NewsGuard

November 2022

Fighting Misinformation with Journalism, not Algorithms

Independent Research on the Effectiveness of Using Human-Curated News Reliability Ratings to Mitigate False News

Contents

Misinformation: A critical concern

CONTEXT

Misinformation has become an increasingly significant issue in recent years, with Russian state-financed propaganda fueling the war in Ukraine, vaccine misinformation undermining public health efforts during the COVID-19 pandemic, and hoaxes about rigged election results threatening democracy across the world.

Fifty percent of Americans see "made-up news" as a "very big problem" in the country today, outranking other critical issues including terrorism, climate change, crime, and racism.[1] Across the world, 54% of those surveyed in the 2022 Reuters Digital News Report said that they worry about identifying the difference between what is real and fake on the internet when it comes to news.[2]

Research suggests that individuals are not well-equipped to identify and avoid false content. The Organization for Economic Cooperation and Development (OECD) Programme for International Student Assessment (PISA) 2018 Insights and Interpretations report found that fewer than **1 in 10 students (9%) in OECD countries are able to distinguish between fact and opinion.[3]**

[1]https://www.journalism.org/2019/06/05/many-americans-say-made-up-news-is-a-critical-problem-that-needs-to-be-fixed/
[2]https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022/dnr-executive-summary
[3]https://www.oecd.org/pisa/PISA%202018%20Insights%20and%20Interpretations%20FINAL%20PDF.pdf

NewsGuard was founded to combat the global rise in misinformation by providing individuals, brands, and democracies with more context to understand the reliability of their sources of news and information. Its approach is based on research measuring how credibility ratings and Nutrition Labels for news sources are effective and scalable ways to inform news consumers and have a meaningful impact on reader behavior. As of November 2022, NewsGuard has rated more than 8,400 websites, representing 95% of engagement with news and information online in the U.S., the U.K., Canada, France, Germany, Italy, and Austria.

This white paper outlines the current research studying misinformation and potential solutions. It first documents how concerned people are about false news and what kind of a threat it poses. The paper then summarizes the evidence assessing the approach of labeling sources with credibility ratings. It ends with a review of the research that has studied NewsGuard's model, specifically, resulting in the conclusion that NewsGuard offers a workable solution to countering widespread misinformation online.

People are concerned about misinformation

Misinformation poses a challenge to citizens of all ages, education levels, and political perspectives, and it has damaging effects for a range of societal sectors including government, the media, and public health. Moreover, researchers using NewsGuard's data suggest that some topics have more coverage in fake news than in legitimate news, such as *getting vaccinated*, *back to school*, and *pandemic in Italy*.[4]

False narratives and propaganda outlets have played a central role in fueling Russia's war in Ukraine. State-financed disinformation online has become a significant instrument of the conflict, and is increasingly widespread: NewsGuard's Russia-Ukraine Disinformation Tracking Center identified more than 280 websites spreading false information about the war as of October 2022, with more polluting the internet every day.[5] Americans struggle to identify misinformation in relation to the war in Ukraine, with researchers finding that **50% of Americans report being uncertain about the veracity of at least one false claim regarding Ukraine.[6]**

Misinformation has also played a part in prolonging the COVID-19 pandemic. Researchers at Indiana University estimate that **online misinformation has been responsible for a 20% decrease in vaccine uptake across states.[7]** The 2022 Reuters Institute Digital News Report found that people regularly come across misinformation online related to COVID-19.

[4]https://link.springer.com/article/10.1007/s10796-022-10329-7
[5]https://www.newsguardtech.com/special-reports/russian-disinformation-tracking-center/
[6]https://www.covidstates.org/reports/misperceptions-about-the-war-in-ukraine-and-covid-19-vaccines
[7]https://www.nature.com/articles/s41598-022-10070-w

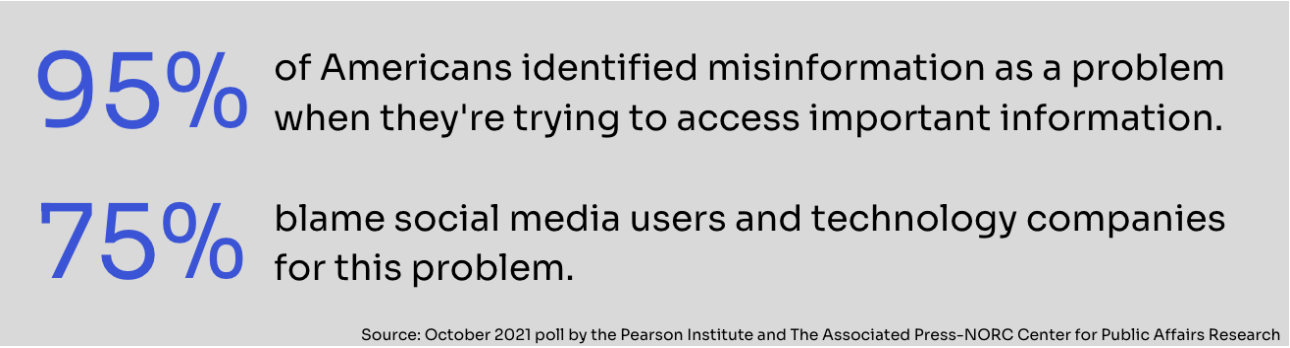
Proportion of people who had seen false or misleading content about COVID-19 in the prior week:[8]

Source: Reuters Institute Digital News Report (2022)

Misinformation has also played a role in recent political elections. Researchers using NewsGuard's source reliability data found that anti-establishment networks disseminated content from a large number of NewsGuard Red-rated sites during the German Federal Election in 2021, proliferating anti-vaccination, anti-lockdown, and anti-climate protection content specifically.[9] False content has also been a feature of recent American elections, with further research showing that politicians in the U.S. have been increasingly sharing links to untrustworthy news sites since the election of President Biden.[10]

[8]https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022/dnr-executive-summary
[9]https://www.medienanstalt-nrw.de/fileadmin/user_upload/NeueWebsite_0120/Zum_Nachlesen/BTW22_Political_Advertisement.PDF
[10]https://academic.oup.com/pnasnexus/article/1/4/pgac186/6695314

Citizens in the U.S. and abroad report feeling anxiety about the spread of misinformation online



- A 2021 study by the Pearson Institute and The Associated Press-NORC Center for Public Affairs Research showed that 95% of Americans identified misinformation as a problem while trying to access important information, and 6 in 10 are at least somewhat concerned that their friends or family members have been part of the problem.[11] A further 75% blame social media users and technology companies.

- The 2022 Reuters Institute Digital News Report found that although two-thirds of Americans access news in some form every day, only 26% of them trust the news.[12] Trust in news has also fallen across much of Europe. In the U.K. for example, only 34% of respondents trust the news, and that figure is lower in France, at 29%.

[11] https://apnews.com/article/coronavirus-pandemic-technology-business-health-misinformation-fbe9d09024d7b92e1600e411d5f931dd?utm_source=newsletter&utm_medium=email&utm_campaign=newsletter_axioslogin&stream=top
[12] https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022/dnr-executive-summary

Individuals' own news consumption and sharing habits reflect their concerns

- The 2022 Edelman Trust Barometer reported that concerns about fake news and misinformation are now at an all-time high of 76%. Trust in search engines is as low as 59%, followed by traditional media at 57%, and social media at 37%.[13]

- An Ipsos poll conducted for BuzzFeed News found that **75% of American adults were fooled by false news headlines.**[14]

- The 2019 Reuters Institute Digital News Report found that **67% of Americans are concerned** about what is real and what is fake online.[15]

Hoax news stories and the websites that spread them are not the only sources of confusion when it comes to current events. Research also shows that people have difficulty telling when an article is intended as satire rather than news reporting. In 2019, communications researchers for Nieman Lab published the findings of a months-long study examining stories — some true, some false — that spread widely on social media.[16] It turned out that many of the highly engaged false stories actually came from satirical sources like The Onion and the Babylon Bee, but in many cases, people did not realize such stories were intended as satire.

[13]https://www.edelman.com/trust/2022-trust-barometer
[14]https://www.buzzfeednews.com/article/craigsilverman/fake-news-survey
[15]https://reutersinstitute.politics.ox.ac.uk/sites/default/files/inline-files/DNR_2019_FINAL.pdf
[16]https://www.niemanlab.org/2019/08/maybe-you-know-that-artice-is-satire-but-a-lot-of-people-cant-tell-the-difference

False information has broad effects

The consequences of misinformation can be far-reaching. False articles spread widely — in many cases, more than accurate ones. Research published in *Science* in March 2018 found that falsehoods on Twitter traveled faster than accurate stories, reaching more people (between 1,000 and 100,000 people) than true stories (seldom reaching more than 1,000 people).[17] The same was true of Facebook, according to a BuzzFeed News analysis published November 2016, which found that the 20 top-performing false stories relating to the 2016 election received more total engagements than the 20 top-performing true stories.[18] These falsehoods have widespread effects across different sectors of society — from politics to the pandemic.

As one might expect, an institution whose credibility is heavily impacted by misinformation is journalism, with a 2019 Reuters/Ipsos poll for the Columbia Journalism Review finding that "the press" was the most mistrusted among a group of institutions in which Americans reported having little trust — more mistrusted than Congress or the Executive Branch.[19]

[17]https://science.sciencemag.org/content/359/6380/1146
[18]https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook
[19]https://www.cjr.org/special_report/how-does-journalism-happen-poll.php

People are also concerned that sharing inaccurate information will affect their personal reputation, reinforcing the need for tools that empower users to make informed decisions about the types of content they can trust and share. A 2020 study, “Why do so few people share fake news? It hurts their reputation” by Sacha Altay, Anne-Sophie Hacquin, and Hugo Mercier, showed that, across four experiments, sharing fake news affects one’s reputation “in a way that is difficult to fix, even for politically congruent fake news”[20] — false news that aligns with one’s political affiliation.

"To maintain a good epistemic reputation people and media outlets must avoid sharing fake news because their audience keeps track of how accurate the news they share has been in the past."

- Sacha Altay, Anne-Sophie Hacquin, Hugo Mercier, "Why do so few people share fake news? It hurts their reputation," published November 2020

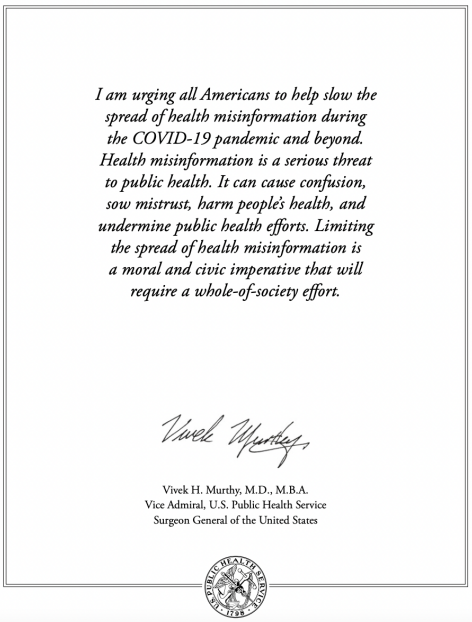
False news also undermines trust in democratic institutions. A 2019 Pew Research Center survey, “Many Americans Say Made-Up News Is a Critical Problem That Needs To Be Fixed,” found that 68% of Americans said made-up news greatly affects Americans’ confidence in government, and 51% said it greatly impacts the ability of politicians to get work done.[21]

[20]https://journals.sagepub.com/eprint/YEGAUV6ASVFGP9AIX7XV/full
[21]https://www.journalism.org/2019/06/05/many-americans-say-made-up-news-is-a-critical-problem-that-needs-to-be-fixed/

Health misinformation in the age of COVID-19

Health misinformation also represents a major category of online falsehoods — a problem exacerbated by the pandemic.

Misinformation about vaccines and COVID-19 spread so widely during the pandemic that the U.S. Surgeon General Vivek Murthy was compelled to issue an advisory about the problem in 2021. The report, titled "Confronting Health Misinformation," outlined Murthy's advice on building a healthy information environment, particularly in light of increasing health misinformation amidst the "rapidly changing information environment" of social media platforms and new technologies.[22] He highlighted the need to "equip Americans with the tools to identify misinformation, make informed choices about what information they share, and address health misinformation in their communities," and "implement product design and policy changes on technology platforms to slow the spread of misinformation."



A series of nineteen surveys conducted by the COVID States Project from April 2020 to July 2021 indicated that not only do people believe false claims about COVID-19 vaccinations, but they are more likely to jeopardize their health because of the misinformation.[23] Mauricio Santillana, assistant professor at Harvard Medical School, explained that the survey also found that the most deaths occurred in regions where misinformation was more prominent from September of 2020 to February of 2021.

Despite the widely acknowledged prevalence of health misinformation in the context of COVID-19 circulating on major social media platforms, research by global think tanks and media organizations has found that the platforms have done little to combat this problem perpetuated by their networks. Preliminary research from global civic organization Avaaz on Facebook, Instagram, YouTube,

[22]https://www.hhs.gov/sites/default/files/surgeon-general-misinformation-advisory.pdf
[23]https://www.hks.harvard.edu/faculty-research/policy-topics/health/vaccine-misinformation-affecting-our-health

and Twitter shows that none of the large tech platforms are effectively combating misinformation, with principal findings as follows:[24]

- Facebook is the biggest "emitter" of COVID-19 disinformation: Across the four platforms, 68% of the total interactions with COVID-19 disinformation, as documented by Avaaz, occurred on Facebook.
- YouTube is the worst of the four platforms when it comes to acting on content: The video hosting platform failed to take action on 93% of the fact-checked content analyzed in the study, with Twitter also performing poorly with 74% of false content left unchecked.
- Lack of labeling: An analysis of Avaaz's sample of COVID-19 disinformation found that YouTube and Twitter seem to focus solely on removing content, rather than labeling false content — a complicated approach especially when dealing with political content.

The Institute for Strategic Dialogue (ISD) echoed these concerns in its October 2021 case study on Facebook's failure to tackle COVID-19 disinformation.[25] The findings note that, "despite detailed policies on mis- and disinformation and promises to enforce them, social media platforms are failing to tackle prominent groups and individuals who spread false claims about COVID-19 and vaccines online."

[24]https://www.politico.eu/wp-content/uploads/2021/06/30/Brief_-None-of-the-Big-Tech-platforms-effectively-curbing-Covid-Disinformation-as-World-confronts-third-wave-1-1.pdf
[25]https://www.isdglobal.org/isd-publications/ill-advice-a-case-study-in-facebooks-failure-to-tackle-covid-19-disinformation/

NewsGuard's data demonstrate similar alarming statistics on the infiltration of health misinformation into news outlets

According to a 2021 NewsGuard analysis based on data from the more than 7,000 news and information sites that account for 96% of online engagement in the U.S.,12.35% of the websites people rely on for news publish false health advice.[26] As of November 2021, these sites accounted for more than 34 million engagements (shares, likes, comments, etc.) on social media over a 90-day period, collectively representing more engagement than major news websites such as Business Insider and the Washington Post.

NewsGuard's Coronavirus Misinformation Tracking Center[27] is a constantly updated resource tracking the websites that have perpetuated false claims about COVID-19 or vaccines. As of July 2022, NewsGuard had found that more than 500 websites have promoted more than 75 distinct false claims about the pandemic, such as:

According to NewsGuard data and ratings of 7,000+ sites:

- 12.35% of the websites people rely on for news publish false health advice.
- As of November 2021, these sites accounted for more than 34 million engagements (shares, likes, comments, etc) on social media over a 90-day period.

- The French government authorized euthanasia in the middle of the crisis
- Colloidal silver can cure COVID-19
- Government food stamps would be denied to those who refuse the COVID-19 vaccine

[26]https://www.statnews.com/2019/07/26/health-websites-are-notoriously-misleading-so-we-rated-their-reliability/
[27]https://www.newsguardtech.com/special-reports/coronavirus-misinformation-tracking-center/

The phenomenon has caused such concern among medical professionals and institutions that the editors-in-chief of the world's major cardiovascular scientific journals published a letter in the Journal of the American Heart Association in January 2019, warning that medical misinformation was causing "significant harm to society and individuals," adding that "human lives are at stake."[28]

"Without exaggeration, significant harm, to society and individuals, derives from the wanton spread of medical misinformation."

- Joint letter by the editors in chief of major cardiovascular journals globally

Research finds that exposure to COVID-19 vaccine misinformation reduces the uptake of the vaccine, which in turn increases the risk of death as a result of the coronavirus. In a February 2021 study, "Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA," researchers from Imperial College London found that exposing people who stated that they would definitely accept a vaccine to recent misinformation induced a decline in intent to receive the vaccine of 6.2 percentage points in the U.K. and 6.4 percentage points in the U.S.[29] The CDC's "Morbidity and Mortality Weekly Report," issued in September 2021, found that after Delta had become the most common variant of COVID-19, fully vaccinated people were more than ten times less likely to die from the virus than the unvaccinated.[30]

Even before the pandemic brought vaccine misinformation into the spotlight, anti-vaccine rhetoric had broad impacts. In April 2019, the CDC released a statement announcing that measles cases were at their highest rate since the disease was declared eliminated in 2000, citing misinformation about vaccine safety as partially to blame for this outbreak.[31]

[28]https://www.ahajournals.org/doi/10.1161/JAHA.118.011838
[29] https://www.nature.com/articles/s41562-021-01056-1
[30]https://www.cdc.gov/mmwr/volumes/70/wr/mm7037e1.htm?s_cid=mm7037e1_whttps://www.cdc.gov/mmwr/volumes/70/wr/mm7037e1.htm?s_cid=mm7037e1_w
[31]https://www.cdc.gov/media/releases/2019/s0424-highest-measles-cases-since-elimination.html

Research summary

CONSUMERS WORRY ABOUT FALSE NEWS

2022 Reuters Institute

3 out of 10 Americans trust the news

2019 IPOS/BuzzFeed News

7 out of 10 American adults were fooled by false news headlines

2018 Gallup

9 out of 10 social media users want social media sites to integrate NewsGuard

2022 Edelman Trust Barometer

8 out of 10 people are concerned about misinformation and fake news

2021 Pearson Institute

6 out of 10 people are concerned that their friends and family have been part of the problem

2019 PISA

1 out of 10 students in OECD countries are able to distinguish between fact and opinion

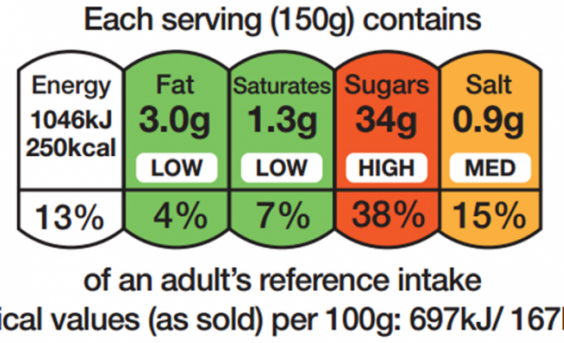
Nutrition Label format: Rating at the source level

AN EFFECTIVE MODEL FOR BOTH FOOD AND NEWS

The Model: Nutrition Labels promote healthy eating

Nutrition labels have been used to steer consumers toward healthier food and drink options for decades.[32] Since the introduction of food labels, a large body of research has been produced to examine the effectiveness of labeling, and to determine which format and design has the greatest impact.

A meta-analysis of such studies, published in 2015 by two researchers for the Organization for Economic Co-operation and Development (OECD), found that food labeling increases the number of people choosing healthier food options by about 18%, while decreasing the number of calories consumed by about 4%.[33] The researchers also noted that, of the different systems studied, traffic light labels were the most effective. Such color-coded systems have been endorsed by medical authorities such as the British Medical Association, adopted in countries such as the UK and France,[34] and supported by local citizens (80% of British parents supported the label format in a 2007 online survey conducted by Netmums).[35]



The UK Health Ministry recommends a traffic-light system for communicating nutritional value
Image credit: The British Nutrition Foundation

[32]https://www.ncbi.nlm.nih.gov/books/NBK209859/
[33]https://onlinelibrary.wiley.com/doi/abs/10.1111/obr.12364
[34]http://www.leparisien.fr/societe/marisol-touraine-annonce-pour-avril-le-logo-qui-classe-les-aliments-par-couleur-15-03-2017-6762966.php
[35]http://news.bbc.co.uk/2/hi/health/6397187.stm

Could Nutrition Labels promote healthy *reading*?

NewsGuard uses its Red-Green traffic light system based on nine apolitical credibility and transparency criteria to assign ratings to sources of information, rather than sources of food. March 2022 research from Cambridge University Press evaluated the effect of signaling the credibility of a fictional media outlet based around nine credibility criteria, much like NewsGuard's approach. It found that participants who were told a source had a high credibility rating reported increased favorability and trust towards that source, demonstrating the effectiveness of source reliability ratings in influencing individual news consumption beliefs and behavior.[36]

This concept of rating and labeling the quality of information online has also been investigated by Alan Dennis and Antino Kim, two researchers at Indiana University.

Professor Dennis and Assistant Professor Kim conducted a study published in March 2019 that looked at the effects of altering the appearance of news articles on social media. Specifically, they examined how highlighting the source of an article in a social media post impacted whether users saw the story as believable, and how much users then engaged with the post.[37]

[36]https://www.cambridge.org/core/journals/journal-of-experimental-political-science/article/when-do-sources-persuade-the-effect-of-source-credibility-on-opinion-change/48ECC9B706B2C3BA733936BE184917CC
[37] https://misq.org/says-who-the-effects-of-presentation-format-and-source-rating-on-fake-news-in-social-media.html

The researchers found that...

- Source ratings **directly impacted the believability of unknown sources**, and;
- Source **believability directly impacted how much people read, shared, and liked** each story.

These results aligned with those of an earlier study conducted by the two researchers in January 2017, in a project that compared different reputation rating formats to assess their ability to influence users' belief of news articles.[38]

The researchers tested three different source rating structures:

- **Expert ratings** produced by reviewers who fact-checked articles. The fact-checks were then aggregated to provide an overall source rating.
- **User article ratings,** where users rated the credibility of individual articles. Article ratings were aggregated to yield an overall source score.
- **User source ratings,** where users directly rated the overall sources.

In their conclusion, the authors wrote:

- "Our results show that **presenting source reputation ratings directly influences the extent to which users believe articles on social media**."
- They found this influence to be particularly strong for low ratings: **lower source ratings decreased believability**, but high ratings did not have as much of an effect.
- In comparing the three rating methods, the researchers found that methods i and ii – **expert ratings** and user article ratings – **had a greater effect** than the third method: user source ratings.

[38]"Behind the Stars: The Effects of News Source Ratings on Fake News in Social Media," accessed at https://www.researchgate.net/publication/322240112_Behind_the_Stars_The_Effects_of_News_Source_Ratings_on_Fake_News_in_Social_Media

Kim and Dennis underscored these conclusions in a December 2019 column they published in The Conversation, along with co-authors Professor Patricia Moravec from the University of Texas at Austin and Professor Randall Minas at the University of Hawaii.[39] In the commentary, headlined "Rating news sources can help limit the spread of misinformation," the four researchers described the aforementioned study examining the effectiveness of source ratings, writing that "helping users mistrust inaccurate material at the moment they encounter it can help curb the spread of disinformation."

What we learned indicates that expert ratings provided by companies like NewsGuard are likely more effective at reducing the spread of propaganda and disinformation than having users rate the reliability and accuracy of news sources themselves."

– Professors Dennis, Kim, Moravec and Minas, December 2019 commentary in The Conversation

The importance of the timing of interventions designed to counter false information — disputing claims before users come across them rather than after — has been articulated by researchers who advocate for "prebunking" rather than "debunking."

In their article "Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence," published in May 2017 in PLOS One, Stephan Lewandowsky, a psychologist at the University of Bristol, John Cook, a researcher at the Center for Climate Change Communication at George Mason University, and Ullrich Ecker, a cognitive psychologist at the University of Western Australia, explain how, because pre-existing beliefs impact how people respond to novel information, warnings about misinformation are more effective when they are administered before misinformation is encountered rather than after.[40] By providing users with Red icons warning of the untrustworthiness of a source before they open or share its articles, NewsGuard serves as a "pre-bunker" rather than a debunker.

[39]https://theconversation.com/rating-news-sources-can-help-limit-the-spread-of-misinformation-126083
[40]https://pubmed.ncbi.nlm.nih.gov/28475576/

Previous research conducted in 2013 by researchers at Princeton University highlighted how proactive warnings about harmful content can influence user behavior. The researchers observed the impact of more than 25 million browser security warnings served to internet users when visiting harmful sites using Mozilla Firefox and Google Chrome. They found that after being presented with warning messages describing phishing or malware risks, users clicked through to fewer than a quarter of the dangerous webpages, demonstrating the clear effectiveness of a warning system in reducing click-through rates, among other metrics.[41]

In their 2019 article in The Conversation, Kim, Dennis, Moravec, and Minas concluded with a reference that recognized NewsGuard as a promising solution. “What we learned indicates that expert ratings provided by companies like NewsGuard are likely more effective at reducing the spread of propaganda and disinformation than having users rate the reliability and accuracy of news sources themselves,” they wrote. “That makes sense, considering that, as we put it on Buzzfeed, “crowdsourcing ‘news’ was what got us into this mess in the first place.”

[41]https://www.researchgate.net/publication/262285806_Alice_in_warningland_A_large-scale_field_study_of_browser_security_warning_effectiveness

NewsGuard's Approach

RELIABILITY RATINGS AND NUTRITION LABELS

Is NewsGuard's approach the right one?

May 2022 research from researchers at NYU and Princeton found NewsGuard's browser extension had a strong and statistically significant effect among participants who routinely see news from Red-rated sources in their social-media feeds, who were much more likely to read news from reliable Green-rated sources after having installed the NewsGuard extension.[42] Princeton researcher Andy Guess commented, "Some people may not be aware if information they encounter online comes from outlets with questionable news-gathering practices, which explains why simple color-coded labels can move the needle and nudge them toward more reliable sources. This is relevant for debates around content moderation and platform governance policies and suggests potential benefits to presenting contextual information about the sources of news links shared on social media."[43]

"Because the rating tool was applied to the vast majority of news articles participants saw in their news feeds, the source rating tool did not produce known, unintended consequences associated with previous efforts to combat online misinformation."

– 2018 Gallup/Knight Foundation Survey, "Assessing the effect of news source ratings on news content" report

Ahead of launching the first news-rating browser extension of its kind, the Knight Foundation, the Ford Foundation, and the Bill and Melinda Gates Foundation tested the then-novel idea of labeling news sources.[44] April 2018 research by the three groups tested the effectiveness of having Green or Red ratings done by a team of experienced journalists appear next to website domains that appear on Facebook news feeds so that news consumers would be advised to proceed with caution whenever they were presented with news from an unreliable source.

[42]https://csmapnyu.org/research/news-credibility-labels-have-limited-average-effects-on-news-diet-quality-and-fail-to-reduce-misperceptions
[43]https://spia.princeton.edu/news/credibility-cues-may-improve-news-diets-misinformation-consumers
[44]https://knightfoundation.org/reports/assessing-the-effect-of-news-source-ratings-on-news-content

The research found:

- **The news source rating tool worked as intended.** Perceived accuracy increased for news headlines with a green source cue and decreased for headlines with a red cue.

- Participants also indicated they were **less likely to read, like, or share news headlines with a red source cue.**
- The source rating tool was **particularly effective for participants who correctly recalled that experienced journalists devised the ratings,** compared with those who did not recall that information.
- **The source rating tool was effective across the political spectrum**. The perceived accuracy of news articles with a red source cue decreased similarly among Republicans and Democrats, with the sharpest decline occurring when the headlines had a clear political orientation that matched the users' political beliefs.
- **No backfire effect:** Because the rating tool was applied to the vast majority of news articles participants saw in their news feeds, the source rating tool did not produce known, unintended consequences associated with previous efforts to combat online misinformation. The experiment did not produce evidence of an "implied truth effect," an increase in perceived accuracy for false stories without a source rating when other false stories have a source rating, or a "backfire effect," a strengthening of one's false beliefs following a factual correction.

"If the results of the Gallup survey hold true, those NewsGuard ratings hold real power."

– "Here's What Happens When News Comes with a Nutrition Label," WIRED magazine

Does NewsGuard work in practice?

After launching the browser extension, NewsGuard and the Knight Foundation commissioned a Gallup survey to assess how the tool worked when installed on personal computers.[45] The study, conducted in November 2018, found:

- **91% find the NewsGuard Nutrition Labels helpful.**
- **90% generally agree with the ratings and respondents trusted the ratings** more because NewsGuard ratings are done by "trained journalists with varied backgrounds."
- 89% of users of social media sites and **83% overall want social media sites and search engines to integrate NewsGuard ratings and reviews into their news feeds and search results.**
- **78% said they would recommend NewsGuard to** a friend or relative.
- **69% would trust social media and search companies** more if they took the simple step of including NewsGuard in their products.
- **63% would be less likely to share news stories from Red-rated websites, and 56% would be more likely to share news from Green-rated websites**.

Gallup researchers concluded: "The positive results among people who accepted Gallup's invitation to download the NewsGuard browser extension suggest a desire for more information about the sources of news people see online, such as in their social media news feeds and in their search results. The news source rating tool offers a scalable solution to identify which news sources adhere to the basic journalistic standards of accuracy and accountability citizens expect and deserve."

[45]https://www.newsguardtech.com/wp-content/uploads/2019/01/Gallup-NewsGuards-Online-Source-Rating-Tool-User-Experience.pdf

The findings of this survey echoed the results of earlier, qualitative research on the effect on Facebook users having access to NewsGuard in their news feeds. This research, which was done by a team at Indiana University who have done considerable research in the area of misinformation and disinformation on social media,[46] used a focus group of Indiana University students (and should thus be read as representing that demographic).[47] The key findings reflecting the views of the participants are as follows:

- **85% found the NewsGuard Red and Green icons useful** and 64% also found the full "Nutrition Label" write-ups useful.
- **72% reported they would be less likely to click on news from Red-rated websites,** and 77% would be more likely to click on news from Green-rated websites.
- **83% said the integration of NewsGuard into the Facebook news feed would positively influence their perception of Facebook.**
- When asked who they would trust to produce the ratings, virtually all respondents trusted ratings produced by NewsGuard as a group of independent experts, but would not trust such ratings if made by technology companies such as Facebook.

[46]Members of the team have also authored the academic papers "Behind the Stars: The Effects of User and Expert Reputation Ratings on Users' Belief in Fake News on Social Media (2017)" and "Says Who? The Effects of Presentation Format and Source Rating on Fake News in Social Media" (2017)"

[47]NewsGuard Wisdom Springboard Focus Group Report published October 18, 2018, by Alan Dennis, Professor, John T. Chambers, Chair of Internet Systems, Antino Kim, Assistant Professor, Patricia Moravec, Ph.D. Candidate at the Kelley School of Business, Indiana University

Offering NewsGuard at scale: Direct integration and "middleware"

To benefit from the "prebunking" service NewsGuard offers, users today need to subscribe to NewsGuard and install its desktop browser extension. While the extension has gained a devoted following since its launch in August 2018, its reach is still limited to the media savvy internet user concerned about online misinformation, or the public librarian who uses the tool to teach media literacy skills. Moreover, browser extensions generally suffer from limited uptake: According to a 2019 ZDNet article,[48] only 13 of the 188,620 extensions available on the Chrome Web Store at the time had more than 10 million users — a low number considering there are more than 2 billion global users of the Chrome browser.[49]

Thus, for NewsGuard's ratings to have an impact at scale, they need to be made available to users right where they encounter information online — directly in their social media feeds or search results. One method for doing so is by having the online platforms adopt a "middleware" approach. This term, coined by the Stanford Working Group on Platform Scale,[50] refers to "software and services that would add an editorial layer between the dominant internet platforms and consumers." The Stanford Working Group authors elaborated on this concept in their December 2020, white paper:

> "We view middleware as an opportunity to introduce competition and innovation into markets currently dominated by the principal internet platforms. There is enormous pressure on the platforms to filter from their domains not just illegal content, but also material that is deemed politically harmful, such as conspiracy theories, fake news, and abusive content.
>
> As noted in the Introduction of this report, this kind of political content curation is done routinely by traditional media companies, but it is altogether different to give these duties to dominant internet platforms—private companies that unilaterally control vast swaths of communication through highly non-transparent means. Indeed, those platforms have often been reluctant to play this role.
>
> A competitive middleware sector would help solve this problem by outsourcing content curation to other organizations that enable consumers to tailor their feeds to their own explicit preferences. At the same time, middleware, in our view, could be a superior alternative to structural remedies imposed by either courts or regulators, in that it would directly respond to consumer preferences and market actors."

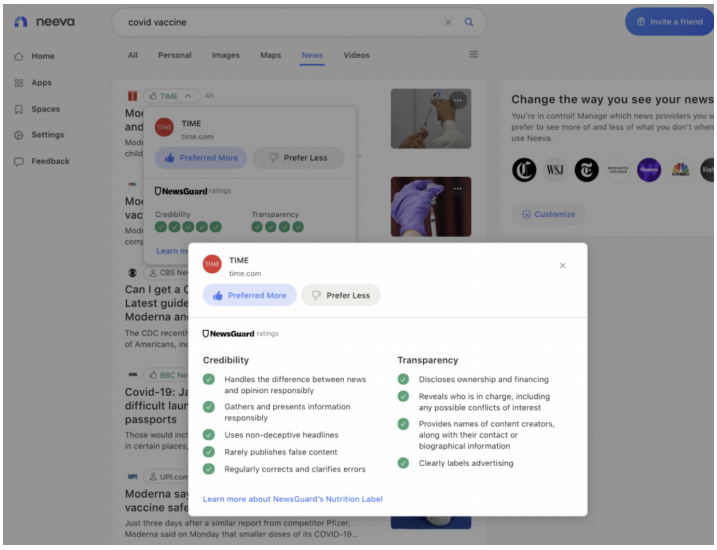
[48]https://www.zdnet.com/article/half-of-all-google-chrome-extensions-have-fewer-than-16-installs/
[49]https://techcrunch.com/2016/11/10/google-says-there-are-now-2-billion-active-chrome-installs/
[50]https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/platform_scale_whitepaper_-cpc-pacs.pdf

In practice, this would require large companies in the search, social, and digital media industries to open up their APIs to enable safety tech providers to create tools, services and solutions to online harms integrated on the platforms themselves. These services could then be offered to users on an opt-in basis so that users can decide how they'd like their social and digital media platforms to be curated. The kinds of services that could be developed and offered to users in-platform include, but aren't limited to:

- Information filters on social media sites and search rankings that give users more control over their online environment;
- Tools to help users avoid engaging with accounts spreading online hate and harassment;
- Opt-in labeling or blocking of harmful content so that users can decide what they feel comfortable reading, trusting, and sharing with friends and family. (Indeed, the lead author of the Stanford report, Stanford professor Dr. Francis Fukuyama, endorsed NewsGuard as one form of a middleware service in subsequent reports promoting this approach).[51,52]

[51]https://muse.jhu.edu/article/787834/pdf
[52]https://muse.jhu.edu/article/797796

Neeva, the first ads-free, private subscription search engine service founded by former YouTube and Google executives, uses NewsGuard as a “middleware” solution, integrating labels directly into the search results it provides to its consumers. There is no censorship of sources; instead, Neeva consumers will have access to information about the sources in their search results so that they can decide which sources are generally trustworthy and which ones are not.[53]



Research from Sciences Po in Paris indicates that although Big Tech companies such as Facebook have mechanisms in place to deprioritize repeat offenders flagged for spreading misinformation, Facebook’s policy does not reduce engagement with these accounts in the long run.[54] The study stresses the importance of using external agents to monitor misinformation interventions and their impact.

Middleware can be enabled by digital platforms voluntarily or in order to comply with emerging laws and regulations requiring that platforms take reasonable steps to reduce the online harms they cause, such as the "safety by design" requirements of the Online Safety legislation being prepared in the UK and the "Empowering Users" commitment by the platforms in the European Commission's Code of Practice on Disinformation. Giving users access to middleware tools such as NewsGuard has several advantages:

- **Boosting transparency:** Tech platforms currently use black-box algorithms to pick and choose what users can see on social media and search. By giving users greater control over the algorithms underpinning their online experience, as well as tools to help users determine who is behind the information they encounter, government, platforms, and the safety tech industry can work together to increase transparency online.

[53]https://www.newsguardtech.com/press/newsguard-and-neeva-team-up-for-first-ever-independent-assessments-of-news-sources-in-search/
[54]https://misinforeview.hks.harvard.edu/article/measuring-the-effect-of-facebooks-downranking-interventions-against-groups-and-websites-that-repeatedly-share-misinformation/

- **Promoting a competitive industry**: With a small number of companies dominating the digital media sector, there is clearly a need to encourage newer entrants and create more product innovation. Social and digital media companies need new tools to help users navigate online harms, and a middleware approach provides an opportunity to create an ecosystem of safety tech firms, thereby creating jobs, investment, and skills.

- **Putting power back in users' hands:** To further inform and empower the public to make its own decisions about which sources to trust, underpinning all of this must be a greater emphasis on online media literacy. Building new tools that help users identify and avoid online harms can increase user awareness of, and resilience to, misinformation and other online harms.

NewsGuard labels are more effective than Wikipedia descriptions

NewsGuard's Nutrition Labels are unique in that they are the only source of vetted, detailed background information on more than 8,400 different news websites that cover 95% of engagement with news and information in the U.S., the U.K., Canada, France, Germany, Italy, and Austria. The Nutrition Label format is analogous to the format of a Wikipedia entry. However, Wikipedia entries are not subject to the same level of editing rigor, nor do they cover many obscure and lesser-known news sources reviewed by NewsGuard. A July 2018 academic study by researchers at Indiana University compared the effectiveness of NewsGuard Nutrition Labels vs. Wikipedia entries when displayed next to news articles on Facebook.[55]

- The study found that while **NewsGuard's Nutrition Labels were likely to cause readers to trust reliable news sources and to distrust unreliable sources**, Wikipedia entries alone had almost no effect on users' behavior or trust in news.
- For users who clicked through to read a full Nutrition Label or Wikipedia entry, the study found **Nutrition Labels had a strong effect on users' trust in a news source.**
- **Green Nutrition Labels had nearly double the effect of positive Wikipedia entries** in making users trust their news sources.
- **Red Nutrition Labels had 133X the effect of negative Wikipedia entries on users' trust in news**. Wikipedia's impact was not statistically significant, while NewsGuard's impact was. ($p \geq .01$)

[55]July 2018 study conducted by Professor Alan Dennis, et al., of Indiana University

Research shows that using AI to combat misinformation often doesn't work

Artificial intelligence has been advanced as a potential solution for detecting misinformation at scale, but researchers warn that such technology fails to capture the evolving nuances of misinformation, which often parades as legitimate news. Anjana Susarla, Associate Professor of Information Systems at Michigan State University, wrote in an article for The Conversation that AI systems are:

- ill-equipped to cope with the shifting tactics of misinformation purveyors;
- fail to register nuance and context for words and how meanings may change over time; and
- worst of all, amplify harmful existing racial and gender stereotypes and biases.[56]

"The best way to combat the spread of fake news may be to depend on people," she wrote. "The societal consequences of fake news – greater political polarization, increased partisanship, and eroded trust in mainstream media and government – are significant." Professor Susarla also suggested that social media sites like YouTube and Facebook could "voluntarily decide to label their content, showing clearly whether an item purporting to be news is verified by a reputable source," demonstrating a clear use case for NewsGuard's Reliability Ratings and Nutrition Labels.

Furthermore, AI can actually be weaponized to *create* fake news, thus putting technology "in an arms race with itself."[57] Research from Tal Schuster, an MIT artificial intelligence scientist in 2019 found that efforts to detect fake news using AI are "not as advanced as they would appear," particularly since the leading methods of detection rely on pattern detection that "can itself be exploited by malicious actors."[58] Reporting that AI required extremely specific examples of human-generated or machine-generated claims to evaluate in order to be trained, the researchers concluded that "a lot more work will be needed to move AI beyond pattern recognition and toward something that can stand up to algorithms in malicious hands."

[56]https://theconversation.com/how-artificial-intelligence-can-detect-and-create-fake-news-95404
[57]https://www.vice.com/en/article/d3wd3k/gfycat-fighting-ai-porn-deepfakes-fake-news
[58]https://www.zdnet.com/article/head-fake-mit-says-fake-news-detection-is-not-what-it-appears/

If asked, people would pay for ratings and Nutrition Labels

At the request of NewsGuard licensees in the internet provider and mobile phone provider industries, NewsGuard commissioned consumer research to determine the value of having its ratings integrated into products that carry news.[59] The survey, conducted between June and July of 2019, questioned a representative sample of adults in the U.S., finding that:

- **86% of respondents think false information or misinformation is a problem** on the internet today (65% think it is a significant problem; 20% think it is a minor problem)
- **70% would find such a service valuable** (34% would find it very valuable; 36% would find it somewhat valuable)
- **59% would trust their provider more just by making these ratings and labels part of their consumer offering** (35% would trust their provider a lot more; 24% would trust their provider a little bit more)
- If asked by an internet or mobile phone provider to pay extra for this service...
    - **36% would be very likely or fairly likely to pay $1.95 per month**
    - **35% would be very likely or fairly likely to pay $2.49 per month**
    - **34% would be very likely or fairly likely to pay $3.49 per month**

[59]NewsGuard-commissioned YouGov survey conducted between June 27, 2019 and July 1, 2019. The survey was carried out online, and the sample size was 1,262 US adults.

Perspectives from NewsGuard browser extension users

In August 2021, NewsGuard conducted a user feedback survey and held focus groups with its browser extension users. Results from the 320 respondents were as follows:

- 96% of surveyed users agree that NewsGuard's source rating tool makes it easier to understand the reliability of online news.
- 9 out of 10 users would recommend NewsGuard to a friend.
- 97% of surveyed users said they believe NewsGuard's ratings are accurate.

Respondents reported that NewsGuard helped them feel "protected from junk news," educated them on unfamiliar news sources in a fact-based and neutral manner, and made it easier for them to quickly vet the reliability of websites without having to conduct independent research themselves. They said:

- "I have a few relatives that reference political sites I don't recognize most of the time. Practically always, NewsGuard has evaluated them, and when they are untrustworthy, I feel confident ignoring the viewpoint as far as truth value."
- "It is simultaneously easy to get a 'quick' rating but also go in depth to understand how NewsGuard got that rating in the first place." "NewsGuard does a really good job of neutrality. I feel I can post a Nutrition Label in response to a dubious claim and it doesn't generate the 'Oh, they're biased!' responses that you often get."
- "I've never had a solid and consistent source for evaluating news sources before. NewsGuard is a game changer for research and everyday 'well-informed living.'"
- "My time (and patience) is limited. NewsGuard helps me cut through the smoke and mirrors of partisanship and get to what's relevant."
- "My favorite thing about NewsGuard is the fact that the ratings are produced by humans."
- "Clear non-partisan site ratings help me find good information for those who may not hold the same political views as me."

What do educators think?

NewsGuard has been adopted by more than 800 public libraries and schools across the world, where the browser extension is installed on computers for students and visitors.

In a small survey of staff at two of the first libraries to adopt NewsGuard — the Los Angeles Public Library and the Hawaii State Public Library system — respondents overwhelmingly reported that the tool was useful for them and their patrons.

"NewsGuard is far and away the best resource we have found that lets people evaluate just what the internet is bringing to their screens. That NewsGuard uses upfront criteria and highly qualified professionals in evaluating websites — and makes these recommendations without hindering access in any way — is just a real benefit to anyone using the internet today."

– Sean Farrell, Director of the Library of Hattiesburg Petal and Forrest County, a NewsGuard library partner in Mississippi

Specifically:

- **91% of librarians strongly agreed or agreed that NewsGuard helps *them* personally** better understand news and information websites.
- **89% of librarians strongly agreed or agreed that NewsGuard helps their patrons** or students better understand news and information websites.
- **80% said they would like social media and search engines to include NewsGuard** ratings and write-ups in their news feeds and search results.

Can NewsGuard have an impact on the media industry?

One of NewsGuard's founding missions was to "restore trust in the media." To that end, **more than 2,100 of the news websites NewsGuard has engaged with have improved their practices** to meet one or more of NewsGuard's nine criteria. For example, GateHouse Media, a large publishing company operating more than 140 local news brands, from Austin, Texas to Akron, Ohio, began to more prominently label sponsored content to comply with the "clearly labels advertising" criterion.[60] Al-Jazeera disclosed its control by the Qatar government after engaging with NewsGuard analysts. Likewise, Reuters and the Times of London improved their transparency practices.

NewsGuard ratings also impact the media industry by restoring digital advertising revenue to trustworthy media outlets that have historically been excluded or overlooked by advertisers. Major advertisers are now using NewsGuard's BrandGuard product to allow their millions of dollars in programmatic advertising to support legitimate online news organizations while not funding hoax healthcare or propaganda sites.

[60]https://www.reuters.com/article/us-media-newsguard/newsguards-real-news-seal-of-approval-helps-spark-change-in-fake-news-era-idUSKCN1PQ5FV

How NewsGuard is transforming the world of advertising

NewsGuard's applications extend beyond supporting media literacy efforts and informing users about the quality of their news diets. Brands, advertising agencies, and ad-tech platforms leverage NewsGuard's human-vetted data to ensure that advertisers proactively invest in quality journalism and avoid funding misinformation through programmatic advertising placements on hoax websites.

In one case study, a **top-five programmatic advertiser** used NewsGuard to enhance its brand safety operation and **flagged or blocked 6.9 million ad impressions on hoax and misinformation websites** (including foreign propaganda sites, "health" sites pushing anti-vaccine content, and more).[61]

In another example from May 2021, NewsGuard worked with an advertising agency within IPG Mediabrands to leverage its Reliability Ratings for an IPG client — **a Fortune 500 company** that had previously removed much of its advertising from news content using broad keyword blocking and a narrow allow list.[62]

Working with NewsGuard, the brand **added 1,259 highly trusted news sites to its allow list**, including 589 sites with perfect 100/100 trust scores. Campaign performance improved on key metrics simply from advertising on more credible news sites. Specifically, the company **expanded the reach of its campaign by 20%** by adding new, trusted sites to its allow list. These trusted news sites enabled **143% higher click-through rates** by reaching engaged, attentive news audiences. The advertiser was also able to lower costs, resulting in **9% lower CPMs** (cost per 1,000 impressions) overall.

These interventions are important because research suggests that **52% of adults in the U.S. and 51% in the U.K. believe that brands always know exactly which websites their advertising appears on.**[63] Forrester found that **73% of business-to-consumer marketers said that the spread of misinformation and disinformation will impact their marketing strategy in 2022,**[64] and recommends that marketers incorporate NewsGuard's data into inclusion lists as a matter of priority.[65]

---

[61]https://www.newsguardtech.com/press/newsguards-first-year/
[62]https://www.newsguardtech.com/press/newsguard-case-study-advertising-solution/
[63]https://www.forrester.com/report/the-trust-imperative/RES164983?utm_source=forbes&utm_medium=pr&utm_campaign=b2cm
[64]https://www.forrester.com/blogs/the-entire-ad-supply-chain-is-at-fault-for-disinformation/?utm_source=forbes&utm_medium=pr&utm_campaign=b2cm
[65]https://www.forrester.com/report/funding-truth-in-the-misinformation-age/RES177496

Crafting Research for Target Consumers About Misinformation and NewsGuard

A TEMPLATE FOR FURTHER RESEARCH

Companies seeking to perform further research to gauge their customers' concerns about misinformation and their interest in a service such as NewsGuard can use the following template language to describe NewsGuard's service when they design such a study.

For accurate results, researchers should avoid simply asking consumers how they would value a service that rates the reliability and trustworthiness of news sites. Instead, researchers should describe at least generically the NewsGuard approach of journalists rating journalistic sites based on journalistic criteria to explain the process for providing ratings and Nutrition Labels. (Research shows that without explaining the journalistic nature of the process, consumers assume that a social media or other digital platform company is doing the work using a secret, non-transparent algorithm, and there is broad public distrust of these companies and this kind of process).

- **Recommended baseline question to assess the seriousness of the problem:**

  "*To what extent, if at all, do you agree with the following statement: "Online misinformation and hoaxes are a problem?"*

  Responses can be measured using a Likert scale or a sliding scale ranging from "strongly disagree" to "strongly agree" or from "not a problem at all" to "a significant problem."

- **Recommended language for brief description and question about the ratings and Nutrition Labels:**

  NewsGuard relies on experienced journalists with varied backgrounds to assign ratings to news sources, using nine basic, apolitical journalistic criteria of credibility and transparency. With this understanding of how this

team of journalists rates journalistic sites using basic, apolitical journalistic criteria, how useful would these ratings and Nutrition Labels be to help people avoid misinformation and hoaxes?

- **Recommended language for more detailed description and question about the ratings and Nutrition Labels:**

  A company called NewsGuard uses analysts who are experienced journalists with varied backgrounds to assign ratings and create a Nutrition Label for news websites. These ratings are based on nine apolitical journalistic criteria of credibility and transparency. These tell people how well each news site follows basic journalistic practices, providing Green and Red icons and Nutrition Labels indicating which sites are generally reliable and which sites are not generally reliable. With this understanding of how this team of journalists rates journalistic sites using basic, apolitical journalistic criteria, how useful would these ratings and Nutrition Labels be to help people avoid misinformation and hoaxes?

Current research/resources referencing NewsGuard

Examples of books and research citing NewsGuard include:

- **"Social media sharing of low-quality news sources by political elites," published by researchers at** Graz University of Technology, the University of Bristol, and the University of Western Australia

  This paper uses NewsGuard's credibility ratings to quantify the extent to which politicians on both sides of the political spectrum share untrustworthy, Red-rated domains on social media in the U.S., U.K., and Germany. The researchers found that Republican politicians in the U.S. have been increasingly sharing links to untrustworthy sources at a higher rate than Democrat politicians, and that this divergence has widened since the election of President Biden. They stress that information shared by political elites shapes citizen and media discourse, illustrating the importance of sharing high-quality sources to preserve democracy. The researchers also find that politicians in the U.K. and Germany share very low counts of links to untrustworthy domains overall, and that the rapid decline of information quality shared by Republican politicians is a phenomenon unique to the U.S.[66]

- **News Credibility Labels Help Improve The Reading Choices of People With The Worst News,**[67] published by researchers at NYU's Center for Social Media and Politics (CSMaP) in May 2022.

  Between May and June 2020, researchers found that NewsGuard's browser extension had a strong and statistically significant effect among participants with the worst news diets, who were much more likely to read news from reliable Green-rated sources after having installed the extension.[68] For the heaviest consumers of misinformation, there was a statistically significant increase in the reliability of their news diets after using NewsGuard's browser extension in the 2-3 week treatment period.

[66]https://academic.oup.com/pnasnexus/article/1/4/pgac186/6695314
[67]https://csmapnyu.org/research/news-credibility-labels-have-limited-average-effects-on-news-diet-quality-and-fail-to-reduce-misperceptions
[68]https://csmapnyu.org/research/news-credibility-labels-have-limited-average-effects-on-news-diet-quality-and-fail-to-reduce-misperceptions

Research data illustrated that 94% of participants in the group who installed NewsGuard felt neutral or positive toward the extension, and 41% liked the extension "a little" or "a lot." Researchers added that NewsGuard is "likely the most comprehensive and transparent" web extension that offers expert ratings on the credibility of news sources.

- **"Political Advertisement and Coordinated Behavior on Social Media in the Lead-Up to the 2021 German Federal Elections,"**[69] published by researchers at the University of Vienna, the University of Urbino Carlo Bo, and the University of Sassari.

  Researchers found that in the run-up to the German Federal Election in 2021, all the main political parties used Coordinated Link Sharing Behavior, which describes public pages and groups in a network repeatedly sharing the same links simultaneously to maximize engagement and reach.

  NewsGuard supplied the researchers with the 18 domains that spread misinformation about the German election as part of our German Federal Election Misinformation Tracker, who revealed that 15 out of 18 of these domains were used by these coordinated networks.

  Further analysis revealed that anti-establishment networks and fan-groups of the far-right Alternative for Germany party (AfD) disseminated the largest number of NewsGuard Red-rated sites using this technique, proliferating anti-vaccination, anti-lockdown, and anti-climate protection content specifically.

- **"News consumption and social media regulations policy,"**[70] published by researchers at La Sapienza University in June 2021.

  In this paper, researchers from La Sapienza investigated the opposing content moderation methods employed by social media platforms Twitter and Gab, with respect to approximately three million pieces of news-related content. According to the study, Twitter users tend to interact with misinformation less than users on Gab, suggesting that platforms with content regulation policies help limit the spread of misinformation. The study also found that users on Twitter and Gab tend to surround themselves with users that share common beliefs. Researchers

[69]https://www.medienanstalt-nrw.de/fileadmin/user_upload/NeueWebsite_0120/Zum_Nachlesen/BTW22_Political_Advertisement.PDF
[70]https://www.researchgate.net/publication/352244765_News_consumption_and_social_media_regulations_policy

used NewsGuard's Source Credibility ratings to distinguish between legitimate and misinformation sources, and found that the lack of content policies on Gab were associated with the circulation of misinformation and the proliferation of harmful narratives on the platform.

- **"When Do Sources Persuade? The Effect of Source Credibility on Opinion Change,"[71]** published by Cambridge University Press in The Journal of Experimental Political Science.

  Researchers Bernhard Clemm von Hohenberg and Andrew Guess constructed a fictional news credibility company "MediaCheckup" that uses nine criteria to measure news credibility. They also created a fictional media outlet and asked the treatment groups how favorable, trustworthy, and biased they judged the source to be, with some participants told that the source was credible and others told that it was not credible according to source rating criteria.

  The group that was told that the source had a high credibility rating reported increased favorability and trust towards the fictional media outlet, perceptions that remained six days after learning about the credibility of a source.

- **"The Technical Feasibility of Middleware as a Remedy to the Online Harms of Platform Moderation at Scale,"[72]** published by a researcher at U.C. Berkeley.

  This paper explores the scope of middleware to curate and moderate content online. It cites the NewsGuard browser extension as a successful example of use-case targeted middleware, adding that "the manual moderation by trained professionals allows for sophisticated categorization, including controlling for satire (something that is challenging for Machine Learning approaches)."

- **"Does Fake News in Different Languages Tell the Same Story? An Analysis of Multi level Thematic and Emotional Characteristics of News about COVID 19,"[73]** published by researchers at the University of North Carolina and Fairfield University.

[71]https://www.cambridge.org/core/journals/journal-of-experimental-political-science/article/when-do-sources-persuade-the-effect-of-source-credibility-on-opinion-change/48ECC9B706B2C3BA733936BE184917CC
[72]https://sbarrington.com/researchprojects/2021/12/22/middleware-for-algorithmic-moderation-at-scale
[73]https://link.springer.com/article/10.1007/s10796-022-10329-7

This paper uses state-of-the-art modeling techniques to investigate whether fake news in English and Mandarin use similar thematic, emotional, and semantic features. They use NewsGuard's Covid-19 Misinformation Tracker Center data to identify sources of English fake news and find that English and Chinese misinformation both show high levels of anger, as well as being identical in terms of topic concentration and uncertainty. However, there are some differences: Chinese misinformation sources evoke higher levels of sadness, whereas English ones evoke more anxiety.

- **"Understanding Engagement with U.S. (Mis)Information News Sources on Facebook,"**[74] published by researchers at New York University and Université Grenoble Alpes in November 2021.

  This research studied how often Facebook users engage with (mis)information from U.S. news providers on Facebook, using NewsGuard data and data from Media Bias/Fact Check to generate a list of news publishers' official Facebook pages and categorize them based on reliability and political leaning. Researchers found that in absolute terms, Far-Right misinformation sources accumulate more engagement than non-misinformation sources of the same partisanship (68.1 % of overall Far-Right engagement, followed by 37.7 % on the Far-Left).

- **"The Big Lie and Big Tech: Misinformation Repeat Offenders and Social Media in the 2020 U.S. Election,"**[75] published by the Carter Center in October 2021.

  This report by the Carter Center details the role played by "repeat offenders" — media outlets known to repeatedly publish false and misleading information — in spreading election fraud narratives in online echo chambers during the 2020 election. The Carter Center used NewsGuard to compile a list of "repeat offender" sources that often amplify misinformation. The Center analyzed 2.93 million posts in 883 Facebook groups engaged in partisan political discourse and found repeat-offender content in 76% of all groups —

[74]https://dl.acm.org/doi/pdf/10.1145/3487552.3487859
[75]https://www.cartercenter.org/resources/pdfs/news/peace_publications/democracy/the-big-lie-and-big-tech.pdf

and in 97% of right-leaning groups — between Aug. 17, 2020, and Jan. 20, 2021.

- **"The Advantage of the Right in Social Media News Sharing,"**[76] a report authored by researchers at the University of Pennsylvania, Fondazione Bruno Kessler, and the University of North Carolina at Chapel Hill, published in August 2021.

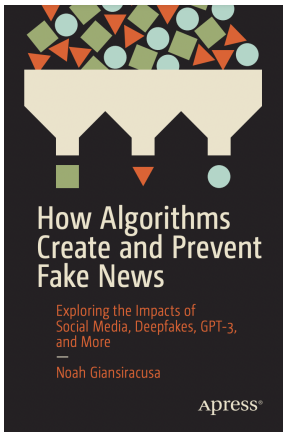
  This paper presents the findings of an analysis of the news sharing habits of Twitter users during mass mobilizations against racial injustices in the wake of the killing of George Floyd in the summer of 2020. The researchers analyzed the credibility and ideological slant of sources shared on the platform and found that, while there was no evidence that unreliable sources received greater visibility online during this period, there was evidence that sources espousing conservative views were more widely shared.

- **"Flow of Online Misinformation During the Peak of the COVID-19 Pandemic in Italy,"**[77] a report authored by researchers at Ca' Foscari University of Venice, European Centre for Living Technology, IMT School For Advanced Studies Lucca, the Institute of Informatics and Telematics, and the National Laboratory for Cybersecurity, published July 2021.

  In this study, researchers used NewsGuard data and Twitter activity of Italian users during the COVID-19 pandemic to assess the impact of online misinformation.

- **"How Algorithms Create and Prevent Fake News,"**[78] Noah Giansiracusa, published 2021.

  "How Algorithms Create and Prevent Fake News" is a broad account of the various ways that data-driven algorithms have distorted reality and rendered the truth harder to grasp. The book cited NewsGuard's January 2021 report on ad placement and revenue on sites publishing misinformation as evidence of the scale of



[76]https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3895410
[77]https://epjds.epj.org/articles/epjdata/abs/2021/01/13688_2021_Article_289/13688_2021_Article_289.html
[78]https://www.apress.com/gp/book/9781484271544

fake news as a lucrative business that rakes in ad dollars from major companies — often without their knowledge.

- **"Cognitive defense of the Joint Force in a digitizing world,"**[79] Nicholas D. Wright, published 2021.

  In a 2021 report on how to defend Joint Force members from misinformation threats, author and Georgetown professor Nicholas D. Wright recommended that all members of the armed forces be given access to NewsGuard. The report describes how adversaries and other destabilizing forces use online misinformation to sow discord and undermine defense forces, highlights the impact of new technologies including AI and deepfakes, and advocates for a more balanced commercial spending approach on offensive and defensive intelligence technology.

- **"Twitter and Facebook posts about COVID-19 are less likely to spread false and low-credibility content compared to other health topics,"**[80] a report by researchers at George Washington University, University of Pittsburgh, University of Memphis, the University of Maryland, College Park, and John Hopkins University, last updated March 2021.

  This research used NewsGuard's ratings to analyze the presence of low-credibility sources in hundreds of millions of Twitter and Facebook posts during the first months of the COVID-19 pandemic. The analysis found that posts about COVID-19 actually linked to proportionally more trustworthy sources than posts about other health topics beyond the virus.

- **"Political audience diversity and news reliability in algorithmic ranking,"**[81] a report by researchers at the universities of Dartmouth, Indiana, and South

[79] https://static1.squarespace.com/static/5c646d66815512fdad0ed1d3/t/61291b17c03bee1a4a46b6ea/1630083866594/cognitive+defense.pdf
[80] http://arxiv-export-1b.library.cornell.edu/abs/2007.09682
[81] https://arxiv.org/abs/2007.08078

Florida, last updated March 2021.

Researchers used NewsGuard's Reliability Ratings to develop improved content recommendation algorithms and study how to improve the diversity of information circulating on social media. The study demonstrated that websites with more extreme and less politically diverse audiences have lower journalistic standards, and found that partisan audience diversity is a valuable signal of higher journalistic standards that should be incorporated into algorithmic ranking decisions.

- **"Cracking Open the News Feed: Exploring What U.S. Facebook Users See and Share with Large-Scale Platform Data,"**[82] a report authored by researchers at Princeton University and the New York University Center for Social Media and Politics in January 2021.

  Researchers analyzed engagement data covering millions of web links shared on Facebook to describe how and by which categories of U.S. users different types of news are seen and shared on the platform. The research focused on articles from low-credibility news publishers, credible news sources, purveyors of clickbait, and news specifically about politics, which was identified using NewsGuard data as a signal. Results supported findings that more fake news is shared by older users and conservatives, and that both viewing and sharing patterns suggest a preference for ideologically congenial misinformation.

- **"'Thought I'd Share First' and Other Conspiracy Theory Tweets from the COVID-19 Infodemic: Exploratory Study,"**[83] a report by researchers at Los Alamos National Laboratory published in April 2021.

  Researchers used Twitter data to explore methods to characterize and classify four COVID-19 conspiracy theories and to provide context for each of these conspiracy theories through the first five months of the pandemic. They used NewsGuard data on domain credibility and websites related to COVID-19 myths as features in classification models. Research showed that misinformation tweets demonstrate more negative sentiment when compared to non-misinformation tweets, and that theories evolve over

[82]https://journalqd.org/article/view/2586
[83]https://publichealth.jmir.org/2021/4/e26527/authors

time, incorporating details from unrelated conspiracy theories as well as real-world events.

- **"Facebook's Algorithm: A Major Threat to Public Health,"**[84] a report published by international non-profit and campaign group Avaaz, in April 2020.

  Avaaz licensed NewsGuard data to show how 82 websites spreading health misinformation racked up 460 million estimated views on Facebook during the COVID-19 pandemic. Avaaz found that content from the top 10 websites spreading health misinformation reached four times as many views on Facebook as equivalent content from the websites of 10 leading health institutions, such as the WHO and CDC.

Iffy Quotient

- **"Iffy Quotient: A Platform Health Metric for Misinformation,"**[85] an ongoing project by the University of Michigan's Center for Social Media Responsibility.

  Social media sites and search engines have become the de facto gatekeepers of public communication, a role once occupied by publishers and broadcasters. With this new role come public responsibilities, including limiting the spread of misinformation. Researchers at the University of Michigan developed the Iffy Quotient using NewsGuard data as a way to measure the progress of media platforms at meeting their public responsibilities. The term "Iffy" is used to describe sites that frequently publish misinformation. The researchers behind the metric put out periodic reports summarizing trends in information spread over the prior months, including:

- **"Racism, election overtake COVID-19 as 'iffy' news on popular social sites,"**[86] November 2020.

  The Iffy Quotient researchers found that content about racism, protests, and riots published by "iffy" websites appeared on Facebook and Twitter

[84]https://secure.avaaz.org/campaign/en/facebook_threat_health/
[85]https://csmr.umich.edu/media/docs/UMSI-CSMR-Iffy-Quotient-Whitepaper-v2.pdf
[86]https://news.umich.edu/racism-election-overtake-covid-19-as-iffy-news-on-popular-social-sites/

nearly three times as much as COVID-19, according to NewsWhip. They also found that URLs of stories related to the U.S. presidential election outpaced those related to the pandemic. At the same time, the pandemic was the lead topic on the two platforms among news stories from "OK" news sites — those that adhere to basic standards of credibility and transparency.

- **"People 'fly to quality' news on social sites when faced with uncertainty,"**[87] July 2020.

  In this report, the researchers found a reduction in the Iffy Quotient that coincided with the onset of the pandemic: fewer of the popular URLs on both Facebook and Twitter came from "iffy" sites. They found a small correlation with an associated surge in sharing of articles from mainstream sources, which they said might be interpreted as a flight to quality in uncertain times. Researchers described this "flight to quality" as similar to the behavior people exhibit when financial markets are volatile and they gravitate towards investing in gold or more conservative investments.

- **"U-M's Iffy Quotient shows steady drop of questionable information on social media, partners with NewsGuard for better data,"**[88] July 2019.

  In July 2019, researchers reported a drop in the Iffy Quotient of unreliable websites shared on Facebook and Twitter, with Facebook dropping to 7.2% on July 1 from 12.2% on October 1, 2018, and Twitter falling slightly from 11.1% to 10.9%. The Center also announced its partnership with NewsGuard in an effort to make the Iffy Quotient more transparent and comparable over time, using NewsGuard's Reliability Ratings data to help researchers track information reliability based on NewsGuard's rating of a website.

German Marshall Fund

- **"Social Media Engagement with Deceptive Sites Reached Record Highs in 2020,"**[89] a policy insights post by the German Marshall Fund of the United States, published in January 2021.

[87]https://news.umich.edu/people-fly-to-quality-news-on-social-sites-when-faced-with-uncertainty/
[88]https://news.umich.edu/u-ms-iffy-quotient-shows-steady-drop-of-questionable-information-on-social-media-partners-with-newsguard-for-better-data/
[89]https://www.gmfus.org/news/social-media-engagement-deceptive-sites-reached-record-highs-2020

The GMF's "Digital New Deal" project partnered with NewsGuard to monitor two kinds of deceptive sites: sites that repeatedly publish provably false content and sites that fail to gather and present information responsibly. Research found that on Twitter, shares by verified accounts of content from deceptive sites reached an all-time high in the fourth quarter of 2020. Such content received 47 million verified account shares, nearly one-third of the total 155 million verified account shares of links to U.S.-based sites.

- **"Health Sites Built Coordinated Networks of Facebook Pages to Spread False Content, Increase Ad Revenue,"**[90] research by the German Marshall Fund of the United States' Digital New Deal, published in September 2020.

  The research identified the ownership and revenue structure of five health websites — all of which repeatedly publish false content, according to NewsGuard — that seem to promote health misinformation on Facebook through a coordinated network of pages. Together, these five sites received 71.1 million interactions on Facebook in the first 11 months of 2020.

- **"New Study by Digital New Deal Finds Engagement with Deceptive Outlets Higher on Facebook Today Than Run-up to 2016 Election,"**[91] research by the German Marshall Fund of the United States' Digital New Deal published in October 2020.

  Research in association with NewsGuard found that the level of engagement with articles from outlets that repeatedly publish verifiably false content had increased 102 percent since the run-up to the 2016 election. In addition, engagement with another set of sites that fail to gather and present information responsibly — especially Fox, Daily Wire, and Breitbart — had grown 293 percent. Interactions with articles from both kinds of deceptive sites had increased by 242 percent between the third quarter of 2016 and the third quarter of 2020.

- **"The Rise of the Parapolitical Sites as the Leading False-Content Producers,"**[92] a report published in January 2020 by the German Marshall

[90]https://adigitalnewdeal.org/work/health-sites-built-coordinated-networks-of-facebook-pages-to-spread-false-content-increase-ad-revenue-2/
[91]https://www.gmfus.org/news/new-study-digital-new-deal-finds-engagement-deceptive-outlets-higher-facebook-today-run-2016
[92]https://www.gmfus.org/news/rise-parapolitical-sites-leading-false-content-producers

Fund, discussing findings from research conducted in association with NewsGuard and Newswhip data.

The GMF found a shift in the type of false content publications that were dominating social circulation, from the overtly political sites to new "parapolitical" health and lifestyle sites that are less overtly political and share misinformation in a more nuanced manner. Researchers found that not only did the number of public interactions with the top ten sites that repeatedly publish false content rise from 2016 to 2019, but also, in 2019, the top ten sites publishing false content contained fewer sites publishing explicitly political content than in 2016, and more parapolitical sites focused on health and wellness.

- **"New Report: U.S. Information Ecosystem Woefully Unprepared Amidst Coronavirus and 2020 Election Disinformation,"**[93] published in March 2020.

  The analysis, performed in conjunction with NewsGuard data, showed that overall user interaction with the top ten outlets categorized by NewsGuard as repeatedly sharing false content had not only increased since 2016, but that 8 out of 10 of these unreliable outlets were pushing misleading or outright false articles about the coronavirus. Report authors concluded that a new policy response is urgently needed to supplement disparate platform-driven responses, instead of leaving it to fact-checkers, journalists, and researchers to call out bad behavior.

[93]https://www.gmfus.org/news/new-report-us-information-ecosystem-woefully-unprepared-amidst-coronavirus-and-2020-election