MITIGATING MEDICAL MISINFORMATION:

A WHOLE-OF-SOCIETY APPROACH TO COUNTERING SPAM, SCAMS, AND HOAXES

By Joan Donovan, PhD, Brian Friedberg, Gabrielle Lim, Nicole Leaver, Jennifer Nilsen, and Emily Dreyfuss

tasc Technology and Social Change Project

HARVARD Kennedy School
SHORENSTEIN CENTER
on Media, Politics and Public Policy

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Media manipulation is a specific and pernicious problem exacerbating what the World Health Organization has termed the "infodemic."[1] Though it can often result in the spread of harmful medical misinformation, media manipulation is understudied and therefore currently remains a major vulnerability for public health. During the SARS-CoV-2 (COVID-19) pandemic, media manipulation campaigns focused on pushing harmful medical misinformation have targeted health experts, undermined legitimate advice, and sowed mistrust at a time when reliable information is needed the most. By grabbing the public's attention during an information and health crisis, media manipulation that spreads medical misinformation can result in people not following expert advice, which, in turn, can exacerbate the spread of the coronavirus and result in illness and death.

Media manipulation is a sociotechnical process, whereby motivated actors leverage conditions or features within an information ecosystem to manipulate the public using the press, networked technology systems, and influential organizations to advance their agenda. It is undertaken by both state and non-state actors. It can have widespread impacts on societies across the globe.

---

[1] World Health Organization, "An ad-hoc WHO technical consultation managing the COVID-19 infodemic: call for action," WHO, September 15, 2020, https://www.who.int/publications/i/item/9789240010314.

This brief addresses how the public health sector, along with a coalition of civil servants, media workers, technology companies, and civil society organizations, can understand and respond to the problem of medical media manipulation, specifically how it spreads online. Here we present a supplementary research-and-response method in correspondence with the World Health Organization (WHO)'s already suggested framework for dealing with the infodemic, with a focus on media manipulation.[2]

Compiled by an expert team of researchers at Harvard's Shorenstein Center on Media, Politics and Public Policy, this brief describes the intentional and directed phenomenon of media manipulation, provides case studies about media manipulation campaigns that spread medical misinformation, and offers actionable insights for how to track and counter their harms.

The following recommendations are grounded in a robust model: the Media Manipulation Life Cycle, which allows researchers and health officials to understand how, when and why to intervene in manipulation campaigns. All forms of media manipulation follow a stereotyped life cycle, which, once understood, can be recognized and mitigated. That life cycle has five stages:

1. **Campaign planning and origins**

2. **Seeding the campaign across social platforms and the web**

3. **Responses by industry, activists, politicians and journalists**

4. **Mitigation**

5. **Adjustments by manipulators to the new environment**

To understand exactly how to respond to a campaign, we provide a response matrix that public health stakeholders can use when encountering harmful media manipulation. An included situational analysis worksheet will help public health advocates determine the best course of action to take and when to undertake it.

This advice is based on years of research into how dangerous misinformation spreads, what techniques work to contain it, and how stakeholders can stop media manipulation campaigns before they can grow. This advice is designed to work within any cultural context. It is fluid, and reliant on the method of research known as *investigative digital ethnography*, which takes

---

[2] *An Ad Hoc WHO Technical Consultation: Managing The COVID-19 Infodemic: Call For Action, 7-8 April 2020* (Geneva: World Health Organization; 2020), https://www.who.int/publications/i/item/9789240010314.

into account the differences in geography, culture, language, law and demographic diversity, so that these recommendations can be tailored to specific environments as per the needs of the locale and situation.

Key actionable recommendations for public health officials include advice on how to:

- **survey, collect and document misinformation;**

- **determine where in the media manipulation life cycle misinformation falls;**

- **boost timely, relevant, local and accurate content;**

- **engage trustworthy press to correct misinformation when necessary and not before;**

- **build early warning networks across civil society, journalism and public health institutions;**

- **coordinate counter-messaging with civil society networks and technology companies;**

- **monitor the impacts of mitigation efforts and adjust messaging as needed.**

Along with short-term recommendations, we also recommend long-term strategies to combat media misinformation. These include:

- **understanding how political cleavages, wedge issues and socioeconomic issues affect the spread of disinformation;**

- **monitoring the community's perception of trust in the media and authorities to understand how best to address them;**

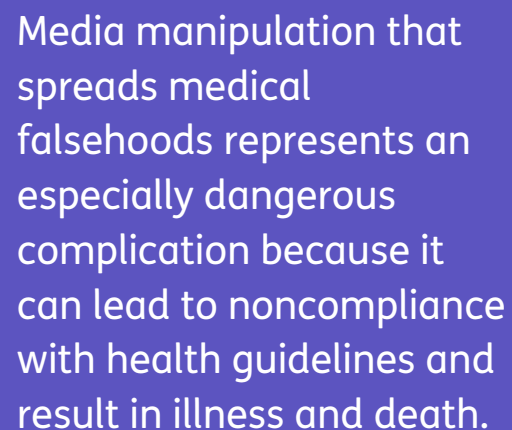- **continuously calling for transparency by technology companies.**

This brief calls for a "whole-of-society" networked response to medical media manipulation, where stakeholders in the private sector, public sector and civil society are able to work toward a common goal – to provide accurate health information amid an infodemic. What follows is a detailed policy report on how these groups can work together to minimize the deleterious effects of media manipulation on societies across the world.

BACKGROUND, CONTEXT, AND SCOPE

Introduction

As the SARS-CoV-2 (COVID-19) pandemic surges across the globe, so too have hoaxes, rumors and dangerous misinformation. As this false information intermingles with true and accurate content, the glut of information creates confusion. The World Health Organization (WHO) has classified this phenomenal complication of global health communications as an "infodemic": the rapid and far-reaching spread of both accurate and inaccurate information about a disease.[3] When misinformation in such conditions is left unchecked, as we saw with HIV Ebola virus disease, it poses a threat to public safety, economic security and global stability, enabling hate crimes and xenophobia, engendering distrust in journalism and medical institutions and, in some cases, causing death from disease. Not all misinformation within an infodemic is the same; some is the result of the harmful media phenomenon called *media manipulation* – the sociotechnical process where motivated actors leverage specific conditions or features within an information ecosystem to generate public attention through press coverage for events that would otherwise go uncovered or to create a false perception of public outrage.[4] These campaigns are not always intended to spread disinformation (which is intentional misinformation), but often they do. Media manipulation is not unique to the COVID-19 pandemic, but it is on the rise, and when the tactics of media manipulation are used to spread medical misinformation, the whole of society suffers the consequences.

Within the infodemic, media manipulation that spreads medical falsehoods represents an especially dangerous complication because it can lead to noncompliance with health guidelines and result in illness and death. For example, the disinformation campaign known as

> Media manipulation that spreads medical falsehoods represents an especially dangerous complication because it can lead to noncompliance with health guidelines and result in illness and death.

---

[3] World Health Organization, "An ad-hoc WHO technical consultation managing the COVID-19 infodemic: call for action."

[4] Becca Lewis and Alice E. Marwick, "Media Manipulation and Disinformation Online," Data & Society, Data & Society Research Institute, May 15, 2017, https://datasociety.net/library/media-manipulation-and-disinfo-online.

Operation Denver, which was run by the KGB at the height of the AIDS pandemic, may have contributed to the delayed response to the disease in South Africa, resulting in 300,000 additional deaths.[5] A more recent example is the viral slogan, "vaccines cause autism," one of the most influential memes of the modern era. Its popularity as a meme was due in part because of the timing of Andrew Wakefield's academic paper published in *The Lancet* (1998) coinciding with the widespread adoption of the internet in people's homes.[6] According to Heidi Larson, an anthropologist studying vaccine hesitancy globally, because this slogan was simple, repeatable, and sticky, social media became the perfect delivery mechanism for this memetic form of medical misinformation.[7]

In recent years, state and non-state actors alike have engaged in media manipulation by taking advantage of networked communication technologies to amplify false information, harassment and politically motivated false narratives. Newsrooms, technology companies, civil society, politicians, educators and researchers have been working to address and mitigate the related and resulting harms.

However, the range of policy options to counteract media manipulation are varied and their success is contingent on local contexts including laws, access to technology, institutional trust, freedom of the press and availability of resources to different stakeholders. To further complicate matters, media manipulation campaigns are a cross-sector problem, and actions from key stakeholders are uncoordinated, often resulting in ineffective solutions which can also exacerbate the problem. For example, while scientists were assessing COVID-19, populations were awaiting guidance on treatment. When politicians in the United States began hyping hydroxychloroquine before clinical trials, prescriptions increased dramatically, resulting in a global shortage.[8] During the COVID-19 pandemic there have been instances of health professionals or experts becoming media manipulators themselves,[9] engaging in pushing false information or attempting to inject the scientific literature with falsehoods,

---

[5] Mark Kramer. "Lessons from Operation 'Denver,' the KGB's Massive AIDS disinformation campaign," MIT Press Reader, May 26, 2020, https://thereader.mitpress.mit.edu/operation-denver-kgb-aids-disinformation-campaign/.

[6] A. J. Wakefield, S. H. Murch, A. Anthony, J. Linnell, D. M. Casson, M. Malik, M. Berelowitz, et al., "RETRACTED: Ileal-Lymphoid-Nodular Hyperplasia, Non-Specific Colitis, and Pervasive Developmental Disorder in Children.," *The Lancet* 351, no. 9103 (February 28, 1998): 637–41, https://doi.org/10.1016/S0140-6736(97)11096-0.

[7] Heidi Larson, *Stuck: How Vaccine Rumors Start and Why They Don't Go Away (*Oxford University Press, 2020), 10.

[8] Joan Donovan, "Social-media companies must flatten the curve of misinformation," *Nature,* April 14, 2020, https://www.nature.com/articles/d41586-020-01107-z.

[9] Jane Lytvynenko, Ryan Broderick, and Craig Silverman, "These are the fake experts pushing pseudoscience and conspiracy theories about the coronavirus pandemic," *BuzzFeed News,* May 21, 2020, https://www.buzzfeednews.com/article/janelytvynenko/coronavirus-spin-doctors.

which further complicates the public's ability to trust authoritative sources. (See our included case study in Appendix A on the *Plandemic* documentary for more information.)

The COVID-19 pandemic has exposed how social and technical conditions can combine to amplify potentially harmful narratives around health, which may be adopted or believed. What's more, this will not be the last time. Professor Larson suggests that to counter misinformation about vaccines, facts are an important aspect of understanding safety, but they are rarely convincing on their own.[10] There are ways to mitigate the adverse effects of media manipulation, but they require cross-sector coordination so that the public receives timely, local, relevant, and redundant information.

This research brief addresses how the public health sector, along with a coalition of civil servants, media workers, technology companies, and civil society organizations (CSOs), should understand and respond to the problem of media manipulation. Though there are many recommendations available for how to deal with the infodemic writ large, fewer resources exist to address media manipulation specifically. This brief intends to fill that gap by focusing directly on this intentional and directed phenomenon. In addition, recommendations included here are tailored only for cases of media manipulation, which are distinct from general health (risk) communications. Though we do include some potential long-term strategies, they should not be taken for overall public communication guidance.

LIMITATIONS AND ASSUMPTIONS

Because of the multi-platform and multi-stakeholder nature of media manipulation, limiting response to just one party or organization is often inadequate. Therefore, the recommendations below assume a "whole-of-society" networked response where the private

---

[10] "About Vaccine Confidence Project," Vaccine Confidence Project, accessed June 15th, 2020, https://www.vaccineconfidence.org/.

sector, public sector, and civil society are able to work toward a common goal.[11,12,13] However, depending on external factors and local context (e.g., legal and regulatory constraints, domestic information ecosystem, political feasibility, and available resources and skills), several of the recommendations below will be outside the scope of some health authorities and governments. As such, we have delineated possible responses into two columns (*see* Table 1): responses by the public health sector, and responses by external stakeholders, such as technology companies, media companies, and CSOs.

Furthermore, it is important to note that many of the actions discussed in this document assume a response from social media companies to quarantine misinformation. However, due to a dearth of legal regulations and uneven enforcement on the part of social media companies, there is no guarantee that private sector companies will respond to recommendations such as reporting content or asserting public pressure or that, if they do, it will happen quickly enough. In those cases, other interventions may be necessary. However, continual dialogue with technology companies has led to some gains in the past, such as Facebook, Twitter, and YouTube's redesign that directs users to credible health authorities (e.g. clear links to the United States Centers for Disease Control and Prevention, WHO, or local public health agencies), Reddit's quarantining of potentially harmful COVID-19-related content, and Twitter's removal and flagging of false and potential harmful tweets by celebrities and political leaders.[14]

Lastly, media manipulation operations, as intended, are deceptive in nature, often covert, and not always clear cut. Because of this, a degree of contextual interpretation is required when assessing a threat and formulating a response. Assumptions will likely have to be made, as all stakeholders are operating with uneven access to information. Continuous research through monitoring and evidence collection are therefore necessary to update any working assumptions and to evaluate whether interventions are effective.

---

[11] For two country case studies on "whole-of-nation" strategies to combat disinformation, see Taiwan and Sweden who have been battling state-sponsored influence operations from China and Russia (see Lien Y-T [Taiwan] and Cederberg G [Sweden]).

[12] Gabriel Cederberg, "Catching Swedish Phish: How Sweden is Protecting its 2018 Elections," Harvard Kennedy School Belfer Center, August 2018, https://www.belfercenter.org/sites/default/files/files/publication/Swedish%20Phish%20-%20final2.pdf.

[13] Yi-Ting Lien, "Why China's COVID-19 Disinformation Campaign Isn't Working in Taiwan," *The Diplomat*, March 20, 2020, https://thediplomat.com/2020/03/why-chinas-covid-19-disinformation-campaign-isnt-working-in-taiwan/.

[14] Elliott Hannon, "Twitter, Facebook Delete World Leaders' Misleading Coronavirus Posts. Could Trump Be Next?" *Slate*, March 31, 2020, https://slate.com/news-and-politics/2020/03/twitter-facebook-delete-brazil-bolsonaro-maduro-misleading-coronavirus-posts.html.

Challenges and opportunities for countering media manipulation

The recommendations below may be applied to different contexts and regions, and additional criteria must be factored in when evaluating whether an intervention is required and what measures are the most effective. For example, interventions must be in accordance with international human rights law, protect individual privacy, and ensure civil liberties are not unduly infringed. Furthermore, interventions must be localized by country to account for differences in governance, information ecosystems, language, diversity and quality of information sources, existing domestic and international laws, and available resources devoted to countering misinformation. While the full range of salient factors will likely differ from region to region, the following content outlines some of the high-level conditions that will inform the effectiveness and risks of potential interventions across environments.

When dealing with media manipulation and misinformation, simply presenting facts may not be effective.[15,16,17] Instead, counter-messaging strategies must consider the entire experience of COVID-19 and how it differs by age, region, language, and culture.

1. Levels of Internet adoption and other communication technologies

Internet adoption varies from region to region and, while media manipulation campaigns may begin online, they can easily be picked up by broadcast and print media. As such, the diversity of information sources and news consumption must be taken into account when formulating any counter-messaging. For example, Cofacts, a local fact-checking chatbot, was developed to fight in-app misinformation and hoaxes for a popular messaging app, Line.[18] Similarly, WhatsApp has rolled out an in-app chatbot, localized to each country and run by a local

---

[15] Fact checking remains a contested means for addressing false information. For more information on the effectiveness and challenges of fact checking, see Li et al. and Walter et al.

[16] Lianing Li and Michael W. Wagner, "When are readers likely to believe a fact-check?" *Brookings*, May 27, 2020, https://www.brookings.edu/techstream/when-are-readers-likely-to-believe-a-fact-check/.

[17] Nathan Walter, Jonathan Cohen, R. Lance Holbert, and Yasmin Morag, "Fact-Checking: A Meta-Analysis of What Works and for Whom," *Political Communication*, 37:3 (2020): 350-375, DOI: 10.1080/10584609.2019.1668894.

[18] "Cofacts Tutorial," Cofacts, accessed June 15, 2020, https://g0v.hackmd.io/s/rkVVQDmqQ.

organization.[19] Meanwhile, in South Africa, where 85% of the adult population listens to the radio for news, Africa Check utilizes the radio for messaging.[20]
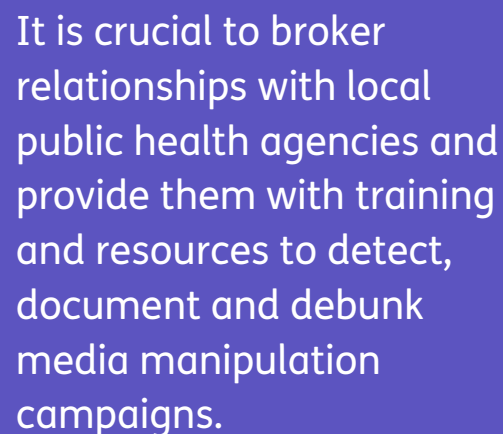
Key questions and factors to consider

- What is the primary means of peer-to-peer communication?
- What are the primary news sources?
- How do individuals and groups seek information and what technology do they rely on to do this?
- What percentage of the population is Internet connected and at what speeds?
- What apps, websites and devices are most used?
- How do the answers to the above questions differ among different demographic groups and languages?

2. Existing trust in local, national and international institutions

Increasingly, research has shown that mistrust is correlated to the spread and harmfulness of misinformation as well as the effectiveness of fact-checking. As such, communication about misinformation must consider not only why populations are mistrustful of certain institutions, but also the degree of perceived legitimacy and credibility for those institutions by different demographics. In highly partisan environments, this is a difficult task and requires careful coalition-building or shared points of trust.

For example, in a recent COVID-19 survey in the US, Gallup reported a significant partisan split between political parties.[21] In addition, a poll from academic researchers at Harvard, Northeastern and Rutgers confirms a partisan

> It is crucial to broker relationships with local public health agencies and provide them with training and resources to detect, document and debunk media manipulation campaigns.

[19] "IFCN Fact Checking Organizations on WhatsApp," Whatsapp, accessed June 15, 2020, https://faq.whatsapp.com/general/ifc-n-fact-checking-organizations-on-whatsapp.

[20] Alexios Mantzarlis, "South Africa's fact-checking site is turning to radio to break through the noise," *Poynter*, June 7, 2016, https://www.poynter.org/fact-checking/2016/south-africas-fact-checking-site-is-turning-to-radio-to-break-through-the-noise/.

[21] Zacc Ritter, "Republicans Still Skeptical of COVID-19 Lethality," *Gallup*, May 26, 2020, https://news.gallup.com/poll/311408/republicans-skeptical-covid-lethality.aspx.

split in trust along party lines, but state governments, the US CDC, hospitals, doctors and scientists remained trustworthy to both parties.[22] Crafting messaging that appeals to both groups through experts, local governments and medical professionals is more likely to be effective.

Therefore, it is crucial to broker relationships with local public health agencies and provide them with training and resources to detect, document and debunk media manipulation campaigns.

Similarly, trust in media sources must also be considered. In countries where mainstream media is controlled by the government, these outlets may have lost credibility with their intended readership due to years of biased reporting. As such, people may have turned to social media and messaging apps for their news, even though their trust in social media is even lower. Recognizing such nuances is necessary in crafting a cogent messaging strategy.

Key questions and factors to consider
- Which media outlets, politicians and organizations are considered the most trustworthy and by which demographics?
- Is this trust delineated between partisan lines, religious beliefs, ethnic identities, age or other demographics?
- Are there organizations or individuals beyond media, government and technology that are willing to engage in building information resilience and countering misinformation?
- What are the risks of crafting counter-messaging through low-trust institutions or organizations?

3. Finding partners, networks, and coalitions

Because most people's information sources are diverse, and trust in leaders, the government, and other institutions varies, it is important to identify potential partners, networks, and coalitions doing work on misinformation in your region. In certain environments, prebunking with accurate information may be warranted, which can help seed the community with accurate data that can possibly protect them from future misinformation. But, it is important that such proactive messaging comes from trusted sources, and is explicitly focused on

---

[22] Matthew A. Baum, Katherine Ognyanova, David Lazer, John Della Volpe, Roy H. Perlis, James Druckman, and Mauricio Santillana, "The State Of The Nation: A 50-State COVID-19 Survey Report #2," *The State Of The Nation: A 50-state Covid-19 Survey,* May 2020, https://covidstates.net/COVID19%20CONSORTIUM%20REPORT%20MAY%202020.pdf.

accurate information and does not include any mention of falsehoods.[23] Previous research has found that false information, when repeated, risks becoming sticky and more believable,[24] and even correcting that information may not dissuade a person's false beliefs.[25] Furthermore, individuals do not consume news from a single source.[26] Thus, it is important to consider how different individuals, organizations and agencies can collaborate to provide a timely, local, relevant and redundant communication strategy so that facts can rise above the noise.

In highly partisan environments, this will likely be difficult. Fact-checking organizations have been politicized,[27] and human rights advocates warn of "fake news" laws across the globe that are intended to silence dissent and government critique.[28] Coalitions in civil society have formed to counter these trends; for example, women's rights, anti-corruption, government transparency, journalism and filmmaking have come together to advocate for media freedom and accurate, responsible reporting. In such information ecosystems, key partners may be limited in the actions they can take, though broad partnerships between CSOs internationally may be possible. Exploring potential coalitions beyond organizations and individuals traditionally associated with the media, technology and government (such as local community groups, women's organizations, neighborhood associations, educational groups, LGBTQ advocacy groups and more) may therefore prove useful in effective coalition-building.

---

[23] Jon Roozenbeek, Sander van der Linden, and Thomas Nygren, "Prebunking interventions based on "inoculation" theory can reduce susceptibility to misinformation across cultures," *Harvard Kennedy School (HKS) Misinformation Review*, 1:2 (2000): https://misinforeview.hks.harvard.edu/article/global-vaccination-badnews/.

[24] Lisa K. Fazio, David G. Rand, and Gordon Pennycook, "Repetition increases perceived truth equally for plausible and implausible statements," Psychon Bull Rev 26 (2019): 1705–1710, https://doi.org/10.3758/s13423-019-01651-4.

[25] Lisa K. Fazio, Nadia M. Brashier, Keith Payne, and Elisabeth J. Marsh, "Knowledge Does Not Protect Against Illusory Truth," Journal of Experimental Psychology: General 144:5 (2015): 993–1002, http://dx.doi.org/10.1037/xge0000098.

[26] Andrew Guess, "(Almost) Everything in Moderation: New Evidence on Americans' Online Media Diets (Working Paper)," Princeton Center for the Study of Democratic Politics, accessed June 15, 2020, https://csdp.princeton.edu/publications/almost-everything-moderation-new-evidence-americans-online-media-dietsl.

[27] Adi Robertson, "Facebook fact-checking is becoming a political cudgel," The Verge, March 3, 2020, https://www.theverge.com/2020/3/3/21163388/facebook-fact-checking-trump-coronavirus-hoax-comment-politico-daily-caller.

[28] Molly Quell, "More Countries Pass 'Fake News' Laws in Pandemic Era," June 5, 2020, http://www.courthousenews.com/more-countries-pass-fake-news-laws-in-pandemic-era/.

Key questions and factors to consider

- Who would be most trusted to share proactive messaging that could function as an inoculation or "prebunk" against potential misinformation?
- What organizations or individuals are willing to participate and work with one another in improving access to accurate and credible information?
- What sectors are they from and what are their unique skills and audiences?
- Are there researchers and academics within universities who can collaborate on data collection, research and analysis?
- Are there additional partners beyond media, technology and the government that can help?

4. Availability of information and censorship

Countering media manipulation also requires a thorough understanding of existing censorship and information controls. This may include domain blocking at the Internet service provider level, content take-down orders, criminalization of certain types of information (e.g., hate speech, sedition, offensive content), and the degree of government transparency, press freedom and diversity of voices represented in the information ecosystem. In recent years, the rise in criminalization and regulation of "fake news" and false information has given some governments a legislative tool to take down content they disagree with, arrest oppositional and dissident voices, and target outlets that publish criticism of the government.

Malaysia's now-repealed Anti-Fake News Act, for example, was one of the first laws to explicitly target "fake news," and it came under heavy criticism from human rights organizations, journalists, and media outlets for its ability to selectively censor content and individuals the government did not approve of.[29] Likewise, Egypt has increased its crackdown on journalists through widespread arrests for the crimes of "spreading false information" and "misuse of social media."[30] Both Malaysia and Egypt, along with Hungary,[31] Nigeria,[32]

---

[29] Gabrielle Lim, "Securitize/Counter-Securitize: The Life and Death of Malaysia's Anti-Fake News Act," *Data and Society*, March 25, 2020, https://datasociety.net/library/securitize-counter-securitize/.
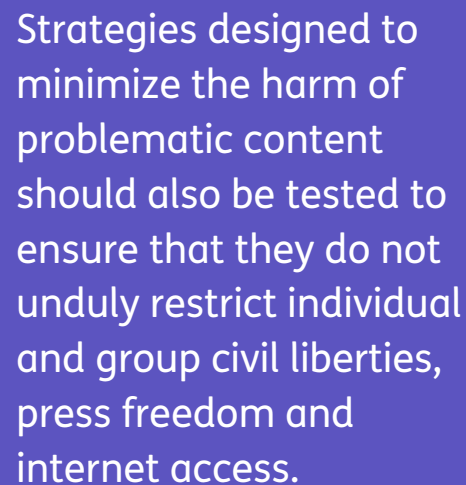
[30] "The Rise of Digital Authoritarianism in Egypt: Digital Expression Arrests from 2011-2019," Open Technology Fund, October, 24 2019 https://www.opentech.fund/news/rise-digital-authoritarianism-egypt-digital-expression-arrests-2011-2019.

[31] *"Defining Fake News in Hungary during Coronavirus Crisis," Deutsche Welle*, May 22, 2020, https://www.dw.com/en/defining-fake-news-in-hungary-during-coronavirus-crisis/av-53536776.

[32] Danielle Paquette, "Nigeria's 'fake news' bill could jail people for lying on social media. Critics call it censorship," *The Washington Post*, November 25, 2019, https://www.washingtonpost.com/world/africa/nigerias-fake-news-bill-could-jail-people-for-lying-on-social-media-critics-call-it-censorship/2019/11/25/ccf33c54-0f81-11ea-a533-90a7becf7713_story.html.

Singapore,[33] and Russia,[34] use "national security" as justification for such laws. However, there is little evidence that such laws deter the spread of misinformation or improve national security.

As such, calls for novel legislation or regulatory measures to target the spread of media manipulation must be assessed in relation to the wider political context, the level of government accountability, existing legislation that criminalizes content or the use of the Internet, and how such legislative and regulatory tools may be abused. Strategies designed to minimize the harm of problematic content, such as those outlined in the response matrix below, should therefore also be tested to ensure that they do not unduly restrict individual and group civil liberties, press freedom and internet access.

> Strategies designed to minimize the harm of problematic content should also be tested to ensure that they do not unduly restrict individual and group civil liberties, press freedom and internet access.

Key questions and factors to consider

- Could existing mechanisms of censorship or media control be used against individuals or marginalized groups (e.g., people living with HIV/AIDS, and the LGBTQ community)?
- Can censorship or other information controls be used to stifle criticism of government handling of a medical emergency (e.g., conditions at a hospital, or quality of patient care)?
- What is the relationship between media outlets and the government? Is there a diverse set of media outlets with varied ownership?
- Will content removals result in an information vacuum? If so, can high-quality and relevant information pertaining to the health issue supplement any content removals?

---

[33] Thum Ping Tjin and Kirsten Han, "Singapore's "Fake News" Bill: The FAQ," *New Naratif,* April 9, 2019, https://newnaratif.com/research/singapores-fake-news-bill-the-faq/.

[34] "Russia internet freedom: Thousands protest against cybersecurity bill," *BBC News,* March 10, 2019, https://www.bbc.com/news/world-europe-47517263.

THE MEDIA MANIPULATION LIFE CYCLE

As explained in the Introduction, media manipulation is defined as the sociotechnical process by which motivated actors leverage specific conditions or features within an information ecosystem to generate public attention and press coverage for events that would otherwise go uncovered, or to create a false perception of public outrage. To provide a clear picture of how this process works and how health officials can document it, we will now describe the life cycle of media manipulation campaigns.[35]

Media manipulation campaigns exhibit patterns, where private or semi-private communication channels are used to generate campaigns, public and semi-public platforms are used to disseminate misinformation, and news sites with low or no editorial oversight store the content. The positive and negative attention these campaigns generate on social media make them a "newsworthy" phenomenon that mainstream press is often compelled to cover.

Media manipulation is distinct from media control, which occurs at the top level by the state and private sector.[36] Furthermore, it is not inherently good or bad, nor is every media manipulation campaign reliant on disinformation or deception. Rather, we identify it as an insurgent strategy to both raise awareness and, in some cases, invoke an institutional response. For the purposes of this research brief, however, we focus on media manipulation campaigns that employ harmful medical misinformation.
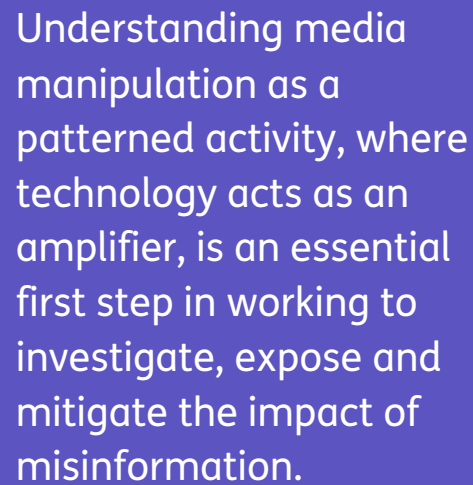
Defining, detecting, documenting, and debunking misinformation and media manipulation online is a global challenge, especially as attacks cross professional sectors – such as journalism, public health, academia, and technology. Critically, misinformation requires both people and technology to circulate. Therefore, understanding media manipulation as a patterned activity, where technology acts as an amplifier, is an essential first step in working to investigate, expose and mitigate the impact of misinformation. We differentiate this method from reactive debunking or fact-checking models, which do not account for

---

[35] This model is patterned after DataONE's "Data Life Cycle", which recommends universal standards for the management and preservation of data for use in multiple sectors of academia, business and civil society. More information here: https://old.dataone.org/data-life-cycle.

[36] Noam Chomsky, *Media control: the spectacular achievements of propaganda* (New York: Seven Stories Press, 2002).

patterned, and sometimes predictable, insidious behavior of motivated actors carrying out media manipulation campaigns.

The media manipulation life cycle model analyzes the order, scale and scope of manipulation campaigns by following media artifacts (online posts of text, video, and images) through space (both geographically and across the web) and time (misinformation often spreads on the heels of a breaking news event or other popular trends). For every case of media manipulation, researchers should map actors, behavior, content, and technological design to analyze how misinformation is distributed and how it gains legitimacy in different communities.[37,38] In practice, this method maps the spread of misinformation as it moves across the open web and social media platforms.
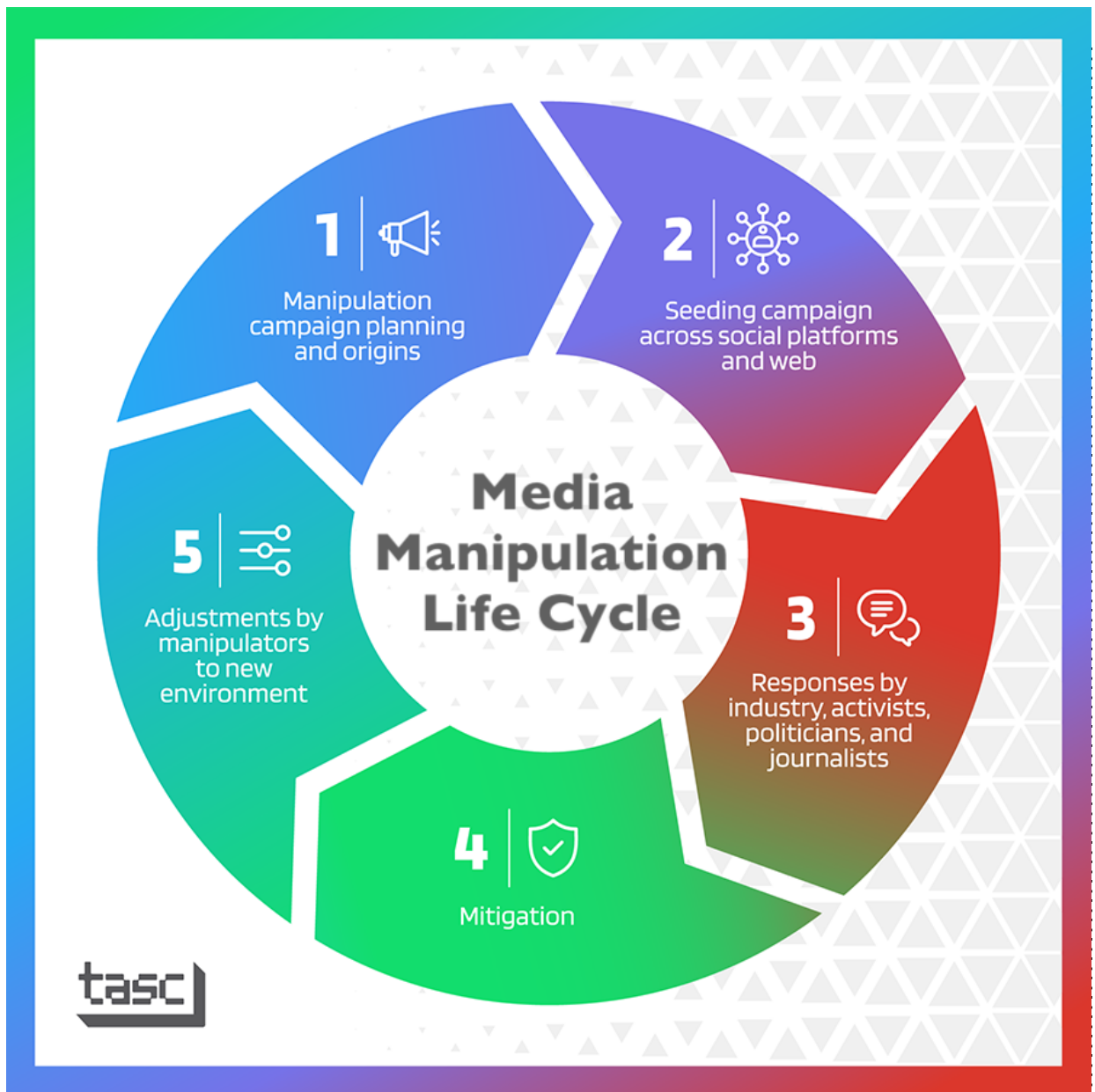
> Understanding media manipulation as a patterned activity, where technology acts as an amplifier, is an essential first step in working to investigate, expose and mitigate the impact of misinformation.

When health officials document a media manipulation campaign in a case study, each stage of the life cycle should be described with supporting evidence for all claims. This method provides a framework for health professionals to identify the stages of a media manipulation campaign and offers suggestions for how to react in concert with other professional sectors (see Figure 1 on the following page).

---

[37] Joan Donovan, "The Lifecycle of Media Manipulation," *The Verification Handbook, European Journalism Centre,* https://datajournalism.com/read/handbook/verification-3/investigating-disinformation-and-media-manipulation/the-lifecycle-of-media-manipulation .

[38] Camille François, "Actors, Behaviors, Content: A Disinformation ABC," Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression, September 20, 2019, https://science.house.gov/imo/media/doc/Francois%20Addendum%20to%20Testimony%20-%20ABC_Framework_2019_Sept_2019.pdf.

Fig. 1. The media manipulation life cycle

Source: The media manipulation life cycle by the Technology and Social Change Project.

In some instances, providing documentation about the origins of misinformation can be enough to discredit it but, in most cases, a networked response that includes content moderation and counter-messaging from authoritative and trusted sources is also necessary. While technology companies have exhibited hesitation in changing platform policies, pressure applied by journalists, governments, civil society, the private sector and other partners has resulted in some action against, for example, anti-vax organizing,[39] COVID-19 misinformation,[40] election interference,[41] and hate speech.[42]

While these stages may be mapped chronologically, most manipulation campaigns are not "discovered" in this order. Instead, when researching, health officials and partners should look for any one of these stages of action and then trace the campaign backward and forward through the life cycle. To properly map a case to the life cycle model, the researcher must first identify which stage it is currently in and then focus on obtaining additional data to clarify earlier stages in the life cycle, and the point at which each stage transitions to the next.

Stage 1: Planning the manipulation campaign

The initial stage of a manipulation campaign is generally limited to conversations by a small group of campaign participants who develop narratives, images, videos or other material to be spread online as "evidence." Campaign planning may happen on social media, message boards, in private online forums or through private messaging on encrypted communication apps. These private or semi-private conversations may not always be legally or ethically accessible to researchers, journalists, medical professionals or law enforcement officials.

If a media manipulation campaign is to succeed, however, it must expand beyond private spaces and enroll hundreds, if not thousands, of other accounts to influence trending and search algorithms, which is often what draws attention from influencers and journalists. Usually, the sources of the campaign are hidden among the array of participants, so researchers must look for campaign participants acting as central nodes in the early amplification of campaigns and the materials that are being shared.

---

[39] "Pinterest, Facebook Fight Back Against Anti-Vaccine Content," *Healthline*, September 8, 2019, https://www.healthline.com/health-news/how-social-media-sites-are-trying-to-stop-anti-vaccine-content.

[40] Spandana Singh and Koustubh "K.J." Bagchi, "How internet platforms are combating disinformation and misinformation in the age of COVID-19," *New America: Open Technology Institute*, June 1, 2020, http://newamerica.org/oti/reports/how-internet-platforms-are-combating-disinformation-and-misinformation-age-covid-19/.

[41] "Civic Integrity," Twitter, https://about.twitter.com/en_us/advocacy/elections-integrity.html.

[42] "Hate Speech," Community Standards, Facebook, https://www.facebook.com/communitystandards/hate_speech.

We identify Stage 1 ending when a confined group of active social media accounts begins posting a particular piece of disinformation in a coordinated manner.

Stage 2: Seeding the campaign across social platforms and the web

The second stage of a manipulation campaign involves the execution of campaign plans, when narratives, slogans, images, videos, or other materials are strategically spread on fringe news websites, social media, or video broadcasting platforms. Campaign participants will attempt to dominate conversations on platforms where they believe they can reach a target audience. This can sometimes be on a single platform, such as the closed environment of WhatsApp, in Facebook pages, or a particular Twitter hashtag, or across the open web through the strategic use of keywords. The rationale is to reach as many individuals as possible so as to achieve a critical mass in conversation that will lead to a campaign becoming newsworthy or result in a false perception of massive public concern.

This stage is referred to as *seeding*, spreading misinformation that larger narratives may grow upon. Often, the seeding phase involves manipulators employing a tactic called "news spamming," where they leave clues, evidence or other materials in the replies of politicians, activists, social media influencers or journalists to garner attention and gain amplification power. Crucially, this is where we see most media manipulation attempts fail as newsworthy individuals often do not take the bait. If they do engage with the content, it can quickly scale to national attention.

We identify Stage 2 ending when a particular piece of disinformation has spread beyond a core group of media manipulation campaign operators, resulting in trending topics on social media, uptake by influential social media accounts, and coverage by fringe websites with little or no editorial oversight.

Stage 3: Responses by industry, activists, politicians and journalists

The third stage of a media manipulation campaign provokes an observable institutional response, which involves reactions from civil society, political figures, government agencies, and mainstream and independent press. Responses include public statements by representatives from social media platforms, activist campaigns drawing attention to malicious behavior by campaign participants, official political statements, or reports addressing disinformation.

With health misinformation in particular, rumors can easily spread without the backing of any activist movement or communication strategy; that is to say, rumors can be authorless and impactful. In fact, most convincing media manipulation campaigns often become detached from their source of origin at this stage and seem to be both untraceable (due to the overflow of attention to the campaign) and intractable (due to level of exposure).[43]

We identify Stage 3 ending when social media platforms or government institutions make policy changes or take actions in an effort to mitigate the spread of a particular piece of disinformation by a media manipulation campaign, and/or they take action against individuals who amplified it. In many cases, no action is taken and public attention wanes, while newly converted believers may begin planning their next campaign.

Stage 4: Mitigation

The fourth stage of a manipulation campaign involves mitigation and major alterations to the availability of information based upon responses from technology companies, the government, the press, or civil society. These mitigations include removal of accounts sharing misinformation, banning of particular types of content, removal or downranking of search terms or hashtags, and banned words in chat rooms, message boards and so on. There are moderation practices of both human and automated content that can be used to challenge the spread of a media manipulation campaign, but during the pandemic, social media companies began to over-rely on automation, which opened up new tactics for manipulators to spread misinformation.[44]

To address COVID-19 misinformation, social media sites and a range of other websites – including those for government, businesses and news organizations – have added pop-ups or banners that link to accurate sources. This consistent curation across the open web and platforms has redirected millions of information seekers to factual materials, but in the case of search and trending algorithms, the situation is much less straightforward. There, misinformation circulates freely alongside factual evidence – a problem that the WHO identified as part of the ongoing infodemic.[45]

---

[43] Joan, Donovan and Brian Friedberg, "Source Hacking: Media Manipulation In Practice," Data and Society, September 4 2019, https://datasociety.net/wp-content/uploads/2019/09/Source-Hacking_Hi-res.pdf.

[44] Sarah T. Roberts, "Digital Humanity: Social Media Content Moderation and the Global Tech Workforce in the COVID-19 Era," *Flow*, March 19, 2020, https://www.flowjournal.org/2020/03/digital-humanity/.

[45] Joan Donovan, "Here's how social media can combat the coronavirus 'infodemic'," *MIT Technology Review*, March 17, 2020, https://www.technologyreview.com/2020/03/17/905279/facebook-twitter-social-media-infodemic-misinformation/.

If a manipulation campaign measurably ends after institutional response and mitigation efforts, it is considered defunct. In cases where a campaign persists despite mitigation efforts, it is in Stage 5.

Stage 5: Adjustments by manipulators to the new environment

The fifth stage of a manipulation campaign involves how manipulators adapt according to mitigation efforts and resulting changes in the information ecosystem. While certain content may be banned or accounts spreading disinformation removed, manipulators will often find ways to circumvent these changes, including creation of new accounts, adoption of coded language, alteration of audio/visual material and iteration on narratives already identified as problematic by platforms.
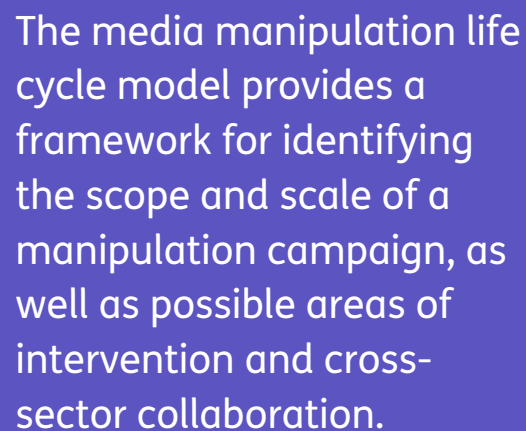
A campaign is considered lower risk when particular pieces of disinformation are no longer being widely spread due to mitigation efforts. This could mean, however, that campaign operators have scaled back an operation or re-entered Stage 1 and will redeploy a new campaign by adjusting to changes in the information ecosystem.

POLICY RECOMMENDATIONS AND IMPLEMENTATION CONSIDERATIONS

Life-cycle interventions

In accordance with World Health Assembly resolution A73/CONF/COVID-19 Response OP 7.6 adopted in May 2020, WHO Member States are mandated to: "Provide the population with reliable and comprehensive information on COVID-19 and the measures taken by authorities in response to the pandemic, and take measures to counter misinformation and disinformation and as well as malicious cyber activities."[46] In light of this call to action, it is imperative that countries identify when to intervene, what policy measures and harm mitigation strategies are available, and the various challenges and opportunities that journalists, civil servants, social media platforms, CSOs, and public health experts may face in countering online misinformation and disinformation.

The media manipulation life cycle model provides a framework for identifying the scope and scale of a manipulation campaign, as well as possible areas of intervention and cross-sector collaboration. Below, we detail the general guidelines for documenting and responding to media manipulation, factoring in the multitude of interventions by the public health sector and relevant stakeholders that can help counter misleading or harmful content and amplify trusted sources.

> The media manipulation life cycle model provides a framework for identifying the scope and scale of a manipulation campaign, as well as possible areas of intervention and cross-sector collaboration.

---

46 "COVID-19 response," Seventy-third World Health Assembly, draft resolution, A73/CONF./1Rev.1, World Health Organization, May 18, 2020, https://apps.who.int/gb/ebwha/pdf_files/WHA73/A73_CONF1Rev1-en.pdf.

General guidelines for every stage

There are some key principles that should be applied at every stage of the life cycle.

1. **Manipulation campaigns thrive when timely, relevant, local, and redundant information is not available.** The lack of authoritative information about particular subjects – information vacuums of reputable sources known as "data voids"[47] – are exploited by manipulation campaign operators. Repeating reliable information across multiple information channels, including TV and radio, will help reinforce the facts during a public health crisis.[48] Redundant messaging, including sloganeering such as "flatten the curve," is key to spreading a shared definition of the situation.

2. **If you have to address misinformation directly, use the "fact–fallacy–fact" communication strategy.** State the facts first, identify the false claim being debunked, then reiterate the facts again. Pairing this with inoculation efforts, where the logical fallacy supporting the misinformation is also explained, will help information seekers understand how they are being manipulated.[20] Keep all messaging respectful, and make sure not to be dismissive, antagonistic or belittling when addressing false beliefs directly, as this can work against the reception of the accurate information.[49]

3. **Consider how to triage misinformation as it moves across platforms.** Misinformation on one platform will likely spread to others, especially if newsworthy individuals begin to share it. Attempts to debunk misinformation should employ a multi-platform campaign if the misinformation has already migrated. Strategic communications must consider various approaches to different communication infrastructures and potentially employ advertising tools to reach key demographics.[35]

4. **Interventions for each stage may be carried forward to the next stage and can be thought of as a continuous process.** For example, recommendations from Stage 1 and 2 can still be applied in Stages 3, 4 and 5.

---

[47] danah boyd and Michael Golebiewsk, "Data Voids," Data & Society, October 29, 2019, https://datasociety.net/library/data-voids/.

[48] Phillip Ball and Amy Maxmen, "The Epic Battle against Coronavirus Misinformation and Conspiracy Theories," *Nature*, May 27, 2020, https://doi.org/10.1038/d41586-020-01452-z.

[49] Marianna Spring, "How should you talk to friends and relatives who believe conspiracy theories?" *BBC*, December 21, 2020, https://www.bbc.com/news/blogs-trending-55350794.

Stakeholders

Public health sector. Individuals in the health sector working on communications and outreach, or who have a public-facing role that requires communicating official statements and other information

News. Journalists, editors, publishers, reporters, and other media workers who engage in both investigating and reporting and communicating information to the public

Civil society organizations. CSOs include nongovernmental organizations, academic institutions, think tanks, faith-based organizations, professional associations, unions or other labor organizations, and domestic and international communities with shared interests and goals.

Government. Government agencies outside the public health sector that may be responsible for coordinating health or crisis-related communications (e.g., Prime Minister's office, transportation agency, elections commission or local government officials)

Technology companies. Includes social media companies, search engines, messaging apps, file-sharing platforms, and other online communications applications that may be involved in the media manipulation campaign

Media Manipulation Response Matrix

Stages	Response by Public Health Sector	Networked Response
1 📢 **Manipulation campaign planning and origins** Campaign is in its planning stage; limited participation; not recognized by mainstream news outlets. It is important to remember that a majority of media manipulation attempts fail to gain widespread attention	• Survey the different misinformation themes and trends related to your issue by filling out a "Situational analysis worksheet" (see Appendix C) • Collect and document data where you see misinformation (e.g., screenshots, hyperlinks, keywords and hashtags) • Flag and report to platforms • Continue to monitor for new developments • Prepare an internal debunking report based upon data collection (see Appendix C)	• CSOs, the government and tech companies should alert researchers and news agencies as a warning or advisory • The government, news agencies and CSOs should send out a warning to your networks and ask for assistance with monitoring, if resources are available • Tech companies, news agencies and CSOs should monitor for any new developments, including tracking how many users are active and if the content is spreading to other platforms • Publicly validating the campaign may produce the "Streisand effect", where attempts to stifle attention increase it
2 🦠 **Seeding campaigns across platforms and web** Small network of accounts spread content related to the campaign; yet to be addressed by tech companies or news organizations	• Boost timely, local and relevant, and accurate content from reputable sources based on the audience's preferred modes of communication • Alert relevant government agencies, but avoid having civil servants amplify the misinformation (i.e., do not publicly address the campaign) • Establish an outreach strategy and official messaging that can be easily shared by journalists and CSOs across social media and private messaging apps	• Government, news agencies and CSOs should encourage tech companies to quarantine or remove the harmful health-related misinformation[a] • Tech companies should apply interstitial notifications (e.g., flags or labels) that require users to acknowledge the presence of misinformation • News agencies and CSOs should report to reputable fact-checking organizations who can continue to monitor, collect data, and inform tech companies • Journalists must avoid reporting on the potential campaign in the mainstream press as this will amplify it • CSOs, news agencies and tech companies should check for evidence of coordination and algorithmic manipulation, such as the use of fake accounts, bots, search engine optimization or foreign influence

[a] Depending on the platform, different features to remove or create friction in accessing content may be available. For example, Reddit allows some content to continue existing on their platform, but they may "quarantine" it by flagging the content as potentially false or offensive and require the user to acknowledge and click through before accessing the content.

3

Responses by tech industry, activists, politicians and journalists

Campaign has been detected and reported by the press, politicians, government or NGO(s) and is moving across different platforms

- Directly address the media manipulation campaign and debunk any false or misleading claims on multiple local and national channels (e.g., official health agency websites, government websites, news agencies, social media)
- Alert and educate frontline health workers who may need to counter the misinformation directly with their patients
- Formalize messaging with government agencies and representatives who will be doing interviews on the issue (consistent and redundant messaging makes all the difference as media manipulators often leverage subtle interagency discrepancies)

- News agencies, CSOs and government agencies should officially debunk using the "fact–fallacy–fact" strategy, and share with trusted partners
- Tech companies must remove individuals and groups who repeatedly violate community standards or often post banned content deemed dangerous to the public, and demonetize accounts generating revenue from misleading information
- Tech companies must provide transparency on service removal and other changes to the information environment as rumors of censorship can also compel interest in a media manipulation campaign

4

Mitigation

Tech companies, government, journalists, or civil society take actions to mitigate the spread of a campaign's content, messaging, and effects.

- Monitor platforms and news outlets to see if accurate and relevant information appears when performing a search of the keywords associated with the misinformation in question
- Where available, use social media measurement tools (e.g., CrowdTangle, Netlytics, Google Trends and others available in your location) to determine if the popularity or spread of the misinformation has declined. You may see an uptick in attention to the issue as it is being debunked, but if attention persists over time, then it is evidence of adaptation

- Academics, CSOs, government and news agencies should publish reports on the harms and impact of the misinformation campaign, and demystify the tactics used by manipulators
- CSOs, media, government and tech companies should collaborate with existing fact-checking databases, such as FEMA and Google, to debunk major trends[b,c]

5

Adjustments by manipulators to the new environment

Actors behind the campaign begin to adapt according to changes in the information ecosystem

- Continue to monitor relevant keywords, hashtags, accounts and online communities that participated in propagating the false information
- Monitor shifts in narratives and messaging from these groups as misinformation tends to resurface as new opportunities arise in breaking news

- Tech companies should publish archives and transparency reports justifying service removals and account deletions
- CSOs should coordinate on problematic areas where public susceptibility to medical misinformation is highest, including treatments, social measures and vaccines and, crucially, the political and economic fallout from public health guidelines

[b] Federal Emergency Management Agency, "Coronavirus rumor control," *FEMA.gov*, June 8, 2020, https://www.fema.gov/coronavirus/rumor-control.

[c] "Fact Check Explorer," *Google Fact Check Tools*, https://toolbox.google.com/factcheck/explorer.

BEYOND THE LIFE CYCLE

Mitigating the adverse effects of media manipulation during an infodemic requires both short-term reactive actions, such as the recommendations above, as well as implementation of long-term strategies by health officials and stakeholders. Media manipulation is the new normal, and countering it will require building coalitions, providing updated media and digital literacy education, enforcing well-defined, consistent and transparent policies on technology platforms, and taking governmental actions to address the sociopolitical grievances that drive audience receptivity to false and deceptive information.

Along with the five key actions spelled out in the WHO's infodemic framework,[50] for long-term planning to counter the spread of medical misinformation, we recommend that the public health sector focus on:

- developing and maintaining a coalition of CSOs, media outlets, educators and government agencies who can pool their various skills and resources to detect and monitor media manipulation and address and mitigate it, if necessary. This must include comprehensive training that emphasizes not just the technical means of media manipulation but the social, political and economic vulnerabilities as well;

- understanding and monitoring existing wedge issues,[51] political cleavages,[52] and socioeconomic conditions that may be leveraged by manipulators during a health emergency;

- using messaging on inoculation to reduce susceptibility to false information;[53]

- identifying barriers and bottlenecks to receiving accurate health information in different regions and demographics, and how best to address them;

---

[50] "Managing the COVID-19 infodemic: a call for action," *World Health Organization*, September 15, 2020, https://www.who.int/publications/i/item/9789240010314.
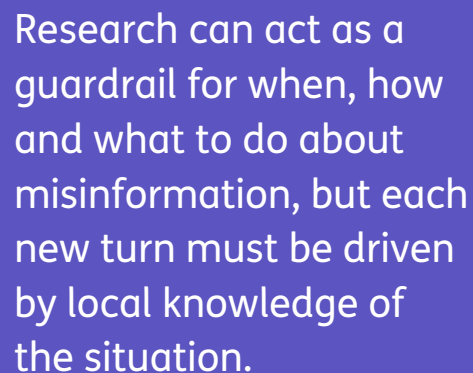
[51] Fredel M. Wiant, "Exploiting Factional Discourse: Wedge Issues in Contemporary American Political Campaigns." *Southern Communication Journal* 67, April 1, 2009 https://doi.org/10.1080/10417940209373236.

[52] S. Fabbrini, "Cleavages: Political," *International Encyclopedia of the Social & Behavioral Sciences*, 2001, accessed via Science Direct, July 31, 2020, https://www.sciencedirect.com/science/article/pii/B0080430767011098.

[53] Jon Roozenbeek, Sander van der Linden, and Thomas Nygren, "Prebunking interventions based on "inoculation" theory can reduce susceptibility to misinformation across cultures," *The Harvard Kennedy School (HKS) Misinformation Review* 1, iss. 2 (January 2020), https://doi.org/10.37016//mr-2020-008.

- establishing guidelines for retractions and corrections, so that decisions about potential misstatements by stakeholders are carried out quickly;

- training journalists on health and scientific topics and developing a speakers' bureau for media interviews. This effort must also include open access to new scientific findings;

- calling for technology companies to enforce existing policies regarding the creation and distribution of medical misinformation and political disinformation, which often function to support one another.

Rising global discontent over social media has urged the United Nations and other global organizations to reconsider not just the content flowing through online information ecosystems, but the infrastructure, policies and people who run them. Researchers working with journalists have found ways to share information and best practices on reporting, evidence collection and narrative framing, and platforms are responding (albeit often too late).[54] Research can act as a guardrail for when, how and what to do about misinformation, but each new turn must be driven by local knowledge of the situation.

> Research can act as a guardrail for when, how and what to do about misinformation, but each new turn must be driven by local knowledge of the situation.

Media manipulation within a health or medical emergency may represent a narrow view of the larger public reckoning on misinformation, political disinformation and Internet governance, but it cannot be disentangled from them. An effective long-term strategy for mitigating the harms caused by medical misinformation should therefore consider all aspects of the information ecosystem – local, regional and international – as regulation takes shape. However, for as long as misinformation remains financially profitable and politically expedient, it will continue to move like digital wildfire causing the most damage under conditions that foster its growth. Therefore, it is the duty of each one of us to carry our own water, rather than wait for the fire next time.

---

[54] Sam Shead, "Facebook, Twitter and YouTube Pull 'false' Coronavirus Video after It Goes Viral," *CNBC*, July 28, 2020, https://www.cnbc.com/2020/07/28/facebook-twitter-youtube-pull-false-coronavirus-video-after-it-goes-viral.html.

APPENDIX A: CASE STUDY

Distributed amplification: Media manipulation and the Plandemic documentary

This case study traces how *Plandemic*, a 26-minute trailer video about coronavirus conspiracy theories, went viral in May 2020 because of distributed amplification – a tactic whereby participants rapidly and widely spread campaign materials across multiple platforms. In response to its high viewership, major social media platforms moderated *Plandemic* and prepared for the full-length video. The platforms' efforts slowed the spread of *Indoctornation,* the anticipated 75-minute movie. *Indoctornation* failed to achieve the virality *Plandemic* had.

STAGE 1: Manipulation campaign planning and origins

The race for a coronavirus vaccine has many people concerned about its safety. Prior to the pandemic, these fears were rooted in government distrust, distaste for pharmaceutical corporations, fallacies about historic vaccination harms and/or preferences for natural remedies. Now, there is an added worry that a vaccine will be approved before its side-effects are fully understood.

These fears have been repeated across social media, and they are coalescing with QAnon conspiracy theories. *Plandemic*, a conspiracy-based video, resonated with these groups and has been viewed tens of millions of times.

Plandemic's 26-minute trailer was released on May 4, 2020, while its 75-minute feature film (*Plandemic: Indoctornation*) followed on August 18, 2020. By misquoting physicians and researchers, and citing conspiracy theorists, *Plandemic* argues that coronavirus was planned (hence the title), vaccines are harmful, masks "activate" coronavirus, and the ocean has "healing microbes." Its anti-vax messaging connects with vaccine-hesitant communities and the general distrust in the coronavirus vaccine.[55]

The video features Judy Mikovits, a discredited scientist with a PhD in biochemistry and molecular biology. She was fired from Whittemore Peterson Institute, the laboratory where she conducted research – and wrote a now-retracted paper in *Science* — on chronic fatigue

---

[55] Alec Tyson, Courtney Johnson, and Cary Funk, "U.S. public now divided over whether to get COVID-19 vaccine," *Pew Research Center*, September 17, 2020, https://www.pewresearch.org/science/2020/09/17/u-s-public-now-divided-over-whether-to-get-covid-19-vaccine/.

syndrome.[56] Mikovits has spoken at anti-vax conferences since 2014, and she published "Plague of corruption" in April 2020, which "frames Dr Mikovits as a truth-teller fighting deception in science."[57]

On May 4, 2020, producer Mikki Willis uploaded the video to YouTube, Facebook, Vimeo, and *Plandemic's* website (plandemicvideo.com).[58] The website was a tool for the video's spread. It "encouraged viewers to download the video from the documentary's website and post it across video platforms",[59] 'Media Matters' Alex Kaplan noted. This strategy distributed amplification, and coached participants to re-upload banned content in an effort to circumvent platform mitigation efforts.

STAGE 2: Seeding campaign across social platforms and the web

Their instructions worked. Kaplan says, "you could see the impact: dozens of re-uploads of the video on YouTube – even after they tried to remove the original – and shares of it on other platforms."[60]

In addition to re-uploads, shares across communities made *Plandemic* go viral. Erin Gallagher, an independent researcher, "found that posts referencing it appeared most often in Facebook groups devoted to QAnon, anti-vaccine misinformation, and conspiracy theories in general".[61]

Gallagher noted that *Plandemic* "spread from YouTube to Facebook thanks to highly active QAnon and conspiracy-related Facebook groups with tens of thousands of members which

[56] Jane Lytvynenko, Ryan Broderick, and Craig Silverman, "Coronavirus pseudoscientists and conspiracy theorists," *Buzzfeed News*, May 21, 2020 https://www.buzzfeednews.com/article/janelytvynenko/coronavirus-spin-doctors.

[57] Davey Alba, "Virus conspiracists elevate a new champion," *The New York Times,* May 9, 2020, https://www.nytimes.com/2020/05/09/technology/plandemic-judy-mikovitz-coronavirus-disinformation.htm.

[58] Sheera Frenkel, Ben Decker, and Davey Alba, "How the 'Plandemic' movie and its falsehoods spread widely online," *The New York Times*, May 20, 2020, https://www.nytimes.com/2020/05/20/technology/plandemic-movie-youtube-facebook-coronavirus.html.

[59] Alex Mahadevan, "Here's how 'Plandemic' went viral," *Poynter*, May 12, 2020 https://www.poynter.org/fact-checking/2020/heres-how-plandemic-went-viral/.
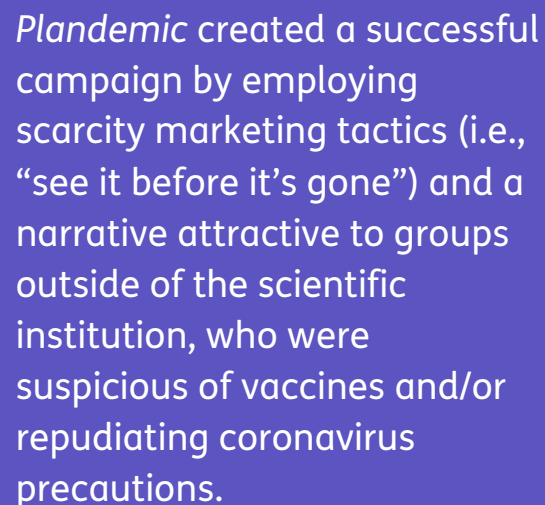
[60] Ibid.

[61] Casey Newton, "How the 'Plandemic' video hoax went viral," *The Verge*, May 12, 2020, https://www.theverge.com/2020/5/12/21254184/how-plandemic-went-viral-facebook-youtube.

caused a massive cascade." She further emphasized that these platforms "were instrumental in spreading viral medical misinformation."[62]

George Washington University's David Broniatowski discussed *Plandemic's* resonance – and hazard – across normally unaligned groups: "The danger with movies like this is that they can weave all of the disparate streams into a common narrative, building a coalition for political and collective action, even when the reasons for this coalition aren't universally shared," he said.[63]

Plandemic created a successful campaign by employing scarcity marketing tactics (i.e., "see it before it's gone") and a narrative attractive to groups outside of the scientific institution, who were suspicious of vaccines and/or repudiating coronavirus precautions. Professor Joan Donovan explained to NBC that "[in] knowing they will be removed from the major platforms, they create a hype cycle around the piece of content, which would probably only get marginal engagement if it was uploaded to a regular website."[64]

> *Plandemic* created a successful campaign by employing scarcity marketing tactics (i.e., "see it before it's gone") and a narrative attractive to groups outside of the scientific institution, who were suspicious of vaccines and/or repudiating coronavirus precautions.

STAGE 3: Responses by industry, activists, politicians and journalists

On May 7, 2020, BuzzFeed reported on *Plandemic* and its falsehoods, signifying the conspiratorial video had made it to mainstream media. Moreover, BuzzFeed's article was shared to Occupy Democrats and 62 other Facebook pages – introducing *Plandemic* to groups who may not have seen it.[65]

---

[62] Ibid.

[63] Brandy Zadrozny and Ben Collins, "As '#Plandemic' goes viral, those targeted by discredited scientist's crusade warn of 'dangerous' claims," *NBC News*, May 7, 2020, https://www.nbcnews.com/tech/tech-news/plandemic-goes-viral-those-targeted-discredited-scientist-s-crusade-warn-n1202361.

[64] Ibid.

[65] Frenkel, Decker, and Alba, "How the 'Plandemic' Movie and Its Falsehoods Spread Widely Online."

The Atlantic, NPR, Wall Street Journal, Science, The Economist, USA Today and *BBC News* are among the publications that published *Plandemic* articles. Advocacy groups – including global health, pro-vaccine and pro-science groups – also covered *Plandemic.*

New York Times' Davey Alba wrote a whole article describing Mikovits' new fame, calling her "a darling of far-right publications like *The Epoch Times* and *The Gateway Pundit.*" Prior to the *Plandemic* and "plague of corruption," Mikovits' online mentions were rare, but by April, she was mentioned about 800 times a day, spiking as high as 14 000 a day.[66]

MIT Technology Review's Abby Ohlheiser laid out the tactics that drove views, describing how activists pursued interviews with "mainstream YouTubers" and "latched on to existing trends, encouraged their fans to amplify their messages, and built presences on every social platform they can find."[67] Its removal made *Plandemic* more popular – leading to censorship claims, increased attention, trending hashtags and media coverage.[68]

STAGE 4: Mitigation efforts

By May 6, 2020, Facebook, YouTube and Vimeo had removed *Plandemic*. In statements to *The Washington Post,* Facebook explained "[s]uggesting that wearing a mask can make you sick could lead to imminent harm, so we're removing the video"; YouTube said it prohibits "content that includes medically unsubstantiated diagnostic advice for covid-19"; and Vimeo emphasized it "stands firm in keeping our platform safe from content that spreads harmful and misleading health information. The video in question has been removed by our Trust & Safety team for violating these very policies."[69] As for Twitter, the platform "removed the hashtags #PlagueofCorruption and #PlandemicMovie from its searches and trends sections."[70]

---

[66] Davey Alba, "Virus Conspiracists Elevate a New Champion," *New York Times*, May 9, 2020, https://www.nytimes.com/2020/05/09/technology/plandemic-judy-mikovitz-coronavirus-disinformation.html.

[67] Abby Ohlheiser, "Facebook and YouTube Are Rushing to Delete 'Plandemic,' a Conspiracy-Laden Video," *MIT Technology Review*, May 7, 2020, https://www.technologyreview.com/2020/05/07/1001469/facebook-youtube-plandemic-covid-misinformation/.

[68] Ibid.

[69] Travis M. Andrews, "Facebook and other companies are removing viral 'Plandemic' conspiracy video," *The Washington Post*, May 7, 2020, http://www.washingtonpost.com/technology/2020/05/07/plandemic-youtube-facebook-vimeo-remove/.

[70] Ibid.

Mashable reported that Twitter allowed the link because users were correcting the information and including the video for content.[71]

Meanwhile, fact-checking organizations corrected baseless arguments. To counter Mikovits' claims, *Science* stated that vaccines save lives and that there is no evidence that coronavirus was planned, is mask-activated or can be cured by nature.[72] *Plandemic* has also been fact-checked by PolitiFact, [73] FactCheck,[74] MedPage Today,[75] and Snopes.[76]

STAGE 5: Adjustments by campaign operators

The instructions to re-upload *Plandemic* were a call to action. Viewers reposted clips from major platforms onto lesser-known websites. Different actors translated the video and added subtitles in other languages, thus making it into a truly global source of vaccine misinformation.[77]

---

[71] Matt Binder, "Twitter takes action against sequel to coronavirus conspiracy film 'Plandemic,'" *Mashable*, August 18, 2020, https://mashable.com/article/twitter-plandemic-sequel/.

[72] Martin Enserink and Jon Cohen, "Fact-checking Judy Mikovits, the controversial virologist attacking Anthony Fauci in a viral conspiracy video," *AAAS: Science*, May 8, 2020, https://www.sciencemag.org/news/2020/05/fact-checking-judy-mikovits-controversial-virologist-attacking-anthony-fauci-viral.

[73] Daniel Funke, "Fact-checking 'Plandemic': a documentary full of false conspiracy theories about the coronavirus," *Politifact: The Poynter Institute*, May 7, 2020, https://www.politifact.com/article/2020/may/08/fact-checking-plandemic-documentary-full-false-con.

[74] Angelo Fichera, Saranac Hale Spencer, D'Angelo Gore, Lori Robertson and Eugene Kiely, "The falsehoods of the 'Plandemic' video," *Factcheck.org*, May 8, 2020, https://www.factcheck.org/2020/05/the-falsehoods-of-the-plandemic-video/.

[75] Mikhail Varshavski, "Point-by-point 'Plandemic' smackdown," *MedPage Today*, May 14, 2020, https://www.medpagetoday.com/infectiousdisease/covid19/86487.

[76] Team Snopes, "Fact check collection: Snopes debunks 'Plandemic,'" *Snopes*, May 11, 2020, https://www.snopes.com/collections/plandemic/.

[77] Jane Lytvynenko, "After the Plandemic video went viral in the US, it was exported to the rest of the world," *BuzzFeed News*, June 1, 2020, https://www.buzzfeednews.com/article/janelytvynenko/coronavirus-plandemic-translation.

Poynter reported that one such site, BitChute, had *Plandemic* clips with 64 000+ views by May 14, 2020.[78] Less than a week later, BitChute's top search result for "Plandemic" had 1.6 million+ views, according to *CBC*.[79]

Beyond migrating to minor platforms, *Plandemic* campaign operators adjusted by releasing a second video, *Indoctornation,* three months after the first.

Unlike the first film, *Indoctornation* was expected. It was promoted "at least 887 times on Facebook, from pages with hundreds of thousands of followers,"[80] writes *The Verge's* Casey Newton*.* To market *Indoctornation*, the creators again relied on its anticipated removal.

Prior to its release, LinkedIn deleted an account advertising *Indoctornation*.[81] Once *Indoctornation* was released, Facebook prohibited users from posting the link. Twitter allowed the link, but warned it was "potentially spammy or unsafe."[82] Mashable reported that Twitter users who try to use the link are met with CDC information.[83]

By taking proactive action, major platforms were able to avoid a repeat of *Plandemic*, which garnered tens of millions of views. That being said, both *Indoctornation* and *Plandemic* continue to live on smaller platforms, easily accessible to those searching for it in multiple languages.

---

[78] Aiyana Ishmael, "'Plandemic' recirculates after platforms say they took it down," *Poynter*, May 18, 2020, https://www.poynter.org/fact-checking/2020/plandemic-recirculates-after-platforms-say-they-took-it-down/

[79] Andrea Bellemare, Katie Nicholson, and Jason Ho, "How a debunked COVID-19 video kept spreading after Facebook and YouTube took it down," *CBC*, May 21, 2020, https://www.cbc.ca/news/technology/alt-tech-platforms-resurface-plandemic-1.5577013.

[80] Casey Newton, "Platforms successfully stopped a COVID conspiracy video from going viral," *The Verge*, August 19, 2020, https://www.theverge.com/interface/2020/8/19/21373820/plandemic-indoctornation-facebook-youtube-twitter-removal-block-covid-hoax-block.

[81] Ibid.

[82] Ibid.

[83] Binder, "Twitter Takes Action against Sequel to Coronavirus Conspiracy Film 'Plandemic.'"

APPENDIX B: CASE STUDY

Hidden virality: Vaccine misinformation campaigns targeting Bill Gates

For the past decade, conspiracy theorists have targeted the philanthropic work of Microsoft founder Bill Gates, accusing him of numerous atrocities from forced population control to medical testing on children.[84],[85] While there are multiple false and baseless allegations that target Gates, this case study breaks down the life cycle of a recent media manipulation campaign linking him with anti-vaccination COVID-19 conspiracies.

With COVID-19, Gates became a familiar target and was accused of everything from creating the virus to using vaccines for demonic purposes. The throughline of all these conspiracies was that Gates was using his wealth and political influence to encourage population control at a global level.[86] This conspiracy theory was spread via anti-vaccination activists and online pundits and influencers using coordinated campaigns on social media, where their audiences click, like, and share this misinformation as a mode of participation.

Social media has given people the unprecedented ability to reach new audiences instantly and at a low cost. Anti-vaccination activists have taken advantage of this new capacity to broadcast and mobilize many new adherents. Vaccination hesitancy, listed by the WHO as one of the top 10 threats to global health,[87] is defined as beliefs or attitudes used to justify reluctance or refusal to inoculate. This issue has emerged as a major challenge for medical professionals and policy makers. Both vaccine hesitancy and anti-vaccination conspiracy

---

[84] Katheryn Joyce, "The Long, Strange History of Bill Gates Population Control Conspiracy Theories," *Type Investigations*, May 12, 2020, https://www.typeinvestigations.org/investigation/2020/05/12/the-long-strange-history-of-bill-gates-population-control-conspiracy-theories/.

[85] Daniel Funke, "Anti-vaxxers spread conspiracy about Bill Gates and India's polio vaccination," *PolitiFact,* April 23, 2020, https://www.politifact.com/factchecks/2020/apr/23/facebook-posts/anti-vaxxers-spread-conspiracy-about-bill-gates-an/.

[86] Katheryn Joyce, "The Long, Strange History of Bill Gates Population Control Conspiracy Theories."

[87] "Ten threats to global health in 2019," *World Health Organization*, https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019.

theories have increased internationally in the 21st century, leading to outbreaks of previously controlled diseases.[88]

Driving this rise in vaccination hesitancy are anti-vaccination activists who grew their community online, using social media to spread disinformation about supposed harmful side effects of vaccines, and sharing tactics to evade vaccination.[89] These attitudes are reinforced by a vast network of disinformation, which includes so-called natural health practitioners, misleading websites, debunked studies, and online influencers who are micro-celebrities within this movement.[90]

With the pandemic, conspiracy communities and the anti-vaccination movement have merged online, often sharing similar content and leveraging search and trending algorithms to their advantage. As Bill Gates took on a public-facing role for his work supporting global immunization and a vaccine for COVID-19, the Gates Foundation became a key scapegoat for various reasons. While taking shape in the US, this narrative spread internationally, resulting in an outpouring of vitriol for Bill Gates and increased skepticism of COVID-19 vaccination work.

Stage 1: Manipulation campaign planning and origins

In the case of the development of the Gates COVID-19 conspiracy, origins of this campaign can be traced to spurious claims and studies,[91] conspiratorial websites,[92] online forums, and small-scale social media discussions. These visible traces of conversation hint at the motivations of the manipulators, exhibiting denial or high degrees of skepticism about the origin and nature of the virus as well as distrust or hatred for Gates and the work of his foundation. After Gates participated in a Reddit AMA, a false quote regarding microchips and

---

[88] Peter Hotez, "America and Europe's new normal: the return of vaccine-preventable diseases," *Nature*, no. 85 (February 2019): 912-914, https://www.nature.com/articles/s41390-019-0354-3.

[89] Jan Hoffman, "How Anti-Vaccine Sentiment Took Hold in the United States," *New York Times*, September 23, 2019, https://www.nytimes.com/2019/09/23/health/anti-vaccination-movement-us.html.

[90] Joan Donovan, "Covid Hoaxes Are Using a Loophole to Stay Alive—Even after Content Is Deleted," *MIT Technology Review*, April 30, 2020, https://www.technologyreview.com/2020/04/30/1000881/covid-hoaxes-zombie-content-wayback-machine-disinformation.

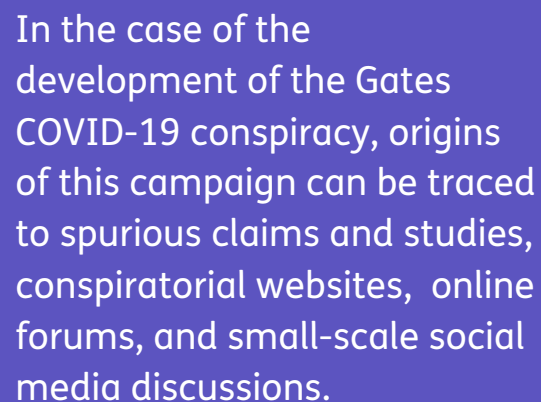[91] Katheryn Joyce, "The Long, Strange History of Bill Gates Population Control Conspiracy Theories."

[92] Alex Kasprak, "Did Bill Gates 'Admit' Vaccinations Are Designed So Governments Can 'Depopulate' the World?" *Snopes*, March 10, 2017, https://www.snopes.com/fact-check/bill-gates-vaccinations-depopulation/.

vaccination went viral, amplified by anti-vaccination influencers and the far-right press in the United States.[93]

Stage 2: Seeding the campaign across social media and the open web

After campaign participants connected Gates with COVID-19 during the planning stage, accounts on social media platforms began sharing documents supporting the conspiracy theory connecting microchips and vaccination. Materials purporting to be evidence – such as Microsoft copyright/patent numbers and a TED talk given by Gates in 2015 about pandemics – were spread in YouTube videos, commentary in Facebook groups, and new and already popular hashtags on Twitter and Instagram were used to increase visibility. Junk and hyperpartisan news sites began reporting on these conspiracy theories as facts, which were then shared on social media by influential manipulators as proof of the validity of their claims. This deployment of preexisting disinformation (Gates' plans to use vaccines with population control) within new frames (Gates was responsible for COVID-19) helped reinforce preexisting vaccine hesitancy among groups sharing the disinformation, where explanations of the political and profit motives of Gates outpaced any fact-based discussion about the potential COVID-19 vaccine.
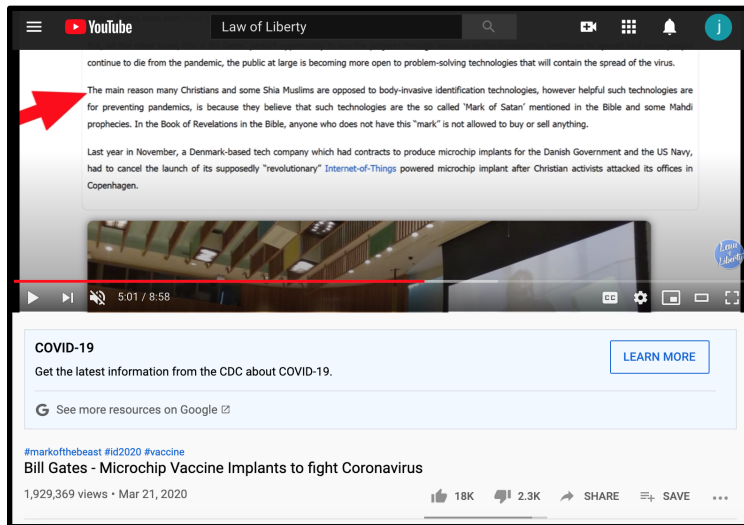
We focus here on two key pieces of content that were heavily circulated in this initial first wave of attention. Both pieces of content exhibit characteristics of "hidden virality," which refers to situations where content is distributed through a wide network of viewers while avoiding content moderation or mainstream media coverage.[94]

In the case of the development of the Gates COVID-19 conspiracy, origins of this campaign can be traced to spurious claims and studies, conspiratorial websites, online forums, and small-scale social media discussions.

---

[93] Jane Lytvynenko, "Here's A Timeline Of How A Bill Gates Reddit AMA Turned Into A Coronavirus Vaccine Conspiracy," *Buzzfeed News*, April 18, 2020, https://www.buzzfeednews.com/article/janelytvynenko/conspiracy-theorists-are-using-a-bill-gates-reddit-ama-to.

[94] Britt Paris and Joan Donovan, "Deepfakes and Cheap Fakes: The Manipulation of Audio and Visual Evidence," *Data & Society*, September 18, 2019, https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal-1.pdf.
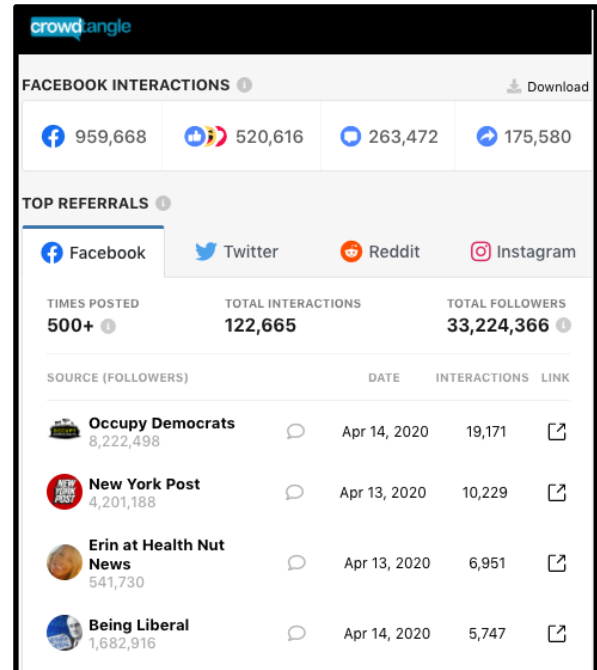
"Microchip" YouTube Video

A YouTube video entitled "Bill Gates: Microchip Vaccine implants to fight Coronavirus" in late March 2020 seems to have kicked off the conspiracy theory about Gates using forced vaccines to tattoo people with the "mark of Satan."[95] Using CrowdTangle, a metrics company owned by Facebook, we measured 1.33M interactions and 307K shares of the "Microchip" video on Facebook.

Stage 3: Responses to the media manipulation campaign

The second piece of viral content, published on April 14, 2020, was a misleading story from the *New York Post* with the headline "Roger Stone: Bill Gates may have created coronavirus to microchip people." This article received impressive engagement on Facebook with nearly 960K interactions and 175K shares across a number of left wing, right wing, and anti-vaccine pages and groups. While it is technically true that Roger Stone stated this, the statement itself is demonstrably false. Yet, because the article attributes the statement to its author, it has not received a fact-checked label by tech companies and is not considered medical misinformation. This kind of tactic is routinely used by manipulators to closely adhere to social media companies' terms of service and skirt any labelling.

---

[95] Jane Lytvynenko, "Here's A Timeline Of How A Bill Gates Reddit AMA Turned Into A Coronavirus Vaccine Conspiracy."

Surprisingly, the top sharer is the Facebook page Occupy Democrats, an American left-wing media organization, which shared the *New York Post* article with a pithy caption critical of Trump and his associates.

The New York Post Headline and associated data from Crowdtangle, a social media metrics company owned by Facebook.

In July of 2020, Gates responded to the allegations, telling CNN host Anderson Cooper that the conspiracy theories targeting him were due to "a bad combination of pandemic and social media and people looking for very simple explanations of who is the bad guy is here." In the interview he explains that "people like myself and Dr. Fauci become the target."[96]

Stage 4: Changes to the information ecosystem

International fact-checking services[97,98] have thoroughly debunked and explained the conspiracies. Major international press outlets provided critical coverage, amplifying the

[96] CNN Town Hall with Anderson Cooper, "Bill Gates' Message to Covid-19 Conspiracy Theorists - CNN Video." Uploaded to *CNN Business*, July 24, 2020. Accessed August 3, 2020. https://www.cnn.com/videos/business/2020/07/24/bill-gates-conspiracy-theories-coronavirus-covid-19-social-media-town-hall-vpx.cnnbusiness.

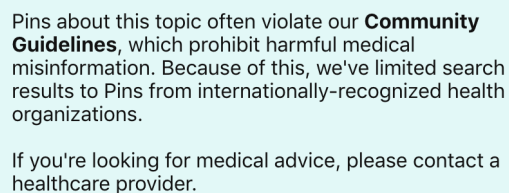[97] Alex Kasprak, "Did Bill Gates 'Admit' Vaccinations Are Designed So Governments Can 'Depopulate' the World?"

[98] Rappler, "The Cases of COVID-19, the Disease Caused by the Novel Coronavirus, Is Not Real and Was Planned by US Billionaire Bill Gates to Push the Use of Vaccines and Microchip Implants," *Poynter*, July 10, 2020, https://www.poynter.org/?ifcn_misinformation=the-cases-of-covid-19-the-disease-caused-by-the-novel-

formal debunking and explaining the context[99] in which this disinformation emerged. As this misinformation spread, it became politicized and more entrenched across the political divides. More mainstream audiences were exposed to the disinformation due to the various responses, and now queries on major platforms for "Bill Gates vaccine" also return conspiracies alongside factual information, a hallmark of the infodemic.

In the case of the Gates COVID-19 conspiracy theories, major platforms implemented no fundamental adjustments. Currently, Facebook has placed a fact-checking label over the "mark of Satan" YouTube video, but YouTube has not taken any action, even though the video has been viewed nearly 2 million times. *The New York Post* article continues to circulate on Facebook and Twitter without labels.

More broadly, Facebook, Twitter, Google Search, YouTube, Pinterest, and numerous other websites have introduced a warning at the top of results for COVID-19 that links users directly to WHO or authoritative content from the US Center for Disease Control.

While some major social media platforms have limitations on speech regarding anti-vaccination claims in their terms of service, Pinterest has gone to greater lengths to limit vaccine misinformation and actively curates content for this search query. Pinterest provides the following disclaimer to those searching for the keyword "vaccines."

Pins about this topic often violate our **Community Guidelines**, which prohibit harmful medical misinformation. Because of this, we've limited search results to Pins from internationally-recognized health organizations.

If you're looking for medical advice, please contact a healthcare provider.

Pinterest disclaimer

While tech companies are not obligated to protect Gates or his foundation's reputation, the sheer amount of health misinformation paired with his name is astounding. As information seekers in this media environment are increasingly encouraged to "do your own research," there are thousands of posts and other online media dissuading people from taking a COVID-19 vaccine, especially if it were to come from research funded by the Gates Foundation.

---

coronavirus-is-not-real-and-was-planned-by-us-billionaire-bill-gates-to-push-the-use-of-vaccines-and-microchip-implants.

[99] Katheryn Joyce, "The Long, Strange History of Bill Gates Population Control Conspiracy Theories."

Stage 5: Adjustments by manipulators to the new information environment.

The persistence of conspiracy theories about Gates' motives is an example of how ideologically motivated groups spread medical disinformation online in the form of targeted attacks on individuals. Across most platforms, any searches for "Bill Gates and vaccines" will surface this health misinformation among the top queries. Despite waves of critical press and fact checking, the conspiracy remains active on most platforms with new conspiracist content produced daily. Media manipulators pushing the false narratives continue to build new audiences, while also iterating as new themes gain traction in trending algorithms. Others are profiting from the chaos and uncertainty caused by the pandemic and are marketing false cures and treatments alongside these very same conspiracy theories.[100]

Conclusion

After a year of these conspiracy theories circulating, and with COVID-19 vaccines actually being injected into arms in 2021, many of the conspiracies persist and adapt to breaking news involving Gates and his foundation. Whether they are directly attributable to the well-documented vaccine hesitancy[101] experienced in the US and other countries is unclear on a case-by-case basis, but the sticking power of the microchip conspiracy, in particular, points out that the real world consequences of medical media manipulation campaigns are far greater than the harassment of a single individual. Gates may have been the target of these conspiracy theories, but it is the general public that pays the bigger price for this kind of medical misinformation.

---

[100] Joan Donovan, "Social-Media Companies Must Flatten the Curve of Misinformation," *Nature*, April 14, 2020, https://doi.org/10.1038/d41586-020-01107-z.

[101] Cary Funk and Alec Tyson, "Intent to Get a COVID-19 Vaccine Rises to 60% as Confidence in Research and Development Process Increases," *Pew Research Center*, December 3, 2020, https://www.pewresearch.org/science/2020/12/03/intent-to-get-a-covid-19-vaccine-rises-to-60-as-confidence-in-research-and-development-process-increases/.

APPENDIX C: SITUATIONAL ANALYSIS WORKSHEET

Individuals (Users/accounts, influencers, entrepreneurs, experts, politicians, journalists, activists, technologists, pundits, etc.)	**Organizations** (Groups, alliances, advocates, research centers, coalitions, platform companies, lobbyists, public relations agencies, professionalization organizations, cybersecurity firms, fact-checking organizations, civil society organizations, etc.)
Wedge issues (Contested politicized positions around identity, authority and justice, often centered on the distribution of resources, rights and representation)	**Popular media and popular culture narratives** (News organizations, websites, news reports, editorials, social media influencers, advertisements, web forums, etc.)

Non-human actors (Search algorithms, recommendation systems, online advertising, technological features, communication infrastructure, mobile technology, bots, malware, phishing, sock puppets, policies and terms of service, etc.)	**Targets of disinformation campaigns** (Groups and/or identity categories that are likely to be targeted by disinformation and media manipulation campaigns)
Networked factions (Online political groups that have a shared affiliation for a specific candidate or particular issue, i.e., virtual communities, Facebook groups, message boards dedicated to specific issues or political positions)	**Space and time** (Geographical scale of local, regional, national or global differences in laws, historical trajectory or cultural factors that create differences in the distribution and understanding of information)
Political and economic factors (Funding, private/public divides, government regulation, corporate policies, legal precedents, who has the power and resources to do something)	

APPENDIX D: GLOSSARY

For an expanded list of terms and definitions, please visit the [Media Manipulation Casebook](#).

Adaptation: the way that campaign operators adjust tactics to get around changes to the information ecosystem created by mitigation efforts

Bots: bots typically refer to social media accounts that are automated and deployed for deceptive purposes, such as to artificially amplify a message, game a trending or recommendation algorithm or inflate an account's engagement metrics

Case study: the documentation of a media manipulation event or disinformation campaign

Civil society: a wide array of organizations, including community groups, nongovernmental organizations (NGOs), labor unions, indigenous groups, charitable organizations, faith-based organizations, professional associations and foundations

Disinformation: the creation and distribution of intentionally false information, usually for political ends

Harassment: behavior (often unlawful) towards an individual or group of people that causes mental, physical or emotional distress. Harassment includes but is not limited to unwanted threats, insults and offensive language.

Hoax: an act designed to dupe or trick people

Infodemic: the overabundance of information – some accurate, some not – that is spreading alongside the COVID-19 pandemic[35]

Information ecosystem: the totality of news, entertainment, social media and other sources available to a community, and the infrastructure that supports it.

Manipulation campaign: an explicitly or tacitly organized effort to attract undue media attention to amplify disinformation or extremism rhetoric

Media manipulation: the sociotechnical process where motivated actors leverage specific conditions or features within an information ecosystem to generate public attention through press coverage for events that would otherwise go uncovered or to create a false perception of public outrage

Medical misinformation: Medical misinformation refers to incorrect or unverified information about the form and function of the human body, and/or misperceptions of health practitioners and medical science.

Misinformation: information whose inaccuracy is unintentional and spread unknowingly

Mitigation: interventions undertaken by stakeholders to stop the spread of misinformation or disinformation. This can take the form of responses such as fact-checking, to changes to technological systems as well as policy implementations.

Networked factions: a loosely organized group that shares some political opinions and primarily (though not exclusively) congregates online. A faction may act in unison as a political force to reach specific ends and then dissolve.

Rumors: widely disseminated unverified information with no clear source

Seeding: the strategic spread of text or content on social media to attract public attention

Strategic amplification: best practices for ensuring responsibility and accountability when producing news content and its social impact

Streisand effect: when the attempt to hide something amplifies it and makes it more visible

Wedge issues: contested politicized positions around identity, authority and justice, often centered on the distribution of resources, rights and representation. Wedge issues will usually split along partisan lines and will be presented as binary positions – for or against.

APPENDIX E: METHODS

The recommendations contained in this research brief are built upon the research of the Technology and Social Change Research Project at Harvard Kennedy School's Shorenstein Center on Media, Politics, and Public Policy. This team, led by Joan Donovan, takes an interdisciplinary approach to researching the impact of technology on society.

In order to conduct our research, we use the methods of digital investigative ethnography, drawn from anthropologists and communication scholars Gabriella Coleman[102] and Sahana Udupa,[103] to detect, document and debunk media manipulation disinformation campaigns.[104] Digital investigative ethnography is a method that combines principles from anthropology, sociology and communication studies and situates research in spaces marked by distinct patterns, beliefs and cultures.[105] It takes into account the cultural aspects that define communities, including geography, history, language, diversity and legal systems. An ethnographer engages with the subjects to varying degrees and, in the case of digital ethnography, with the traces they leave across the information ecosystem. Observing online communities properly takes time, and the ethnographic process requires a commitment to observation during breaking news events and also during the downtime in between. This investigative ethnographic method merges the pointed search for specific information that defines journalistic and legal investigation, with the long-term observation that defines ethnography.[106]

In the detection phase, we begin by looking closely for the suspected misinformation across different websites and social media platforms. If we suspect that a website, post or video

---

[102] Gabriella Coleman, *Hacker, Hoaxer, Whistleblower, Spy: The Many Faces of Anonymous*. Verso; 2014 (https://www.versobooks.com/books/2027-hacker-hoaxer-whistleblower-spy, accessed 25 January 2021).

[103] Sahana Udupa, "Extreme speech. Nationalism in the digital age: fun as a metapractice of extreme speech," International Journal of Communication, 2019;13:3143–63, https://ijoc.org/index.php/ijoc/article/view/9105.

[104] Joan Donovan, "Toward a militant ethnography of infrastructure: cybercartographies of order, scale, and scope across the occupy movement," *Journal of Contemporary Ethnography*, https://doi.org/10.1177/0891241618792311.

[105] Sarah Pink, Heather Horst, John Postill, Larissa Hjorth, Tania Lewis, and Jo Tacchi, "Digital ethnography: principles and practice," Sage Publications; October 2015 (https://uk.sagepub.com/en-gb/eur/digital-ethnography/book243111.

[106] Brian Friedberg, "Investigative digital ethnography: methods for environmental modeling," *The Media Manipulation Casebook*, October 20, 2020, https://mediamanipulation.org/sites/default/files/2020-10/Investigative_Ethnography_v1.pdf.

contains misinformation, we then investigate if the account or website is legitimate and representing itself accurately. We use various open-source investigation tools available on the open web to document signs of imposters, including assessing the history of the accounts in question, the degree of automation on a specific topic, and the networks of actors sharing the suspicious materials. Once we have established and documented a pattern of malicious behavior through process tracing, we create a timeline of events for deeper analysis.[107]

In the analysis phase, we debunk the misinformation by placing it into the life-cycle model to assess how the campaign grew and what actions were taken to mitigate its spread. Patterned after data life-cycle models that describe how data should be gathered and used,[108] the media manipulation life-cycle is the product of three years of research on how journalists, civil society groups, health professionals and technologists grapple with media manipulation and disinformation campaigns.[109] The life-cycle draws together materials collected by a researcher to analyze timelines, behavioral patterns, and the broader context of the disinformation, including how it attaches to hot button wedge issues, which are contested political issues that often have to do with the redistribution of rights, resources or representation. Situated in the emerging field of Critical Internet Studies,[110] this research methodology combines social science and data science to create a new framework for studying sociotechnical systems and their vulnerabilities.[111]

The life-cycle model was developed to give a common frame for journalists, researchers, technologists and members of civil society to understand the origins and impacts of disinformation.[112] The policy recommendations in this document are informed by a corpus of

---

[107] PM Krafft and Joan Donovan, "Disinformation by design: the use of evidence collages and platform filtering in a media manipulation campaign," *Political Communication*, 37:(2)194-214, https://doi.org/10.1080/10584609.2019.1686094.

[108] Joan Donovan, 2020. "The Life Cycle of Media Manipulation," *The Verification Handbook 3*, 2020. https://datajournalism.com/read/handbook/verification-3/investigating-disinformation-and-media-manipulation/the-lifecycle-of-media-manipulation.

[109] Joan Donovan and danah boyd, "Stop the presses? Moving from strategic silence to strategic amplification in a networked media ecosystem," *American Behavioral Scientist*, September 29, 2019; 65(2):333-350, https://doi.org/10.1177/0002764219878229.

[110] Amelia Acker and Joan Donovan, "Data craft: a theory/methods package for critical internet studies," *Information, Communication, and Society*, 22(11):1590–609, https://doi.org/10.1080/1369118X.2019.1645194.

[111] Matt Goerzen, Elizabeth Anne Watkins, Gabrielle Lim, "Entanglements and exploits: sociotechnical security as an analytic framework," *9th USENIX Workshop on Free and Open Communications on the Internet (FOCI)*, https://www.usenix.org/conference/foci19/presentation/goerzen.

[112] Joan Donovan, 2020. "The Life Cycle of Media Manipulation."

research conducted by the Shorenstein Center's discussion papers, reports and policy briefs on COVID-19 and vaccination hesitancy,[113] social media governance[114] and misinformation,[115] as well as engagement with global NGOs and communities of health professionals on the frontlines of the infodemic. We have also solicited reviews of the response matrix from collaborators in the health field and incorporated feedback from these health professionals, including the CDC and WHO. While we present two case studies in this document, more are available on our research platform at The Media Manipulation Casebook: www.mediamanipulation.org.

---

[113] David Lazer, Alexi Quintana, Katherine Ognyanova, Matthew A. Baum, John Della Volpe, James Druckman, Roy H. Perlis, Mauricio Santillana, Hanyu Chwe, Matthew Simonson, Jon Green, and Jennifer Lin, "The state of the nation: a 50-state COVID-19 survey," *Harvard Kennedy School: Shorenstein Center,* September 2020, https://shorensteincenter.org/wp-content/uploads/2020/09/COVID19-CONSORTIUM-REPORT-12-APPROVAL-SEP-2020.pdf.

[114] Tom Wheeler, Phil Verveer, and Gene Kimmelman, "New digital realities: new oversight solutions, *Harvard Kennedy School: Shorenstein Center,* August 20, 2020, https://shorensteincenter.org/new-digital-realities-tom-wheeler-phil-verveer-gene-kimmelman/.

[115] Gene Kimmelman, "Key elements and functions of a new digital regulatory agency," *Harvard Kennedy School: Shorenstein Center*, February 12, 2020, https://shorensteincenter.org/key-elements-and-functions-of-a-new-digital-regulatory-agency/.

ABOUT THE AUTHORS

Joan Donovan

Joan Donovan, PhD, is the Research Director of Harvard Kennedy's Shorenstein Center on Media, Politics and Public Policy and the Director of the Technology and Social Change Project (TaSC). Her research specializes in Critical Internet Studies, Science and Technology Studies, and the Sociology of Social Movements.

Brian Friedberg

Brian Friedberg is the Senior Researcher at TaSC at the Shorenstein Center. Blending academic research and open-source intelligence techniques, Brian is an investigative ethnographer, focusing on the impacts alternative media, anonymous communities, and unpopular cultures have on political communication and organization.

Gabrielle Lim

Gabrielle Lim is a Researcher with TaSC at the Shorenstein Center, as well as an Associate with Citizen Lab at the Munk School, University of Toronto. Her research focuses primarily on information controls and security.

Nicole Leaver

Nicole Leaver is a recent Master of Arts in Law and Diplomacy graduate at the Fletcher School, Tufts University, specializing in national security and international information and communication. Her current research examines contemporary US foreign policy, influence operations and human rights.

Jennifer Nilsen

Jennifer Nilsen graduated from Harvard Kennedy School with her MPP in 2020. She is currently a Research Fellow with TaSC, where her research focuses primarily on medical misinformation, vaccine hesitancy, and cloaked science.

Emily Dreyfus

Emily Dreyfuss is a Senior Editor at TaSC and a journalist who covers the impacts of technology on society, with a focus on social media and information systems. She was a 2018 Nieman Berkman Klein fellow, a senior editor at WIRED, and the editorial director of Protocol.

ACKNOWLEDGMENTS