

May 2023

CIVIC INFORMATION HANDBOOK

Karen Kornbluh and Adrienne Goldstein

in collaboration with UNC Center for Information, Technology, and Public Life

TABLE OF CONTENTS

Introduction	3
Definitions and terminology	4
The Model: Online Networked Information Campaigns	5
Building the pipeline	6
Bringing the narrative to life	10
Flooding the zone	13
Activation	17
Actors and Goals	19
Who are the actors behind coordinated deceptive campaigns?	19
Who are the targeted audiences?	19
The Case Study of Texas Wind Turbines	21
Why Does This Work?	25
Psychological and sociological drivers of mis- and disinformation	25
Building Civic Infrastructure	28
The Case Study of Ukraine’s Social Media Mastery	33
Conclusion	35
Appendix: Building Resilience Against Online Attacks	36
About and Acknowledgements	38

INTRODUCTION

In our current online world, civic information—important information needed to participate in democracy—is too often drowned out by viral falsehoods, including conspiracy theories.

Often, this is not an accident. Carefully orchestrated social media campaigns exploit social media tools, like algorithmic amplification and micro-targeting, to manipulate users and the information environment. These campaigns leverage the inherent platform product design to promote narratives, sell products, persuade users, and even provoke users to act for political, economic, or social purposes.

As a result, today’s civic leaders must play a more active role in the amplification of fact-based information. As Dr. Anthony Fauci said, “we’ve gotta be out there — scientists and the general public and those who understand the facts — talking about true and correct information.”¹

*Today’s civic leaders must play a more active role in the
amplification of civic information.*

To be sure, the playing field is not even. Social media platform tools are better suited for campaigns seeking to manipulate and agitate users than to empower and inform. Platforms and regulators must get involved to fix the design flaws that allow false and misleading information to flourish in the first place.² Policymakers should update and enforce civil and human rights laws for the online environment, compel radical transparency, update consumer protection rules, insist that industry make a high-level commitment to democratic design, and create civic information infrastructure through a new PBS of the Internet. In the absence of such policy reform, amplifiers of civic information may never be able to beat out the well-resourced, well-networked groups that intentionally spread falsehoods. Nonetheless, there are strategies for helping civic information compete.

This handbook aims to:

1. Educate civic information providers about coordinated deceptive campaigns

...including how they build their audiences, seed compelling narratives, amplify their messages, and activate their followers, as well as why false narratives take hold, and who the primary actors and targeted audiences are.

2. Serve as a resource on how to flood the zone with trustworthy civic information

...namely, how civic information providers can repurpose the tactics coordinated deceptive campaigns use in transparent, empowering ways and protect themselves and their message online.

This handbook will function as a media literacy tool, giving readers the skills and opportunity to consider who is behind networked information campaigns and how they spread their messages.

¹ Morning Joe, “[Dr. Fauci: Disinformation can be dangerous to the health of the nation](#),” MSNBC, December 13, 2022.

² Karen Kornbluh and Ellen P. Goodman, “[Safeguarding Digital Democracy: Digital Innovation and Democracy Initiative Roadmap](#),” German Marshall Fund of the United States, March 24, 2020, 4.

Its focus is limited to how information spreads on social media, but modern networked information campaigns work across an entire ecosystem of on- and offline tactics. Information campaigns use radio, mail, email, print media, television, and face-to-face communication.³

Definitions and terminology

An array of terms are applied to viral falsehoods, including fake news, misinformation, disinformation, malinformation, propaganda, and, in the national security context, information operations, hybrid threats, and hack and leak. Mis-, dis- and malinformation, as defined by Claire Wardle and Hossein Derakhshan, are three of the most prevalent today:

- *Misinformation* – “Information that is false, but not created with the intention of causing harm.”⁴
- *Disinformation* – “Information that is false and deliberately created to harm a person, social group, organization, or country.”⁵
- *Malinformation* – “Information that is based in reality, used to inflict harm on a person, organization, or country.”⁶ Examples include leaks, harassment, and hate speech.

While it is important to distinguish between the intentional and unintentional spread of falsehoods, discussions around mis- and disinformation tend to center the veracity of a specific narrative or piece of content. What is missing, oftentimes, is a focus on how false or misleading narratives are deployed by deceptive actors to accomplish a strategic goal. In addition to “what is true?”, we need to ask the question: who benefits? For that reason, in this handbook we often use the terms “networked information” or “coordinated deceptive” campaigns:

- *Networked information campaigns* – a combination of grassroots efforts and a central organizer who frames issues, coordinates energies, and sets goals to spread any type of information – civic or false.
- *Coordinated deceptive campaigns* – a subset of networked information campaigns that spread false or misleading information.

³ Alice E. Marwick and Rebecca Lewis, “[Media manipulation and disinformation online](#),” Data & Society Research Institute, 2017; Mark Kumleben, Samuel Woolley, and Katie Joseff, “[Electoral Confusion: Contending with Structural Disinformation in Communities of Color](#),” June 2022; Yochai Benkler, Robert Faris, and Hal Roberts, “Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics,” Oxford University Press, October 18, 2018.

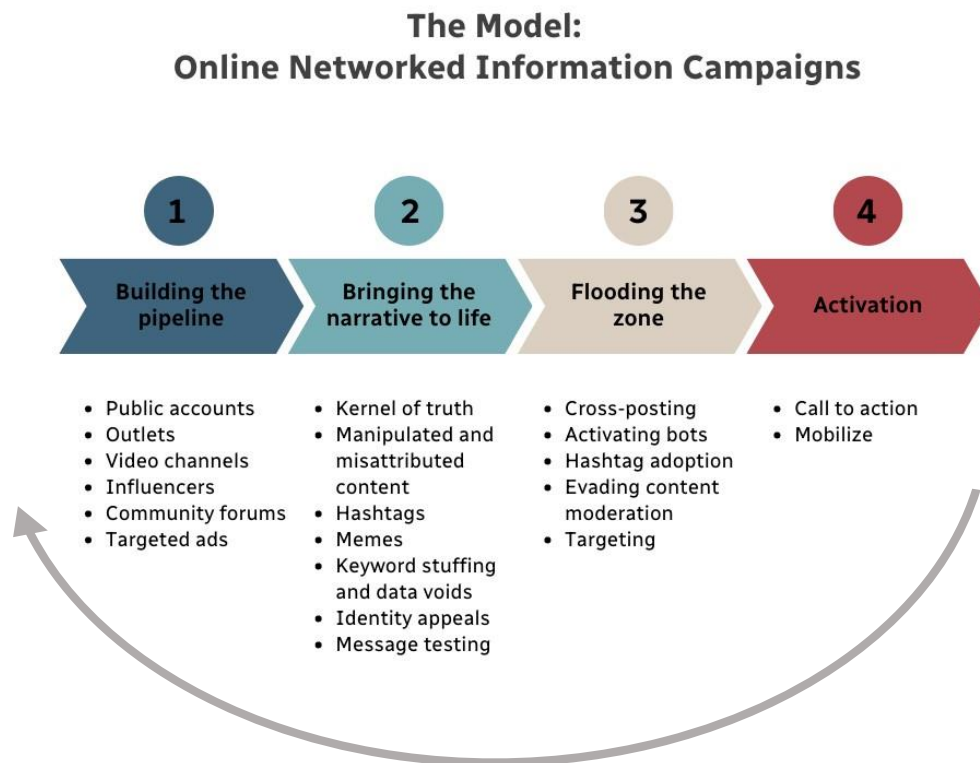
⁴ Claire Wardle and Hossein Derakhshan, “[Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making](#),” Council of Europe Report, 2017, 20.

⁵ Wardle and Derakhshan, “[Information Disorder](#),” 20.

⁶ Wardle and Derakhshan, “[Information Disorder](#),” 20.

THE MODEL: ONLINE NETWORKED INFORMATION CAMPAIGNS

The most salient and prominent online conspiracy theories rarely go viral on their own. Successful ones reach their target audiences by deploying a series of tactics that leverage the platforms' design loopholes through several steps, involving all the tools and infrastructure social media has to offer.



Successful networked information campaigns begin by building massive digital audiences, often through large public accounts and partisan, opinion-oriented outlets. Campaigners then seed narratives that appeal to their targeted audiences and flood the zone with their message. Finally, successful networked information campaigns activate their audiences through opportunities to become involved and refine their messages based on signals from their audiences.

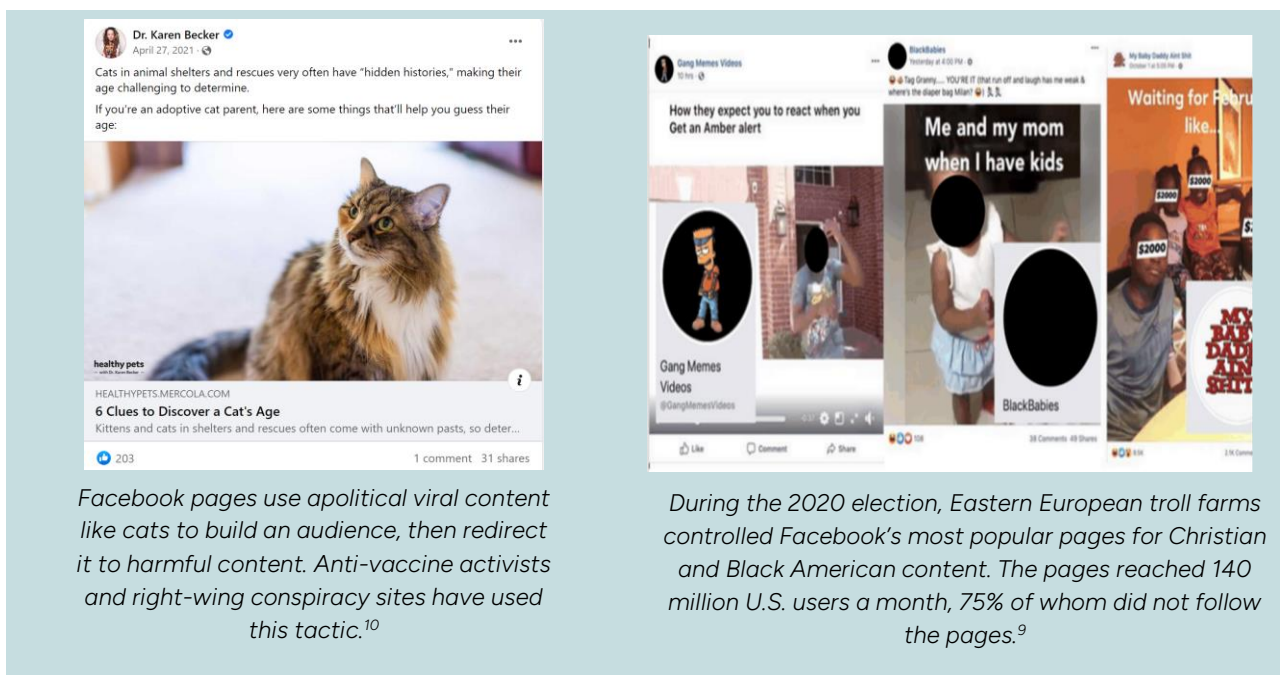
Although coordinated deceptive campaigns deploy these tactics in ways that are inherently manipulative, similar tactics can be used by providers of civic information to share factual information that enables people to engage more fully in their communities. For example, whereas manipulated and misattributed content is deceptive by its very definition, influencers, hashtags, and message testing can be used for civic information campaigns.

Building the pipeline

The first step in coordinated deceptive campaigns involves building up an ongoing audience, or “pipeline.” The biggest networked, manipulative campaigns are successful because they put in the effort to build audiences using outlets, public accounts, channels, and influencers across platforms. Each of these grows followers using distinct tactics.

Public accounts: Disguise their intentions or identities through innocuous content to build their audiences

Deceptive information campaigns grow audiences for public accounts – such as Facebook pages and Twitter, Instagram, and TikTok profiles – by disguising their intentions and identities. They run non-political, attractive content like cute cats on their accounts only to later leverage their followers for political and economic gain. Some coordinated deceptive campaigns pay to grow their audiences, either through advertisements or by paying other accounts to cross-post their content.⁷ Platforms also reward accounts for posting engaging content, showing users content from accounts they do not follow if someone they follow comments or reshares posts.⁸



Facebook pages use apolitical viral content like cats to build an audience, then redirect it to harmful content. Anti-vaccine activists and right-wing conspiracy sites have used this tactic.¹⁰

During the 2020 election, Eastern European troll farms controlled Facebook's most popular pages for Christian and Black American content. The pages reached 140 million U.S. users a month, 75% of whom did not follow the pages.⁹

Coordinated deceptive campaigns may operate profiles that appear superficially independent but are in fact centrally coordinated. For example, The Daily Wire posted its most interacted with article from the second half of 2021 51 times across 17 different Facebook pages.¹¹

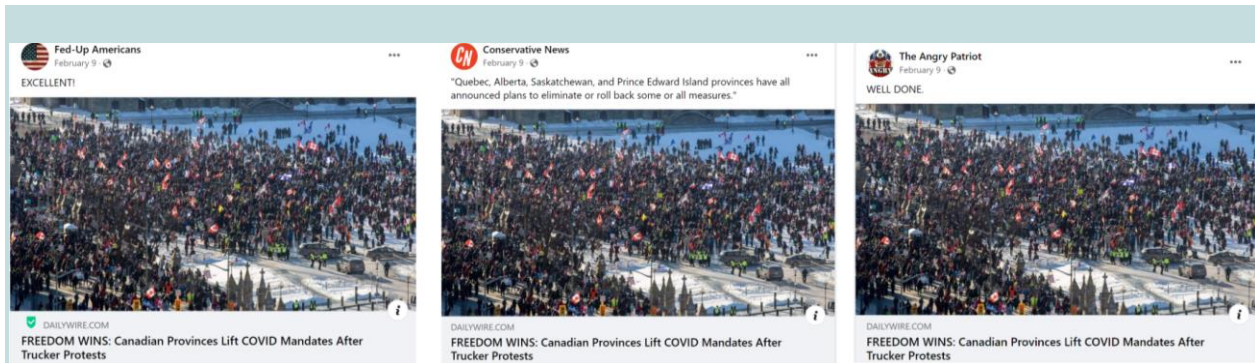
⁷ Adrienne Goldstein and Eli Weiner, “[Health Sites Built Coordinated Networks of Facebook Pages to Spread False Content, Increase Ad Revenue](#),” German Marshall Fund of the United States, December 9, 2020.

⁸ Karen Hao, “[Troll farms reached 140 million Americans a month on Facebook before 2020 election, internal report shows](#),” September 16, 2021.

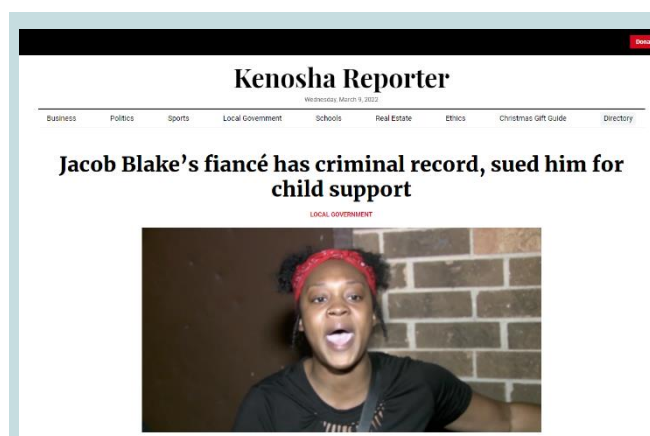
⁹ Karen Hao, “[Troll farms reached 140 million Americans a month on Facebook before 2020 election, internal report shows](#),” September 16, 2021.

¹⁰ Davey Alba, “[Those Cute Cats Online? They Help Spread Misinformation](#),” New York Times, December 1, 2021.

¹¹ Original GMF Digital research, using social media interaction data from NewsWhip.



The Daily Wire posts its content across a network of official outlet pages, many of which are not clearly connected to the site based on their page name and profile picture alone.¹²



The Tow Center for Digital Journalism at Columbia Journalism School discovered a network of 1,300 “pink slime” news sites designed to look like independent local newspapers which are, in fact, run by political operatives to seed false or misleading claims. These sites lie dormant most of the time but can be activated when there is a major event such as the Jacob Blake shooting in Kenosha, Wisconsin.¹⁶

Outlets: Mimic news outlets while acting in manipulative ways

Deceptive outlets, which we have called ‘trojan horse outlets’, use the trappings of independent journalism while eschewing journalistic standards of transparency to spread misleading narratives.¹³ News rating service NewsGuard evaluates whether outlets meet basic journalistic standards,¹⁴ and many of the highest-engagement online outlets are rated as failing key criteria.¹⁵

Outlets are effective in part because social media sharing presents all content in similar formats, which strips news stories from signals of journalistic integrity. From an article’s URL and thumbnail alone, it is unclear if the news site has separate sections for news and opinion, a masthead, and bylines and datelines.

Video channels: Build audiences for videos across social media platforms

¹² Judd Legum and Tesnim Zekeria, “[The dirty secret behind Ben Shapiro’s extraordinary success on Facebook](#),” Popular Information, June 25, 2020.

¹³ Karen Kornbluh and Ellen P. Goodman, “[Safeguarding Digital Democracy: Digital Innovation and Democracy Initiative Roadmap](#),” German Marshall Fund of the United States, March 24, 2020, 4.

¹⁴ NewsGuard, “[Rating Process and Criteria](#).”

¹⁵ For example, in the first three quarters of 2022, six of the top twenty outlets for Facebook interactions failed the NewsGuard standard of “gathers and presents information responsibly,” by which they mean the outlet references multiple sources and does not egregiously distort or misrepresent information. Original GMF Digital research, using social media interaction data from NewsWhip.

¹⁶ Priyanjana Bengani, “[Hundreds of ‘pink slime’ local news outlets are distributing algorithmic stories and conservative talking points](#),” Columbia Journalism Review, December 18, 2019; Davey Alba and Jack Nicas, “[As Local News Dies, a Pay-for-Play Network Rises in Its Place](#),” New York Times, October 18, 2020.


Videos can be particularly useful for deceptive information campaigns. Most social platforms do not highlight what videos are trending, making it harder for fact-checkers to intervene, and they are not easily searchable in the way that text or even photos are. Streaming services such as Twitch are especially difficult to moderate. YouTube channels use tools such as the subscribe button and recommendations to grow their audiences, though YouTube videos typically go viral thanks to amplification of the link on sites like Facebook and Twitter.¹⁷

Influencers: Influencers build audiences around their individual personas

Influencers often have public profiles across several platforms and may be influential in offline spaces as well as online. They develop intimate, one-sided, relationships with their audiences by sharing personal stories, posting selfies, and giving followers broad access to their lives, “mak[ing] sure the audience identifies with them, much in the way a friend would.”¹⁸ Followers are more likely to accept information as true if it is shared by friends and family.

Influencers build their audience by activating their existing supporters, giving them opportunities to participate in the organizing work of raising the visibility of the influencer’s account. “Follow-for-follow” campaigns are an example of this, in which influencers cross-promote one another’s accounts to their own followers.¹⁹

Influencers make money by selling or promoting products, soliciting donations, taking part in social media profit-sharing partnership programs, and creating content for others (see “[Targeting through paid influencer promotion](#)”).



Anti-vaccine influencer Joseph Mercola appeals to a broad audience using an everyman demeanor. He operates on Facebook, Instagram, Twitter, Telegram, Reddit, TikTok, an email list, and websites. Mercola translates content into nearly a dozen languages. Some of his accounts focus on topics like pets or fitness. In 2017, Mercola's net worth was in excess of \$100 million.²⁰

¹⁷ U.S. Senate Select Committee on Intelligence, “[Report on Russian Active Measures Campaigns and Interference in the 2016 U.S. Election, Volume 2: Russia’s Use of Social Media with Additional Views](#),” October 8, 2019, 58.

¹⁸ Becca Lewis, “[Why influence matters in the spread of misinformation](#),” Medium, November 20, 2018.

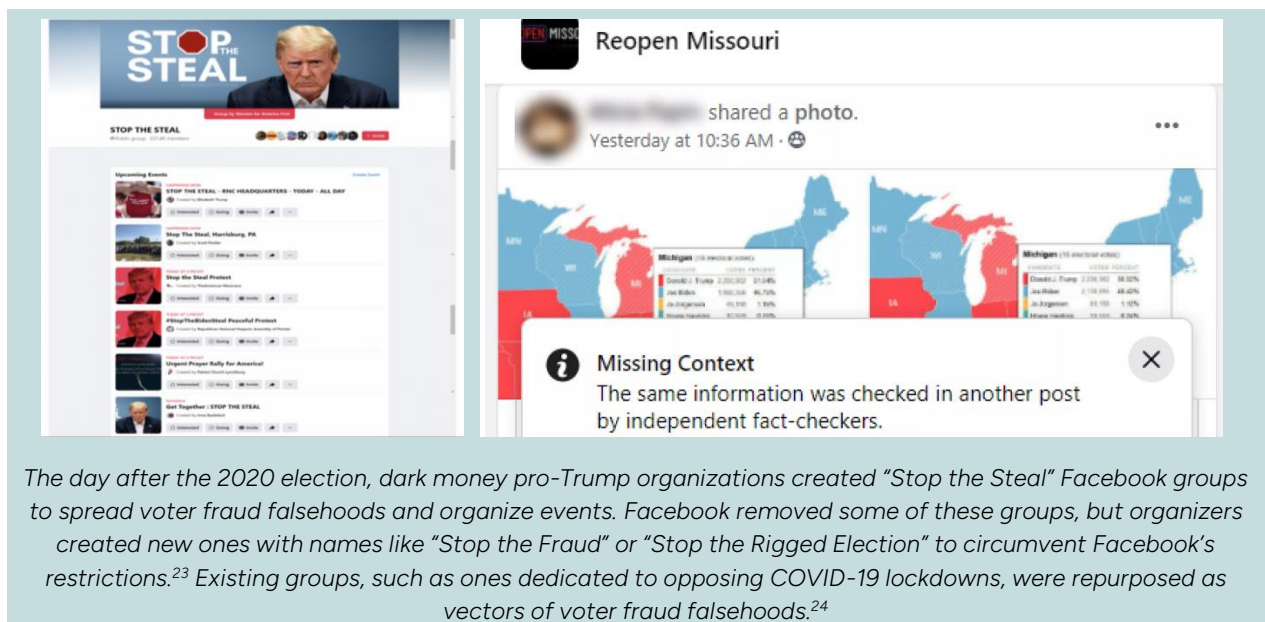
¹⁹ For example, “Trumptrains” were a way to mass-amplify messages and build up followers for pro-Trump Twitter users. Users posted lists of Twitter handlers, emojis and usually a meme or GIF, and the “train cars” operated as follow-for-follow networks. The result was explosive follower growth for everyone involved. Erin Gallagher, “[Trump Trains](#),” Medium, September 15, 2019; Karen Kornbluh and Ellen P. Goodman, “[Safeguarding Digital Democracy: Digital Innovation and Democracy Initiative Roadmap](#),” German Marshall Fund of the United States, March 2020, 21.

²⁰ Sheera Frenkel, “[The Most Influential Spreader of Coronavirus Misinformation Online](#),” New York Times, July 24, 2021.

Community forums: Assemble like-minded users and facilitate mobilization using platform tools

Public and private community forums such as Facebook groups, Reddit threads, Twitter Communities, and WhatsApp groups are a key vector of recruitment for extremist movements. The most active political groups on Facebook have been rife with hate, bullying, harassment and misinformation, and they grew large quickly by leveraging platform tools.²¹

Facebook recommends groups to users based on their interests and helps owners find new members. Facebook tools allow users to build invitation lists from those who are members of similar groups or express related interests and to automatically invite them, which has helped networked information campaigns in the past. In 2021, Facebook stopped recommending health and political groups and slowed the growth of newly created groups.²²



The day after the 2020 election, dark money pro-Trump organizations created “Stop the Steal” Facebook groups to spread voter fraud falsehoods and organize events. Facebook removed some of these groups, but organizers created new ones with names like “Stop the Fraud” or “Stop the Rigged Election” to circumvent Facebook’s restrictions.²³ Existing groups, such as ones dedicated to opposing COVID-19 lockdowns, were repurposed as vectors of voter fraud falsehoods.²⁴

Another crucial medium for Stop the Steal organizing was the Reddit copycat website thedonald.win. The site got its start as a subreddit, r/The_Donald, where it amassed 790,000 subscribers before Reddit banned it in mid-2020. Much of this original audience migrated to thedonald.win, where users organized violent January 6th protests.²⁵

Targeted ads: Using data about users to persuade them to become part of the pipeline

²¹ Shannon Bond and Bobby Allyn, “[How the ‘Stop the Steal’ movement outwitted Facebook ahead of the Jan. 6 insurrection](#),” NPR, October 22, 2021; Karen Kornbluh, “[Written Testimony for Hearing on ‘Social Media Platforms and the Amplification of Domestic Extremism & Other Harmful Content](#),” U.S. Senate Committee on Homeland Security and Governmental Affairs, October 28, 2021.

²² Tom Alison, “[Changes to Keep Facebook Groups Safe](#),” Meta Newsroom, March 17, 2021; Jeff Horwitz and Justin Scheck, “[Facebook Increasingly Suppresses Political Movements It Deems Dangerous](#),” Wall Street Journal, October 22, 2021.

²³ Brian Fung and Donie O’Sullivan, “[‘Stop the steal’ groups hide in plain sight on Facebook](#),” CNN, January 15, 2021.

²⁴ Ashley Gold, “[Facebook groups are turning into election disinformation vectors](#),” Axios, November 5, 2020.

²⁵ Select Committee to Investigate the January 6th Attack on the United States Capitol, “[Final Report](#),” December 22, 2022, 527. [Example post](#) from thedonald.win.

On social media platforms, ad buyers can target users based on demographic information, demonstrated interest in certain topics, or even with lookalike audiences.²⁶ This allows for microtargeting and siloed conversations.²⁷

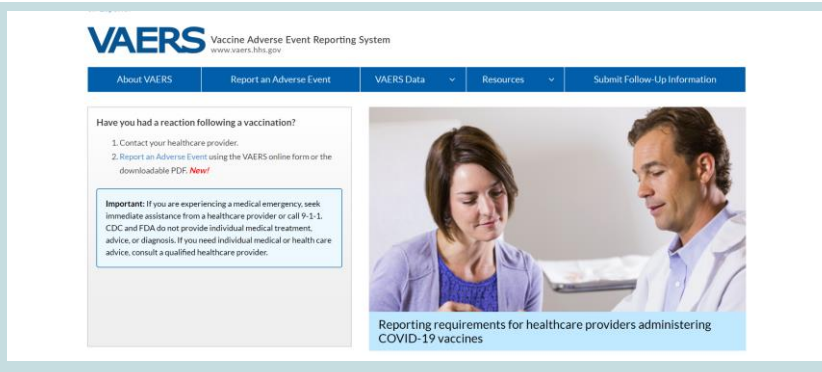
Off social media platforms, advertisers collect or purchase data using web tracking (cookies), location tracking, and behavioral data (clicks, impressions, or engagement on content).

Bringing the narrative to life

Coordinated deceptive campaigns push a specific narrative, for example, that vaccines are dangerous.²⁸ To do so, they deploy real-world stories or controversies, cherry-pick headlines or grains of truth, elevate anecdotes without context as “evidence,” and cater their message to narrow identity groups. Campaigns take advantage of what people are already inclined to believe and test out various messages to see what sticks. The following tactics are common ways coordinated deceptive campaigns bring a narrative to life:

Kernel of truth: A cherry-picked detail from a reputable source—such as an article from a responsible journalism outlet, a court filing, a personal anecdote, or a leak—presented with insufficient or misleading context

Kernels of truth can lack important context and obscure the big picture. These kinds of details sometimes originate from individuals and gain traction through bottom-up grassroots amplification. Other times, a central figure strips from its context to be misleading. In either event, coordinated deceptive campaigns prime their audiences with a broad narrative which a kernel of truth, including the audience’s lived experience, can then support.²⁹



Anti-vaccine activists took data from a government database on the possible side effects of vaccines to spread falsehoods about the COVID-19 vaccine.³⁰

²⁶ Some social media companies, such as Meta, have banned the targeting of users by categories such as health conditions, race, political causes, specific sexual orientations, and religion. However, advertisers have been able to continue use proxies such as “Gospel music,” “Hispanic culture,” and “Anime movies,” as proxies. This demographic targeting can facilitate racial discrimination in employment, housing, and credit card opportunities. Jon Keegan, [“Facebook Got Rid of Racial Ad Categories. Or Did It?”](#), The Markup, July 9, 2021.

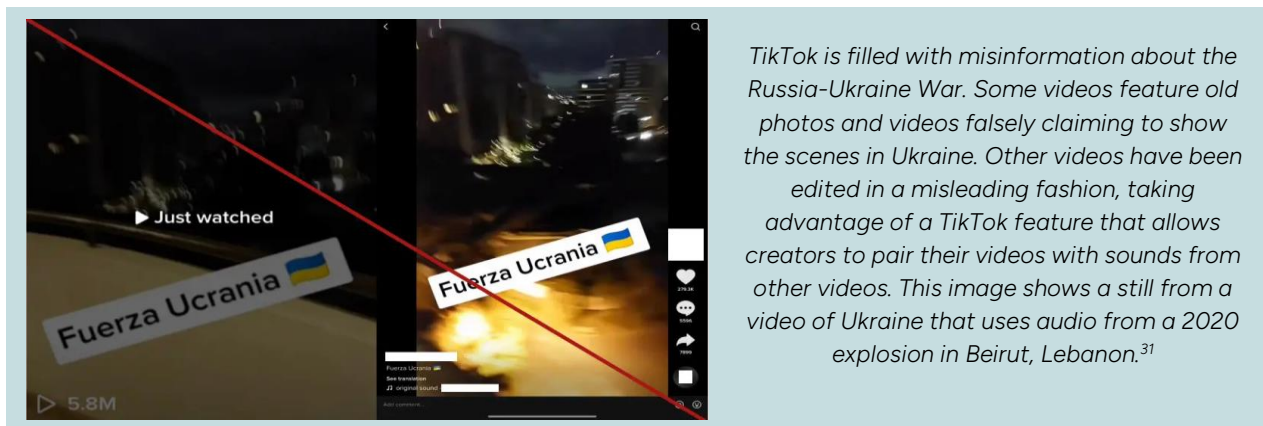
²⁷ Kornbluh and Goodman, [“Safeguarding Digital Democracy,”](#) 16.

²⁸ Ronnie Das and Wasim Ahmed, [“Rethinking Fake News: Disinformation and Ideology during the time of COVID-19 Global Pandemic,”](#) IIM Kozhikode Society & Management Review, 11(1), 2022, 154.

²⁹ Michael Grass, [“What is participatory disinformation,”](#) Center for an Informed Public, University of Washington, May 26, 2021.

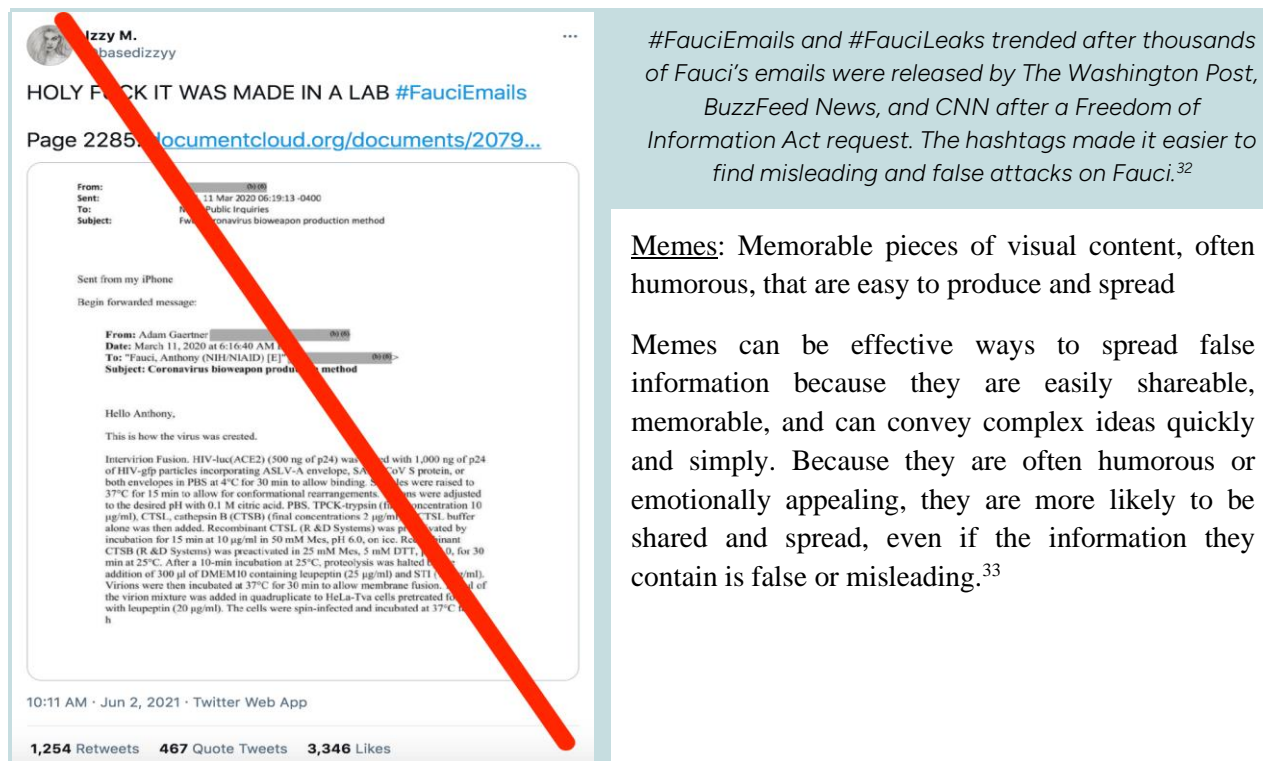
³⁰ Meredith Wadman, [“Antivaccine activists use a government database on side effects to scare the public,”](#) Science, May 26, 2021.

Manipulated and misattributed content: Images, audio, and other content that is deceptively altered or shared in a misleading manner



Hashtags: Viral slogans that place individual posts within a broader context and connect them to other posts about the same topic

Hashtags can be instrumental to gaming the algorithm by giving followers a common set of language, which can help elevate content in a platform's trending list.



Memes: Memorable pieces of visual content, often humorous, that are easy to produce and spread

Memes can be effective ways to spread false information because they are easily shareable, memorable, and can convey complex ideas quickly and simply. Because they are often humorous or emotionally appealing, they are more likely to be shared and spread, even if the information they contain is false or misleading.³³

³¹ Sheera Frenkel, "TikTok Is Gripped by the Violence and Misinformation of Ukraine War," New York Times, March 5, 2022.

³² Natasha Korecki and Sarah Oweremohle, "Attacks on Fauci grow more intense, personal and conspiratorial," Politico, June 4, 2021.

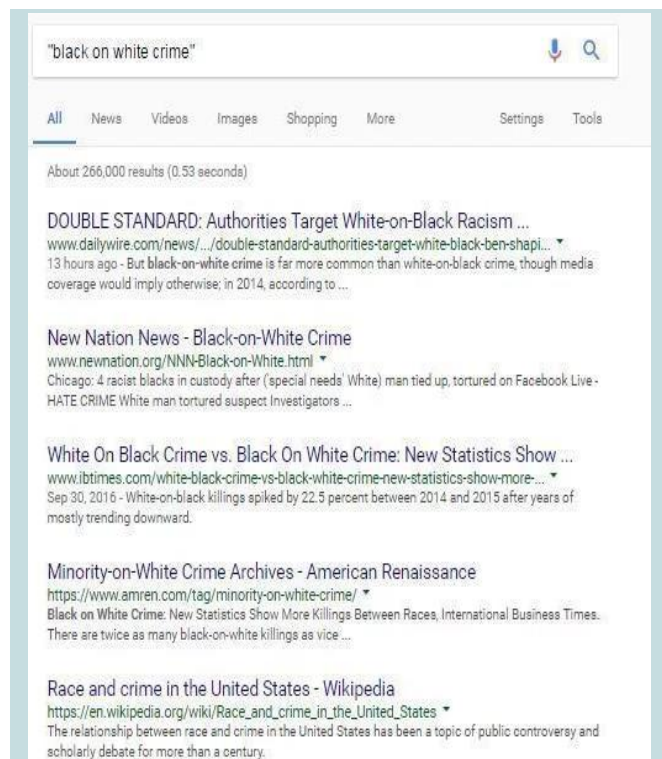
³³ Media Manipulation Casebook, "Definitions: Meme," Joan Donovan, Emily Dreyfuss, and Brian Friedberg, "How Memes Led to an Insurrection," The Atlantic, September 13, 2022; Marc Dupuis and Andrew Williams, "The Spread of Disinformation on the Web: An Examination of Memes on Social Networking," Institute of Electrical and Electronics Engineering, 2019.



Keyword stuffing and data voids: Search engine manipulation tactics that game search result rankings

Keyword stuffing: Adding popular keywords to unrelated websites to promote content in search engine rankings. This elevated ranking can create the illusion that a site reflects the general consensus and is supported by the scientific community or independent journalism.³⁵

Exploiting data voids: Data voids, a concept coined by Michael Golebiewski and danah boyd, describe obscure terms which, when entered into a search engine, return deeply problematic, few, or no results.³⁶ Data voids can lead searchers to sites filled with false information because those sites rank highly in search results in the absence of high-quality, trusted sites using the search terms. Users can stumble upon data voids or be directed to them by malicious actors who know there will be no counter-content.³⁷



Dylann Roof Googled "black on white crime" in 2015 and it introduced him to white supremacist extremist content. There was a data void around the search term because no one outside of white supremacist communities used the phrase. White supremacists also keyword stuffed the phrase on Wikipedia pages so they would appear with

³⁴ Image source: Jack Goodman and Flora Carmichael, "Covid-19: What's the farm of 'funny' anti-vaccine memes?," BBC, November 29, 2020; Sara Fischer and Alison Snyder, "How memes became a major vehicle for misinformation," Axios, February 23, 2021.

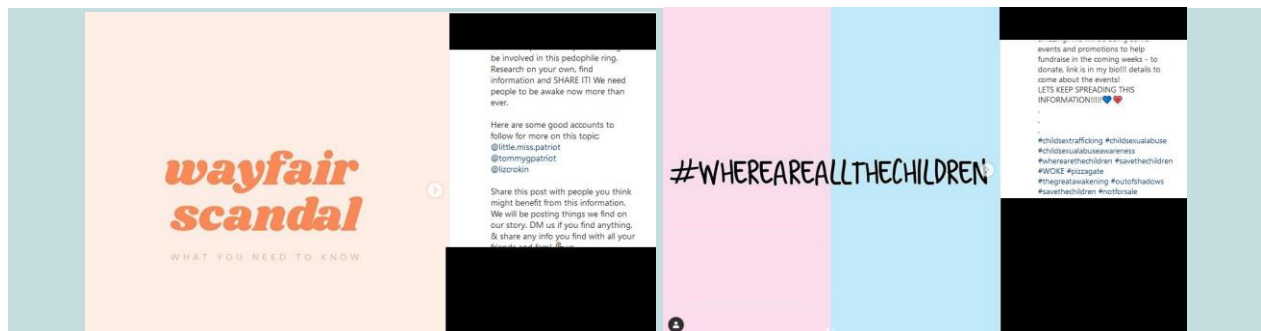
³⁵ Panagiotis Takis Metaxas, "Web Spam, Social Propaganda and the Evolution of Search Engine Rankings," Web Information Systems and Technologies, 2010, as quoted by Tucker et al., "Social Media."

³⁶ Michael Golebiewski and danah boyd, "Data Voids: Where Missing Data Can Easily Be Exploited," Data & Society, May 2018, 5.

³⁷ Michael Golebiewski and danah boyd, "Data Voids: Where Missing Data Can Easily Be Exploited," Data & Society, May 2018, 4.

searches for “black on white crime.” This image shows the search results for “black on white crime” in January 2017, where two of the top five results were white supremacist websites.³⁸

Identity appeals: Campaigns morph their messages to appeal to specific identities



On Instagram, lifestyle influencers advanced a softened version of the QAnon conspiracy theory using appealing aesthetics and messaging focused on children. “Pastel QAnon,” a term coined by Marc-André Argentino to describe the trend, was effective at reaching women by using the visual cues of a mainstream, feminine aesthetic found across social media.³⁹

Message testing: Optimize the effectiveness of content and susceptibility of audiences

Campaigns publish multiple versions of the same content, allowing them to determine what spreads fastest online. This tactic, used by coordinated deceptive campaigns and mainstream outlets alike, allows actors to test and refine the effectiveness of messaging.⁴⁰



Ben Shapiro posts the same article with slightly different framing in the caption.

Flooding the zone

Once brought to life, these narratives get distributed through the pipeline to reach large audiences. Social media virality can have a compounding effect; on many platforms, algorithms boost content they recognize as popular, adding it to the timelines of users who may not follow any pages or belong to any groups within the pipeline. Each of the following tactics amplify a narrative so it will flood users’ online experience:

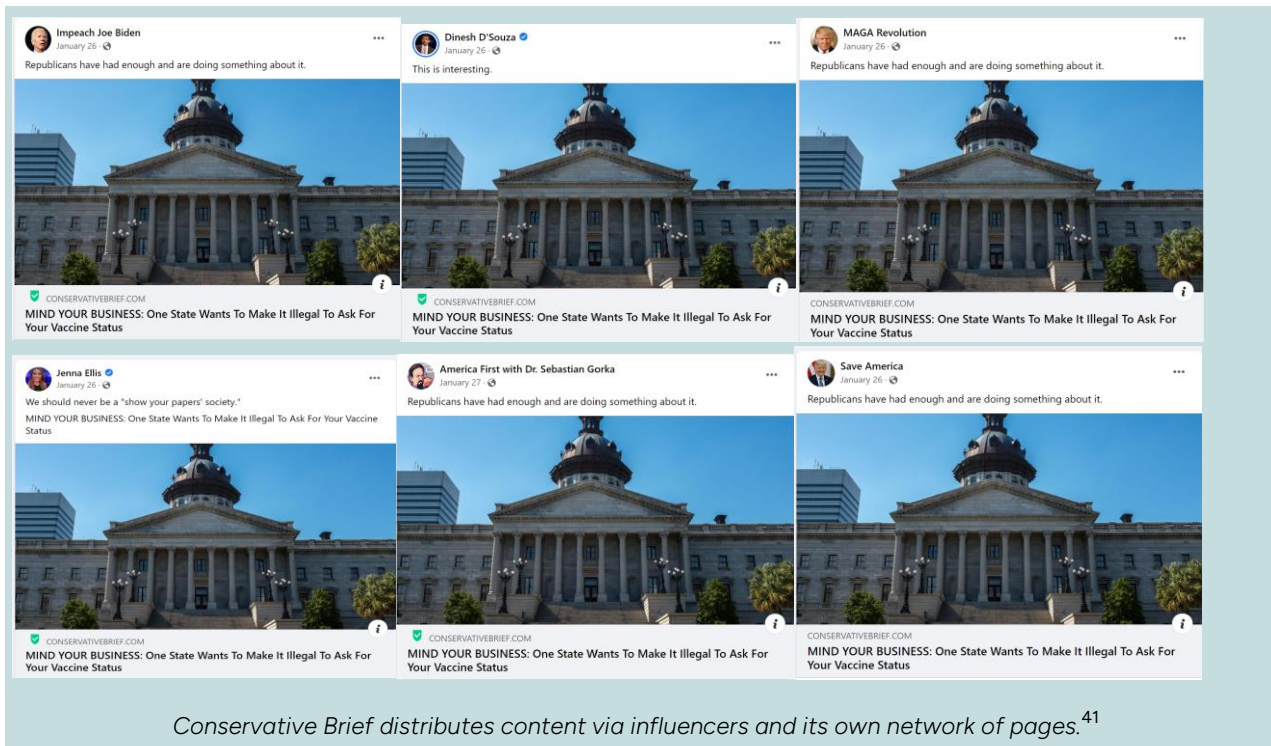
³⁸ Rebecca Hersher, “[What Happened When Dylann Roof Asked Google For Information About Race?](#),” NPR, January 10, 2017.

³⁹ Kaitlyn Tiffany, “[The Women Making Conspiracy Theories Beautiful](#),” The Atlantic, August 18, 2020; Marc-Andre Argentino, @ [MAArgentino](#) Twitter thread, September 1, 2020.

⁴⁰ Sheera Frenkel, “[The Most Influential Spreader of Coronavirus Misinformation Online](#),” New York Times, July 24, 2021.

Cross-posting: Posting content across many accounts and platforms

Coordinated deceptive campaigns often post content across multiple pages, accounts, and social media platforms to game the algorithm and create the appearance of broad, grassroots support.

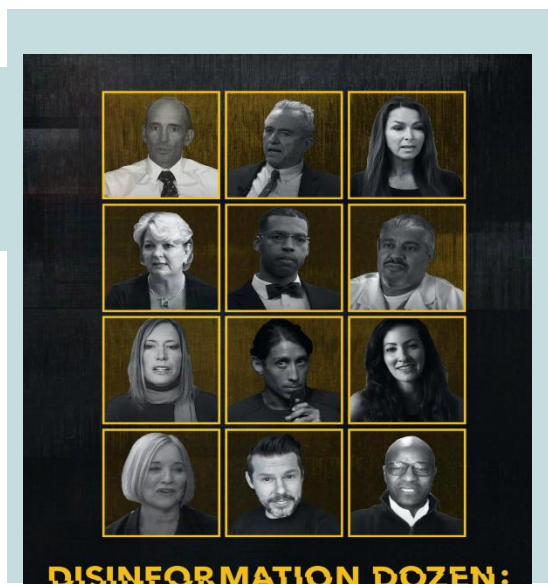


Influencers also cross-post each other's content on topics of mutual interest.

The "Disinformation Dozen," twelve anti-vaccine influencers, were responsible for 65% of the shares of anti-vaccine misinformation on social media.⁴²

Activating bots: Coordinating automated accounts to deceive users or manipulate algorithms

Bots are often used to artificially amplify a message, game a trending algorithm, or boost engagement metrics.⁴³



⁴¹ Judd Legum, "[How an obscure far-right website with 3 employees dominates Facebook in 2022](#)," Popular Information, February 23, 2022.

⁴² Center for Countering Digital Hate, "[The Disinformation Dozen: The Sequel](#)," April 28, 2021.

⁴³ Media Manipulation Casebook, "[Definitions: Bot](#)."

A study of June 2017 Twitter bot activity found they produced approximately one-quarter of all original tweets referencing climate change on a typical day. The graph shows that, in the weeks before and after the United States announced its withdrawal from the Paris climate agreement, 40% of the most active accounts posting about climate change were bots, although they comprised only 15% of the most influential accounts.⁴⁴



Hashtag adoption: Shape the narrative and coordinate efforts

Online activists use a common set of hashtags or keywords. When these hashtags become popular enough, many social media algorithms will promote the topic in their targeted audience’s feeds or display it on a list of trending topics. Narratives with effective hashtags allow supporters to help the story trend and inflate its perceived popularity.

Evading automated content moderation: Replacing specific keywords or phrases that automated content moderation tools are likely to flag

Many social media sites rely on automated content moderation tools to flag the use of specific words or phrases as a first step in removing or reducing the spread of false content. To evade these tools, influencers use “Algospeak”—code words, turns of phrase or emojis—to prevent their posts from being removed or downranked.⁴⁵

⁴⁴ Thomas Marlow, Sean Miller, and J. Timmons Roberts, “[Bots and online climate discourses: Twitter discourse on President Trump’s announcement of U.S. withdrawal from the Paris Agreement](#),” *Climate Policy*, January 15, 2021.

⁴⁵ Taylor Lorenz, “[Internet ‘algospeak’ is changing our language in real time, from ‘nip nops’ to ‘le dollar bean’](#),” *Washington Post*, April 8, 2022; Zoe Kleinman, “[Anti-vax groups use carrot emojis to hide Facebook posts](#),” *BBC*, September 16, 2022.

⁴⁶ Patricio Durán and Tomás Lawrence, “[Constitutional Vote in Chile Targeted by Coordinated Hashtag Campaign Before Election](#),” *Media Manipulation Casebook*, December 16, 2021.

Targeting audiences: Information operations match identity appeals to specific subgroups

Race plays a substantial role in the targeting of coordinated deceptive content. For example, an analysis of 5.2 million tweets from the Russian-funded Internet Research Agency troll farm found that presenting as a Black activist was the most effective predictor of disinformation engagement.⁴⁷ Targeting by race, ethnicity, or national origin preys on what Mark Kumleben, Samuel Woolley, and Katie Joseff have termed “structural disinformation,” or “systemic issues related to the broader information environment, born out of long-term efforts to control minority groups’ access to and understanding of country’s electoral and media systems.”⁴⁸



*Children's Health Defense, an anti-vaccine group, produced Medical Racism: The New Apartheid, a documentary which alleges that COVID-19 vaccines are a cover to conduct medical experiments on the Black and Latino communities.*⁵⁰

Code word	What it means
Swimmers	People who have been vaccinated
Dance Party/ Dinner Party	Group of people who oppose covid vaccines
Drank beer	Took the vaccine
Pizza or Pizza King	Pfizer
Moana	Moderna
Cross country trip where we spent the night with two dancers	Spent time with two people who had been vaccinated
Glitter	A false notion that taking the covid vaccine makes the person “shed” the vaccine to the unvaccinated.

*Examples of “Algospeak” code words used by anti-vaccine activists and their meanings.*⁴⁹

Targeting through paid microtargeted ads: Paid targeting of users based on their interests

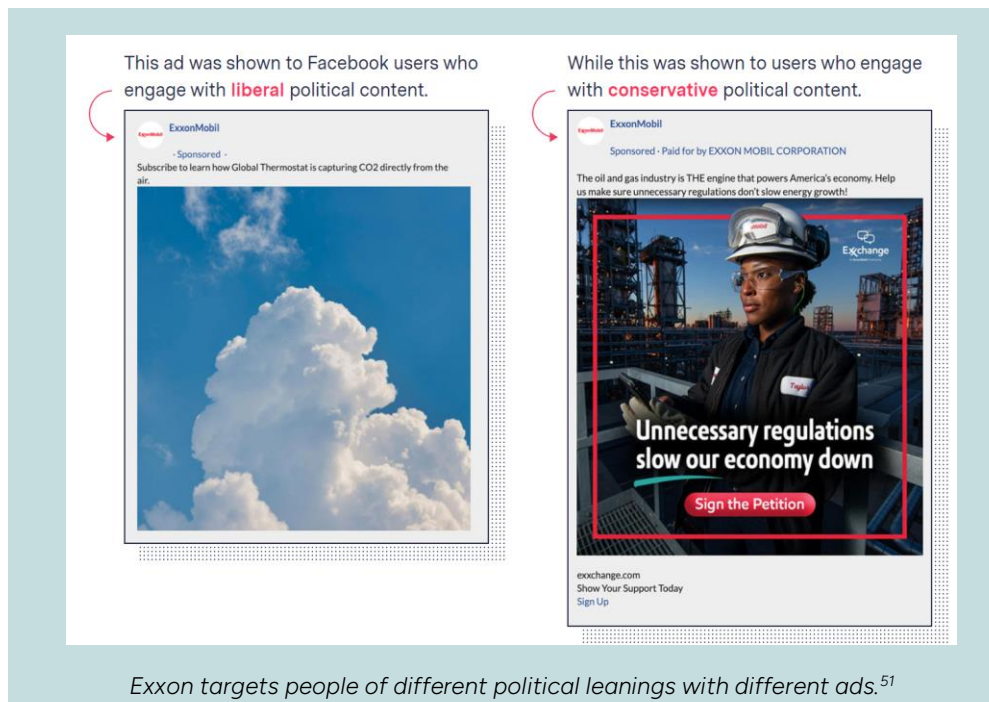
Targeting tools allow advertisers to send different ads to people based on their personalized profiles.

⁴⁷ Deen Freelon, Michael Bossetta, Chris Wells, Josephine Lukito, Yiping Xia, and Kirsten Adams, “[Black trolls matter: Racial and ideological asymmetries in social media disinformation](#),” Social Science Computer Review 40, no. 3, 2022, 560-578.

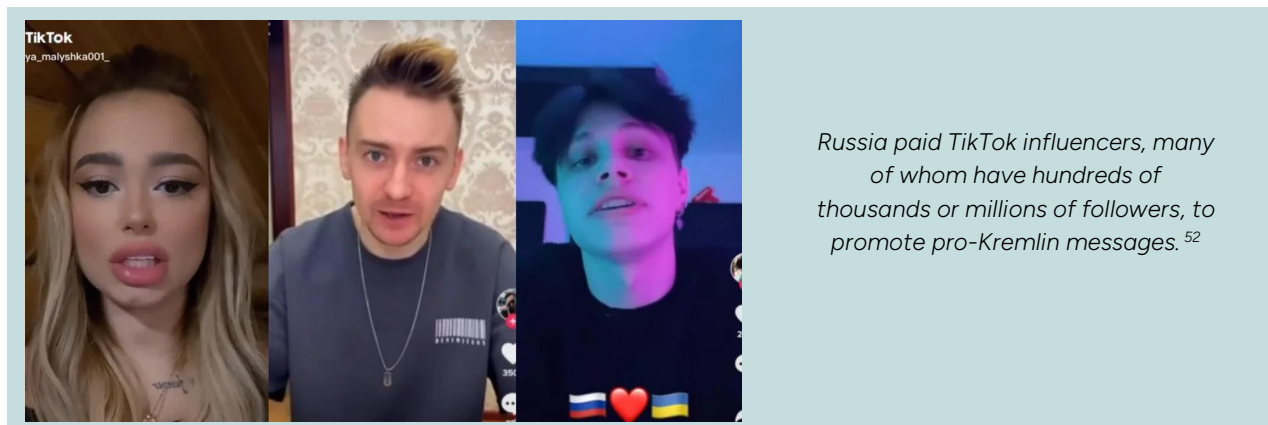
⁴⁸ Mark Kumleben, Samuel Woolley, and Katie Joseff, “[Electoral Confusion: Contending with Structural Disinformation in Communities of Color](#),” June 2022, 2.

⁴⁹ Ben Collins and Brandy Zadrozny, “Anti-vaccine groups changing into ‘dance parties’ on Facebook to avoid detection,” NBC News, July 21, 2021; Image source: Al Tompkins, “[How anti-vaxxers avoid being detected by Facebook](#),” Poynter, July 26, 2021.

⁵⁰ Will Stone, “[An Anti-Vaccine Film Targeted to Black Americans Spreads False Information](#),” NPR, June 8, 2021.



Targeting through paid influencer promotion: Leverage influencers' trust among their audiences



Activation

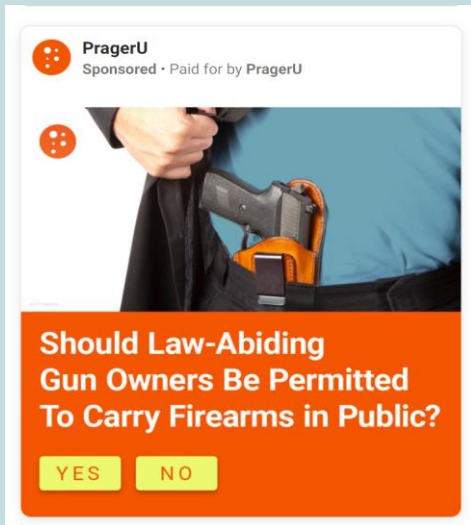
Campaigns give audiences actions to take. These actions are ends in themselves and also strengthen audience enthusiasm and loyalty and bolster pipelines for future narrative campaigns. Typical forms of activation include:

Call to action – subscribe and engage: Give audience ways to join the distribution pipeline

Users follow and subscribe or provide data (via data trackers and by filling in their information). In doing so, they become part of the distribution pipeline for future narrative campaigns.

⁵¹ Jeremy B. Merrill, "[How Facebook's Ad System Lets Companies Talk Out of Both Sides of Their Mouths](#)," The Markup, April 13, 2021.

⁵² David Gilbert, "[Russian TikTok Influencers Are Being Paid to Spread Kremlin Propaganda](#)," Vice, March 11, 2022.



Sites such as PragerU and the Daily Wire create ads with polls, which often ask respondents to enter an email address in order to submit their response. These polling sites contain data-gathering trackers that allow the sites to build new lookalike audiences.⁵³

Call to action – build community: Invite others to follow, join groups, or subscribe

Campaigns offer supporters ways to stay involved and facilitate future organizing or activism.



Stop the steal groups ballooned thanks to a small number of super inviters who used platform tools to amplify the groups. Just 0.3% of group members created 30% of invitations.⁵⁴ They created invitation links based on Facebook's suggestions or pulled from lists of other groups members that could be shared on- or off-platform.

Mobilize: Organize digital grassroots troops

Coordinated deceptive campaigns use social media including closed networks such as Facebook groups, messaging services, or alternative platforms to mobilize supporters to engage in on- and offline activism. Offline, campaigns may encourage their audiences to attend a protest or event or vote. Offline mobilization can, in turn, feed into online networked information campaigns, when photos and videos of in-person events are posted online.

Supporters of the 2022 anti-vaccine Canadian trucker convoy used Facebook groups and Telegram channels to organize and provide funding for in-person protests.⁵⁵

⁵³ Corin Faife, “[How The Daily Wire Uses Facebook’s Targeted Advertising to Build Its Brand](#),” The Markup. August 10, 2021.

⁵⁴ Ryan Mac, Craig Silverman, and Jane Lytvynenko, “[Facebook Stopped Employees From Reading An Internal Report About Its Role In The Insurrection. You Can Read It Here](#),” BuzzFeed News, April 26, 2021.

⁵⁵ Mark Scott, “[Ottawa truckers’ convoy galvanizes far-right worldwide](#),” Politico, February 6, 2022; Ryan Broderick, “[How Facebook twisted Canada’s trucker convoy into an international movement](#),” The Verge, February 19, 2022.

ACTORS AND GOALS

Who are the actors behind coordinated deceptive campaigns?

- Foreign adversaries: State actors, including China, Iran, and Russia, set up fake social media accounts and newsrooms and used existing state media sites to spread false information. However, domestic coordinated deceptive campaigns now often dwarf foreign-operated ones.⁵⁶
- Scammers: Cybercriminals use falsehoods as bait in scams and phishing schemes.⁵⁷
- Profiteers: Deceptive tactics are used to sell products, subscriptions, and tickets to events.⁵⁸
- Political candidates and campaigns: Candidates and political elites promote false claims that support their policy positions and engage their supporters.⁵⁹
- Activists: People with sincere beliefs in conspiracies and false narratives work to promote those campaigns.⁶⁰
- Industry: Companies spread false and misleading information where doing so supports their business. For example, the oil industry has spread disinformation about climate change.⁶¹

Who are the targeted audiences?

Campaigns often tap into their audience's beliefs, cultures, anxieties, and identities. For example, writing about "black on white crime" preys on white fear.

Sometimes, targeted audiences are narrow and selective, such as specific elites or racial and identity groups. Other times, actors target their messages to a broad swath of the population.

- General public: Some campaigns target as broad an audience as possible in the hopes that pushback—such as fact-check replies—will boost the original content.⁶²
- Social identity groups: Some campaigns target social identities, such as race, ethnicity, religion, gender, or sexual orientation. These identities are used to unite members of the identity group and divide them from others.⁶³

⁵⁶ Less than 1% of reports of election misinformation submitted to the Election Integrity Partnership, a non-partisan coalition of researchers tracking efforts to delegitimize the 2020 election, related to foreign interference. Center for an Informed Public, Digital Forensic Research Lab, Graphika, & Stanford Internet Observatory, "[The Long Fuse: Misinformation and the 2020 Election](#)," Stanford Digital Repository: Election Integrity Partnership. v1.3.0, 2021.

⁵⁷ Colleen Tressler, "[Coronavirus: Scammers follow the headlines](#)," Federal Trade Commission Consumer Advice, February 10, 2020.

⁵⁸ Cat Zakrzewski, "[The Technology 202: 'Pandemic profiteers' are using deceptive tactics to peddle products and subscriptions, according to a memo submitted to the FTC](#)," Washington Post, July 16, 2021.

⁵⁹ Adrienne Goldstein and Eli Weiner, "How the Disinformation Supply Chain Created a Deceptive Narrative about the Texas Blackout," GMF, February 19, 2021; Avaaz, "[Facebook's Climate of Deception: How Viral Misinformation Fuels the Climate Emergency](#)," May 11, 2021.

⁶⁰ Max Rizzuto and Jared Holt, "[DC anti-mandate rally leveraged to broaden audience for anti-vax narratives](#)," DFRLab, February 2, 2022.

⁶¹ Emily Atkin, "[How Exxon duped 'The Daily'](#)," HEATED, November 17, 2021.

⁶² Keach Hagey and Jeff Horwitz, "[Facebook Tried to Make Its Platform a Healthier Place. It Got Angrier Instead](#)," Wall Street Journal, September 15, 2021.

⁶³ Kumbleben, Woolley, and Joseff, "[Electoral Confusion](#)," Protect Democracy, 25-26.

-
- Conspiratorial thinkers: New narratives may build on existing conspiracy beliefs (like QAnon or medical pseudoscience) to gain audiences.⁶⁴
 - Elites: Media manipulators target political elites, institutions, and influencers to reach larger audiences in a practice called “trading up the chain.”⁶⁵
 - Political subgroups: Scammers and profiteers rely on political passions to promote their content.⁶⁶
 - Activists: Like political beliefs, activism in one area (like anti-lockdown participation) can be channeled by other campaigns (like anti-vaccination falsehoods).
 - Specific regions: Election coordinated deceptive campaigns in particular may focus on specific states or cities to suppress the vote or spread falsehoods about candidates.⁶⁷

Coordinated deceptive campaigns often target more than one category at once, such as conservative elites, older Black voters, or anti-vaccine mothers in California. The targeting often maps onto institutional, intersectional identities.

Bringing all this together, understanding a networked information campaign requires answering:

- What are the goals?
- Who are the targeted audiences?
- What tactics are employed?

⁶⁴ The Virality Project, “[Memes, Magnets, and Microchips: Narrative dynamics around COVID-19 vaccines](#),” Stanford Digital Repository, 2022, 82.

⁶⁵ Alice Marwick and Rebecca Lewis, “[Media Manipulation and Disinformation Online](#),” Data & Society Research Institute, 2017, 38, Kate Starbird, [@katestarbird Twitter thread](#), May 6, 2021.

⁶⁶ Craig Silverman and Lawrence Alexander, “[How Teens In The Balkans Are Duping Trump Supporters With Fake News](#),” BuzzFeed News, November 3, 2016.

⁶⁷ Kevin Collier, “[Disinformation via text message is a problem with few answers](#),” NBC News, September 13, 2022.

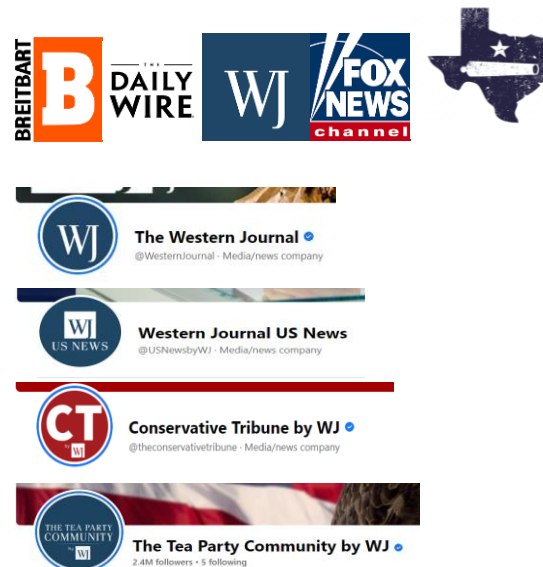
THE CASE STUDY OF TEXAS WIND TURBINES

Activists, industry, and political candidates worked together during the February 2021 Texas winter storms to target conservatives, anti-renewable energy activists, and Texans with the false narrative that frozen wind turbines were responsible for widespread power outages. Some wind turbines indeed went offline due to the storm, but two-thirds of the state's shortfall in power generation came from failures at gas and coal power plants. Despite this, the false claim about wind turbines soon went viral on social media, accruing millions of views and interactions, and top Texas officials and members of Congress parroted the claim. The false narrative deflected blame from the systemic causes of the power outages—an independent electric grid that made it difficult to import electricity, a failure to winterize power sources, failure to address the causes and effects of climate change—and onto renewable energies.⁶⁸

Building the pipeline

Outlets: A collection of outlets—including Breitbart, Daily Wire, Texas Scorecard, Western Journal, and Fox News—built large audiences across social media while eschewing journalistic standards between 2016 and late 2020.⁶⁹

Public accounts: The outlets operated their own Facebook pages and pro-fossil fuel pages such as “Friends of Coal” amplified the narrative. The Western Journal cross posts its content across over a dozen Facebook pages.



Video channels and influencers: Influencers across platforms with millions of combined followers and subscribers shared content about the outages and created new videos and posts.



⁶⁸ Mandi Cai, Erin Douglas and Mitchell Ferman, “How Texas’ power grid failed in 2021 – and who’s responsible for preventing a repeat,” Texas Tribune, February 15, 2022.

⁶⁹ Karen Kornbluh, Eli Weiner and Adrienne Goldstein “New Study by Digital New Deal Finds Engagement with Deceptive Outlets Higher on Facebook Today Than Run-up to 2016 Election,” German Marshall Fund of the United States, October 12, 2020.

Community forums: Political and issue-specific interest groups created Facebook share groups to articles and memes.



Environmentalists Against Wind Turbines

Public group · 2.4K members

Join group

Targeted ads: Deceptive outlets had been building up their audiences through ads in the months preceding the power outages.

The Western Journal
Sponsored · Paid for by The Western Journal

The presidential election is quickly approaching. Who do you think will be president for the next four years?

Share your thoughts by completing this short questionnaire.

THE WESTERN JOURNAL

WRJTBALJOURNAL.COM
Presidential Election 2020
Let Your Voice Be Heard!

Learn more

Bringing the narrative to life

Misattributed media: In the early stage of the crisis, social media posts featured an image of a helicopter spraying hot water on wind turbines in Sweden in 2014 to falsely blame green energy for the energy shortfalls in Texas.

Oilfield Rando
@Oilfield_Rando

Welp time to go add some more externalities to the green energy

11:28 AM · Feb 13, 2021 · Twitter for iPhone

511 Retweets 108 Quote Tweets 1,906 Likes

Luke Legate
@lukelegate

A helicopter running on fossil fuel spraying a chemical made from fossil fuels onto a wind turbine made with fossils fuels during an ice storm is awesome.

7:05 PM · 14 Feb 2021

30,741 Retweets 89,720 Likes

Meme: As it was reshared, the image of the wind turbine became a meme, making fun of renewable energies and those who advocate for them.

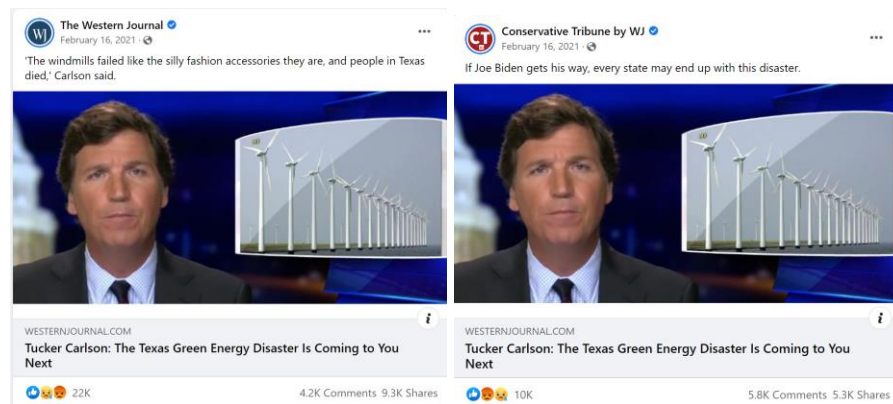
Kernel of truth: The Austin-American Statesman, a reputable local newspaper, published an article about the frozen wind turbines on February 14 and provided context on the scope of the outages. However, deceptive outlets selectively cited the American Statesman article to feed their false narrative that green energy was the main culprit for the outages.

Austin American Statesman
SERVING OUR COMMUNITY SINCE 1871

NEWS

Frozen wind turbines hamper Texas power output, state's electric grid operator says

Message testing: The Western Journal tested different captions for the same article across its Facebook pages.

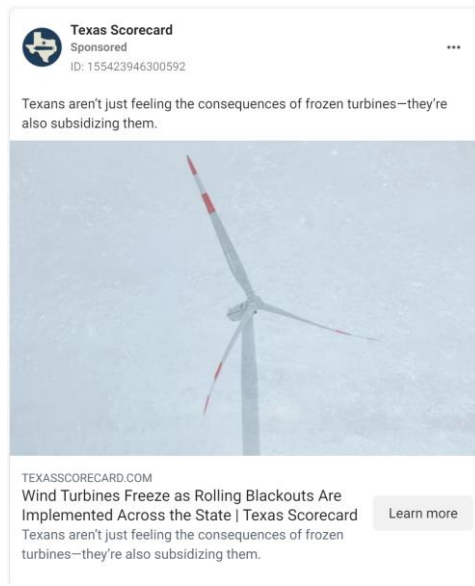


Flooding the zone

Cross-posting: The official Daily Wire Facebook page and three of the site's top influencers shared the article, often pointing their readers to the false tweet of a helicopter spraying wind turbines.

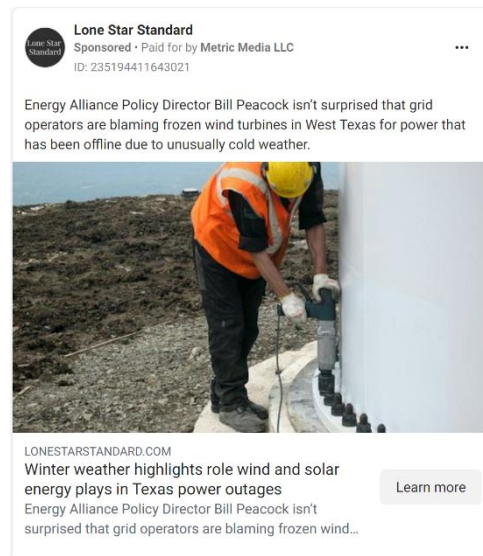


Targeting through paid microtargeted ads: Texas Scorecard ran a Facebook advertising campaign using frozen wind turbines as a hook.



Activation

Call to action – subscribe and engage: The Lone Star Standard, one of 1,300+ sites run by political operatives but designed to look like local, independent journalism, used the power outages to target Facebook users in Texas and grow its following. This larger audience could be targeted in future narrative campaigns.⁷⁰



⁷⁰ Lone Star Standard, [Meta Ad Library](#), March 18, 2021 – March 31, 2021.

WHY DOES THIS WORK?

Psychological and sociological drivers of mis- and disinformation

Successful coordinated deceptive campaigns take advantage of humans' psychological and sociological vulnerabilities. These biases and predispositions make us more likely to form or accept false views and create barriers to knowledge revision, even after the false view has been corrected.

We rely on two distinct systems for processing information. System 1 is intuitive and quick, which means that it relies on mental shortcuts, while System 2 is deliberative and analytical.⁷¹

System 1 shortcuts can be important and useful. Academic literature suggests that humans do not have the capacity to process everything and so we must rely on mental shortcuts (accuracy-effort trade-off theory) and that mental shortcuts can be particularly helpful for decision-making in specific situations with high uncertainty and redundancy (ecological rationality theory).⁷² Coordinated deceptive campaigns take advantage of System 1 shortcuts, meaning that anyone can be vulnerable to a well-targeted message.

Some examples of System 1 shortcuts include:⁷³

Illusory-truth effect: The illusory-truth effect describes the tendency to perceive repeated information as more truthful than new information.⁷⁴ Repetition can increase people's perceptions of the truthfulness of false statements, even when they know that such statements are false.⁷⁵ Even trained Facebook content moderators embraced fringe views after repeated exposure to the videos and memes they were supposed to moderate.⁷⁶

Barriers to belief revision: False information continues to influence people's thinking even after they receive a correction and accept the correction as true. This also persists for those who can recall the correction.⁷⁷

⁷¹ Daniel Kahneman, *Thinking, Fast and Slow*, Farrar, Strauss, and Giroux, 2011.

⁷² Anastasia Kozyreva and Ralph Hertwig, "[The interpretation of uncertainty in ecological rationality](#)," *Synthese* 198, 1517-1547, 2021.

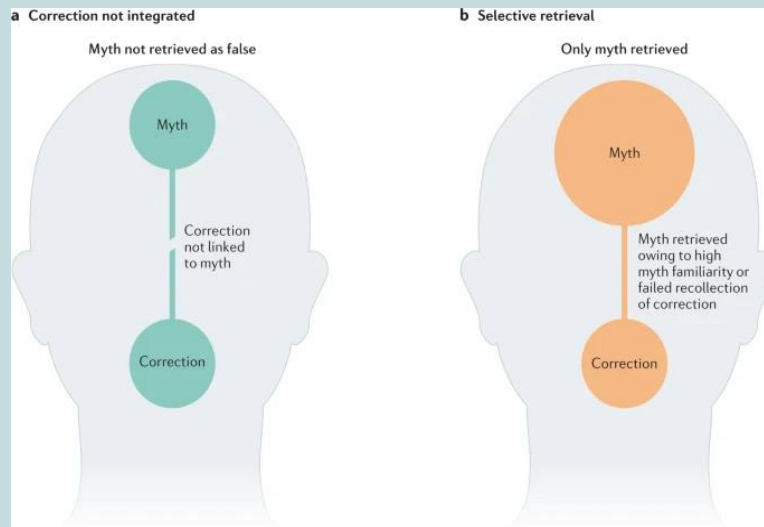
⁷³ Ullrich K. H. Ecker, Stephan Lewandowsky, John Cook, Philipp Schmid, Lisa K. Fazio, Nadia Brashier, Panayiota Kendeou, Emily K. Vraga, and Michelle A. Amazeen, "[The psychological drivers of misinformation belief and its resistance to correction](#)," *Nature Reviews Psychology*, 1, 13-29, January 12, 2022.

⁷⁴ Aumyo Hassan and Sarah J. Barber, "[The effects of repetition frequency on the illusory truth effect](#)," *Cognitive Research: Principles and Implications*, 6 no. 38, May 31, 2021

⁷⁵ Hassan and Barber, "[The effects of repetition frequency on the illusory truth effect](#)"; Maria S. Zaragoza and Karen J. Mitchell, "[Repeated Exposure to Suggestion and the Creation of False Memories](#)," *Psychological Science*, 7, 294-300, September 1996.

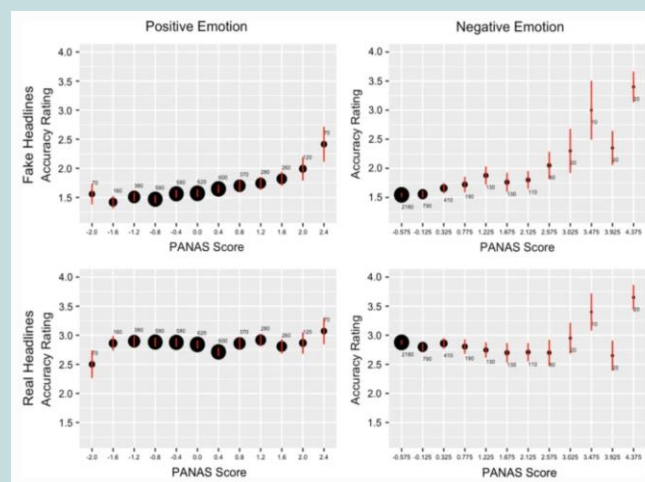
⁷⁶ Casey Newton, "[The Trauma Floor](#)," *The Verge*, February 25, 2019.

⁷⁷ Ecker et. al., "[The psychological drivers of misinformation belief and its resistance to correction](#)."



Left, an example of a barrier to belief revision. Misinformation continues to influence a person's thinking even after receiving a correction because the correction is not linked to the myth. Right, the impact of the illusory-truth effect. A myth dominates a person's thinking over a correction because the former is represented in the memory more strongly than the latter. This is either because of greater repetition of the myth (illusory-truth effect) or because the correction is not recalled.⁷⁸

Appeals to emotion: Greater emotionality, of both positive and negative emotions, predicts increased belief in fake news and decreased truth discernment.



There is a positive relationship between the perceived accuracy of a fake news headline and the emotionality of content. This is true for both positive and negative emotions.⁷⁹

System 2 explanations for why people believe mis- and disinformation and are likely to share it online focus on the social benefit of sharing the content. As researcher Alice Marwick notes, “people do not share fake news stories solely to spread factual information, nor because they are ‘duped’ by powerful partisan media. Their worldviews are shaped by their social positions and their deep beliefs, which are often both partisan

⁷⁸ Ecker et. al., “[The psychological drivers of misinformation belief and its resistance to correction.](#)”

⁷⁹ Cameron Martel, Gordon Pennycook and David G. Rand, “[Reliance on emotion promotes belief in fake news,](#)” Cognitive Research: Principles and Implications 5 no. 47, October 7, 2020.

and polarized. Problematic information is often simply one step further on a continuum with mainstream partisan news or even well-known politicians.”⁸⁰

Examples of System 2 explanations include:

Collective storytelling: Stories use connections in the human experience to make sense of the world.⁸¹ Storytelling can link facts and events together in distorted but appealing narratives.

Identity signaling: People sometimes share a story because it is a useful way to build or reinforce an identity, which takes precedence over the truth value of the story.

Collective identity: People share stories as a way of building collective identity among the group engaging with a specific narrative. Coordinated deceptive campaigns aim to reinforce cultural identity and create discord.⁸²

⁸⁰ Alice Marwick, “[Why do People Share Fake News? A Sociotechnical Model of Media Effects](#),” Georgetown Law Technology Review, 2018, 474.

⁸¹ Michael F. Dahlstrom, “[The narrative truth about scientific misinformation](#),” Proceedings of the National Academy of Sciences 118, 15, April 9, 2021.

⁸² Renee DiResta, Kris Shaffer, Becky Ruppel, David Sullivan, Robert Matney, Ryan Fox, Jonathan Albright, and Ben Johnson, “[The Tactics & Tropes of the Internet Research Agency](#),” U.S. Senate, October, 2019.

BUILDING CIVIC INFRASTRUCTURE

Many of the tools used to spread deceptive narratives can be repurposed in transparent, empowering ways to boost civic information and build more trust in fact-based information. Civic information providers must play an active role in using digital platforms to amplify content to targeted audiences.

Joining and leading civic information campaigns can look a great deal like participating in a deceptive narrative campaign. The same tactics used to build a pipeline, develop a narrative, flood the zone, and activate followers can be used to promote community-strengthening, trustworthy information.

- First, **build your pipeline** by engaging with other community leaders and groups. Seek out collaborations and help train other trusted voices to use these social media tools.
- As you get to know your audience, **develop narratives** and draft new content on topics where you have the expertise or expert partners to do so.
- **Flood the zone**, re-sharing and amplifying quality information from your own sources and other trusted accounts.
- **Activate** your audience with opportunities to be part of the distribution of information.

Building the pipeline

Take stock of existing online accounts and sites and create new ones where gaps exist.

- Build on- and offline relationships with journalists, government officials, and other sources of accurate information that you can amplify. Develop influence.
- Map out networks, communities, advocates, and methods for collaboration. What other partners could join in amplifying civic information?
- Create organic online groups, public accounts, video channels, community forums, outlets, and pages where they are helpful and needed.
- Cross-pollinate ideas and messages about key civic issues that are relevant across different groups.
- Connect with local influencers to have them share key messages, either for free or through a disclosed paid partnership.



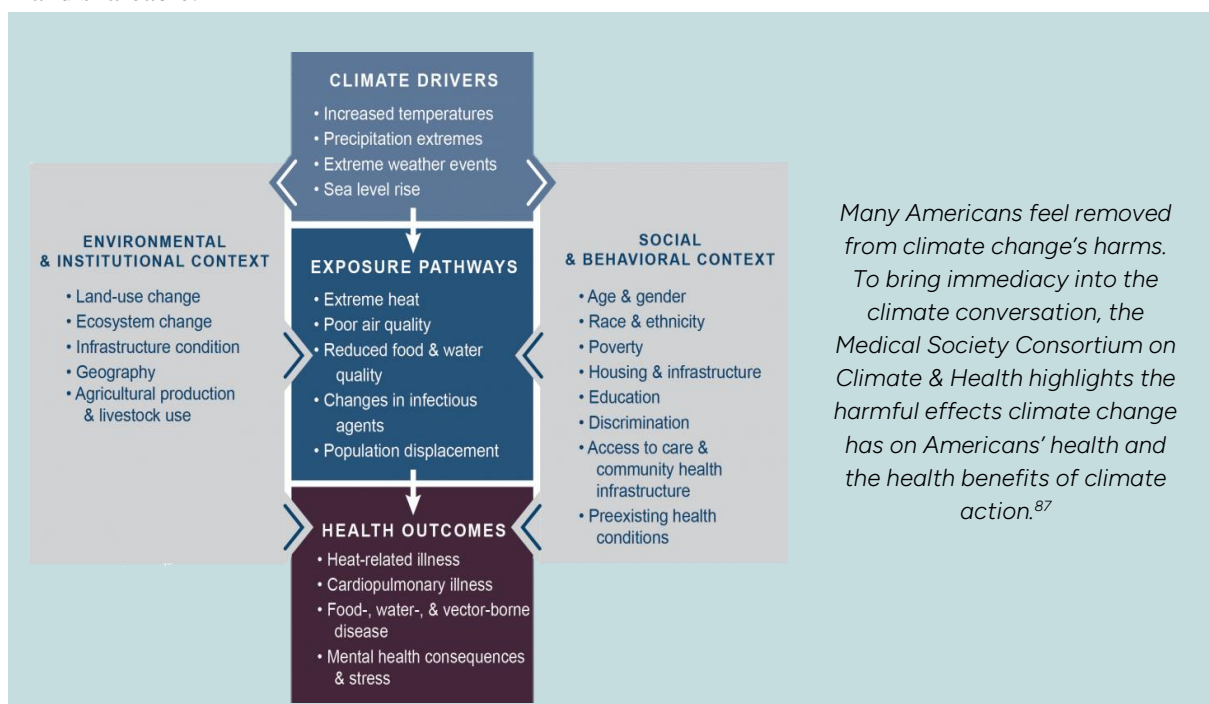
A George Mason University Center for Climate Change Communication survey revealed that TV weathercasters are trusted sources of information about climate change. Leveraging that trust, as well as weathercasters' broad access to the public and excellent science communication skills, the Center joined the Climate Matters partnership to distribute weekly climate change reporting materials to weathercasters.⁸³

Bringing the narrative to life

⁸³ George Mason University Center for Climate Change Communication, "[Climate Matters, Helping TV weathercasters and journalists report local climate change stories.](#)"

Translate civic information into compelling narratives using emotion, identity, and trusted messengers.

- Define your target audiences. Who already trusts you? What shared identities can you use to build communities and establish trust?
- Identify topics of interest, emerging trends, narratives that are starting to be seeded, and stories you want to share.
- Be cognizant of your biases. Understand your personal biases and how they may impact the messages and messengers you find trustworthy.
- Prime positive identities: where deceptive campaigns might draw on racialized fears, you can develop positive collective identities around a shared purpose. For example, you can prime a shared national or regional identity, race, ethnicity, gender, or a social or professional group.⁸⁴
- Share resources that pre-bunk the tactics used by coordinated deceptive campaigns.⁸⁵ Pre-bunking entails teaching audiences about the common characteristics seen across false narratives such as emotional language, scapegoating, or false comparisons between unrelated items, and showing them the critical thinking skills that will help them resist disinformation.⁸⁶
- Message testing: Try out different narratives that make factual information emotionally compelling and shareable.

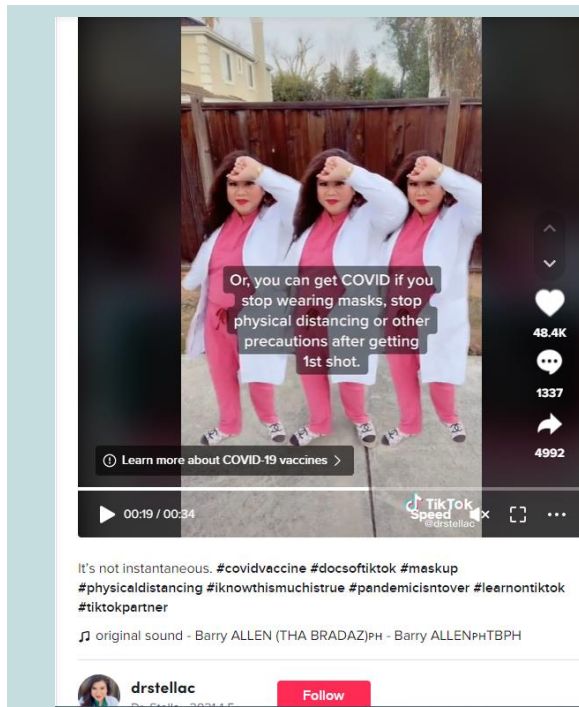


⁸⁴ Matthew S. Levendusky, "Americans, not partisans: Can priming American national identity reduce affective polarization?," *The Journal of Politics* 80, no. 1, 2018.

⁸⁵ Jon Roozenbeek, Sander van der Linden, Beth Goldberg, Steve Rathje, and Stephan Lewandowsky, "Psychological inoculation improves resilience against misinformation on social media," *Science Advances* vol. 8, no. 34, August 22, 2022.

⁸⁶ David Klepper, "'Pre-bunking' shows promise in fight against misinformation," AP News, August 24, 2022.

⁸⁷ George Mason University Center for Climate Change Communication, "Program on Climate & Health." Image source: U.S. Global Change Research Program, "The Impacts of Climate Change on Human Health in the United States: A Scientific Assessment," 2016.



Drstellac uses viral TikTok trends in her pre-bunks of false COVID claims. Above, she explains why COVID infection can occur between vaccine doses and how to prevent as part of a pre-bunk of false claims that COVID infection after vaccination proves the shots are ineffective.⁸⁸

Flooding the zone

Amplify content to your targeted audiences.

- Cross-posting: post frequently and post across social media channels, via email, on message boards. The repetition of accurate messages is necessary for engendering long-lasting resistance to disinformation.⁸⁹
- Amplify the posts from trusted communities and networks—get those algorithms working!
- Ask trusted messengers, including social media influencers, to promote your civic information campaigns.
- Create a hashtag to facilitate surfacing content and coordinating messages.
- Target particular audiences by purchasing ads or paying to boost content, disclosing who paid for the ad.
- Familiarize yourself with platform content moderation policies and practices. Marginalized social media users face disproportionate content moderation and removal, especially when discussing topics such as race, sexual orientation, gender identity, disability rights, or sexual assault.⁹⁰ Consider using “Algospeak” code words or eschewing specific terms but be cautious to avoid creating additional confusion for those unfamiliar with the code word.⁹¹

⁸⁸ Dr. Stella C, [@drstellac TikTok video](#), January 5, 2021; Laura Garcia and Tommy Shane, “[A guide to prebunking: a promising way to inoculate against misinformation](#),” First Draft News, June 29, 2021.

⁸⁹ Brendan Nyhan, Ethan Porter, and Thomas J. Wood, “[Time and skeptical opinion content erode the effects of science coverage on climate beliefs and attitudes](#),” Proceedings of the National Academy of Sciences, 119 (26), 2022.

⁹⁰ Hibby Thach, Samuel Mayworm, Daniel Delmonaco, and Oliver Haimson, “[\(In\)visible moderation: A digital ethnography of marginalized users and content moderation on Twitch and Reddit](#),” New Media & Society, July 18, 2022; Elena Botella, “[TikTok Admits It Suppressed Videos by Disabled, Queer, and Fat Creators](#),” Slate, December 4, 2019.

⁹¹ Roger J. Kreuz, “[What is ‘algospeak’? Inside the newest version of linguistic subterfuge](#),” The Conversation, April 13, 2023.



Colorado paid microinfluencers (those with 5,000–100,000 followers) to promote COVID-19 vaccines. The disclosure of the partnership with the Colorado Department of Public Health and Environment promotes trust.⁹²



Climate communications and advocacy groups such as Climate Nexus, Climate Advocacy Lab, and Digital Climate Coalition coordinate narrative development, digital deployment, advocate training, and network building on climate issues. Above, a picture of a Climate Advocacy Lab workshop.

Activation

Give ways for followers to stay involved.

- Provide opportunities for people to join your distribution pipeline, such as social media links or an email list sign-up.
- Invite your personal network to join the distribution pipelines.
- Give your audience actionable items, such as polls, petitions, and volunteer sign ups to expand your reach.
- Ask your pipeline to amplify factual posts.
- Track your effectiveness as you share your own content and refine messaging according to engagement metrics and audience feedback.
- Participate in or host opportunities to engage offline. With permission, use photos and videos from the events to increase online engagement.

⁹² James Anderson, “[US turns to social media influencers to boost vaccine rates](#),” AP News, August 10, 2021.



BallotReady, an organization dedicated to helping voters, provides resources for voters to host a BallotParty. A BallotParty is an in-person or virtual event where a host guides their invitees through information on the voting process and the candidates on their ballot.⁹³



Colorado Informed created an interactive poll that helps users learn about different voting methods, registration deadlines, and ballot drop box locations.⁹⁴

Amplifying civic information is a team sport. Subject matter experts, journalists, content creators, organizers, grassroots groups, and passionate individuals all play a vital role in ensuring civic information reaches targeted audiences.⁹⁵ They will be more effective if they network together.

⁹³ BallotReady, “[BallotParty Host Toolkit](#),” 2022.

⁹⁴ Colorado Informed, “[Voting Made Easy](#),” 2022. The German Marshall Fund partnered with The Colorado Forum and COLab to identify ways to promote civic information in the leadup to the 2022 elections.

⁹⁵ For a model of the interplay between members of the fact-checking community, see journalist and Nobel Peace

Prize laureate Maria Ressa’s #FactsFirstPH pyramid: Maria Ressa, “[Supporting fact-checking communities with Nobel laureate Maria Ressa](#),” Google News Initiative, September 21, 2022; Maria Ressa, “[We’re All Being Manipulated the Same Way](#),” The Atlantic, April 6, 2022.

THE CASE STUDY OF UKRAINE'S SOCIAL MEDIA MASTERY

The Ukrainian effort to counter Russian propaganda and showcase the horrors of the war has been a master class in using the engine of social media to promote information.

Building the pipeline

Ukrainian President Volodymyr Zelenskyy, a former television star, and other government officials embraced social media accounts to share their messages. Civil society organizations such as Promote Ukraine, a nongovernmental media hub based in Brussels, post information online about the war. They translate reports from the ground into English and hold news conferences to amplify stories. Ordinary Ukrainians described their firsthand experiences of the war using social media.

Bringing the narrative to life

Ukraine's social media strategy has embraced the principle of "show don't tell." Zelenskyy's accounts exhibit his own personal bravery by filming himself in his office and the streets of Kyiv. Citizens have posted photos and videos of the war's destruction.⁹⁶ Some of this content is crucial evidence of Russian war crimes.

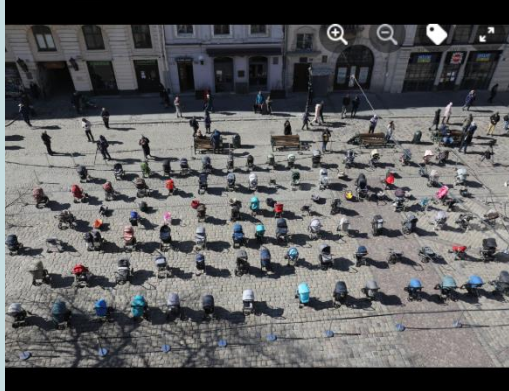


Flooding the zone

Ukrainian officials have paraded an increasingly tired and hoarse President Zelenskyy before world leaders and maintained a steady stream of videos of high-level officials pre-bunking Russian disinformation attempts.

Additionally, the Ukrainian people have taken to using Telegram, YouTube, TikTok, and even Twitter—which is not a popular platform within Ukraine—to disseminate information and updates on military and diplomatic success.

⁹⁶ Megan Specia, "[Like a Weapon': Ukrainians Use Social Media to Stir Resistance](#)," New York Times, March 25, 2022.



109 kids. 109 kids have been killed by Russians since the day one of full-scale Russian invasion of Ukraine.

Empty prams on Rynok Square today symbolize the lives of little angels. They are now defending Ukraine'sky instead of decisive actions of the world.

This is a terrible price of war that Ukraine is paying today. We call on all adults around the world to stand as one shield to protect Ukrainian children and give them future.

Urge the governments of other countries to close the sky over Ukraine.

In-person events can serve as visceral, memorable pieces of social media content that can be reshared easily.

Activation

To mobilize support, Ukraine's online campaigns have contained calls to action. They have asked for volunteers to join its decentralized "IT Army" in the cyberwar against Russia, volunteer fighters, cryptocurrency donations, political support for foreign aid, and medical materials.⁹⁷

⁹⁷ Lorenzo Franceschi-Bicchieri, "[Inside Ukraine's Decentralized Cyber Army](#)," VICE, July 19, 2022; Sara Brown, "[In Russia-Ukraine war, social media stokes ingenuity, disinformation](#)," MIT Sloan School of Management, April 6, 2022.

CONCLUSION

Civic information providers must take an active role in the amplification of the fact-based information that is critical to democracy, public health, and the environment. The tactics outlined above form a media literacy guide to understand how coordinated deceptive campaigns disseminate their messages, who the actors are, who they target, and why they are successful. This will serve civic leaders as they combat viral falsehoods and can function as a template for the distribution of their information online.

Civic information providers can and should play a more active role online, but platforms and regulators must be involved in order to fix the design flaws that allow false and misleading information to flourish in the first place.

Civic information providers can and should play a more active role online, but platforms and regulators must be involved in order to fix the design flaws that allow false and misleading information to flourish in the first place. The debate should move beyond the focus on “disinformation” and the accuracy of individual pieces of content. The current, after-the-fact platform whack-a-mole content moderation strategy is ineffective and gives rise to concerns about censorship. Instead, platforms and regulators should apply the norms and laws developed over many decades for the offline issues of consumer protection, civil rights, media, election, and national security law and renew them for the online world.⁹⁸

⁹⁸ For more information on policy reform, see: Karen Kornbluh and Ellen P. Goodman, “[Safeguarding Digital Democracy: Digital Innovation and Democracy Initiative Roadmap](#),” German Marshall Fund of the United States, March 24, 2020.

APPENDIX: BUILDING RESILIENCE AGAINST ONLINE ATTACKS

Trusted validators sometimes become targets of conspiracies or harassment. Below are practical steps on how to protect yourself and what to do if you are being harassed.

Online harassment maps onto existing power structures, privilege, and erasure. Women and people who are LGBTQ+, non-white, immigrants, have a disability, or belong to an ethnic or religious minorities are frequently targeted by attacks—especially if they belong to more than one marginalized group.

Protect yourself

Secure your accounts

- Implement two-factor authentication and use unique and strong passwords on your credit card, cell phone provider, utilities, bank, and social media accounts. Conduct regular data backups.
- Adjust personal social media account settings to the most private settings, remove addresses or specific locations from accounts, and avoid discussing personal information that could be used against you. Where possible, create separate professional accounts for social media.

Manage online footprints

- Be aware of your digital footprint through services like DeleteMe or by searching for yourself on Google or DuckDuckGo.
- Install a secure Virtual Private Network (VPN) to privatize your network traffic and ensure that attackers cannot find you using your IP address.

Strengthen community

- Connect with others in your online network. Have a plan to speak out and support one another in the event of online harassment.
- Identify a person who can monitor accounts if you are being harassed, have been doxed, or are experiencing a similar privacy-based emergency. This could be colleagues, friends, or family members.
- Tell your family and friends about the risks of online harassment. Online attackers often target family and friends, so it is important that they learn how to reduce outside access to your social media accounts.

Have a plan

- Carry out a risk assessment; different activities have different online risks and certain issues are likely to attract more online abuse than others. Once you have identified the potential attackers, become familiar with the actors and their tactics.⁹⁹
- Become familiar with the Terms of Service of the platforms you use and learn how to file a takedown request if your information gets posted.
- Leaders of organizations should develop policies to protect their employees, such as a ban on giving out personal phone numbers and addresses.

⁹⁹ Committee to Protect Journalists, “[Basic preparedness: Risk Assessment](#),” September 10, 2018.

-
- Continue to educate yourself on ways in which you may be targeted. This is particularly important if you identify as a member of a historically marginalized group. What are the ways your community has been targeted in the past? In what ways may that show up today? Who is your support system to reach out to if you face online harassment?

What to do if you are being harassed

Understand and document abuse

- Try to determine who is behind the attack and what their motives are. Many emails can be tied to a real person using an IP address. Understand that not all accounts attacking you are real people—some may be automated accounts or people paid to harass others online.
- Create a system for documenting abuse, especially anything that you feel is especially threatening and could lead to a physical attack. Documentation should include screenshots of the offensive message or image, the date, time, and name or handle of the harasser, and the date and time of any instance where the abuse is reported to a social media platform.
- Consider compiling an incident log and timeline on an encrypted word processing platform.¹⁰⁰

Prioritize your physical and mental well-being

- Stay away from the online attacks. Avoid responding to trolls, since this is often what they want and can worsen the situation. Consider blocking or muting accounts that are causing problems and disabling replies to your posts.
- Consider going offline. This could include locking down all accounts for a period of time, particularly based on your tolerance for risk and harassment, or ask a friend, colleague, or family member to monitor your accounts while you are offline.
- Lean on your community for support and healing. This is important for all groups, but particularly historically marginalized individuals experiencing attacks on their livelihood online. Make space and time to connect with your community as you disengage from harmful online behavior.

Alert institutions

- Let your credit card companies, cell phone provider, utilities provider, and bank that you are a target of online attacks.¹⁰¹
- If you are concerned about physical attacks, contact security, police, or seek support and safety within your community.

¹⁰⁰ Equality Labs, “[Anti-Doxing Guide for Activists Facing Attacks](#),” Medium, September 2, 2017.

¹⁰¹ Equality Labs, “[Anti-Doxing Guide for Activists Facing Attacks](#).”

ABOUT & ACKNOWLEDGEMENTS

The views expressed in GMF publications and commentary are the views of the authors alone.

As a non-partisan and independent research institution, The German Marshall Fund of the United States is committed to research integrity and transparency.

About the Authors

Karen Kornbluh is the managing director of GMF's Digital Innovation and Democracy Initiative.

Adrienne Goldstein is a research assistant with GMF's Digital Innovation and Democracy Initiative.

About GMF Digital

The German Marshall Fund's Digital Innovation and Democracy Initiative (GMF Digital) works to support democracy in the digital age. GMF Digital leverages a transatlantic network of senior fellows to develop and advance strategic reforms

that foster innovation, create opportunity, and advance an equitable society.

About GMF

The German Marshall Fund of the United States (GMF) is a nonpartisan, nonprofit, transatlantic organization headquartered in Washington, DC, with offices in Ankara, Belgrade, Berlin, Brussels, Bucharest, Paris, and Warsaw.

Acknowledgements

Thank you to Yael Eisenstat for advising on this project. Yael is currently Vice President at the Anti-Defamation League, where she heads the Center for Technology and Society. Thank you to the University of North Carolina Center for Information, Technology, and Public Life (CITAP), including Heesoo Jang, Kathryn Peters, and Daniel Kreiss, for their insights and feedback.

Cover photo credit: optimarc | Shutterstock

This report was made possible with the generous support of:

