

The Media Manipulation Casebook

Code Book

Last updated: April 20, 2021

Version 1.2



Table of Contents

About the Media Manipulation Casebook	3
What is Media Manipulation?	3
Selection Criteria	4
Novelty	4
Scale	4
Evidence	4
Contributors	5
Cautionary Notes and Limitations	5
Methodology and Theory	6
Research Methods	6
Coding Process	6
The Media Manipulation Life Cycle	7
Variables	11
Case Name	11
Region	11
Date	11
Strategy	11
Tactics	13
Network Terrain	15
Vulnerabilities	17
Attribution	19
Targets	20
Observable Outcomes	21
Mitigation	22
Campaign Adaptation	23

About the Media Manipulation Casebook

The Media Manipulation Casebook (the Casebook) is a research repository consisting of documented attempts to manipulate on- and offline media ecosystems. It is intended for researchers, journalists, policymakers, and other members of civil society to better understand how sociotechnical information ecosystems can be gamed and manipulated and the outcomes of such actions. It is led by Dr. Joan Donovan and developed and maintained by the Technology and Social Change Project at the Harvard Kennedy School Shorenstein Center for Media, Politics, and Public Policy.

Each case study is coded according to a set of predefined variables and includes a chronological description of the campaign in question using the Media Manipulation Life Cycle model.

The Casebook is still in its expansion phase and welcomes collaboration. For further inquiry, please contact us at manipulation@hks.harvard.edu.

Changes in Version 1.2

The following new values were added: Open collaboration tools (Category: Network Terrain); Copy-pasta (Category: Tactics); Hijacked accounts (Category: Tactics); Public relations or marketing firm (Category: Attribution); State actor (Category: Attribution); and Trolls (Category: Attribution).

For the previous version (v1.1), click [here](#).

What is Media Manipulation?

We define media manipulation as a process where actors leverage specific conditions or features within an information ecosystem in an attempt to generate public attention and influence public discourse through deceptive, creative, or unfair means. Media is a reference to *artifacts of communication* and not simply a description of news. Although much has been written about the harmful effects of media manipulation and is often attributed or linked to so-called “bad actors,” it is not inherently good or bad. Activists, constrained by heavy censorship in traditional media, for example, may rely on media manipulation in the digital space to circumvent such information controls. However, violent extremists may likewise use the same platforms and tactics to mainstream hateful or dangerous speech. Furthermore, media manipulation is a broad term in that it can be used to define a variety of other terms, such as disinformation, information operations, or influence operations. This is intentional as it allows for a wider variety of cases to be analyzed.

Note that we differentiate media manipulation from media control, which occurs at the top-level by the state and private sector. Media control would instead refer to activity like ISP-level content blocking, government censorship agencies, media ownership, or distribution and licensing regimes.

Selection Criteria

The Casebook aims to provide users with a wide range of cases from around the world that illustrate how specific conditions and features of technology and society can be used to manipulate the information ecosystem with the goal of generating wider press coverage for issues or events that would otherwise go uncovered or to create a false perception of an issue. With this in mind, cases are selected based on three main criteria: their *novelty* of tactics and strategies; the *scale* of the operation and its resulting outcomes and institutional responses; and whether there is enough high-quality *empirical evidence*. To identify whether or not a case should be included we ask the following questions:

Novelty

Goal: To increase diversity of information within the Casebook

- Does this case expand the diversity of cases with regard to strategies and tactics?
- Does this case introduce a new tactical mix in executing a strategy?
- Are the social and technical vulnerabilities being exploited in new or different ways?
- Are the targets or campaign operators or participants involved in the case novel?
- Did the campaign operators or participants adjust their tactics in response to institutional actions (ex. user ban, content removal, account suspension)?

Scale

Goal: To include cases with observable outcomes and reach

- Does this case stand out because of the scale of media coverage?
- Does this case stand out because of high engagement (ex. retweets, comments, likes)?
- Does this case span multiple platforms?
- Did this case invoke institutional response (ex. political response, civil society response, change in platform governance)?

Evidence

Goal: To ensure cases are backed by high-quality sources and multiple indicators

- Is there enough evidence to support analyses of the tactics and strategies employed?
- Are the secondary sources reliable or drawn from high-quality investigations? (ex. academic research or credible investigative reporting)

- Is the data collection and analysis repeatable by other researchers?

Contributors

Cases are researched, written, and coded by members of the Technology and Social Change Project at the Harvard Kennedy School's Shorenstein Center on Media, Politics, and Public Policy. In addition, researchers and scholars from a wide range of experiences and backgrounds, including the fields of sociology, political science, and science and technology studies, have also contributed to the Casebook.

The Casebook is still in its expansion phase and welcomes future contribution and collaboration. For further inquiry, please contact us at manipulation@hks.harvard.edu.

Cautionary notes and limitations

Firstly, media manipulation campaigns are difficult to detect, trace, and attribute, due largely to their ephemeral and covert nature of their planning and execution. Depending on the availability of evidence, some stages will be thinner in description than others. This does not imply there was no activity in that stage, but rather there was no credible evidence available. For example, Stage 1, which documents campaign planning, is often conducted in private. Researchers who primarily rely on open source means of gathering data may therefore be unable to ascertain what actually happened during this period. Similarly, some cases' variables will be coded with "Unclear" where there is not enough evidence. For example, attribution, which is notoriously difficult to pinpoint, will often be coded as "Unclear," implying there is not enough evidence to ascertain with a high-level of confidence who the operators of a campaign are.

Secondly, the casebook is neither an exhaustive nor representative collection of media manipulation and disinformation. Therefore, statistical analyses that take the case as the unit of analysis may not be robust to selection biases. Analyses that attempt to extrapolate trends or quantify elements or features of media manipulation are not advised.

Lastly, as with many cases of media manipulation, new and emerging evidence may change a case's analysis and findings. Any errors should be brought to our attention at manipulation@hks.harvard.edu and will be greatly appreciated. Changes made to a case will be identified with a note explaining why and the date of the change.

Methodology and Theory

Research methods

The means for documenting case studies require a variety of methods in order to establish a chronological account of a campaign, the tactics and strategies employed, the actors involved, and the outcomes. As such, a variety of indicators from different sources (ex. commercial threat reporting, platform reporting, independent research, investigative journalism, website scraping, etc.) is used to triangulate the findings. In addition, each case study varies in the methods used to detect and document the operations, depending on the availability of evidence. Methods may therefore include a mix of quantitative content analysis, network mapping, and ethnographic studies. Where original research and analysis was undertaken, source material is presented and cited, if practicable.

Examples of evidence include:

- Screenshots (ex. JPG, PNG)
- Archived links (ex. Web Archive, Archive.is, Perma.cc, Local HTML Archives)
- Graphs and charts (ex. Graphml files, Gephi files and images, Tableau files)
- Larger datasets (ex. API-drawn data, Factiva or Lexis Nexis exports, other scraped platform data)
- Internet infrastructure (ex. WHOIS information, IP addresses)
- Digital forensics (ex. Malware analysis)
- Credible and well-researched external sources (ex. investigative journalism, peer-reviewed article, research paper)

Coding process

The variables and values used to code each case are emergent from the existing cases and have been developed after multiple rounds of review. They were determined by identifying the most salient traits that are also applicable to all the cases. These codes function as a way to organize and filter case studies based on a common feature while offering comparability and nuance across the different cases.

Each case goes through a minimum of two rounds of coding to ensure each code corresponds to the evidence and text in the description of the life cycle. The initial round is conducted by the primary author of the case, followed by a member of the TaSC team. Any discrepancies in codes are discussed after between the two coders. If there are still any outstanding discrepancies, a third round of coding by a different member of the TaSC team is conducted, and any differences resolved between all three coders.

To view all the variables and values used, see the [Variables](#) section below.

The media manipulation life cycle

The media manipulation life cycle (MMLC) forms the basis of the Casebook. Patterned after data life cycle models that describe how data should be gathered and used,¹ the MMLC model was developed to give a common framework for journalists, researchers, technologists, and members of civil society to understand the origins and impacts of disinformation and its relation to the wider information ecosystem.² It is the product of three years of digital ethnography and field research by Dr. Donovan and her team on how journalists, civil society groups, and technologists grapple with media manipulation and disinformation campaigns.³ Situated in the emerging field of Critical Internet Studies,⁴ this research methodology combines social science and data science to create a new framework for studying sociotechnical systems and their vulnerabilities.⁵

Each case is written according to the five stages defined in the MMLC (details below), allowing researchers to analyze the order, scale and scope of the campaign in question, as well as the actors involved, platforms used, vulnerabilities (both social and technical) exploited, and outcomes. This semi-structured format allows for comparability across cases while providing any necessary context and nuance.

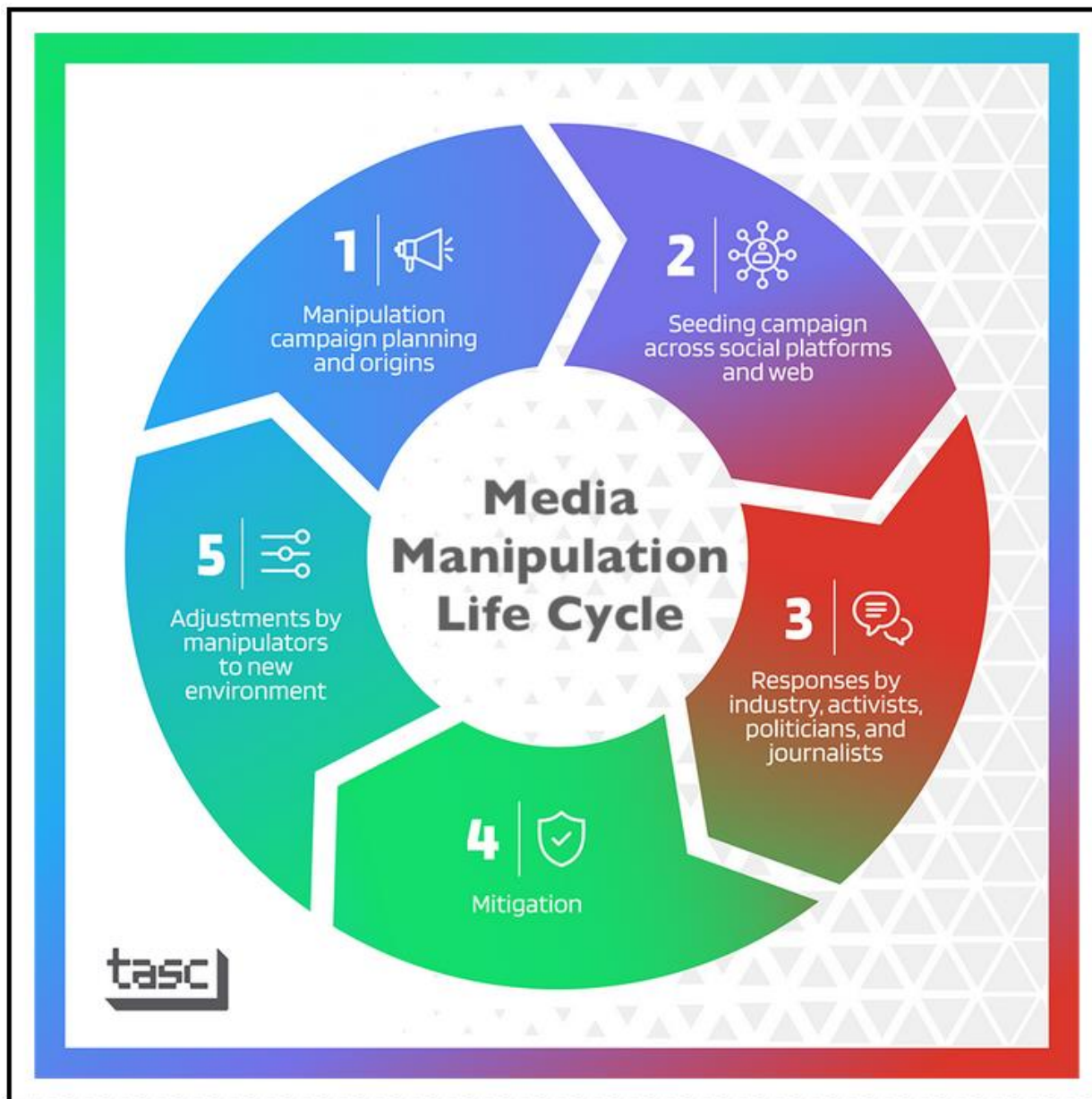
¹ Joan Donovan, “The Lifecycle of Media Manipulation,” in *Verification Handbook for Disinformation and Media Manipulation*, ed. Craig Silverman (European Journalism Centre, 2020), <https://datajournalism.com/read/handbook/verification-3/investigating-disinformation-and-media-manipulation/the-lifecycle-of-media-manipulation>.

² Donovan, “The Life Cycle of Media Manipulation.”

³ 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), <https://doi.org/10.1177/0002764219878229>.

⁴ Amelia Acker and Joan Donovan, “Data Craft: A Theory/Methods Package for Critical Internet Studies,” *Information, Communication & Society* 22, no. 11 (September 19, 2019): 1590–1609, <https://doi.org/10.1080/1369118X.2019.1645194>.

⁵ Matt Goerzen, Elizabeth Anne Watkins, and Gabrielle Lim, “Entanglements and Exploits: Sociotechnical Security as an Analytic Framework,” in *9th USENIX Workshop on Free and Open Communications on the Internet (FOCI)* (Santa Clara, CA, 2019), <https://www.usenix.org/conference/foci19/presentation/goerzen>.



Stages of the Media Manipulation Life Cycle

Stage 1: Manipulation campaign planning and origins

Stage 1 documents the campaign planning process and its origins. Depending on the availability of source materials and evidence, this stage would typically describe the platforms and technologies used by campaign operators to plan and coordinate, as well as the social and technical circumstances that facilitated the campaign's genesis. This may include evidence from private and semi-private chat applications (ex. WhatsApp or Telegram), less popular platforms (ex. Discord, 4chan, other message boards), and more mainstream platforms (ex. Twitter or Facebook). Due to the clandestine nature of campaign planning, it is not always feasible, ethical, or legal for researchers to obtain such evidence. However, when it is possible, available evidence that elucidates the campaign planning stage should be included.

Stage 2: Seeding the campaign across social platforms and web

Stage 2 documents the tactics and relevant technologies used to execute the campaign. It details the dissemination and propagation of content relevant to the operation. Typically, this stage 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, a particular Twitter hashtag, or across the open web through the strategic use of keywords. The rationale is often to reach as many individuals as possible so as to achieve a critical mass in conversation that will lead to a campaign becoming newsworthy, result in a false perception of massive public concern, sway public opinion, recruit followers, or a number of other off- and online responses.

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

After content has been seeded, the campaign moves on to Stage 3, which documents how institutional actors (ex. civil society organizations, politicians, political parties, mainstream media outlets) amplify, adopt, or extend the campaign. The third stage of the operation is usually a turning point indicating whether the campaign was effective in gaining attention through amplification or if led to an observable outcome. Responses may include public statements by representatives from social media platforms, activist campaigns drawing attention to malicious behavior by campaign participants, official political statements, critical reporting in the mainstream press, or political adoption of an idea or narrative pushed by the campaign.

Stage 4: Mitigation

The fourth stage of a manipulation campaign documents actions by tech companies, government, journalists, or civil society to mitigate the spread of a campaign's content and messaging and its effects. This may include actions from civil society (ex. debunking and

research), technology companies (ex. user ban, account deletion, content removal), media organizations (ex. fact-checking and investigative reporting), or the government (ex. draft bills, regulatory changes, take down orders).

Stage 5: Adjustments by manipulators to new environment

The fifth stage of a manipulation campaign involves how the operators and campaign participants adapt according to mitigation efforts described in Stage 4 and the 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 by creating new accounts, adapting coded language, altering audio/visual material, and iterating on narratives already identified as objectionable by platforms.

Note that while there may be no evidence of tactical or strategic adaptation, that does not imply the operators did not adapt. Media manipulation campaigns are often covert and as such the operators may have evolved to become better at hiding their tracks. However, if Stage 5 includes successful tactical adaptation or redeployment, a new cycle may begin (i.e. Stage 5 actions turn into Stage 1).

Variables

The following variables and values are used to code each case study. They are emergent from the cases and determined after multiple rounds of review. As new cases are included in the Casebook, the variables and values may change. In this case a new version of the Codebook will be released, and any changes noted.

Case Name

Type: Text

Descriptive case title.

Region

Type: Text

The geographical location where the campaign was most likely carried out based on the evidence available regarding the campaign's origins, participants, or audience. If evidence is inconclusive or unavailable, "Unclear" may be used.

Date

Type: Numerical

Date or date range the campaign was carried out based on available evidence.

Strategy

Type: Categorical (multiple selection allowed)

The plan of action or series of actions designed to achieve an overall goal as observed by the available evidence. Multiple selection is allowed as there may be several strategies in play or working in tandem with one another.

Astroturfing - Astroturfing occurs when campaign operators attempt to create the false perception of grassroots support for an issue by concealing their identities and using other deceptive practices, like hiding the origins of information being disseminated or artificially inflating engagement metrics.

Butterfly attack - Butterfly attacks occur when imposters mimic the patterns of behavior of a social group (usually a group that has to fight for representation). Imposters pretend to be part of the group in order to insert divisive rhetoric and disinformation into popular online conversation or within the information networks used by these groups. Distinct from astroturfing, which tries to falsify grassroots support for an issue, butterfly attacks are designed to infiltrate existing communities, media campaigns, or hashtags to disrupt their operations and discredit the group by sowing divisive, inflammatory, or confusing information.

Coined by Patrick Ryan to describe a series of manipulation campaigns he claims to have orchestrated in 2013, the term butterfly attack is inspired by the mimicry behavior of certain species of butterflies, who impersonate the fluttering patterns of other species to confuse predators.⁶

Gaming an algorithm - Attempting to manipulate an algorithm in order to gain attention. This may include tactics that elevate content into a platform's trending list, being recommended to other users, or placing in the top ten of a search engine's results.

Meme war - The intentional propagation of political memes on social media for the purpose of political persuasion, community building, or to strategically spread narratives and other messaging crucial to a media manipulation campaign.

Muddy the waters - The distribution of information with intent to create confusion during unresolved events that precedes or obscures verified information and consensus. In doing so, the target subject becomes more confusing as credible or authoritative sources are forced to compete with speculation, unfounded claims, or outright false information.

Targeted harassment - Coordinated and organized online harassment of an individual or groups of individuals to threaten, censor, or upset them or to disrupt their operations or behavior.

Trading up the chain - Gaining exposure by placing information or disinformation artifacts in locations that will be taken up and amplified by other systems, individuals, or publications. Typically, information may be introduced on smaller blogs or social media before being reported by mainstream media outlets or politicians and other influential individuals.

Unclear strategy - There is no discernible strategy based on the available evidence.

⁶ Patrick Ryan, "The Butterfly War," October 13, 2017, <https://cultstate.com/2017/10/13/The-Butterfly-War/>.

Tactics

Type: Categorical (multiple selection allowed)

Actions employed in service of the strategy as observed by the available evidence.

Bots - Social media accounts that are automated and deployed for deceptive purposes, such as artificially amplifying a message, to game a trending or recommendation algorithm, or inflate an account's engagement metrics. These accounts are typically centrally controlled or at least in coordination with each other.

Cheap Fake - The use of conventional editing techniques like speeding, slowing, or cutting, footage or images to create a false impression of an individual or an event.

Cloaked science - The use of scientific jargon and community norms to cloak or hide a political, ideological, or financial agenda within the appearance of legitimate scientific research. This may include the use of technical language, difficult-to-understand graphs and charts, or seemingly scientific data presented as empirical evidence to lend credibility to the claims being made. Cloaked science may be seeded onto public preprint servers, in data repositories, journals, or publications with lax review standards, through press releases, or by baiting journalists who may not be able to scrutinize the claims thoroughly.

This definition builds upon Jessie Daniel's research on cloaked websites, which she describes as "sites published by individuals or groups who conceal authorship in order to disguise deliberately a hidden political agenda,"⁷ and is inspired by Sarah Richardson's description of transphobic politics being "cloaked in science."⁸ Science scholar Timothy Caulfield, describing a similar phenomenon – the use of scientific language to mask otherwise unscientific motives (ex. financial gain) – uses the term "scienceploitation."⁹

Note that cloaked science as a tactic is the deliberate use of information masquerading as science and should not be confused with "junk science," which is a term that has been used to discredit scientific findings, claims, and data as fraudulent or misleading in a similar way that "fake news" can be used to dismiss critical news coverage.¹⁰

⁷ Jessie Daniels, "Cloaked Websites: Propaganda, Cyber-Racism and Epistemology in the Digital Era," *New Media & Society*, July 21, 2009, <https://journals.sagepub.com/doi/10.1177/1461444809105345>.

⁸ Sarah Richardson, "Transphobia, Cloaked in Science," *Los Angeles Review of Books* (blog), November 8, 2018, <https://blog.lareviewofbooks.org/essays/transphobia-cloaked-science/>.

⁹ Timothy Caulfield, "Pseudoscience and COVID-19 – We've Had Enough Already," *Nature*, April 27, 2020, <https://doi.org/10.1038/d41586-020-01266-z>.

¹⁰ Jonathan M. Samet and Thomas A. Burke, "Turning Science Into Junk: The Tobacco Industry and Passive Smoking," *American Journal of Public Health* 91, no. 11 (November 1, 2001): 1742–44, <https://doi.org/10.2105/AJPH.91.11.1742>.

Copy-pasta - A portmanteau of “copy,” “paste,” and “pasta,” copy-pasta refers to any block of text that is repeatedly reposted, often on social media, messaging apps, online discussion forums, and comments sections.

Distributed amplification - A call to participants to rapidly and widely spread campaign materials, including propaganda or disinformation.

Evidence collage - Compiling information from multiple sources into a single, shareable document, usually as an image, to persuade or convince a target audience.

Forgery - The creation of a fake document with intent to deceive via distribution.

Hijacked accounts - The unauthorized use of an individual’s account, typically accessed through stolen credentials or hacking. Hijacked accounts can refer to email accounts, social media profiles, messaging apps, or any other account associated with a digital service or product.

Impersonation - Pretending to be another person or member of a social identity group, either by mimicking their behavior or creating a falsified online presence.

Keyword squatting - The strategic domination of unique or under-used keywords on a social media platform or search engine that will return search results and content in favor of the campaign operators’ goals.

Leak - The unauthorized release of sensitive materials or documents.

Memes - Memes, a term coined by Richard Dawkins (1976), are “units of culture” that spread through the diffusion of ideas. Memes are particularly salient online because the internet crystallizes them as artifacts of communication and accelerates their distribution through subcultures. Within media manipulation they typically take on the form of images, gifs, or videos.

Misinfographic - Infographics with false or misleading information. In some cases, they may also be classified as a forgery when they borrow an existing organization's brand aesthetics and logo in order to make it seem as if the content was coming from the organization.

Phishing - Fraudulently posing as a trustworthy entity in a malicious attempt to access confidential information such as usernames, passwords and credit card details, usually by the means of email.

Recontextualized media - Media (image, video, audio) that has been taken out of its original context and reframed for an entirely different purpose or narrative frame.

Swarming - When loosely organized online groups come together for specific objectives or campaigns.

Trolling - Engaging in inflammatory, divisive, or distracting behavior in an online community with the goal of provoking readers or viewers into an emotional, often negative, response (ex. anger, outrage, offense).

Typosquatting - The intentional registration of a domain name that incorporates typographical variants of the target domain name in order to deceive visitors. This may involve misspelling a domain or using a different top-level domain. Typosquatting is a form of cybersquatting, or an attempt to mislead users by fraudulently posing under someone else's brand or copyright.

Viral sloganeering - Repackaging of provocative, revolutionary or reactionary talking points into a short, catchy, and memorable format for social media and press amplification.

Network terrain

Type: Categorical (multiple selection allowed)

The digital platforms and technologies used to carry out the campaign.

4chan (4chan.org) - A website that hosts message boards on a variety of topics from music to politics to anime. The site is broken down into distinct message boards where mostly anonymous users post and reply to threads.

8chan (8ch.net) - A now defunct imageboard owned by Jim Watkins that has been linked to the propagation of white supremacy, neo-Nazism, and manifestos by multiple mass shooters. In 2019, the website was taken offline after termination of its DNS and domain registration. The site reformed and rebranded itself as "8kun" in November 2019.

Bitchute (bitchute.com) – Launched in 2017, Bitchute is a video-hosting platform that positions itself as a “free speech” alternative to YouTube. The site hosts videos promoting violent far-right ideology, medical disinformation, and conspiracy theories. Bitchute links are blocked on Twitter as of August 2020.

Blogs - Self-published websites or web pages, with no editorial oversight, that are usually run by an individual or small group and are regularly updated with new content.

Discord (discord.com) - Discord is an instant messaging software that launched in 2015. Users can send text, video, images, and audio through its channels. This value includes web app, desktop app, and other means of use.

Facebook (facebook.com) - A social networking website that allows registered users to create unique profiles, have public and private conversations, join groups, create events, upload photos, etc. This value includes web app, mobile app, and other means of use.

Gab (gab.com) - A social networking service launched publicly in May 2017. It is known for its right-leaning user base, and as a safe haven for far-right extremists.

Google - google.* Includes all other top-level domains owned by Google (ex. Google.ca, google.com, google.co.uk). This code specifically refers to Google's flagship search engine product and not to the company's subsidiaries or other products.

Instagram (instagram.com) - Acquired by Facebook in 2012, Instagram is a social network platform that enables users to edit, upload and comment on photos and short videos, broadcast live videos, and have private conversations using the chat feature. This value includes web app, mobile app, and other means of use.

Media outlets - News and entertainment publishers that provide news and feature stories to the public and are not owned or controlled by the state. They may be distributed over broadcast (TV and radio), online, or print media. This variable includes independent and alternative media, mainstream corporate press, and publicly funded media that are free from state interference (ex. *BBC* and *NPR*).

Open collaboration tools - Open collaboration tools are services such as Pastebin or Google Docs, which are open-access, easy-to-use services for hosting, crowdsourcing, and sharing information. The openness of these tools presents an opportunity for campaign planners and participants, giving them a frictionless, easy-to-share repository for collaboration, coordination, and information distribution. Depending on the intentions of the campaign, they can be used for advocacy, resource-sharing, and activism, as well as more malicious means, such as housing false or misleading information, unauthorized leaks of personal and private information, or harassment campaign instructions. Other examples of open collaboration tools include Dropbox, Jira, Asana, and Trello.

Open editorial platforms - Platforms that have both an editorial arm and a self-publishing arm for users to publish and post their own articles and other content. Examples include Medium and BuzzFeed Community.

Parler - Parler was marketed as the free-speech alternative to popular social media. The social media site was launched in 2018 with an interface similar to Twitter's, and gained users rapidly after the 2020 presidential election. After the January 6, 2021 attack on the U.S. Capitol, it was removed from the Apple and Google app stores, and dropped by its server host, Amazon, based on its alleged use by people involved in the insurrection.

Preprint server - A preprint server is a data repository that hosts scholarly articles before they have been formally peer-reviewed and published. Typically, the author(s) of a paper will upload a draft version to the preprint server, which generates a publicly accessible URL where the paper and any other relevant data may be viewed, downloaded, or shared. Because the peer review process can often take months or even years, preprints are a quick way to circulate research outputs and data. However, authors of preprints may seek formal publication after uploading to a preprint server. Preprint servers can be discipline-specific (ex. bioRxiv), regional (ex. AfricArXiv), multidisciplinary (ex. arXiv, SocArXiv), or general purpose (ex. Zenodo).

Reddit (reddit.com) - Reddit is a website where users can post information and prompts. These posts and prompts get responses and up or down votes from other users, which ranks the display of the content. The platform is divided into user-created categories called "subreddits." Value includes web and mobile uses.

State-controlled media - Media outlets that are under editorial control or influence by a country's government. The articles and stories produced by these state media outlets may be distributed over broadcast (TV and radio), online, or print media. State-controlled media is a designation that applies when editorial freedom has been taken away by government influence, pressure, or money. The label does not necessarily apply to all media that receives funding from the public. Media organizations that receive public funds but maintain their editorial freedom, such as the *British Broadcasting Corporation (BBC)* or *Canadian Broadcasting Corporation (CBC)*, are not designated as state media.

Twitter (twitter.com) – A micro-blogging platform where users can post "tweets" of a limited character length and video/audio messages. These user accounts can like, comment on, and share other users' messages. Some user accounts are "verified" by the company, which bestows on the accounts special privileges, such as more moderation options. This value includes both web and mobile app use.

Vimeo (vimeo.com) - Vimeo is a video-sharing platform launched in 2004.

YouTube (youtube.com) - An online video-sharing platform launched in 2005. It is a subsidiary of Google.

Vulnerabilities

Type: Categorical (multiple selection allowed)

The social and technical conditions being exploited by the campaign.

Active crisis - A period of time when the normal state of affairs is interrupted by unforeseen events that are troubling and potentially dangerous. Active crises trigger confusion and require

urgent action and immediate attention. Due to the increased media attention and importance of any decisions made during this time, active crises are vulnerable to being exploited by media manipulation.

Breaking news event - Periods of heightened attention to current events of local, national, or international importance in mass media and on social media. During these moments of mass attention, legitimate information and misinformation may be indistinguishable until facts are established and vetted by official bodies. This period of confusion creates opportunities to sow confusion, target individuals, or shape certain narratives.

Data void - Coined and theorized by Michael Golebiewski and danah boyd (2018), this refers to unique topics or terms that result in minimal, low quality, or manipulative information from search engine queries. Data voids are social or technical securities risks depending on the subject matter of the query.

Election period - Refers to the time leading up to an election when candidates have begun campaigning. Depending on the country, there may be legal limits to what constitutes a campaign period.

Lax security practices - A lax security practice is anything that makes the user more vulnerable to security attacks or scams, like phishing. An example of a lax security practice is having a password that can be guessed easily or is repeated across multiple accounts.

Open science - Open science is an approach to scientific inquiry that advocates for collaboration, accessibility, and transparency in an effort to increase the dissemination of scientific knowledge and participation of individuals from diverse backgrounds. Common practices include making research data public, campaigning for open access, and communication strategies that are inclusive of a wide audience.¹¹

Although open science as a movement confers multiple benefits,¹² its openness and inclusivity can be exploited by motivated actors intent on seeding false or misleading content. Digital data repositories and preprint servers, for example, are an outcome of the movement for open science, but because of their lack of peer review they can be misused or abused to spread poor

¹¹ "Open Science Movement | United Nations Educational, Scientific and Cultural Organization," accessed January 3, 2021, <http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/open-science-movement/>; "What Is Open Science? Introduction," Foster Open Science, accessed January 3, 2021, <https://www.fosteropenscience.eu/content/what-open-science-introduction>.

¹² Christopher Allen and David M. A. Mehler, "Open Science Challenges, Benefits and Tips in Early Career and Beyond," *PLOS Biology* 17, no. 5 (May 1, 2019): e3000246, <https://doi.org/10.1371/journal.pbio.3000246>; Martin Lakomý, Renata Hlavová, and Hana Machackova, "Open Science and the Science-Society Relationship," *Society* 56, no. 3 (June 1, 2019): 246–55, <https://doi.org/10.1007/s12115-019-00361-w>.

quality research or disinformation masked as science.¹³ Publicly available data, even if credible and from authoritative sources, can also be manipulated to mislead or undermine scientific consensus.¹⁴

Prejudice - A bias that can result in an injury or detriment to another individual's legal rights or claims, wellbeing, or participation in society. Such preconceived judgements are not informed by facts and often target an individual or group based on race, religion, sexual orientation, age, class, or other demographic identifier.

Public directory - Publicly available information pertaining to individuals, organizations, companies, or any other entity that has been aggregated into an accessible, searchable, and organized format.

Wedge issue - Political or social issues that are divisive in nature and divide social groups. They typically split along partisan lines and are often presented as binary positions – for or against. Politicians, political influencers, and those running for office will often amplify these wedges in popular discourse, in mainstream press, and on social media.

Attribution

Type: Categorical (multiple selection allowed)

The individuals or groups responsible for planning, carrying out, or participating in the campaign based on the available evidence.

Conspiracists - Individuals or groups that actively propagate unfounded or unverified narratives and frames. This often includes speculation, unsubstantiated claims, and explanations predicated on secretive and powerful actors scheming with malicious intent.

Extremists (right wing) - Groups or individuals that espouse right-leaning radical or violent positions, often associated with organized white supremacy or other prejudice-driven ideologies.

¹³ Amy Koerber, "Is It Fake News or Is It Open Science? Science Communication in the COVID-19 Pandemic," *Journal of Business and Technical Communication* 35, no. 1 (January 1, 2021): 22–27, <https://doi.org/10.1177/1050651920958506>; Joan Donovan, Irene Pasquetto, and Jennifer Pierre, "Cracking Open the Black Box of Genetic Ancestry Testing," *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018, <https://doi.org/10.24251/HICSS.2018.218>; Aaron Panofsky and Joan Donovan, "Genetic Ancestry Testing among White Nationalists: From Identity Repair to Citizen Science," *Social Studies of Science* 49, no. 5 (October 1, 2019): 653–81, <https://doi.org/10.1177/0306312719861434>.

¹⁴ Crystal Lee, Tanya Yang, Gabrielle Inchoco, Graham M. Jones, and Arvind Satyanarayan, "Viral Visualizations: How Coronavirus Skeptics Use Orthodox Data Practices to Promote Unorthodox Science Online," CHI '21, May 8–13, 2021, Yokohama, Japan, <https://arxiv.org/pdf/2101.07993.pdf>.

Influencers - Visible pundits, journalists, or public figures who drive conversation around particular topics in broadcast media and online networks.

Partisans - A strong supporter or committed member of a party, cause, or person.

Prankster - Individuals who engage in activity designed to elicit a reaction from a target purely for fun or mischief.

Public relations or marketing firm - A private company that engages in public relations, branding, advertising and sales, or any other type of activity related to marketing, typically in service to a client. Marketing firms and public relations companies have been used in media manipulation campaigns to game engagement metrics, create the false sense of grassroots support (i.e., astroturfing), and amplify specific narratives or pieces of content for their clients.

Networked faction - Tacit coalitions or groups of people who share some, but not all, political positions, primarily congregate online (though not exclusively), and often come together as a swarm to act in unison as a political force. Networked factions maintain these coalitions using shared phrases, hashtags, memes, or similar media. These factions can form and dissolve according to the political context.

State actor – An entity that is a part of, or which operates licitly or semi-licitly on behalf or in service of, a government agency. Within media manipulation, this may refer to state-run media outlets, operatives working for or with an intelligence or security agency (or other government agency), or other parties that are deliberately working to advance a given state’s objectives with the support, encouragement, or compulsion of the state.

Trolls - Individuals who engage in inflammatory, divisive, or distracting behavior in an online community with the goal of provoking readers or viewers into an emotional, often negative, response (e.g., anger, outrage, offense).

Unclear attribution - Attribution, whether referring to the campaign planners or participants, is unclear based on available evidence.

Targets

Type: Categorical (multiple selection allowed)

The individual or group the campaign intended to discredit, disrupt, criticize, or frame in a negative light based on available evidence.

Activist group - Individuals or groups that campaign for social, political, or legal change. They may be formally organized (ex. registered non-governmental organization) or loosely affiliated (ex. advocacy networks).

Individual - A single person.

Political party - A group of people sharing similar ideology or political positions who participate in elections by fielding candidates that will then carry out their goals and policies.

Politician - A person engaged in party politics or occupying public office.

Scientific and medical community - Individuals or groups involved in scientific research, medicine, or healthcare. This may include scientists, researchers, research labs, scientific organizations, health authorities, doctors, nurses, and other healthcare professionals.

Social identity group - Groups defined by some social, physical, or mental characteristics. Examples include race, ethnicity, gender, social class, sexual orientation, or religious beliefs.

Observable outcomes

Type: Categorical (multiple selection allowed)

The results (intended or unintended) of the campaign as observed by available evidence.

Dox - The act of publishing on the internet private or identifying information about a specific individual against their wishes and usually with malicious intent (i.e. retaliation, punishment).

Harassment - Targeted and repeated behavior towards an individual or group of people that causes mental, physical or emotional distress. Harassment includes but is not limited to unwanted threats, insults, touching or offensive language.

Media exposure - Coverage and reporting by journalists in popular or mainstream media.

Misidentification - Erroneously identifying an individual as someone else, intentionally or accidentally.

Political adoption - When a political party or politician adopts or co-opts a phrase, term, or idea for politically motivated purposes.

Recognition by target - When a target of a media manipulation or disinformation campaign acknowledges and responds to the campaign's activities or the operators.

Mitigation

Type: Categorical (multiple selection allowed)

Attempts, measures, and other actions taken by the private sector, government, media organizations, and civil society in an attempt to contain or prevent the continuation of a campaign.

Account suspension - Accounts that have been suspended by a platform or company, preventing the user from log in or using the account.

Blocking - User-instigated action that prevents another account from interacting with them or viewing their content.

Civil society response - Civil society response refers to actions taken by members or groups of civil society in an attempt to mitigate a campaign's harms or spread. We define civil society as groups or organizations engaged in advocating for certain issues, educating the wider public, holding the government accountable, or promoting civil and human rights. They may be formally organized or loosely coordinated and include non-governmental organizations (NGOs), community groups, labor unions, educational organizations, faith-based organizations, professional associations, non-profit think tanks, and foundations.

Civil/Private lawsuit - A legal proceeding by a private party or parties against another in a civil court of law that seeks remedy for a wrongdoing or harm.

Content removal - Content removal is the act of platforms taking down specific pieces of content, like videos, tweets, posts, etc. The platform's terms of service are often a guideline for what can be removed, though these are rarely enforced uniformly or consistently.

Counterspeech - A tactic used for countering hate speech and misinformation by advancing alternative narratives and challenging information.

Criminal investigation - All activities involved in the process of investigating and prosecuting a crime including collecting evidence or information pertaining to a crime, apprehending a suspect, and any subsequent related proceedings such as a trial or sentencing.

Critical press - Press coverage that is critical of a manipulation campaign. Articles may debunk false claims or investigate the origins and motivations of a campaign.

Debunking - Exposing and correcting false or misleading claims. Debunking includes fact-checking efforts, research and investigation, exposés, and other critical content or actions that attempt to correct the false claims.

De-indexing - Removing a link or other content from search results. The content or website in question is still available but will not be included in a search engine's, website's, or platform's results.

Deplatforming - The removal of individuals or groups from a platform, preventing them from using the platform's services even if they try to create new accounts.

Flagging - Reporting harmful or offensive content to an online social media platform or company. Content can be flagged by an algorithm, content moderator, or another user.

Labeling - Labelling refers to the application of informational labels to social media posts, accounts, channels, or other content by the host platform in an effort to give viewers additional context. This may include labeling content that could be potentially sensitive or graphic, affiliated with a nation state, containing false or misleading claims, or at risk of inciting violence.

Media blackout - Self-imposed or state mandated censorship of a certain news topic.

Research and investigation - Individual or coordinated group efforts to establish the origins and impact of a manipulation campaign.

Campaign adaptation

Type: Categorical (multiple selection allowed)

Actions taken by campaign operators and participants in response to the observable outcomes and mitigation attempts. Details of campaign adaptations will be described in case studies when applicable.

Tactical adjustment - The continuation of a media manipulation or disinformation campaign with adjustments to the tactics or new tactics altogether.

Tactical redeployment - The redeployment of a media manipulation or disinformation campaign's tactics.

Unclear or no adaptation - There is no adaptation or redeployment of tactics based on the available evidence.