ranged between 5.8 and 52.1. More than 75% of high-risk sites did not publish any policies on sources and bylines.
- Most domains performed strongly on the **Content pillar**, with the market averaging 70 or above on all ten indicators in this pillar.
- The indicators measuring *Article bias* and *Sensational language* resulted in scores ranging almost 80 points across the sample, with sites scoring as high as 96 out of 100 and as poorly as 11.6. Importantly, however, the data showed that every site in the sample displayed some degree of bias or sensationalism in their content.

- The ten lowest-risk online news outlets demonstrated minimal *Bias*, *Negative targeting* of groups or individuals, and *Sensational language*. They also excelled in disclosing and following their operational policies and practices.
- The ten sites with the lowest level of disinformation risk were NPR, AP News, The New York Times, ProPublica, Insider, USA Today, The Washington Post, BuzzFeed News, Wall Street Journal, and HuffPost.
- The ten riskiest sites published content with far fewer fact-based *Ledes* and far more *Bias, Negative targeting* and *Sensational language* than the rest of the sample.
- The sites with the highest level of disinformation risk were the New York Post, Reason Magazine, RealClearPolitics, The Daily Wire, TheBlaze, OAN, The American Conservative, The Federalist, Newsmax, and The American Spectator.

Introduction: GDI's approach

GDI defines disinformation as "adversarial narratives, which are intentionally misleading; financially or ideologically motivated; and/or, aimed at fostering long-term social, political or economic conflict; and which create a risk of harm by undermining trust in science or targeting at-risk individuals or institutions."

This definition was developed to transcend many of the semantic arguments and other challenges facing the anti-disinformation space. Most definitions of disinformation emphasise its intentional nature, which cannot be directly measured, and the veracity of specific facts, which becomes extremely difficult to assess at scale. However, identifying disinformation requires more nuance than simply evaluating whether an assertion is true or false. Not all false statements are disinformation. For example, the North American Aerospace Defense Command (NORAD) Santa tracker claims to track Santa Claus on his Christmas Eve journey around the world.¹ While this claim is clearly false, it poses no harm and would not generally be considered disinformation. On the other hand, a statement that is technically true can be presented out of context in a misleading and harmful way.

Based on the adversarial narrative approach, GDI has developed a methodology that can quantify the level of disinformation risk on open-web news domains by identifying narratives that are misleading and harmful. The GDI methodology looks at over 80 different signals in combination to generate an overall assessment of disinformation risk for a news website as a whole. The resulting score does not determine whether a site or a specific piece of content is disinformation or not. The summation of all the data collected does, however, allow GDI to measure the risk that a given site may disinform its readers. Such data also gives advertisers and algorithms an evidence-based metric for making various decisions. GDI has developed and iterated on this methodology for digital news sources in more than 20 media markets worldwide, with input from our Technical Advisory Group and our research partners around the world.

The United States Media Market Review was conducted by a team of researchers from the Global Disinformation Lab at the University of Texas at Austin, who were trained by GDI to collect data on a set of indicators in two pillars: the **Content pillar**, based on a sample of content published on the site including news and opinion articles, and the **Operations pillar**, which reflects the operational policies, practices and past behaviour of the media outlet. Table 1 lists each of the indicators in the two pillars, with more detail provided in the full Methodology appendix.

The study was designed to categorise each of the 69 sites as either minimum-, low-, medium-, high- or maximum-risk. These risk ratings were based on where the site's overall index score fell within the distribution of all the scores in the dataset. That means the risk rating can be interpreted as the level of disinformation risk relative to the other domains included in the study.

Accordingly, a rating of minimum risk in this study does not imply that disinformation will never appear on a given site. All newsrooms are vulnerable to disinformation risks, ranging from everyday human error to more nefarious tactics. Conversely, a maximum-risk rating does not imply that specific pieces of disinformation have been identified on a given site. Rather, the index looks at a wide variety of practices and mitigation strategies to holistically measure a given site's overall disinformation risk level.

Table 1. Global Disinformation Index Pillars and Indicators

| Pillar | Indicator | Definition | | | |
|------------|-----------------------------|--|--|--|--|
| | Article bias | Rating for the degree of bias in the article. Biased writing misrepresents facts, is based on faulty logic, and/or fails to include or unfairly engages with different views on the story | | | |
| | Negative targeting | Rating for whether and to what degree the story negatively targets a specific individual or group | | | |
| | Out-group inferiority | Rating for whether the story builds upon or establishes that one group is inferior based on identity and to what degree | | | |
| | Sensational language | Rating for the degree of sensationalism in the article | | | |
| Contont | Sensational visuals | Rating for the degree of sensationalism in the visual presentation of the article | | | |
| Content | Sources | Rating for the quantity and quality of the story's sources | | | |
| | Attribution | Rating for whether the story's statistics, quotations, and external media are clearly attributed to a source | | | |
| | Headline accuracy | Rating for how accurately the story's headline describes the content of the story | | | |
| | Lede present | Rating for whether the article begins with a fact-based lede | | | |
| | Byline information | Rating for how much information is provided in the article's byline | | | |
| | Editorial guidelines | Rating for the number of policies identified on the site (adjusted if there are episodes of editorial interference or conflicts of interest) | | | |
| | | Rating for the degree to which the site is likely to adhere to an ideological affiliation, based on its published editorial positions | | | |
| | | Rating for the number of policies identified on the site (adjusted if the site violates guidelines) | | | |
| | | Rating for the number of policies and practices identified on the site (adjusted if the site violates guidelines) | | | |
| | Accuracy policies | Rating for the number of policies and practices identified on the site (adjusted if the site violates guidelines) | | | |
| | | Rating for the number of policies and practices identified on the site (adjusted if the site practises stealth editing) | | | |
| | Sources and byline policies | Rating for the number of policies and practices identified on the site (adjusted if the site violates guidelines) | | | |
| Operations | Funding | Rating for the number of revenue sources identified on the site (adjusted if there are episodes of editorial interference or conflicts of interest) | | | |
| | | Rating based on whether reader subscriptions or donations are identified as a revenue source | | | |
| | | Rating based on the degree of transparency the site provides regarding its sources of funding | | | |
| | a | Rating based on the number of distinct executive or board level financial and editorial decision-makers listed on the site (adjusted if there are episodes of editorial interference or conflicts of interest) | | | |
| | Ownership | Rating based on the degree of transparency the site provides regarding its ownership structure | | | |
| | | Rating for the number of policies identified on the site | | | |
| | Comment policies | Rating for the mechanisms to enforce comment policies identified on the site | | | |

Ensuring neutrality

GDI implements several safeguards to ensure that site reviews are fair and that scores are not based on whether the research team agrees or disagrees with content on the site. The index does not assess partisanship or the specific political, religious or ideological orientation of the site. Rather, the index indicators focus on disinformation risk factors. The following steps were implemented to maintain nuance and neutrality:

- The researchers engaged for each GDI media market review are locally based media experts.
- Researchers are trained by GDI to employ a highly structured methodology, including how to detect disinformation risk across the political spectrum.
- The **Operations** indicators are reviewed independently by two researchers who then work together to validate each individual data point.

- The **Content** indicators are reviewed independently by three different content analysts whose ratings are averaged.
- Content reviews are conducted using anonymised, plain-text content, from which the name of the media outlet and associated journalists have been removed.
- Content reviews are randomised in GDI's data entry system, so that content from each site is reviewed by a variety of researchers.
- All risk ratings undergo an additional internal Quality Assurance process before any data are provided to advertisers, ad tech companies, or platforms.

The United States media market: Key features and scope

Most U.S. readers currently get their news from online sources, with 85% percent of Americans getting at least some news from a digital device, according to a study from Pew Research.²

Those younger than 50 years of age get their news almost exclusively from online sources. Social media play a major role in this consumption, with 48% of Americans getting their news mostly from social media.³ Access to the internet is widespread in populated urban and suburban areas. But as of 2019, 27% of people in rural areas still lacked broadband internet access.⁴

While Americans have embraced the digital consumption of news, the overall trust in media has continued to decline. According to a Gallup poll, America's trust in the media is lower than ever, with just 36% of Americans saying they had a great deal or fair amount of trust in the media as of 2021. This represents a trend that began in the late 90s and early '00s, when trust in mass media first began to decline.⁵

The ubiquitousness of the internet and smartphones has not only increased Americans digital news consumption but also spurred growth in the online advertising industry. The digital advertising market in the U.S. makes up roughly 53% of total global digital ad spend.⁶ According to survey data reported by Statista, digital ad spend for the U.S. totalled 152.25 billion dollars in 2020. It is estimated that, from 2019 to 2024, the digital advertising market in the U.S. will more than double. Most of these gains are expected in the mobile advertising market. Sometime in early 2017 there was a shift in the industry toward the mobile medium, and this trend has continued ever since.⁷ The U.S. ad market is highly concentrated among three key players: Google, Facebook, and Amazon. Google is the largest contributor and its ad spend makes up 28% of the total market share. Facebook's spend makes up 23.8% of the total market share, and Amazon's makes up 11.3%.⁸ It is not surprising that the top contributors to the digital advertising market in the U.S. are social media or platform companies, given that about half of readers in the United States in 2021 received their news via social media at least occasionally.⁹

The journalism industry has had to adapt to social media's role as a key source and sharing medium for news, with Twitter representing the top social media platform used by U.S. journalists for professional reasons. According to a report by Cision, journalists face three major challenges in the digital media landscape: being expected to fight fake news, publishing more content with fewer resources and time, and driving audience readership with their articles. According to Cision, nearly three in ten journalists file ten or more stories per week.¹⁰

The shift to digital as the primary news consumption in the United States has increased the risk of disinformation. Since the Russian interference in the 2016 elections, disinformation has become a primary issue for journalists and the general public. Half of the public and 71% of the journalists in the U.S. view disinformation as a major problem.¹¹ Foreign actors from Russia, Iran and others launched disinformation campaigns between 2018 and 2020, according to Brookings. These state-backed disinformation campaigns have been increasingly outsourced to third-party firms (e.g., marketing or PR companies).¹²

While disinformation remains an issue in the U.S., legislators are still working on how to address the problem. Several bills have been proposed in Congress; however, the highly politicised nature of the issue has prevented progress.¹³

Disinformation risk ratings

This study looks specifically at a sample of 69 news websites which published articles in English and/or Spanish.

Market overview

The sample was defined based on the sites' reach, (using each site's Alexa rankings, Facebook followers, and Twitter followers), relevance, and the ability to gather complete data for the site. For this study, only articles in English were reviewed.

Table 2. Media sites assessed in the United States (in alphabetical order)

| News outlet | Domain | News outlet | Domain |
|--------------------------|----------------------------|---------------------------|---------------------------------|
| ABC News | abcnews.go.com | ProPublica | www.propublica.org |
| AlterNet | www.alternet.org | RealClearPolitics | www.realclearpolitics.com |
| American Muslim Today | americanmuslimtoday.com | Reason Magazine | reason.com |
| AP News | apnews.com | Slate Magazine | slate.com |
| AsAmNews | asamnews.com | Talking Points Memo | talkingpointsmemo.com |
| Axios | www.axios.com | The American Conservative | www.theamericanconservative.com |
| Bloomberg | www.bloomberg.com | The American Spectator | spectator.org |
| Breitbart | www.breitbart.com | The Atlantic | www.theatlantic.com |
| BuzzFeed News | www.buzzfeednews.com | The Daily Beast | www.thedailybeast.com |
| Catholic News Agency | www.catholicnewsagency.com | The Daily Caller | dailycaller.com |
| CBN News | www1.cbn.com | The Daily Wire | www.dailywire.com |
| CBS News | www.cbsnews.com | The Epoch Times | www.theepochtimes.com |
| CNBC | www.cnbc.com | The Federalist | thefederalist.com |
| CNN | www.cnn.com | The Hill | thehill.com |
| El Nuevo Día | www.elnuevodia.com | The Intercept | theintercept.com |
| Forbes | www.forbes.com | The Nation | www.thenation.com |
| Fortune | fortune.com | The New Republic | newrepublic.com |
| Forward | forward.com | The New York Times | www.nytimes.com |
| Fox News | www.foxnews.com | The Post Millennial | thepostmillennial.com |
| HuffPost | www.huffpost.com | The Root | www.theroot.com |
| ndian Country Today | indiancountrytoday.com | The Washington Post | www.washingtonpost.com |
| Insider | www.insider.com | TheBlaze | www.theblaze.com |
| lacobin | jacobin.com | TIME | time.com |
| Los Angeles Times | www.latimes.com | TMZ | www.tmz.com |
| Mother Jones | www.motherjones.com | Univision | www.univision.com |
| MSNBC News | www.msnbc.com | USA Today | www.usatoday.com |
| National Review | www.nationalreview.com | U.S. News | www.usnews.com |
| NBC News | www.nbcnews.com | VICE | www.vice.com |
| New York Daily News | www.nydailynews.com | Vox | www.vox.com |
| New York Post | nypost.com | Wall Street Journal | www.wsj.com |
| Newsmax | www.newsmax.com | Washington Examiner | www.washingtonexaminer.com |
| Newsweek | www.newsweek.com | Washington Times | www.washingtontimes.com |
| NPR | www.npr.org | WIRED | www.wired.com |
| One America News Network | www.oann.com | Yahoo | www.yahoo.com |
| Politico | www.politico.com | | |

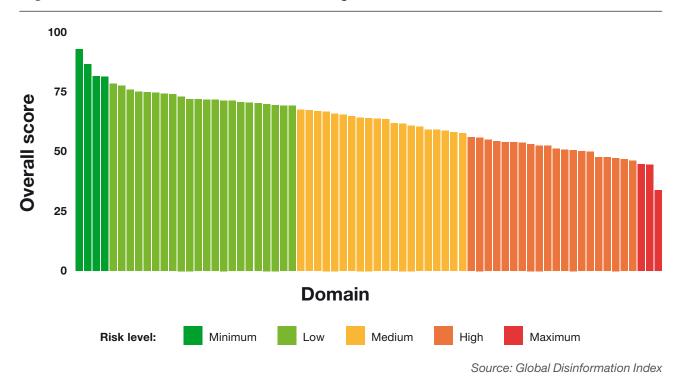


Figure 1. Distribution of disinformation risk ratings

The United States Media Market Review included sites that fell across all five risk categories, with overall scores ranging between 33.9 and 92.9 (see Figure 1). Despite the 59-point gap between the best and worst performers, a plurality of domains were assessed as low-risk (22). Twenty sites fell into the medium-risk category, and another twenty sites were high-risk. Of the remaining seven sites, four were minimum-risk and three were maximum-risk. The sites in this study demonstrated wide disparities between their **Content pillar** and **Operations pillar** scores, receiving average scores of 79 and 47 in these pillars, respectively (see Figure 2). More than 60% of sites scored below 50 on the **Operations pillar**. In contrast, only one site scored below 50 on the **Content pillar**. The average overall score was 63.

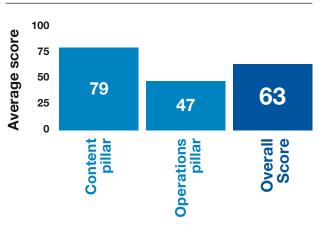


Figure 2. Overall market scores, by pillar

Source: Global Disinformation Index

There were four minimum-risk sites, namely NPR, AP News, The New York Times and ProPublica. These sites each achieved very high scores on the **Content** **pillar** — three of the four scored above 90. Minimum-risk sites avoided content that relied on negative targeting and narratives of out-group inferiority, and consistently published byline information with their content, receiving scores of 90 or above on each of these indicators. In addition, these sites received the four best **Operations pillar** scores in the entire sample. NPR was the only site to score better on the **Operations pillar** than the **Content pillar**, receiving a 94.5 and 91.3 on each pillar, respectively.

Low-risk sites performed on-par with minimum-risk sites in the **Content pillar**, with an average pillar score of 87.6. Similar to the minimum-risk sites, content on the low-risk sites largely avoided negative targeting and narratives that establish group-based inferiority. But lack of transparent policies for attribution, pre-publication fact-checking and post-publication corrections — indicators in the **Operations pillar** — were a consistent risk factor. Even so, none of the low-risk sites scored below the average **Operations pillar** score of 47.

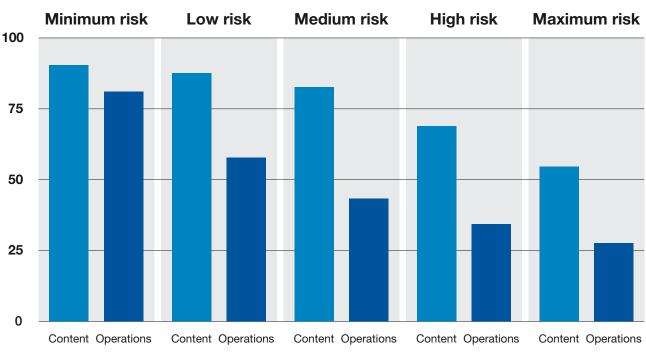


Figure 3. Average pillar scores by site risk rating level

Medium-risk sites received an average **Content pillar** score of 82.6, just 5 points lower than that of low-risk sites. While the medium-risk sites showed slightly more bias than less risky sites, they performed similarly in other indicators. However, **Operations pillar** scores dropped off significantly in this group. As a result, the medium-risk sites demonstrated the greatest disparity between their average **Content pillar** and **Operations pillar** scores (see Figure 3). For sites at this risk level, four of the six **Operations pillar** indicators received average scores below 50 (namely, *Editorial guidelines*, *Accuracy policies*, *Sources and byline policies*, and *Funding*).

Weaknesses in the **Operations pillar** were the predominant risk factor for sites in the high-risk group. Among high-risk sites, average indicator scores for the **Operations pillar** ranged between 5.8 and 52.1. More than 75% of high-risk sites did not publish any sources and byline policies.

The three maximum-risk sites showed significant variability across the pillar and indicator scores. For example, one of the three sites received a score of 70.4 on the **Content pillar** but only an 18.8 on the **Operations pillar**. Maximum-risk sites had an average **Operations pillar** score of just 27.6, drawing down their overall scores. These sites lacked any public policies on bylines and sources, and the editorial guidelines and accuracy policies they published were limited.

Content on maximum-risk sites did not fare much better — notably, all three of these sites received scores below 50 on the *Article bias* indicator. Two of the three maximum-risk sites published content with a significant degree of *Negative targeting*, *Out-group inferiority*, and *Sensational language*, scoring below 60 on each of these **Content pillar** indicators. The same two sites also struggled to start articles with accurate headlines and fact-based ledes, scoring below 60 on these indicators as well.

Pillar overview

Operations pillar

The **Operations pillar** assesses the operational and editorial integrity of a news site. All scores were based on a scale of zero (worst) to 100 (best), as scored by the country reviewers according to the information available on the site and elsewhere online. The operations indicators might be the quickest avenues through which sites could improve their disinformation risk ratings, as they correspond to policies that sites can easily develop, implement and make public. However, failure to follow these policies was taken into account when calculating the **Operations pillar** scores.

In the United States there was a clear need for news sites to enhance their operational checks and balances to mitigate disinformation risk. The average **Operations pillar** score was fairly low (47 out of 100) in comparison with the average **Content pillar** score (79 out of 100). However, performance varied widely in this pillar, with sites receiving **Operations pillar** scores between 18.8 and 94.5. NPR was the only site that scored higher than 90, while more than one third of domains received scores below 40 on this pillar.

Many sites failed to transparently publish and consistently follow policies to guarantee that articles were accurate, well-sourced and attributed with bylines. This is reflected in low average scores for the *Sources* and byline policies and the *Accuracy policies* indicators (31 and 22, respectively). However, the scores for these indicators featured a wide range. Nine sites scored 0 on both indicators, while three sites reached a perfect score of 100 on one of these indicators. Properly developing and consistently implementing such policies contributes to mitigating the risk of spreading disinformation, by publishing factual, well-sourced news and taking appropriate action when corrections are necessary.

Similarly, the *Editorial guidelines* scores were poor on average (39 out of 100), with almost 75% of the domains scoring under 50 on this indicator. Most of the sites failed to publish or consistently observe guidelines that protect the site from editorial interference and conflicts of interest. Such guidelines are key to preventing unethical, ideologically motivated, and misleading practices with respect to the production of content and its publication.

Enhanced disclosure of these domains' guidelines would increase public accountability and trust. Ideally, these guidelines would establish editorial independence and the absence of conflicts of interest. They would also profess adherence to principles like accuracy and factbased communication, accountability and responsibility to the public, fairness, respect, consideration of the impact of journalism on the lives of others, and ethically appropriate treatment of sensitive content. Lastly, these guidelines would require a clear differentiation between news and opinion content. However, transparency alone is insufficient to effectively reduce disinformation risk. Hence, sites would improve in this indicator only if they consistently implemented and followed such guidelines.

The average *Funding* indicator score was 50 out of 100. Despite being the third best scoring indicator in the **Operations pillar**, news sites in the United States still had plenty of room to improve the transparency of their funding sources to increase public accountability, as well as the diversification of their funding structure, which can prevent the editorial process from being dictated by or tied to specific interests. Similarly, the *Ownership* indicator (which scored 68 out of 100) took into account the degree of transparency and diversification of each site's ownership structure. However, diversified funding and ownership structures are not always able to prevent conflicts of interest or editorial interference. Accordingly, the current methodology penalised these indicator scores when episodes of conflicts of interest or editorial

interference involving financial interests or owners had occurred.

The fairly high score for the *Comment policies* indicator (72 out of 100) highlights that most of the assessed U.S. sites limited the risk of spreading disinformation on their user-generated comment sections, either by thoroughly regulating users' comments or by having no comment section at all. Article reactions, comments and other user-generated content can become conduits for promoting disinformation, hate speech, harassment, and other problematic behaviours.

The relatively low scores on the **Operations pillar** in the U.S. media market can be attributed to a lack of public disclosure of and compliance with operational best practices by the news domains in the sample. By making key operational guidelines and policies available to the public and consistently following them, news domains in the United States could significantly improve their overall disinformation risk scores.

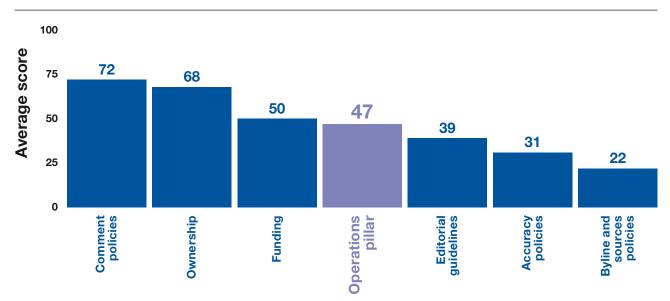
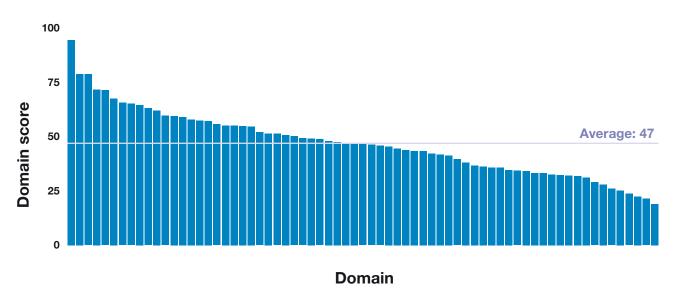


Figure 4. Average Operations pillar scores by indicator





Content pillar

The **Content pillar** focuses on the reliability of the content provided on the site. Analysis for this pillar is based on an assessment of 20 anonymised articles for each domain. These articles were drawn from the most frequently shared pieces of content during the data

collection period and from a sample of content pertaining to topics which present a disinformation risk, such as politics and health. The **Content pillar** indicators for each article were aggregated and normalised, resulting in the final score for each domain. All scores were based on a scale of zero (worst) to 100 (best).

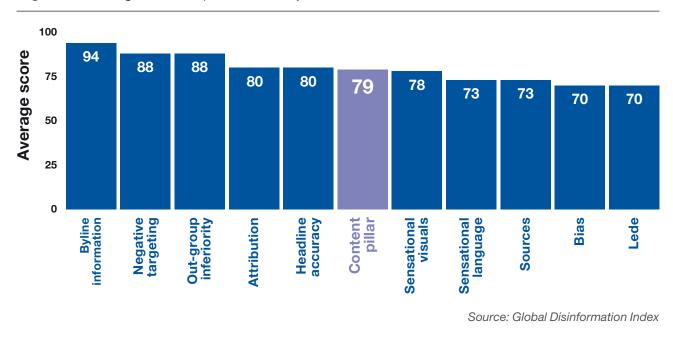


Figure 6. Average Content pillar scores by indicator

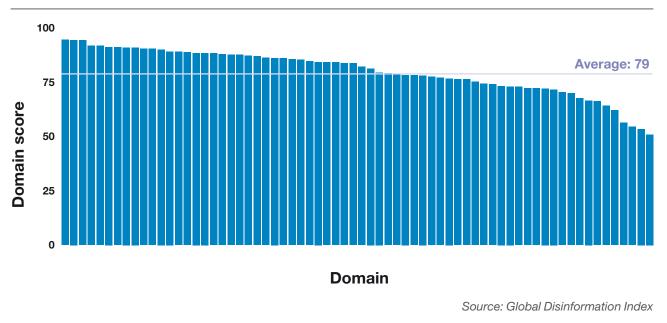


Figure 7. Content pillar scores by site

The majority of the sampled news sites in the U.S. performed strongly in the **Content pillar**. All ten of the indicators for this pillar had an average score of 70 or above. This suggests that the sites in the sample generally tended to avoid adversarial narratives in their articles and followed sound sourcing and attribution practices, along with other journalistic elements, like ledes and bylines. However, 11 sites scored below 70 on this pillar. These domains demonstrated substantial variability, with individual indicator scores ranging from 12 to 100 out of 100. All of these sites were rated high- or maximum-risk overall.

The average *Byline information* indicator score was 94 out of 100, making it the best performing **Content pillar** indicator and suggesting that attributing individual authorship is a common journalistic practice in the United States.

The relatively low scoring *Lede* indicator reached an average score of 70 out of 100. Roughly one third of the sites in the sample did not include a fact-based lede about a third of the time. However, this indicator had one of the broadest ranges of scores across sites, reaching a score of 13 on the lower end. Hence, the use of ledes to introduce readers to the key facts of stories before their analysis or interpretation was heterogeneous among the sites in the sample.

The indicators measuring *Article bias* and *Sensational language* varied substantially, resulting in scores ranging almost 80 points across the sample. The *Article bias* indicator featured the lowest performing **Content pillar** indicator score for any site, with one site scoring 11.6 out of 100 points. Importantly, this indicator defines bias in terms of the construction and presentation of the story or argument, not the perspective of the journalist. The highest score achieved on this indicator was 96.1, indicating that every site in the sample displayed some amount of bias in their content. The same was true of the *Sensational language* indicator, as the highest site score was 95.9 and even minimum-risk sites displayed some degree of sensationalism. However, results varied, with most site scores falling between 60 and 90.

The *Sources* indicator, which scored 73 on average, assesses the quality and quantity of the sources that substantiate the sites' articles. Interestingly, the lowest score in this indicator was higher than the previous indicators, i.e. 50 out of 100. This suggests that the worst performing sites generally avoid entirely baseless

content or low-quality sources. However, no site received a perfect score for this indicator. At times even the bestperforming sites published articles that were reliant on a limited number of sources and/or exclusively social media sources (e.g., posts, live feeds, videos, etc.), and/or did not give all the parties involved a chance to comment.

The use of sensational visual cues to elicit an emotional response from the readers is captured by the *Sensational visuals* indicator, which scored 78 out of 100 points across the sample. Only six sites scored below 60, which meant that even the worst scoring sites avoided constant use of extremely sensational images, data visualisations, videos, thumbnails and social media posts. However, the highest score in this indicator (i.e. 94.45) failed to reach 100, illustrating that even the best performing site in this indicator made use of somewhat sensational visual cues.

The *Attribution* indicator scored slightly above the **Content pillar** average (80 out of 100). In most of their articles, the U.S. sites identified and sourced elements like statistics, quotations, and external media. More than three quarters of the assessed sites used proper attribution in at least 75% of their content. At the same time, there were no sites in the sample that attributed all of these elements in all of their articles. Even for top-performing sites, the source of quotes, statistical claims or external media content was occasionally omitted. Similarly, the Headline accuracy average score of 80.4 suggests that most news sites used fairly accurate headlines and generally avoided clickbait, but none of the assessed sites were without some degree of risk.

The *Negative targeting* and *Out-group inferiority* indicators reached an average score of 87.9 and 88.3 out of 100, respectively. While negative targeting appeared across the market, the frequency and intensity varied significantly from site to site. Few sites employed language that directly undermined an individual, group or institution in a significant portion of the content reviewed. Similarly, most sites avoided content that suggested or claimed that one group of people was inferior, unworthy, or somehow worse than another on the basis of political or ideological perspective, race or ethnicity, gender, or similar characteristics. Overall, divisive language that established a specific target and an us-versus-them narrative seemed to be very limited on the majority of domains. However, this type of content was concentrated on ten sites that scored below 75 on the *Negative targeting* indicator.

Adversarial Narratives

This section explores the data gathered to compute the *Negative targeting* indicator of the **Content pillar**. For each article, the reviewers evaluated the presence and intensity of negative targeting against several types of actors, such as individuals, groups or communities, organisations, institutions, and others. For this particular analysis, only the presence of negative targeting was considered, not its intensity. Out of the total of 1,380 articles reviewed, some level of negative targeting was detected in 65.6% of the articles.

Negative targeting was defined as the use of ridiculing, derogatory, or hateful remarks, along with the promotion

of unsubstantiated doubts or distrust in a specific actor. Criticism was not considered negative targeting, as long as the author used solid reasoning and backed their claims with strong evidence. This assessment was based on the anonymised version of the article text, allowing the reviewers to evaluate the presence and intensity of negative targeting without knowing the author of the article or the site where it was published.

In terms of the targeted actors, the table below shows the percentage of articles that contained some level of negative targeting for each category. Note that a single article can target more than one actor.

| Actors negatively targeted | Percentage of articles |
|--|------------------------|
| Reputation of an individual | 56.6% |
| Institution | 47.8% |
| Democrats or liberals | 38.3% |
| Republicans or conservatives | 38.3% |
| Organisation or company | 28.8% |
| Activists or other collective movements | 22.7% |
| Groups or individuals based on race, ethnicity, or nationality | 10.1% |
| Groups or individuals based on gender | 9.4% |
| LGBTQ+ community | 8.6% |
| Centrists | 8.6% |
| Scientists or science | 6.7% |
| Immigrants | 5.6% |
| Christians or Christianity | 3.1% |
| Other religion | 2.1% |
| Muslims or Islam | 1.2% |
| Jewish people or Judaism | 0.4% |

Table 3. Percentage of articles with presence of negative targeting, by actor targeted

Negative targeting against individuals was the most common category, with 57% of the total 905 articles scorning or denigrating individuals such as politicians, the President of the United States, journalists, TV personalities, or celebrities. Institutions were targeted in 48% of these articles, often in connection with specific issues. For instance, the WHO and CDC, along with local governments (typically in Democratic states), were targeted by articles opposing COVID booster shots or new restrictions to contain monkeypox. Police, Rangers, and Border Patrol Tactical Units were heavily criticised following the Uvalde mass shooting, before any investigation on the timing of their response was carried out. The Supreme Court was negatively targeted because of the decision to overturn Roe v. Wade.

The percentage of articles targeting Democrats or Republicans were extremely close (both groups were targeted in 38% of the articles), reflecting the wellestablished political polarisation in the United States.¹⁴ Organisations were also targeted in 29% of these articles. The NRA received a lot of attention and negative targeting after the Buffalo and Uvalde mass shootings. "Big Tech" and specific companies were heavily criticised for taking a position regarding issues like disinformation, LGBTQ+ rights and climate change. Black Lives Matter was also denigrated in several instances. Activists were targeted in 23% of the articles, especially pro-choice, racial equality or LGBTQ+ rights advocates.

A relatively significant share of these articles targeted minorities based on race, ethnicity, or nationality (10%), and based on gender or their belonging to the LGBTQ+ community (9%). Negative targeting against individuals or groups based on religion was also present (almost 7% of the articles in total).

The ten lowest-risk online news outlets

NPR.org (Risk level: Minimum)

NPR's online news presented a minimum level of disinformation risk, both based on its neutral, factbased content and its transparent and comprehensive operational policies and practices. Some small degree of bias and sensationalism was detected in the content sample, which has the potential to mislead readers. But on the whole, the site appears to have sufficient safeguards in place to prevent disinformation from making its way into the newsroom.

APNews.com (Risk level: Minimum)

The news homepage for The Associated Press had the best **Content pillar** score of all assessed sites. The AP could stand to improve in some of the **Operations** metrics, including transparent and diverse funding. However, readers can rely on the AP for neutral, fair and well-developed reporting.

NYTimes.com (Risk level: Minimum)

The New York Times online was also rated minimum-risk, in large part based on a high degree of transparency all around, from who authors the news to who owns the company and how it makes its money. Content from NYTimes.com was not always free of bias, but it generally avoided targeting language and adversarial narratives.

ProPublica.org (Risk level: Minimum)

A non-profit with an emphasis on investigative journalism, ProPublica was a minimum-risk site in the index based on strong scores across the board. Readers will find in-depth coverage without bias, sensationalism or negative targeting.

Insider.com (Risk level: Low)

Insider was among the highest-scoring sites in the low-risk category. Content on this site was largely free of bias, negative targeting or sensationalism, and the articles used journalistic best practices to familiarise readers with the topic at hand.

USAToday.com (Risk level: Low)

USA Today received a low-risk rating based on strong scores across the board. The site could improve in terms of relying on a wide range of sources and being sure to clearly attribute statistics, quotations and external media. However, the articles reviewed were almost entirely free of divisive or demeaning language. USA Today also ranked amongst the top 20 domains for headline accuracy, suggesting that clickbait is relatively rare on the site.

WashingtonPost.com (Risk level: Low)

The Washington Post publishes some of the strongest editorial guidelines among the assessed sites. This domain largely avoids sensational or negatively targeted reporting — but its content includes occasional bias, and its funding structure could do more to prevent conflicts of interest.

BuzzFeedNews.com (Risk level: Low)

BuzzFeed News – a separate domain from the popular entertainment site known for its quizzes — demonstrated a strong **Content pillar** score based on neutrality and journalistic best practices. Statistics, quotations and external media were properly cited, and its articles frequently employed objective, fact-based ledes. The site scored relatively well on indicators of neutral, unemotional language, but could stand to tone down its sensational visuals.

WSJ.com (Risk level: Low)

Readers of The Wall Street Journal can expect neutral reporting free of content that is either sensationalised or demeaning toward specific groups or individuals. Articles on this site featured a degree of bias similar to The New York Times and The Washington Post; which is to say, not absent, but limited.

HuffPost.com (Risk level: Low)

HuffPost largely featured fact-based, unbiased content free of sensational text or visuals. This domain also refrained from perpetuating divisive narratives via the negative targeting of groups or individuals. The outlet's scores for the **Operations pillar** indicators were imperfect, but better than most.

The ten riskiest online news outlets

NYPost.com (Risk level: High)

The New York Post was rated as high-risk, largely because of its lack of transparency around operational policies and practices. The site published no public auidelines for the use of bylines on its content, the types and number of sources its content relies on, or pre-publication fact-checking and post-publication corrections processes. As a result, even if relevant policies exist, they cannot be factored into the site's risk score. Additionally, content sampled from the Post frequently displayed bias, sensationalism and clickbait, which carries the risk of misleading the site's readers. Importantly, GDI's study did not review specific highprofile stories and attempt to determine whether they were disinformation. Rather, the risk score is based on a robust operational framework and a blind review of a sample of articles from across the site.

Reason.com (Risk level: High)

Reason Magazine's high-risk rating can be attributed to scores of zero on three Operations pillar indicators: the site publishes no information regarding authorship attribution, pre-publication fact-checking or postpublication corrections processes, or policies to prevent disinformation in its comments section. In terms of its content, Reason Magazine did largely refrain from perpetuating in-group out-group narratives or unfairly targeting certain actors via its reporting, but its articles were often biased in their construction and relied on sensationalised, emotional language.

RealClearPolitics.com (Risk level: High)

RealClearPolitics scored poorly on the **Content pillar** due to the prevalence of biased and sensational language, which risks misleading and manipulating readers. Their articles often lacked clear and diverse sources, and there was no information regarding policies on sources and bylines on the site. These factors can make it difficult for readers to double-check the basis of questionable arguments or claims. RealClearPolitics scored well on the **Sensational visuals** indicator due to the fact that almost none of its articles had visual elements (aside from their headline image).

DailyWire.com (Risk level: High)

In addition to bias and sensational language, articles on The Daily Wire featured a high degree of sensational visuals. Combined, such content runs the risk of manipulating readers' emotional responses and disseminating biased interpretations of events, thus garnering a high-risk rating.

TheBlaze.com (Risk level: High)

The Blaze scored as high-risk, receiving fairly even **Content pillar** and **Operations pillar** scores. This domain's content showed the third-highest degree of bias and second-highest prevalence of sensational language among sites in this study. Most articles also failed to use journalistic best practices to familiarise readers with the topic at hand, instead leading with bold claims or emotional appeals.

OANN.com (Risk level: High)

One America News Network (OANN) was also scored as high-risk, but demonstrated a substantial difference between its **Content pillar** and **Operations pillar** scores. OANN's low **Operations pillar** score was largely the result of publishing no information regarding its policies to ensure accuracy (i.e., fact-checking, etc.) or attribute authorship, or about its ownership structure, which is a risk factor for conflicts of interest and/or editorial interference. OANN did moderately well on some of the **Content** indicators, but was one of only a few sites to fail to include a complete byline on most of the articles sampled.

TheAmericanConservative.com (Risk level: High)

The American Conservative had one of the lowest scores in the study for bias, indicating that almost all of the content sampled was either somewhat or entirely biased. Importantly, this indicator does not measure whether the author of an article agrees with one or another side of an issue; it assesses the construction of the story or argument, looking for elements like unsubstantiated claims, logical fallacies, ad hominem attacks, and obvious omissions of pertinent information. In the case of The American Conservative, these features were widespread, putting readers at risk of being consistently misled.

TheFederalist.com (Risk level: Maximum)

The Federalist performed well in a handful of areas, principally a transparent ownership structure free from conflicts of interest. However, the site fell short in other aspects of the **Operations pillar**. It also had one of the lowest **Content pillar** scores in the study, scoring in the 20s for *Article bias* and in the 40s for *Sensational language*. Further, the use of language that demeans, belittles or otherwise targets individuals, groups or institutions was frequent. Taken together, articles written in this way — especially when they appear across a news domain — establish misleading and harmful narratives that amount to disinformation.

Newsmax.com (Risk level: Maximum)

Newsmax received one of the lowest **Operations pillar** scores, putting it in the maximum-risk category. The site lacks transparency around its operational practices across the board. Newsmax performed much better on the **Content pillar**, but the outlet's scores for *Article bias* and *Sensational visuals* fell to the 50s, indicating a significant frequency and degree of misleading arguments and emotional images, videos, and other visual elements.

Spectator.org (Risk level: Maximum)

In content published by The American Spectator, bias, sensationalism, and divisive and targeting language were prevalent, while fact-based ledes and wellmeasured headlines were rare. On the contrary, most of the assessed articles on this domain negatively targeted a group or individual in their title or opening sentences. Frequent hyperbole and generalisations further supported the establishment of adversarial narratives. The site also provided little transparency around its operations, in particular its policies on sources and attribution and its editorial guidelines.

Appendix: Methodology

The Global Disinformation Index evaluated the level of disinformation risk of the U.S. online media market, represented by a sample of 69 news domains, selected on the basis of online traffic and social media followers, as well as geographical coverage and racial, ethnic and religious community representation.

The index was composed of the **Content** and **Operations pillars**. The pillars were, in turn, composed of 16 indicators. The **Content pillar** included indicators that assess elements and characteristics of each domain's content to capture its level of adversariality, credibility, sensationalism, and impartiality. The **Operations pillar's** indicators evaluated the transparency and enforcement of policies and rules that a specific domain followed to ensure the reliability and quality of the news being published.

Site selection

The market sample for the study was developed based on a mix of quantitative and qualitative criteria. GDI created a list of the 120 news websites with the greatest traffic in the media market. This list was internally vetted to gauge relevance and reach. Then the list was reduced to 69 sites, ensuring that the sample provided adequate geographical coverage and racial, ethnic and religious community representation. The final media market sample reflected the set of sites for which complete data could be collected throughout the review process.

Data collection

The **Content** indicators were based on the review of a sample of 20 articles published by each domain. Ten of these articles were randomly selected among a domain's most frequently shared articles on Facebook, typically within a two-month period. The remaining ten articles were randomly selected from a group of the domain's articles covering topics that are likely to carry disinformation narratives.

The sampled articles were anonymised by removing any information that allowed the analysts to identify the publisher or the author of the articles. Each anonymised article was reviewed by three country analysts who were trained on the GDI **Content pillar** codebook. For each anonymised article, the country analysts answered a set of 17 questions designed to evaluate the elements and characteristics of the article text and its headline. After the information was recorded based on the anonymised text, the analysts subsequently reviewed how the article was presented on the domain.

The **Operations pillar** was based on the information gathered during the manual assessment of each domain performed by the country analysts. The country analysts answered a set of 72 questions designed to evaluate each domain's ownership, management, and funding structure, editorial independence, principles and guidelines, attribution policies, error-correction and fact-checking policies, and rules and policies for the comments section. The reviewers answered a set of seven additional questions to capture documented incidents of editorial and ethical violations of the site's stated guidelines. The analysts gathered evidence to support their assessments as they performed each Operations and Enforcement review.

Data analysis and indicator construction

The data gathered by the country analysts for the **Content pillar** were used to compute ten indicators. The **Content pillar** indicators included in the final risk rating were: *Article bias*, *Attribution*, *Byline information*, *Headline accuracy*, *Out-group inferiority*, *Lede present*, *Negative targeting*, *Sensational language*, *Sensational visuals*, and *Sources*. For each indicator, values were normalised to a scale of 0 to 100. The pillar score for each domain was the weighted average of all the scores for all of the pillar's indicators, and ranged from 0 to 100. Table 4 gives the weights.

| Indicator | Weight |
|-----------------------|--------|
| Article bias | 1 |
| Negative targeting | 1 |
| Out-group inferiority | 1 |
| Sensational language | 1 |
| Sensational visuals | 1 |
| Sources | 0.5 |
| Attribution | 0.5 |
| Headline accuracy | 0.5 |
| Lede present | 0.25 |
| Byline information | 0.25 |

Table 4. Content pillar indicator weights

Source: Global Disinformation Index

For the **Operations pillar**, the answers gathered during the Operations and Enforcement reviews by the country analysts were translated into a set of sub-indicators. The six indicators were calculated as the averages of these sub-indicator scores. The resulting **Operations pillar** indicators were: *Accuracy policies, Comment policies, Editorial guidelines, Funding, Ownership*, and *Sources and byline policies*. For each indicator, values were normalised to a scale of 0 to 100. The domain score for the **Operations pillar** was the average score across indicators. The complete list of sub-indicators and indicators for both pillars is given in Table 5

Global Disinformation Index Advisory Panel

GDI's risk assessment framework is developed with the advice and support of an Advisory Panel, including:

- Ben Nimmo (Facebook)
- Camille François (Niantic)
- Miguel Martinez (co-founder and chief data scientist, Signal Al)
- Nic Newman (Reuters Institute of Journalism)
- Olaf Steenfadt (Reporters without Borders)
- Cristina Tardáguila (Lupa)
- Amy Mitchell (Pew Research)
- Scott Hale (Meedan and Credibility Coalition)
- Finn Heinrich (OSF), and
- Laura Zommer (Chequeado)

Table 5. Global Disinformation Index pillars and indicators

| Pillar | Indicator | Sub- indicators | Unit of analysis | Definition | Rationale |
|------------|-----------------------------------|---------------------------------------|---------------------|---|---|
| Content | Article bias | | Article | Rating for the degree of bias in the article. Biased writing misrepresents facts, is based on faulty logic, and/or fails to include or unfairly engages with different views on the story. | Indicative of neutral fact-based reporting or well- rounded analysis |
| | Negative targeting | | | Rating for whether and to what degree the story negatively targets a specific individual or group | Indicative of hate speech, bias or an adversarial narrative |
| | Out-group inferiority | None | | Rating for whether and to what degree the story builds upon or establishes that one group is inferior based on identity and to what degree | Indicative of hate speech, bias or an adversarial narrative |
| | Sensational language | | | Rating for the degree of sensationalism in the article text | Indicative of neutral fact-based reporting or well- rounded analysis |
| | Sensational visuals | | | Rating for the degree of sensationalism in the visual presentation of the article | Indicative of neutral fact-based reporting or well- rounded analysis |
| | Sources | | | Rating for the quantity and quality of the story's sources | Indicative of fact-based reporting and high journalistic standards |
| | Attribution | | | Rating for whether the story's statistics, quotations, and external media are clearly attributed to a source | Indicative of fact-based reporting and high journalistic standards |
| | Headline accuracy | | | Rating for how accurately the story's headline describes the content of the story | Indicative of clickbait |
| | Lede present | | | Rating for whether the article begins with a fact-based lede | Indicative of fact-based reporting and high journalistic standards |
| | Byline information | | | Rating for how much information is provided in the article's byline | Attribution of stories creates accountability for their veracity |
| Operations | Editorial guidelines | Editorial independence | Site | Rating for the number of policies identified on the site (adjusted if there are episodes of editorial interference or conflicts of interest) | Assesses the degree of editorial independence and the policies in place to mitigate conflicts of interest |
| | | Adherence to narrative | | Rating for the degree to which the site is likely to adhere to an ideological affiliation, based on its published editorial positions | Indicative of politicised or ideological editorial decision-making |
| | | Content guidelines | | Rating for the number of policies identified on the site (adjusted if the site violates guidelines) | Assesses the policies in place to ensure that factual information is reported without bias |
| | | News vs. analysis | | Rating for the number of policies and practices identified on the site (adjusted if the site violates guidelines) | Assesses the policies in place to ensure that readers can distinguish between news and opinion content |
| | Accuracy policies | Pre- publication fact checking | | Rating for the number of policies and practices identified on the site (adjusted if the site violates guidelines) | Assesses policies to ensure that only accurate information is reported |
| | | Post- publication corrections | | Rating for the number of policies and practices identified on the site (adjusted if the site practises stealth editing) | Assesses policies to ensure that needed corrections are adequately and transparently disseminated |
| | Sources and byline policies | None | | Rating for the number of policies and practices identified on the site (adjusted if the site violates guidelines) | Assesses policies regarding the attribution of stories, facts, and media (either publicly or anonymously); indicative of policies that ensure accurate facts, authentic media and accountability for stories |
| | Funding | Diversified incentive structure | | Rating for the number of revenue sources identified on the site (adjusted if there are episodes of editorial interference or conflicts of interest) | Indicative of possible conflicts of interest stemming for over-reliance on one or few sources of revenue |
| | | Accountability to readership | | Rating based on whether reader subscriptions or donations are identified as a revenue source | Indicative of accountability for high-quality information over content that drives ad revenue |
| | | Transparent funding | | Rating based on the degree of transparency the site provides regarding its sources of funding | Indicative of the transparency that is required to monitor the incentives and conflicts of interest that can arise from opaque revenue sources |
| | Ownership | Owner- operator division | | Rating based on the number of distinct executive or board-level financial and editorial decision makers listed on the site (adjusted if there are episodes of editorial interference or conflicts of interest) | Indicative of a separation between financial and editorial decision-making, to avoid conflicts of interest |
| | | Transparent ownership | | Rating based on the degree of transparency the site provides regarding its ownership structure | Indicative of the transparency that is required to monitor the incentives and conflicts of interest that can arise from opaque ownership structures |
| | Comment policies | Policies | | Rating for the number of policies identified on the site | Assesses policies to reduce disinformation in user- generated content |
| | | Moderation | | Rating for the mechanisms to enforce comment policies identified on the site | Assesses the mechanism to enforce policies to reduce disinformation in user-generated content |

Risk ratings

The overall index score for each domain was the average of the pillar scores. The domains were then classified on the basis of a five-category risk scale based on the overall index score. The risk categories were defined based on the distribution of risk ratings from the current country study. This dataset was standardised to fit a normal distribution with a mean of 0 and a standard deviation of 1. The standardised scores and their distance from the mean were used to determine the bands for each risk level, given in Table 6.

| Risk level | Lower bound | Upper bound | Standard deviation |
|--------------|-------------|-------------|-----------------------|
| Minimum risk | 80.28 | 100 | > 1.5 |
| Low risk | 68.84 | 80.27 | > 0.5 and ≤ 1.5 |
| Medium risk | 57.41 | 68.83 | > -0.5 and \le 0.5 |
| High risk | 45.97 | 57.40 | > -1.5 and \le -0.5 |
| Maximum risk | 0 | 45.96 | ≤ -1.5 |

Table 6. Disinformation risk levels

Endnotes

1 Wikipedia: <u>https://en.wikipedia.org/wiki/NORAD</u> <u>Tracks Santa</u>.

2 Pew Research: <u>https://www.pewresearch.org/fact-</u> tank/2021/01/12/more-than-eight-in-ten-americans-getnews-from-digital-devices.

3 Pew Research: <u>https://www.pewtrusts.org/en/trust/</u> <u>archive/summer-2019/americas-digital-divide</u>.

4 Pew Research: <u>https://www.pewtrusts.org/en/trust/</u> archive/summer-2019/americas-digital-divide.

5 Gallup: <u>https://news.gallup.com/poll/355526/</u> americans-trust-media-dips-second-lowest-record.aspx.

6 Statista: <u>https://www.statista.com/statistics/242552/</u> <u>digital-advertising-spending-in-the-us</u>.

7 Statista: <u>https://www.statista.com/statistics/242552/</u> <u>digital-advertising-spending-in-the-us</u>.

8 Statista: <u>https://www.statista.com/statistics/242549/</u> digital-ad-market-share-of-major-ad-selling-companies-inthe-us-by-revenue. 9 Pew Research: <u>https://www.pewresearch.org/</u> journalism/2021/09/20/news-consumption-across-socialmedia-in-2021.

10 Cision: <u>https://cmm-custom.prnasia.com/cms-prod/</u> attachment/20220523/Cision_SOTM_2022_Report.pdf.

11 Pew Research: <u>https://www.pewresearch.org/</u> journalism/2022/06/14/journalists-highly-concerned-aboutmisinformation-future-of-press-freedoms.

12 Brookings: <u>https://www.brookings.edu/techstream/</u> how-disinformation-evolved-in-2020.

13 Congress: <u>https://www.congress.gov/bill/117th-congress/house-bill/6971</u>.

14 Dimock, M., & Wike, R. (2021, March 29). America is exceptional in its political divide. The Pew Charitable Trusts. Retrieved September 21 2022, from <u>https://www.pewtrusts.org/en/trust/archive/winter-2021/america-is-exceptional-in-its-political-divide</u>.



www.disinformationindex.org