'We Want You To Be A Proud Boy'

How Social Media Facilitates Political Intimidation and Violence

PAUL M. BARRETT

Contents

Authors

Paul M. Barrett is the deputy director of the NYU Stern Center for Business and Human Rights and an adjunct professor at the NYU School of Law.

Dean Jackson, a former investigative analyst with the House Select Committee to Investigate the January 6th Attack on the U.S. Capitol, is the principal at Public Circle Research and Consulting and a Tech Policy Press fellow.

Justin Hendrix is an associate research scientist and adjunct professor at NYU Tandon School of Engineering and the CEO and editor of Tech Policy Press. Tim Bernard, a Tech Policy Press contributor, helped with research.

Sarah Bernardo and Andy Zeng, summer research assistants at the NYU Stern Center, provided extensive research for this project.

Executive Summary

The main finding of this report is that social science research reveals that social media platforms can be—and often are—exploited to facilitate political intimidation and violence. Certain features of social media platforms make them susceptible to such exploitation, and some of these features should be changed to reduce the danger.

"

The main finding of this report is that social science research reveals that social media platforms can be—and often are— exploited to facilitate political intimidation and violence.

"

Based on a review of more than 400 studies published by peer-reviewed journals and think tanks, the report provides a platform-by-platform survey focusing on the particular features of each site that make it susceptible to exploitation by extremists promoting intimidation and violence and/or seeking recruits for their various causes.

The report emphasizes that neither subjective observation nor social science research indicates that social media platforms are the sole or even primary cause of political intimidation and violence. Other media and irresponsible political leaders play crucial roles. But use of social media can enable or facilitate violence in a fashion that deserves attention and mitigation. Most of this problem—extremism and occasional use of force for political ends—occurs on the political right. But the left is not immune to these pathologies.

The platforms discussed in the following pages range from some of the best known, like Facebook and YouTube, to the more recently ascendent TikTok to those on the right-wing fringe, such as Gab, Parler, and 4chan. Among the features we examine are:

- Facebook's Groups, which helped the sometimes-violent QAnon to grow into a full-blown movement devoted to the delusion that former President Donald Trump has secretly battled "deep state" bureaucrats and Satanic pedophiles.[1]

- Instagram's comments function, which has allowed the Iranian government to threaten dissidents with sexual assault and death as a way of silencing them.[2]

- TikTok's powerful recommendation algorithm, which in one experiment promoted violent videos, including incitement of students to launch attacks at school.[3]

After a case study of January 6 by our collaborators at Tech Policy Press, the report concludes with recommendations for industry and government, which we preview in capsule form here:

Recommendations In Brief

For the social media industry

1 Sound the alarm.
Social media companies first need to end their tendency to deflect responsibility and instead acknowledge the role that their platforms play in facilitating political intimidation and violence.

2 Put more people on the content moderation beat.
Given the volume of posts, use of artificial intelligence is unavoidable, but human judgment remains crucial, especially because a lot of extremist communication is cloaked in code language.

3 Directly confront election delegitimization.
Tech companies need to act aggressively to label and/or remove baseless allegations of election fraud and redirect users to authoritative sources of information.

4 Make design changes to mitigate harm.
Social media companies can reform some of the features discussed in this report to reduce the likelihood that they will facilitate political intimidation and violence.

5 Cultivate academic and civil society researchers.
Tech companies need to stand up to brazen partisan intimidation of researchers and find ways to revive productive information exchanges in the interest of staving off potential political violence.

For the U.S. government

6 Enforce existing laws.
With healthy respect for free speech protected by the First Amendment, the U.S. Department of Justice, which includes the Federal Bureau of Investigation, needs to be vigilant about enforcing criminal laws banning political intimidation and the incitement of violence.

7 Step up protection of election workers.
Of all the public officials subjected to political intimidation in recent years, election workers at all levels are the most vulnerable. We need stronger laws protecting them and more aggressive enforcement.

8 Enhance federal authority to oversee digital industries.
Longer term, the U.S. Congress needs to enhance the federal government's authority to regulate digital industries in a more systematic fashion.

9 Mandate more transparency.
Serious proposals for more vigorous regulation of digital companies begin with the need for greater disclosure of how these businesses make decisions.

1. Introduction

> **The finding that a particular technology facilitates political violence leads to our further conviction that there are steps social media companies can take to diminish their contribution to the scourge.**

Political intimidation and violence cast a cloud over civic life in the United States. Threats against judges, lawmakers, and election workers have become routine. Men driven by white nationalist conspiracy theories shoot Blacks and Jews in supermarkets and synagogues. One of the two major candidates in the approaching presidential election is promising to pardon violent insurrectionists and seek "retribution" against his political foes.[4]

The attempt on Donald Trump's life in July underscored the volatility of the current political environment but should not obscure that Trump himself has fanned the flames of political violence. Other malign influences include the subset of Trump's followers who are inclined to use harassment and force against opponents. Adding fuel to the fire are right-wing cable television networks, radio outlets, and online influencers. Left-wing extremists occasionally threaten or engage in violence, but with nowhere near the frequency of their counterparts on the right.[5]

The contribution of social media to the phenomenon of political intimidation and violence is the topic of this report. More specifically, the paper distills what social science research tells us about the relationship between social media and political violence, both rhetorical and physical. We have reviewed more than 400 studies related to our topic—most published by peer-reviewed academic journals, some by university-affiliated research groups or independent think tanks.

Our main finding is that the research shows that social media is exploited to facilitate political intimidation and violence. What's more, certain features of social media platforms make them particularly susceptible to such exploitation, and some of those features can be changed to reduce the danger.

These conclusions are quite different from saying that social media alone causes political violence. Neither social science nor subjective observation in the form of serious-minded journalism or legal investigation points to a simple causal relationship. But there *is* a relationship.

The finding that a particular technology facilitates political violence leads to our further conviction that there are steps social media companies can

take to diminish their contribution to the scourge. Building on our Center's past work on social media and political polarization, this report concludes with recommendations for such steps.[6]

Well documented episodes

Before delving into the academic studies, it's worth noting the strong anecdotal evidence of social media's dangerous potential. Numerous recent episodes that have been documented by journalists, prosecutors, and congressional investigators indicate that the platforms facilitate political harassment and physical attacks. Consider some examples:

- To incite, organize, and carry out the historic attack on the U.S. Capitol on January 6, 2021, insurrectionists relied on mainstream platforms like Twitter (now X), Facebook, and YouTube; fringe sites like TheDonald.win, Gab, and Parler; and extremist "image boards" like 4chan and 8chan (now 8kun).[7]

- The then-18-year-old perpetrator of the May 2022 massacre of Black shoppers at a Buffalo, N.Y., supermarket used 4chan, Reddit, and YouTube to school himself on racist conspiracy theories, mass shootings, and military-style weapons. In a now-familiar effort to promote and celebrate such violence, he live-streamed the attack via Discord and Twitch, social platforms with roots in the world of online gaming.[8]

- In 2023 and 2024, menacing posts on platforms ranging from X to Truth Social have been linked to a spike in direct death threats to members of Congress, judges, and prosecutors. Similarly, right-wing activists' recent incendiary posts on X have preceded intimidation of faith-based organizations that help recent immigrants to the U.S.[9]

- Elsewhere in the world: Hamas and Islamic State have used Twitter, Facebook, YouTube, and Telegram to promote their terrorist agendas and recruit adherents.[10] Myanmar's government and its militant Buddhist allies exploited Facebook to carry out an ethnic cleansing of Rohingya Muslims.[11] Hindu nationalists in India have employed WhatsApp to persecute Muslims.[12] And supporters of former Brazilian President Jair Bolsonaro took to TikTok and Facebook to spread false election fraud incitement and encourage attacks on government buildings in Brasília in January 2023.[13]

What social science contributes

While these and many similar illustrations show that social media facilitates political intimidation and violence, the expert analysis of social scientists adds depth to our understanding of *how* this relationship works—the mechanics and features of various platforms that make them vulnerable to exploitation. One common feature of social media sites is user anonymity. Among major platforms, only Facebook requires users to identify themselves with real names, and even Facebook doesn't enforce this policy rigorously. The opportunity to post anonymously, Francesco Marone, a research fellow at the Institute for International Political Studies in Milan, has found, "tends to create a disinhibition effect that can, in turn, foster increased hostility and polarization."[14]

Fulfilling their professional imperative to examine complicated social, economic, and technological developments in a systematic fashion, social scientists also provide vital context—historical and geographical—for events that otherwise might be reported in a sensational, narrow, or subjective manner.

Examining the social science is important for another reason: The social media industry itself frequently mischaracterizes the academic findings. Meta, the company whose platforms claim the largest collective user base, has repeatedly invoked what its leadership claims are the inconclusive results of social science to deflect attention from the industry's role in exacerbating political strife. Called to testify before Congress in the wake of January 6, Mark Zuckerberg, Meta's multibillionaire founder and chief executive, tried to dodge accountability. "We did our part to secure the integrity of the election," he told lawmakers in March 2021. "The reality is our country is deeply divided right now, and that isn't something that tech companies alone can fix…. Some people say that the problem is that social networks are polarizing us, but that's not at all clear from the evidence or research."[15]

Zuckerberg employed two misleading rhetorical moves. First, far from successfully securing election integrity, Meta conspicuously failed to stop its Facebook Groups feature from becoming a key venue for the incitement and planning of the "Stop the Steal" campaign to overturn a legitimate election.[16]

Second, no serious analyst contends that "tech companies *alone* can fix" the vicious and sometimes-violent

divisiveness now eroding the American political system. That is a classic straw-man argument, one echoed by Zuckerberg's top corporate lieutenant, Nick Clegg. Meta's global president, Clegg has written: "What evidence there is simply does not support the idea that social media, or the filter bubbles it supposedly creates, are the unambiguous driver of polarization that many assert."[17] Again, no observer worth listening to has contended that social media is the sole or "unambiguous driver" of political polarization or violence.

What social scientists themselves have said and written was summarized by a group of 15 researchers from such institutions as Northwestern and New York Universities and Dartmouth College in a nuanced overview article published in October 2020 in *Science*. "In recent years," they wrote, "social media companies like Facebook and Twitter have played an influential role in political discourse, intensifying political sectarianism." The researchers traced this role to the companies' fundamental advertising/engagement-driven business model: "Social media technology employs popularity-based algorithms that tailor content to maximize user engagement, increasing sectarianism within homogeneous networks, in part because of the contagious power of content that elicits sectarian fear or indignation."[18]

On some occasions, the same dynamics result in social media *facilitating*— or *intensifying* or *enabling*—political intimidation or violence.

What We Mean By 'Social Media'

"Social media" refers to internet sites that allow users to post content and interact with each other.

- **Mainstream social media**
  These are large, advertiser-supported platforms that host user-generated content. They include Facebook, Instagram, YouTube, Twitter (X), TikTok and Reddit. These platforms enforce content-moderation policies with varying degrees of effectiveness.

- **Alt-right (aka "alt-tech") and the "chans"**
  These right-leaning platforms present themselves as bastions of free speech; they typically tolerate hateful and conspiratorial expression. They include Gab, Parler, Rumble, and Donald Trump's Truth Social. The even more extreme chans—4chan, 8chan (now 8kun), and their offshoots—are "image boards" whose anonymous users post graphic memes with text.

- **Encrypted messaging apps**
  Telegram and Meta's WhatsApp began by offering private one-to-one chats but have evolved into platforms that enable sustained communication among large groups of people with barely any oversight. A forthcoming report by the Center will examine the danger (and value) of encrypted apps.

- **Gaming-adjacent sites**
  Discord and Amazon's Twitch are offshoots of the online gaming world which have grown into large social platforms that enable communication on a variety of topics. Their role in hosting extremist incitement was the topic of an earlier report by the NYU Stern Center.

Norwegian massacre, July 22, 2011
Shortly before he killed 77 people, a far-right extremist posted a YouTube video decrying Muslims and multiculturalism.

2. Surveying the Social Science

One logical place to start a survey of the social science on social media and political violence is with a survey of social scientists and policy experts. In May 2021, the Global Network on Extremism & Technology (GNET) published an overview of research combined with a survey of experts in the field. Referring primarily to mainstream social media platforms, it concluded that "internet technology, while not necessarily causing violent extremism, can have multiple and various roles in facilitating radicalization and mobilization to violent extremism."[19]

> " Referring primarily to mainstream social media platforms, the Global Network on Extremism & Technology concluded that 'internet technology, while not necessarily causing violent extremism, can have multiple and various roles in facilitating radicalization and mobilization to violent extremism.' "

The GNET describes itself as "the academic research arm" of the Global Internet Forum on Countering Terrorism, a group funded by Meta, Google, Microsoft, and other large tech companies to share information on terrorism-related content on their platforms. The GNET is based at King's College London. The author of the May 2021 GNET study, Lydia Khalil, a research fellow at the Lowy Institute, an Australian think tank, wrote: "The very establishment of the Global Network on Extremism and Technology, and the greater willingness of the tech industry to acknowledge, however haltingly, that their platforms and technologies are not only exploited by extremist actors but that their affordances have contributed to the rapid spread of extremist ideologies, has progressed our understanding."

Questions for the experts

Khalil reported that 158 experts responded to her survey. Nearly three-quarters were academics; the rest, researchers affiliated with think tanks, the tech industry, or civil society groups. Some 90% "agreed" or "strongly agreed" that internet-enabled communication facilitated the recruitment of individuals to extremist movements. "When asked if the internet has made it easier to plan attacks or mobilize to violence," Khalil found that 84% agreed or strongly agreed. One told her that "the internet and encrypted social media communications in particular have heightened the flow of information, resources, tactical and logistical support and real-time contact, which has in turn removed or flattened earlier barriers to mounting attacks."

Russian invasion of Crimea, 2014
Paving the way for the annexation of Ukraine's Crimea region, Russian military intelligence employed fake personas on Facebook and the Russian VKontakte to spread disinformation.

The GNET survey results are consistent with those of peer-reviewed literature reviews such as one published in 2020 in the *International Journal of Conflict and Violence* that tracked over time how violent extremists have migrated from static websites and password-protected forums to mainstream social media and encrypted messaging apps, using the platforms to spread propaganda, recruit followers, arrange logistics, and raise funds.[20]

Our thesis—emphasizing facilitation, not sole causation—comports with what social scientists generally have found, according to Rajan Basra, a senior research fellow at the International Center for the Study of Radicalization at King's College London. "You're bang on the money," he said in an interview. "There are important learning and instructional effects [from social media] that can shape an ideological worldview favoring political violence."

Some are more skeptical. Thomas Zeitzoff, a political scientist at American University who studies political violence and psychology, told us that it is difficult to disentangle the effects of social media from those of other media, such as cable television, as well as from factors like societal instability, poverty, and levels of violence other than political violence. Of social media and political violence, he said: "There is an effect; it's about the effect's size."

Mainstream platforms

The balance of this section is organized by platform, with an emphasis on what social science research says about platform features that extremists can use to promote political intimidation and violence.

Why Social Science Yields So Few Hard Answers

Social scientists ask useful questions about the relationship between social media and political intimidation and violence, but they haven't come up with a lot of hard answers. There are several reasons for this:

- Data is scarce, in large part because social media companies typically don't provide access to granular information about their users or platform content. In fact, companies like Meta and Twitter (X) have been restricting or eliminating access to tools that have allowed certain types of analysis by outside academics and others.[1]

- Empirical studies of the relationship between social media and political violence are particularly challenging. Let's say you identify 235 convicted terrorists imprisoned in the United Kingdom, as researchers did for a paper published in 2022 by *Studies in Conflict & Terrorism*. By rigorously interviewing these individuals and reviewing their records, you may be able to conclude that social media "is playing an increasingly prominent role in radicalization."[2] But it is difficult to say much more because of the near impossibility of assembling a control group of similar violent extremists who were not exposed to incendiary social media content.

- Sorting out the relative effects of different media is exceedingly difficult. Social media platforms exist in a complicated media ecosystem that includes talk radio, partisan cable television, podcasts, and blogs.[3] If someone's Twitter feed carries a Fox News segment highlighting false election-fraud claims by Donald Trump, is the main source of potential influence the social media platform, Fox News, or Trump himself? "There are just so many variables at play," Darren Linvill, co-director of Clemson University's Media Forensics Hub, told us in an interview.

[1] https://www.techpolicy.press/what-does-crowdtangles-demise-signal-for-data-access-under-the-dsa/

[2] https://www.tandfonline.com/doi/full/10.1080/1057610X.2022.2065902

[3] https://carnegieendowment.org/research/2023/09/polarization-democracy-and-political-violence-in-the-united-states-what-the-research-says?lang=en

#Gamergate, 2014
Misogynists used Twitter, Reddit, and other platforms to harass women advocating more progressive online games, leading to on- and offline threats of rape and murder.

Facebook

Since its launch 20 years ago, Facebook, which is part of Meta, has grown into the largest social media platform in the world, with more than three billion monthly active users.[21]

Connectivity without adequate content moderation

Facebook says its core mission is to "bring people closer together."[22] In this sense, connectivity is the most basic feature of Facebook and most other social media platforms. Making personal, educational, and commercial connections has many beneficial effects.

The persecution of Myanmar's Rohingya Muslim minority has illustrated the dark underside of introducing mass digital connectivity without adequate safeguards. In March 2018, the head of a United Nations fact-finding mission declared that Facebook had played a "determining role" in the ethnic cleansing of the Rohingya population by the country's military and allied Buddhist militants. As many as 10,000 Rohingya were killed, and more than 700,000 fled as refugees to neighboring Bangladesh. Murderous persecution of Rohingya has been continuing in 2024, according to the UN.[23]

While the UN and news organizations have documented the Facebook-fueled intimidation and violence in Myanmar, scholars have put the unnatural disaster into helpful context. Christina Fink, a professor of international affairs at George Washington University, has written a series of articles published by *Asian Survey* explaining the exploitation of Facebook's pervasiveness by ultranationalist Buddhist monks who worked in parallel with military operatives "to persuade people that Muslims are inhuman, violent, and determined to make Myanmar, which is 87% Buddhist, a Muslim-majority nation."[24]

Alison Haynes Stuart, a professor at the Charleston School of Law, builds on the Myanmar example to make the point that "increasingly, social media is used as a tool to foment violence, *particularly in regions of the world where access to the internet is otherwise limited*" (emphasis added). For most people in Myanmar, Facebook is the primary way they get online. In India, ultranationalist Hindu activists have persecuted Muslims using WhatsApp, a messaging app that, like Facebook, is owned by Meta.[25]

As these scholars have noted, the potential danger of vastly expanded digital connectivity is exacerbated when the company in question fails to oversee the situation with sufficient vigor. For years before and during the Rohingya ethnic cleansing, Facebook all but ignored warnings from civil society observers about anti-Muslim agitation. The company had only a handful of Burmese-speaking content reviewers working on an outsourced basis from outside the country.[26] In November 2018, Facebook belatedly acknowledged that it didn't do "enough to help prevent our platform from being used to foment division and incite off-line violence." Since then, the company has beefed up human and automated moderation of Burmese content.[27]

But the problem of increased connectivity without adequate content moderation remains a fundamental feature of Meta's platforms and the social media industry more broadly, especially outside of English-speaking countries. A literature review published in 2023 in *Annals of the International Communications Association* provides a wealth of illustrations of how far-right extremist groups use Facebook and other platforms to connect to members globally to spread anti-immigrant and anti-

Muslim propaganda and coordinate off-line activity, which has included political intimidation and violence.[28]

Pages and Groups

Facebook describes Pages as places on the platform where artists, brands, and organizations "can connect with their fans or customers." People with similar interests can also form Groups to discuss everything from parenting to politics. Groups can be public or private, meaning hidden from public view and accessible by invitation only.[29]

While potentially constructive, Pages and Groups also have less savory uses. A joint investigation published in January 2022 by *ProPublica* and *The Washington Post* revealed that Facebook Groups swelled with hundreds of thousands of posts attacking the legitimacy of Joe Biden's victory between Election Day in November 2020 and the January 6, 2021, siege of the Capitol: "Many posts portrayed Biden's election as the result of widespread fraud that required extraordinary action—including the use of force—to prevent the nation from falling into the hands of traitors."[30]

Social scientists have broadened our understanding of how Pages and Groups function. Stephane Baele, a professor of international relations at UCLouvain in Belgium, has written with colleagues about how right-wing extremist groups use anodyne public Pages to attract a broad audience from which more ideologically committed individuals are invited to private Groups that, in the case of the Canadian branch of the Three Percenters, are affiliated with an offline armed militia. In a paper published in 2020 in *Studies in Conflict & Terrorism*, the researchers also described how the apocalyptic, anti-government, and heavily armed Boogaloo movement in

Islamic State, 2014 and beyond
The terrorist group used tens of thousands of accounts on Twitter and other platforms to promote its ideology, organize attacks, and attract new recruits.

the U.S. migrated from the weapons-focused /k/ board on the fringe 4chan site to dozens of Facebook Groups.[31]

Geographic reach and variety are central contributions of social science, as illustrated by a book released by the academic publisher Routledge in 2019 entitled, *Researching Far-Right Movements*. The authors of a chapter on Thailand show how vigilante groups with names such as the Rubbish Collector Organization have used Facebook Groups to coordinate highly organized online and offline mob campaigns targeting political opponents.[32]

Algorithmic recommendations

Facebook is not a passive host. Its recommendation algorithm steers users to Groups based on their online activity and the activity of people in their social network. These recommendations can point users to what analyst Renée DiResta has called "deeply toxic communities." The former technical research manager at the Stanford Internet Observatory, DiResta described in her 2024 book, *Invisible Rulers*, how by early 2018, Facebook was suggesting QAnon Groups to people who had demonstrated interest in other conspiracy theories.[33]

Facebook's recommendations helped the then-nascent QAnon to grow into a full-blown movement devoted to the delusion that then-President Trump was secretly battling "deep state" bureaucrats, as well as a Satanic cabal of globalist pedophiles. As of October 2020, QAnon supporters reportedly had been arrested for threats against politicians, a break-in at the residence of the Canadian prime minister, an armed standoff near the Hoover dam, kidnapping plots, and at least one murder. That same month, under public pressure, Facebook finally disabled

Group recommendations for political and social issues and banned QAnon Groups. But "the Groups simply reassembled elsewhere" on Facebook and other platforms, DiResta wrote.

Live-stream

The Christchurch, New Zealand, shooter used Facebook Live to live-stream his mass murder of 51 worshipers at two mosques in March 2019—a gesture that promoted his bigoted violence and helped inspire copycat mass murderers. In the 24 hours after the attack, Facebook said that it blocked or removed 1.5 million versions of the video that proliferated on its platform. But copies continued to circulate online, some of them on Facebook itself.[34]

Academic consideration of the hazards of live-streaming add depth to the debate about "ethical design" in social media. Writing in the *Journal of Business Ethics*, Ateeq Abdul Rauf, an associate professor at Information Technology University in Lahore, Pakistan, argued cogently that "public policy makers need to engrain more ethical design responsibilities as part of the law for new product launches and penalize companies that fail to realistically consider the misuse or abuse of their technologies by consumers."[35]

Instagram

Instagram, also a Meta-owned platform, provides a forum for still and video imagery, often accompanied by brief text. With two billion monthly active users worldwide, Instagram is the third-most-popular social media platform, after Facebook and YouTube.[36]

Visual content: memes

As a result of its focus on visual imagery, Instagram provides a wealth of memes, helpfully described by the communications scholar Julia DeCook

as "bite-sized nuggets of politics and culture that are easily digestible." Memes, which appear on most major platforms, often include brief text.

While many memes are innocuous, the form can be used for dangerous ends, as discussed in a number of academic papers. DeCook, a former communications professor at Loyola University in Chicago who now works on policy at the Mozilla Foundation, has written about the social media-savvy Proud Boys and their use of Instagram memes to recruit members and spread their white nationalist, anti-immigrant, and violence-tinged ideology. Five Proud Boys members, including the group's former national chairman, Enrique Tarrio, were sentenced to prison terms for helping to spearhead the January 6 attack.[37]

The Proud Boys memes that DeCook analyzed in a prescient article published in 2018 by *Learning, Media and Technology* include images of Pepe, a green cartoon frog widely recognized as a mascot of alt-right groups (in this incarnation, wearing Proud Boys black-and-gold clothing); Uncle Sam accompanied by text reading, "We want you to be a Proud Boy"; and a mock "Antifa hunting permit."[38] The would-be humorous tone of many memes can camouflage a serious agenda, according to DeCook: "Behind the absurdist facade lies a powerful form of propaganda—an indoctrination that is subconscious, invisible, and violates our very understanding of logic and rational thought."

Lena Clever of the University of Münster and co-authors examined the deployment of memes on Instagram by a German group called Generation Islam, which, according to German authorities, is an affiliate of the international Hizb ut-Tahrir movement. A fundamentalist

---

Ethnic cleansing of Rohingya Muslims, 2017
Myanmar's military and its militant Buddhist allies spread hatred of the Muslim minority via Facebook, leading to thousands of deaths and forcing 700,000 to flee to Bangladesh.

Islamic organization, Hizb ut-Tahrir advocates reestablishment of an Islamic caliphate and encourages terrorism by other Islamist groups. (Hizb ut-Tahrir is banned in Germany; Generation Islam is not.) In their 2023 article in *Social Media + Society*, Clever and her colleagues described how Generation Islam's Instagram memes deploy images of weapons, burning buildings, angry men, and dead terrorists, which are intended to catch viewers' attention, but combine them with textual messages suggesting that a solution lies in community membership and collective action. The goal, according to the authors, is to "inspir[e] adherence to a narrow, radical ideology."

Comments

Comments on Instagram posts can also provide a vector for political intimidation, as illustrated by threats and harassment by apparently government-affiliated accounts in Iran. That was a central finding of research published in 2019 by *New Media & Society.* Researchers studied more than 2.8 million comments on Instagram posts by 18 Iranian dissidents in 2016 and 2017. At the time, Instagram was not blocked by the Iranian government, which has since shut down the platform. Still, many Iranians use social media by means of virtual private networks, which disguise internet protocol addresses and can be used to circumvent government bans.[39]

The *New Media & Society* study found that Iranians whose politics, art, or lifestyles offended their government were besieged by hostile comments on their Instagram posts, including "threats of death, rape, and sexual assault against women." The accounts spewing this intimidation often affiliated themselves directly or indirectly with the government, and some were accompanied by

profile pictures showing men wearing black masks and military garb. Unsurprisingly, a number of the targets of this abuse curtailed the public behavior and expression that had drawn the government's attention.

YouTube

With 2.7 billion monthly average users worldwide, the video-sharing platform YouTube is one of the most popular social media sites. It is owned by Google, which is part of the holding company Alphabet.[40]

Social science on YouTube has focused heavily on the effects of the platform's recommendation engine, which points users to videos related to the one they are watching. But the role of recommendations is most usefully considered in tandem with YouTube's more general feature as an online repository, or library.

Video library

The government of New Zealand identified YouTube as a key contributor to the racist radicalization of the lone terrorist who killed 51 Muslims in Christchurch in March 2019.[41] In response, YouTube vowed to step up its efforts to remove hateful and violent content. But three years later, another loner, this one in New York state, was watching videos on YouTube about mass shootings, crimes committed by African-Americans, and tips for using military-style weapons. Inspired by the writings of the Christchurch terrorist about the great replacement, which he found elsewhere online, the New York loner murdered 10 Black shoppers in a Buffalo, N.Y., supermarket in May 2022.[42]

As these horrendous episodes illustrate, the main way that YouTube facilitates political violence is by serving as a

library of virtually all types of content, including darkly inspirational material. People inclined toward extremism go to YouTube looking for extremist and violent content, wallow in it, share it, and, in the most extreme cases, become violent threats to others. When intimidation or physical violence ensues, YouTube cannot be described as the sole or even primary cause of that brutality, but the platform may have intensified the user's extreme views and facilitated the violence.

Algorithmic Recommendations

Our analysis of the library effect tells a somewhat different story from the YouTube "rabbit hole" phenomenon. Beginning in 2018, journalists and academics reported episodes when YouTube fed users a series of recommendations of increasingly extreme political content, essentially pushing people toward white supremacist rants and Holocaust denialism when those people had not necessarily shown interest in such material. A 2020 article published by *Open Information Science*, entitled, "The YouTube Algorithm and the Alt-Right Filter Bubble," exemplifies this genre with its claim that the platform's recommendation engine "makes YouTube a powerful recruitment tool for neo-Nazis and the alt-right."[43]

A number of more recent studies contest the rabbit hole thesis. Researchers at Dartmouth led a study published in 2023 in *Science Advances* based on behavioral and survey data that found that YouTube users rarely see or follow recommendations to right-wing or extremist videos, unless they seek out such content. "Our findings suggest that YouTube's algorithms were not sending people down 'rabbit holes' during our observation window in 2020," the researchers wrote, noting that the company made changes to its

Charlottesville Unite the Right march, August 12, 2017
Organizers used a Discord chat group and Twitter to plan the white nationalist event, which led to violent clashes and the death of a counter-protester.

technology in 2019 designed to reduce recommendations of conspiracy theories, misinformation, and other problematic content.[44]

But the Dartmouth-led team added an important qualification: Among audiences already dedicated to hateful or conspiratorial ideas, YouTube "continues to play a key role in facilitating exposure to content from alternative and extremist channels." And when users do seek out YouTube channels featuring extremist fare, the platform recommends more such content to them.

TikTok

The Chinese-controlled short-video platform TikTok, famous initially for teenage dance moves, has 1.6 billion monthly average users worldwide, with more than 120 million in the U.S.[45]

For You Page
Many analysts attribute TikTok's rapid rise in popularity to its recommendation algorithm's uncanny ability to serve up videos that users find fascinating. These recommended videos appear on users' For You Page as a result of technology that differs from that of other major platforms. As researchers affiliated with New York University's Center for Social Media and Politics have explained, Facebook, Instagram, and Twitter (X) "are structured around a social graph in which users follow and are followed by other users. In turn, the content we see depends on what's shared in our network." TikTok's For You Page, by contrast "surfaces videos based on algorithmic recommendations from outside of one's social network," based on users' intuited interests.[46] (Scrambling to keep up with TikTok, Meta reportedly has tweaked its Reels video feature to make its algorithm more like that of its Chinese-controlled competitor.[47])

Almost immediately upon becoming available globally in 2018, TikTok became the subject of journalistic exposés finding that in addition to its more lighthearted fare, the platform hosted content supporting neo-Nazis and promoting Islamic State beheadings.[48] A paper published in July 2024 by the Institute for Strategic Dialogue found that TikTok is still hosting "hundreds of accounts which openly support Nazism and use the video app to promote their ideology and propaganda."[49]

A paper published earlier in 2024 by the Global Network on Extremism and Technology shows how TikTok's powerful recommendation engine can bring content promoting intimidation and violence to the For You Page. Responding to a spate of school shootings in her native Brazil, researcher Beatriz Buarque, a fellow at the London School of Economics and Political Science, tested TikTok by establishing an account that mimicked a young Brazilian interested in entertainment, beauty, news, and opinion. After she watched videos on those topics, she "liked" those referring to shaming, discrimination, and violence. "It took exactly 12 days," Buarque wrote, "for TikTok's algorithm to move from discriminatory videos and cyberbullying to videos containing explicit appeals to violence—in some cases, encouraging youngsters to take violent action at school." In all, during the month-long experiment, TikTok served up 157 videos making explicit appeals to violence, including five that applauded violence in schools. One hundred and fifteen of those referred to a website containing videos of people being tortured and killed.[50]

Video effects
TikTok features a suite of video effects that allow users to alter and augment videos. Meant to elicit creativity and keep users on the site longer, these effects can also be exploited by promoters of hatred and violence, as explained by a 2021 paper published by the Institute for Strategic Dialogue (ISD).

The ISD identified 491 extremist-oriented accounts active in June 2021 that posted 1,030 videos promoting white supremacy, the 2019 Christchurch massacre, and the Islamic State terrorist group. Users applied greyscale or sepia "vintage/retro filters" to heighten the effect of imagery of Nazi rallies. Others used a "green screen" effect so they could appear in front of a video—in one case, showing the interior of a Nazi concentration camp as a backdrop, with a superimposed person in a skull mask, mocking Jews. A third use of effects involved a "photo slideshow" employed to display images of Hitler, Mussolini, and other 20th century fascists.[51]

Hashtags
People on TikTok and other platforms commonly use hashtags to label content they post in hopes that the labels will facilitate searches for their material. On TikTok, the hashtags #ForYouPage and #FYP are widely used for this purpose, but ISD researchers identified a range of other hashtags used to amplify extremist content, including #MakeEuropeGreatAgain, #Fascism, and #Hyperborea, a reference to esoteric Nazism. "They tag their hateful or extremist content with these terms with the aim of having more people see their content and engage with their account," the ISD paper stated.

Indian Hindu nationalist attacks on Muslims, 2018 and beyond
Members of the country's Hindu majority use WhatsApp messages and videos to persecute Muslims, including incitement that has led to violence and lynchings.

Sparse Research on Left-wing Extremism

Social scientists have paid relatively little attention to left-wing activity on social media and its role in facilitating political intimidation and violence. This inattention may reflect the liberal political sympathies of the majority of academics and/or evidence that, as the Carnegie Endowment's Rachel Kleinfeld has written, "political violence and spontaneous hate crimes that harm people are being committed vastly more by those on the right" than on the left.

Kleinfeld, an expert on political violence, went on to note that left-wing violence, while relatively rare, nevertheless does occur—for example, the attacks on police officers and commission of significant property damage during some of the 2020 Black Lives Matter protests—and deserves attention.[1] Likewise, the role of social media in left-wing political intimidation and violence merits analysis.

The section on Twitter (X) in this report includes a discussion of how antagonists on the left and right used that platform to bait one another in the build-up to the physical violence during the August 2017 Unite the Right rally. (Illustrating Kleinfeld's point, white supremacist marchers in Charlottesville were overwhelmingly responsible for the physical clashes, injuries, and death of a counter-protester.)[2]

Another study published in 2023 by the *European Journal on Criminal Policy and Research* examined how Twitter (X) users affiliated with the loosely organized leftist network Antifa "dox"—publicly disclose personal information to intimidate, humiliate or otherwise inflict harm —on targets they believe are members of the Proud Boys, Patriot Front, and other right-wing groups.[3]

And a pair of papers published by researchers at Victoria University in Australia and the Institute for Strategic Dialogue found small numbers of posts on Facebook and Twitter reflecting advocacy of violence and antisemitism by left-leaning users in that country.[4]

---

[1] https://carnegieendowment.org/research/2022/09/five-strategies-to-support-us-democracy?lang=en

[2] https://ijoc.org/index.php/ijoc/article/view/10076

[3] https://link.springer.com/article/10.1007/s10610-023-09558-6

[4] https://static1.squarespace.com/static/5d48cb4d61091100011eded9/t/5f-c00ee72dd96f5918fb40a6/160642232; https://static1.squarespace.com/static/5d-48cb4d61091100011eded9/t/6179e7f6c6cdbb44257c1ffa/163537921

Twitter (X)

Renamed X after Elon Musk's takeover in 2022, this microblogging platform has more than 600 million monthly active users worldwide. About 22% of U.S. adults report using it.[52]

Anonymity

Renée DiResta has written: "Twitter's design creates an opportunity for emergent collective behavior in which bystanders everywhere can instantly jump right into open crowds and start brawling." #Gamergate in 2014 provided an alarming illustration. Male online gaming devotees lashed out at individual women critics of the misogyny and violence in digital gaming culture, in some cases threatening rape and murder. Participants in such mobs "are often further emboldened by the cloak of online anonymity," according to DiResta. "There are no consequences for the behavior and minimal potential for de-escalation short of the platform suppressing a trend or suspending accounts of the worst participants—cold comfort, as new members of the crowd will be online again a few hours later to continue the fighting, and still more will appear to complain about 'censorship.'"[53]

Tamar Mitts, an assistant professor of international and public affairs at Columbia University, has illustrated how over the years anonymity has been central to Twitter's role as a tool for radicalizing and recruiting Muslims in the West to join the protean Islamic State terrorist organization active in Syria and Iraq. In a paper published in 2018 in *American Political Science Review*, Mitts wrote: "Potential recruits found it appealing to connect to the organization through Twitter, as the platform enabled the anonymous consumption of radical and extremist ideas, without being exposed to the risk of physically interacting with a recruiter."

The Mitts study is valuable because it underscored that social media has not

Christchurch mosque attacks, March 15, 2019
The killer of 51 Muslim worshipers promoted the violence in a Facebook livestream which was reposted on other platforms and helped inspire other mass shootings.

been the sole radicalizing influence on Muslims in the four countries that were the political scientist's focus—Belgium, France, Germany and the U.K. Mitts found that local anti-Muslim hostility, as measured by the political successes of far-right anti-Muslim parties, correlated with increased recruitment to violent jihad.[54]

Hashtags and trending lists

Mentioned earlier in connection with TikTok, hashtags provide a useful mechanism on Twitter as well for spreading violent ideology. Bradley Wiggins, a professor at Webster University in Vienna, has explained how hashtags, in addition to facilitating searches, commonly appear in lists of trending subject matter that Twitter provides to users in their For You and Trending tabs. These lists of the most heavily discussed subjects at any given time essentially serve as recommendations by the platform. They create an incentive for extremists to get their recruitment and incitement content trending, a first step toward the influencer's goal of making their material "go viral."

In a paper published in 2021 in *New Media & Society*, Wiggins illustrated these observations by analyzing the spread of three hashtags: #boogaloo, #boogaloo2020, and #civilwar2—all popular with the armed anti-government Boogaloo movement in the U.S. "In the Twitter Boogaloo discourse, hashtags are used to further conspiratorial thinking (such as claims of 'crisis actors' used to stage fatal shootings) but also to mobilize and encourage solidarity (as in the deployment of #boogaloo and various iterations)," Wiggins wrote. He examined Boogaloo Twitter activity during the 10-day period around a January 2020 gun-rights rally in Virginia during which many attendees openly carried firearms. Among the examples cited in the paper was a Boogaloo tweet appropriating a common meme of actor John Krasinski, best known

for his role on the television series "The Office," with a white board. In the Boogaloo version, there were two white boards. The first stated, "Government is a monopoly on violence"; the second, "So let's give it some competition."[55]

Adam Klein, an associate professor of communications and media at Pace University in New York, studied how Twitter offered antagonists on both the right and left an online staging ground for physical violence that occurred during the August 2017 Unite the Right rally in Charlottesville, Va. In a paper published by the *International Journal of Communication*, Klein charted tweets by the far-right groups Proud Boys and Oath Keepers as well as by self-described adherents of the anarchist network Antifa.[56]

While emphasizing that people associating themselves with Twitter accounts such as Antifa Berkeley and Antifa NYC were responding to plans on the right to protest the removal of a Confederate statue in Charlottesville, Klein nevertheless pointed out that some of the leftists also appeared to welcome the anticipated violence that, in fact, came to pass. "When you punch a Nazi, the whole world punches with you," one left-wing tweet stated. On the other side, right-wingers tweeted, "Always #ready for war," and "The left is preparing lynch mobs to descend on the Unite the Right rally in Charlottesville, Va. This is going to be fun."

Reddit

With 306 million weekly active unique visits, Reddit consists of thousands of "subreddits," semiautonomous communities dedicated to specific topics and governed by volunteer moderators. Reddit also enforces certain platform-wide content policies.[57]

Subreddits and upvoting

The decentralized structure and do-it-yourself content moderation that make Reddit appealing to millions of users

also can promote an environment suited to extremist discourse and mobilization of like-minded bigots. That's one of the main findings of Michigan State University PhD candidate Tiana Gaudette and co-authors of her 2021 paper in *New Media & Society*. What's more, the study revealed that Reddit's "upvoting" feature can heighten more extreme content. By voting for particular content, users can increase its visibility and engagement level; conversely, content can also be downvoted and diminished.[58]

Gaudette and colleagues analyzed the subreddit /r/The_Donald, which was started in 2015 by Trump supporters and grew to nearly 800,000 users. During 2017, Trump's first year as president, a sample of highly upvoted content tended to feature violence-tinged anti-Muslim and anti-left commentary when compared to a random sample of non-highly upvoted content.

In June 2019, Reddit "quarantined" The_Donald in response to its users posting violent threats against police officers and politicians in Oregon. Quarantining restricts access to a subreddit's content and requires people to opt-in to view the restricted material. In June 2020, Reddit banned the pro-Trump thread altogether for engaging in harassment, hate speech and content manipulation. "Communities and users that incite violence or that promote hate based on identity or vulnerability will be banned," Reddit said at the time, noting that it was shutting down 2,000 subreddits, the vast majority of which were inactive.[59]

Migration

The migration of former users of The_Donald to new online locales illustrates that the collective identity formation and mobilization that can occur on subreddits (and in Facebook Groups) does not necessarily evaporate if a platform intervenes to shut down

Riot the U.S. Capitol, January 6, 2021
Donald Trump used Twitter to summon supporters for a protest he promised "will be wild"; participants congregated in Facebook Groups to whip up enthusiasm for the violent attack.

a toxic venue. Beginning when Reddit quarantined The_Donald, its former moderators led a mass move to an alternative site known as TheDonald.win, which functioned similarly to Reddit but without the latter's rules against hate speech and violence. In the aftermath of the November 2020 presidential election, according to the investigative staff of the House Select Committee to Investigate the January 6th Attack on the U.S. Capitol, TheDonald.win became a hotbed of racist and violent fulmination, including mobilization for the siege of Congress. The site has since been shut down;[60] many of its users have migrated to Patriots.win.

Several scholars have noted the out-migration pattern from Reddit in connection with the sometimes-violent incel movement. Incel (involuntarily celibate) adherents promote a misogynistic notion of male victimization which in extreme cases has led to murderous outbursts. A multi-author paper published by *Terrorism and Political Violence* in 2023 traced the evolution of several incel-focused subreddits—with names like /r/Incel and /r/Braincels—toward increasingly violent rhetoric, culminating in Reddit ultimately shutting them down in 2018 and 2019.

Incel adherents migrated in two directions: toward even more extreme dedicated forums like incels.is and to tamer incel and men's rights subreddits like /r/TheRedPill, which has been quarantined. Of TheRedPill, the authors of the 2023 paper wrote: "Users actively tone down some of the more extreme conversation to avoid having the board shut down after Reddit placed it in quarantine."[61]

Encrypted messaging apps

Encrypted apps like Telegram (900 million monthly average users) and WhatsApp (nearly 3 billion) were designed for private one-to-one communication which is of obvious value to people who want to shield their messages from government surveillance or other eavesdropping. Political dissenters, human rights defenders, and ordinary confidentiality-minded individuals all benefit.[62]

But the same secrecy that insulates legitimate free speech also protects those facilitating political intimidation and violence. What's more, messaging apps have evolved to include group and channel features that enable one-to-many communication, which can turbocharge mass incitement. Our Center will assess the pros and cons of encrypted messaging in the context of elections in a forthcoming report, so here we will note only briefly the dangers the technology presents.

Telegram

Telegram hosts both encrypted and unencrypted channels. Unsurprisingly, bad actors gravitate to the secret ones.

In 2022, the industry-affiliated Global Network on Extremism & Technology published a helpful assessment of the "Terrorgram community"—"a loosely connected network of Telegram channels and accounts that adhere to and promote militant accelerationism. Terrorgram channels are typically neofascist in ideological orientation, and regularly share instructions and manuals on how to carry out acts of racially-motivated violence and anti-government, anti-authority terrorism." The authors noted that Telegram has at times restricted or banned certain Terrorgram channels, but "new groups keep emerging and some attempt to create coalitions through Telegram's permissive stance towards terrorist content."[63]

Graham Macklin, an assistant professor at the Center for Research on Extremism at the University of Oslo, has described Terrorgram and the broader ecosystem that it belongs to as "a 'dark fandom' that venerates and valorizes extreme-right terrorists as 'saints' and 'martyrs' in a manner similar to the heroization of school shooters and serial killers."[64] This repulsive "saints culture" extends in unlikely directions, including its honoring of Ted Kaczynski, the anti-modernity "Unabomber" terrorist who has had a significant influence on leftist eco-militants.[65]

WhatsApp

WhatsApp, owned by Meta, is hugely popular in India, where its vigorous use by Hindu nationalists to persecute Muslims (and Dalits and Christians) has on some occasions turned on allegations about the slaughter, trafficking, or theft of the cattle that Hindus hold sacred. Other WhatsApp groups have trafficked in false accusations of child snatching and organ harvesting. "Although the victims are targeted for different reasons, these incidents have in common mobs of vigilantes who use peer-to-peer messaging applications such as WhatsApp to spread lies about the victims, and use misinformation to mobilize, defend, and in some cases to document and circulate images of their violence," according to a report entitled, *WhatsApp Vigilantes*, published in 2019 by the London School of Economics and Political Science.[66]

Chinmayi Arun, a research scholar and executive director of the Information Society Project at Yale Law School, has added an important qualification, stressing that malign rumors have triggered violence for eons. "The question we should ask is not whether WhatsApp causes violence, but whether (and how far) WhatsApp may have exacerbated the proliferation of lynchings in India."[67]

Alt-right platforms and 'the chans'

White supremacists and others on the far right were relatively early in moving a good deal of their communication

Russian full-scale invasion of Ukraine, February 2022
The Kremlin used thousands of fake accounts on TikTok and other platforms to undermine support for Ukraine.

and recruitment to the internet, as illustrated by the appearance of the neo-Nazi Stormfront website in 1995. Joan Donovan, an assistant professor of journalism and emerging media studies at Boston University, has written[68] with colleagues about how digital sites provided the far right with "new tools to harass and intimidate vulnerable populations." In this section, we will discuss some of the most prominent alt-right (alternative-right) platforms, as well as the even more extreme "chan" sites.[69]

Gab

In a chapter in the 2019 book *Post Digital Cultures of the Far Right*, Donovan and her coauthors show that Gab enjoyed a surge of new users in the wake of the August 2017 Unite the Right march in Charlottesville, Va. after which Twitter (X) and other mainstream platforms banned a number of right-wing users and their content. Three-and-a-half years later, the mainstream platform reaction to the January 6 siege led to a similar user boomlet for Gab.

Lax content moderation
Numerous social scientists have observed that the fundamental feature that makes Gab and other alt-right platforms conducive to facilitating political intimidation and violence is lax or absent content moderation. The free speech that these platforms champion includes hateful expression that sometimes incites violence. Gab's most notorious user was the shooter in the 2018 Tree of Life massacre, who steeped himself in antisemitism and anti-immigrant hatred on the platform and then announced his murderous intentions on the site just before entering the Pittsburg synagogue and killing 11 congregants.[70]

'Cloned' mainstream features
Donovan and co-authors point out that Gab appeals to its users by combining cloned versions of the features of mainstream platforms from which

many of them migrated. These include Facebook's social connectivity and groups, Twitter's micro-blogging, and the news aggregation and "voting" of Reddit. "By cloning features common to larger platforms and consolidating them into a single user experience, Gab's platform is both political and infrastructural," the researchers wrote. In contrast, its antisemitic, racist, and misogynistic fare limit its appeal to advertisers and potential as a commercial venture.[71]

Parler

'Preparatory media'
Parler has played a role similar to Gab. Both provided tools to the organizers of January 6 and the people who actually invaded the Capitol. Writing in the journal *First Monday* about Parler, Luke Munn, a research fellow at the University of Queensland in Australia, categorized the site as featuring "preparatory media" for incitement, mobilization, and legitimization. He quoted the late Yale political scientist David Apter: "People do not commit political violence without discourse. They need to talk themselves into it."[72]

Among the posts on Parler attracting tens of thousands of impressions immediately before January 6 that Munn collected: "CALLING ALL PATRIOTS. DONALD TRUMP HAS CALLED FOR US TO COME TO THE NATIONS [sic] CAPITOL FOR THE LAST STAND AGAINST THE GLOBALISTS." Another one was: "WE JUST ESSENTIALLY GOT VOTED OUT OF OUR OWN COUNTRY YOU FUCKING MORONS!!! …PUT DOWN YOUR FAGGY TRUMP SIGNS AND PICK UP AN AR-15!!!!! WAKE UP CHILDREN, IT'S TIME TO GO TO WAR!!!!!"[73]

4chan and the other chans

Stephane Baele and co-authors provided a useful survey of the chan (derived from "channel") ecosystem,

which has included the likes of 4chan, 8chan (now 8kun), 16chan, Shitchan, EndChan, NeinChan, and Kohlchan. These rudimentary, anonymous "image boards" have hosted the manifestos and livestreams of mass shooters in Christchurch, New Zealand (2019); El Paso, Texas (2019); and Hanan, Germany (2020), among others. Much if not all of the relevant violent content on the chan sites tends to appear on threads labeled "/pol/," short for "politically incorrect."[74]

Havens for 'transgressive fun'
The chans offer their users a virtual gathering place for "transgressive fun," according to Cathrine Thorleifsson, an anthropologist at the University of Oslo. In a 2021 paper published by *Nations and Nationalism*, she noted that "acute amusement in the face of someone else's distress is a sensation driving users to chan forums." It's not really fun and games, though. As Thorleifsson noted, "an affective politics of fear that dehumanizes people of color and minorities as an existential threat to the white race can inspire users to violent action in the real world." One manifestation of this mindset is violent "gamification": the tendency of users to rate and score offline incidents of political violence as if they were video games.[75]

Testing grounds
A final facet of the chans worth noting is that they often serve as the point of origin of white nationalist conspiracy theories that resurface on mainstream platforms, sometimes in sanitized form to avoid content moderation. "Online far-right communities such as 4chan/pol/, 8chan, and Gab often act as a testing ground for these memes before then are deployed on a more ubiquitous social media platform such as Twitter or Facebook," according to a 2023 paper published by *New Media & Society*.[76]

Buffalo supermarket massacre, May 14, 2022
Before killing 10 Black shoppers, the gunman embraced the racist "great replacement theory" he learned about online.

3. January 6: A Case Study

By Dean Jackson and Justin Hendrix

The Select Committee to Investigate the January 6th Attack on the United States Capitol produced a 122-page draft memo on the role of social media in the insurrection which, while not included in the Committee's final report, has since been published online.

> " QAnon and Stop the Steal were both prominent on Facebook, Twitter (X), and other mainstream sites, but other platforms also played an important role. "

The memo criticized key decisions by leaders of major platforms such as Twitter (X) and Facebook and explored the role that alt-right and other, smaller platforms played in encouraging and facilitating offline violence on and before January 6. Committee investigators found that former President Donald Trump's online utterances were a chief source of incitement; that a small number of online influencers and organizers played a disproportionate role in mobilizing the mob at the Capitol; and that key decisions by platform executives on design, policy, and content moderation failed to stem the flood of election conspiracy theories and violent rhetoric.

Nearly four years later, social science has confirmed many of these findings. This case study distills more than 280 academic analyses exploring the relationship between social media and the insurrection.

Stop the Steal

Both mainstream and alternative platforms allowed domestic extremists and prominent figures on the right to accelerate the growth of movements like QAnon and Stop the Steal, contributing directly to the insurrection.

Hal Berghel (2022) cited research by University of Chicago political scientist Robert Pape suggesting that adherence to QAnon and high rates of social media use are strong predictors of both election denial and an appetite for political violence.[77] A joint report by Polaris and the Soufan Group suggested that online rumors about human trafficking, and especially sex trafficking of children, have been a major gateway to QAnon adherence.[78] Claire Seungeun Lee, et al. (2022) compared Trump's online speeches to subsequent hashtags used by QAnon adherents to further link the former President's rhetoric to the movement.[79]

In an analysis of Twitter (X) data leading up to the Capitol attack, Padinjaredath Vishnuprasad, el al. (2024) found that many election deniers were adherents to QAnon before the emergence of the Stop the Steal movement and that networks of rapid retweeters, copy-pasted messages, and the participation of high profile right-wing outlets and commentators were essential to the spread of these narratives.[80] Meta's own internal analysis of Stop the

Attacks on Brazilian government buildings, January 8, 2023
In an echo of the January 6 U.S. Capitol riot, supporters of former Brazilian President Jair Bolsonaro spread election fraud lies on Facebook and TikTok.

Steal activity on Facebook reached a similar conclusion: Just 137 "super inviters" were responsible for 67% of growth in the largest Stop the Steal groups, and "almost all of the fastest growing [Facebook] groups were Stop the Steal during their peak growth." Taken together, their findings support other work such as Rimi Nandy and Jhilli Tewary (2024) which suggested that a nexus of radicalized individuals, influential public figures, and the affordances of social media platforms contributed to the insurrectionists' motivating beliefs.

Alt-right growth

QAnon and Stop the Steal were both prominent on Facebook, Twitter (X), and other mainstream sites, but other platforms also played an important role. Max Aliapoulios, et al. (2021) showed how a combination of endorsements from prominent right-wing figures and "deplatforming" of users on more mainstream services led to growth on the alt-right platform Parler before the insurrection.[81] Wei Zhong, et al. (2024) demonstrated "substantial interaction" between the Proud Boys, QAnon supporters, and white supremacists on Telegram, the largely unmoderated messaging service.

Trump's culpability

Twenty-two studies in our sample discussed former President Trump's culpability in the insurrection. Many focus on Trump's use of social media, especially Twitter (X). Trump capitalized on the new media environment to encourage violence by his supporters in a way that is unique in the history of the presidency. Mohd Razman Achmadi Muhammad and Norr Nirwandy (2021) wrote for the *Journal of Media and Information Warfare* that Trump's use of social media challenges the news media's traditional agenda-setting power, allowing him to bypass traditional media and directly inspire and encourage action from his supporters online.[82] John Allen Hendricks and

Dan Schill (2024) likewise emphasized the "unfiltered and unmediated" nature of Trump's digital communication to the public and its role in inspiring violence at the Capitol.[83]

Helen Harton, et al. (2022) suggested that Trump's online activity, combined with the "ease of online discussion," created dangerous group dynamics in which both an authority figure—the former president—and peer influencers spurred individuals to violence.[84] Barseel AlBzour (2022) used "speech act theory" to analyze Trump's tweets leading up to the insurrection and conclude they were understood by their audience as directives that incited violence both implicitly and explicitly.[85]

Gaps in the literature

Important gaps remain in the literature tracing social media's influence on the insurrection. Studies focusing primarily on the design and affordances of platforms were scarce; features like algorithmic content recommendations, group invitations, or other common aspects of social media were mentioned only rarely and usually with passing references to how social media accelerated polarization, "echo chambers," and "filter bubbles." Internal Meta research leaked by Facebook whistleblower Frances Haugen suggested that Facebook (and Facebook Groups in particular) has had a strong independent effect on the trend toward negative, angry political discourse online.[86]

Rather than delving deep into the role these trends played in producing January 6, though, studies were much more likely to look at user-generated content or the impact of content moderation decisions such as deplatforming. This is in contrast to the findings of the committee investigators' memo, which dedicated significant space to levers "upstream" of content moderation like the "break glass" measures deployed by Facebook to limit the risk of violent incitement and extremist organizing before the 2020 election. More recent

work on platform design codes by University of Southern California scholar and former Meta data scientist Ravi Iyer and others explores this work in contemporary context, but little empirical work has been done on how design choices facilitated January 6.[87] This paucity of empirical work is partially a result of the continuing opacity of the social media industry.

Another key gap is the connection between broadcast and social media. An exception—Muhammed and Nirwandy (2021)—demonstrated that, on Twitter, Trump actively promoted One America News Network (OANN) content and disparaged Fox News in comparison. OANN was a common news source for Capitol rioters. The internet does not exist in isolation, and research on the media environment is too often siloed into the on- and offline.[88]

Historical context needed

Finally, more could be done to situate January 6 historically. While shocking, the insurrection was not unprecedented. As Diren Valayden, et al. argued in the *Journal of Right-Wing Studies* (2024), the insurrection was "one out of 45 protests at state capitols and elsewhere in 32 states on that day."[89] Earlier in 2020, the country witnessed several other incidents in which armed protestors breached state capitols. Moreover, these scholars pointed out, January 6 was precisely the type of event one would expect if right-wing rhetoric were taken at face value ("take our country back," etc.).

In a PhD dissertation submitted to the University of Michigan, Hanah Stiverson (2023) similarly referred to "banal facism" and traced the history of far-right organizing in the United States.[90] Understanding the relationship between political violence and long running historical themes in American life is an essential part of disentangling social media's impact on society from other trends and influences.

Death threats against U.S. public officials, 2024
Menacing posts on platforms ranging from X to Truth Social have been linked to a spike in death threats targeting lawmakers, judges, and prosecutors.

4. Company Responses

> **TikTok reviewed the charging documents in the federal cases of nearly 600 people prosecuted for their role in the January 6, 2021, siege of the U.S. Capitol, finding that TikTok was mentioned in only 10 cases, far fewer than Facebook (255), YouTube (82), and Instagram (59).**

We asked social media companies to respond to our general assertion about the facilitative role that social media plays vis a vis political intimidation and violence—and to our observations about their platforms in particular. It was striking that both Meta (Facebook and Instagram) and YouTube pointed to a 2023 article written by Rachel Kleinfeld entitled, "Polarization, Democracy, and Political Violence in the United States: What the Research Says."[91]

When addressing the role of social media, Kleinfeld, a senior fellow at the Carnegie Endowment for International Peace, actually examined how the technology contributes to *political polarization*, a phenomenon that may provide a foundation for violence, but is not necessarily violent. She concluded that the relevant studies "suggest that all media, not just social media, may be playing a role" in increasing "affective polarization," meaning us-versus-them divisiveness that goes beyond policy differences. Kleinfeld noted that the scholarship on social media and polarization is "extremely mixed" but that the technology has been "singled out for its negative effects because its algorithms and business models have been shown to exacerbate outrage and anger. Social media has also been found to help recruit and provide platforms to extremists."

Far from contradicting the main point of this report, Kleinfeld's conclusions reinforce it.

Militia groups organizing, 2024
After lying low after the January 6 insurrection, armed militia groups have been stepping up organizing and recruiting on Facebook.

More specific company responses

Meta

A company spokesperson said: "We want our platforms to be a place where people can safely express themselves. That's why we take action on content, like hate speech, bullying, or harassment, that violates our Community Standards and why we continue to invest in new technologies and methods to help protect people on our services. In addition, the company pointed us to its:

- responses to human rights reviews that the company has commissioned, including improvements to automated content moderation systems to detect hateful material in Arabic and Hebrew and stronger protections against harassment and political intimidation in the Philippines;[92]

- community standards banning content that incites violence, traffics in hate speech, or promotes dangerous (violent) organizations or individuals.[93]

- election integrity policy, under which it will block new political advertisements during the final week of the November campaign in the U.S. and require advertisers to disclose when they use artificial intelligence to "create or alter political or social issues ads in certain cases," among other steps;[94]

- decision to stop proactively recommending political content on its Instagram and Threads platforms.[95]

YouTube

The company declined to comment on the record but noted:

- its community guidelines, which include a ban on incitement to violence, harassment, and hate;[96]

- its election information panels, which may provide context and authoritative information when a user searches for or watches videos related to political candidates, parties, or voting;[97]

- a 2023 study by researchers at the University of California, Davis who found that a small segment of the U.S. population considers violence, including lethal violence, to be usually or always justified to advance political objectives;[98]

- a 2024 survey by the Pew Research Center finding that "seven-in-ten Americans say elected officials should avoid heated or aggressive language because it could encourage some people to take violent action."[99]

Anti-immigrant riots in Europe, 2024
Social media posts attacking immigrants contributed to violence in Britain, Portugal, and other countries.

TikTok

Lisa Hays, the company's head of safety public policy and senior counsel for the Americas, said that TikTok removes more than 98% of material posted by hateful organizations and individuals and works with experts "to keep ahead of evolving trends….When a video is first uploaded to TikTok, it goes through our content moderation system, which combines both automated systems and moderation teams, and if a member of our community comes across a video, they can report it from a pop-up menu and send it to our moderation team for review." In addition, she said TikTok:

• reviewed the charging documents in the federal cases of nearly 600 people prosecuted for their role in the January 6, 2021, siege of the U.S. Capitol, finding that TikTok was mentioned in only 10 cases, far fewer than Facebook (255), YouTube (82), and Instagram (59);

• does not have a groups feature of the sort that allow "like-minded individuals to plan and coordinate activities" on certain other platforms;

• bans "any violent threats, promotion of violence, incitement to violence, or promotion of criminal activities that may harm people, animals, or property;[100]

• prohibits unverified claims about an election, statements that significantly misrepresent authoritative civic information, and unfounded conspiracy theories and claims that certain events or situations are carried out by covert or powerful groups, such as "the government" or a "secret society."[101]

Reddit

A spokesperson offered written comments, including that Reddit:

• banned the /r/The_Donald subreddit in June 2020, as we noted on page 13;

• employs safety teams to enforce its platform-wide Content Policy, while volunteer content moderators of particular subreddits "are free to set and enforce rules that go beyond our Content Policy";[102]

• prohibits any hate based on identity or vulnerability, as well as content that encourages, glorifies, incites, or calls for violence or physical harm against an individual or group of people;[103]

The spokesperson indicated that posts which included news articles fact-checked as true received more engagement and positive reaction on Reddit than posts with news labeled as false, according to researchers at Ohio State University.[104]

Twitter (X)   Gab   Parler
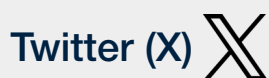
These companies did not respond to our inquiries. Twitter's content rules and enforcement are described here.[105] Gab's terms of service are described here.[106] And Parler's guidelines are described here.[107]

5. Conclusion and Recommendations

> "Rather than retreat from modest reforms made in the recent past, major social media companies should be intensifying efforts to protect against political threats and incitement to violence."

Mainstream technology companies cannot on their own arrest the increase in political intimidation and violence in the United States or anywhere else where this escalation is occurring. Political leaders, hyper-partisan cable television and radio outlets, social media sites that affirmatively encourage hateful incitement, and individuals who for a variety of reasons are drawn to extremist views all bear responsibility for the volatile state of politics in the U.S. and other parts of the world.

But as the work of social scientists makes clear, major tech companies need to do better. Rather than retreat from modest reforms made in the recent past, they should be intensifying efforts to protect against political threats and incitement to violence. In the U.S., the turbulence buffeting the election process makes it all the more important that social media companies take extra precautions, preparing for worst-possible scenarios in hopes that they can be ameliorated if they occur.

While the precise contribution that social media makes to political intimidation and violence is difficult to quantify, the technology does play a key facilitating role.

What follows are practical recommendations to the industry and government that could help reduce that contribution. We (and others) have proposed these and similar ideas for some years now. The terms of debate have shifted, but tangible improvement has been scant. Still, there is no alternative to continued advocacy for constructive change based on the hope for more enlightened corporate and governmental leadership.[108]

Recommendations

For the social media industry

1 | **Sound the alarm.**

To reduce risks, social media companies first need to end their tendency to deflect and obfuscate and instead acknowledge the role that their platforms play in facilitating political intimidation and violence. Bold public statements recognizing the phenomenon and accepting partial responsibility are the necessary precursor to meaningful action. Objections from the platforms' public relations and legal executives would be immediate and loud. They must be overcome. Another likely objection would reflect the fatalistic notion that if dangerous online actors are chased from mainstream platforms, they will simply migrate to free-for-all fringe sites, where they can incite and plot unmolested. Such migration does occur, as some of the research we have distilled demonstrates. But denying bad actors the reach and visibility available on mainstream platforms would still constitute significant progress.

2 | **Put more people on the content moderation beat.**

In 2022 and 2023, most major social media companies laid off "trust and safety" employees—the people who devise and enforce policies aimed at reducing online hatred and incitement.[109] This ill-advised retreat must be reversed. Artificial intelligence-driven content moderation systems have difficulty distinguishing between the ordinary hostility of political disagreement and dangerous extremism. In addition, a lot of extremist communication is cloaked in code language. Human judgment is crucial, and more humans, especially counter-extremism experts, are needed. Moreover, tech companies have tried to do most content moderation on the cheap by outsourcing the critical function to third-party vendors. Content oversight should primarily be done in-house, where employees are better compensated and supervised.

3 | **Directly confront election delegitimization.**

In the U.S. and certain other countries, political threats and actual confrontations often stem from irresponsible efforts to undermine trust in elections. Tech companies need to act aggressively to label and/or remove baseless allegations of election fraud and redirect users to authoritative sources of information. These companies ought to prepare for potential post-election attempts to undermine legitimate results, with plans in place to surge employees into election integrity efforts and adjust recommendation algorithms to prioritize responsibly gathered news, rather than conspiracy theories. Platforms should be particularly vigilant in filtering out threats against election workers and other public figures charged with overseeing voting and certifying the results.

4 | **Make design changes to mitigate harm.**

Social media companies can reform some of the features discussed in this report to reduce the likelihood that they will facilitate political intimidation and violence. Rather than allow anonymity, the companies should require users to verify their identity (with provisions for storing verification data securely and/or erasing it once it's no longer needed). Platforms should monitor groups for the prevalence of content advocating violence, regardless of partisan orientation. Invitations to, and recommendations of, volatile groups could be shut down, as could the groups themselves if they become dangerous. More broadly, recommendation systems should be redesigned to reduce, rather than heighten, sectarianism. Sheer user

engagement, which may reflect hateful and other sensationalistic posts, can be reduced as a criterion for amplification. Introducing "friction"—*Do you want to read that article before sharing it? Fact-checkers have determined that this claim lacks a factual basis*—has been shown to slow down the spread of some problematic content. "Circuit breakers" that briefly delay the viral spread of high-engagement posts while content reviewers consider potential harm would have a similar salutary effect.[110]

5 | Cultivate academic and civil society researchers.

A coordinated right-wing campaign combining legislative investigation and litigation has intimidated researchers specializing in online misinformation and extremism.[111] Social scientists and computational experts who previously had shared information and insights with industry have been bullied into retreat based on a made-up allegation that they were participating in a liberal conspiracy to silence conservative voices. Tech companies need to stand up to this brazen partisan intimidation and find ways to revive productive information exchanges in the interest of reducing harmful content and staving off potential political violence. Unlike most of the researchers, the companies have deep pockets, dedicated legal teams, and the wherewithal to deal with subpoenas and depositions—if they can muster the courage to do so.

For the U.S. Government

6 | Enforce existing laws.

With healthy respect for free speech protected by the First Amendment, the U.S. Departments of Justice and Homeland Security need to be vigilant about enforcing criminal laws banning political intimidation and the incitement of violence. Other executive branch agencies—primarily the Federal Trade Commission, Federal Election Commission, and their state counterparts—also must use their full authority to enforce existing laws against election fraud, voter suppression, cyberattacks, and other offenses relevant to protecting elections and preventing the erosion of democracy. All government agencies should be as transparent as possible about their communication with social media companies to preclude the appearance or reality that they are coercing private organizations and speakers.[112]

7 | Step up protection of election workers.

Of all the public officials subjected to political intimidation in recent years, election workers are the most vulnerable, as they frequently live and work without the benefit of the sort of protections provided to judges, lawmakers, and executive branch officials. To arrest the continued exodus of election workers, governments should raise the stakes for those who seek to intimidate these public servants by hardening existing penalties and introducing new ones that take into account the coordinated disinformation campaigns that lie behind the harassment. Given that real world violence often begins with online threats, prosecutions under such laws may ultimately serve as a deterrent to menacing behavior offline.[113]

8 | Enhance federal authority to oversee digital industries.

Longer term, the U.S. Congress needs to enhance the federal government's authority to regulate digital industries in a more systematic fashion. Our Center has made this recommendation for several years; the idea does not relate specifically to political intimidation and violence, but it would create incentives for social media companies to conduct themselves in a more constructive fashion. Previously, we have recommended expanding the consumer protection authority of the Federal Trade Commission to accomplish sustained oversight of digital industries. This approach would require additional funding, recruitment of technically adept personnel, and explicit Congressional authorization to ensure that major tech companies receive the sort of expert supervision that, for example, the Securities and Exchange Commission provides to the equity markets. An even more ambitious strategy would involve the creation of a new oversight agency with responsibility for social media, artificial intelligence, and other aspects of the digital sector. Two bills have been introduced in the Senate that would create a new digital commission with authority to oversee both competition and consumer protection.[114]

Another legislative strategy Congress should pursue would focus on regulating the sort of platform design features that, as we have shown, social science research has linked to political intimidation and violence. An example of such legislation is the Kids Online Safety and Privacy Act, which passed with overwhelming support in the Senate and is now pending in the House of Representatives. While it targets harm to children and teens, the act provides a model for how design improvements could reduce the risk of physical violence.[115]

9 | Mandate more transparency.

Serious proposals for more vigorous regulation of digital companies begin with the need for greater disclosure of how these businesses make decisions. More transparency will provide better insight into why the technology sometimes goes awry and how to extend useful government oversight. These revelations will benefit the fight against political intimidation and violence as well as the defense of elections. Regardless of whether Congress can muster the will to enhance the authority of the FTC, lawmakers should broaden and deepen their field of vision by passing legislation resembling the Platform Accountability and Transparency Act, a bipartisan measure introduced in the Senate, and the Digital Services Oversight and Safety Act, a similar bill backed by Democrats in the House. The European Union's Digital Services Act also contains transparency provisions worth considering.[116] The U.S. Supreme Court's recent ruling in the consolidated *NetChoice* cases left open the possibility that cautiously crafted disclosure legislation can withstand First Amendment review.[117]

Endnotes

1 https://www.hachettebookgroup.com/titles/renee-diresta/invisible-rulers/9781541703377/?lens=publicaffairs

2 https://journals.sagepub.com/doi/abs/10.1177/1461444818825133

3 https://gnet-research.org/2024/03/04/why-not-bring-weapons-to-school-how-tiktoks-algorithms-contribute-to-a-culture-of-violence-in-brazilian-schools/

4 https://www.nytimes.com/2024/05/19/us/politics/political-violence.html; https://ag.ny.gov/sites/default/files/buffaloshooting-onlineplatformsreport.pdf; https://www.nytimes.com/2018/10/28/us/gab-robert-bowers-pittsburgh-synagogue-shootings.html; https://www.nytimes.com/2024/06/05/us/trump-retribution-justice.html; https://www.cnn.com/2022/12/03/politics/trump-constitution-truth-social/index.html; https://abcnews.go.com/Politics/trump-promises-free-jan-6-rioters-day-back/story?id=108050867

5 https://carnegieendowment.org/research/2022/09/five-strategies-to-support-us-democracy?lang=en

6 https://bhr.stern.nyu.edu/publication/fueling-the-fire-how-social-media-intensifies-u-s-political-polarization-and-what-can-be-done-about-it/

7 https://www.techpolicy.press/read-the-january-6-committee-social-media-report/

8 https://ag.ny.gov/sites/default/files/buffaloshooting-onlineplatformsreport.pdf; https://www.nytimes.com/2022/05/19/technology/mass-shootings-livestream-online.html

9 https://truthsocial.com/@realDonaldTrump/posts/112265455770130939; https://www.nytimes.com/2024/05/19/us/politics/political-violence.html; https://www.nytimes.com/2024/06/02/us/migrants-charities-shelters-threats.html?smid=nytcore-ios-share&referringSource=articleShare&s-grp=c-cb

10 https://www.nytimes.com/2023/10/10/technology/hamas-violent-videos-online.html; https://extremism.gwu.edu/sites/g/files/zaxdzs5746/files/EncryptedExtremism.pdf; https://www.isdglobal.org/wp-content/uploads/2022/06/Under-moderated-Unhinged-and-Ubiquitous-al-shabaab-and-islamic-state-networks-on-facebook.pdf

11 https://www.ohchr.org/en/hr-bodies/hrc/myanmar-ffm/index

12 https://www.nbcnews.com/news/world/hindu-extremists-india-escalate-rhetoric-calls-kill-muslims-rcna12450

13 https://www.washingtonpost.com/world/2022/10/29/facebook-tiktok-brazil-election-disinformation/; https://www.washingtonpost.com/world/2022/10/29/facebook-tiktok-brazil-election-disinformation/

14 https://link.springer.com/content/pdf/10.1057/s41284-020-00274-y.pdf; see also https://www.adl.org/resources/backgrounder/backgrounder-how-can-online-anonymity-affect-hate

15 https://www.rev.com/blog/transcripts/mark-zuckerberg-opening-statement-transcript-house-hearing-on-misinformation

16 https://www.washingtonpost.com/technology/2022/01/04/facebook-election-misinformation-capitol-riot/

17 https://nickclegg.medium.com/you-and-the-algorithm-it-takes-two-to-tango-7722b19aa1c2

18 https://www.science.org/doi/10.1126/science.abe1715

19 https://gnet-research.org/wp-content/uploads/2021/05/GNET-Report-Survey-On-The-Role-Of-Technology-In-Violent-Extremism.pdf

20 https://www.ijcv.org/index.php/ijcv/article/view/3809

21 https://www.pewresearch.org/short-reads/2024/02/02/5-facts-about-how-americans-use-facebook-two-decades-after-its-launch/

22 https://www.facebook.com/notes/393134628500376/

23 https://time.com/5197039/un-facebook-myanmar-rohingya-violence/; https://www.ungeneva.org/en/news-media/press-briefing/2024/05/un-geneva-press-briefing-5

24 https://online.ucpress.edu/as/article-abstract/59/1/177/25073/Myanmar-in-2018The-Rohingya-Crisis-Continues?redirectedFrom=fulltext

25 https://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=1197&context=scjilb

26 https://www.stern.nyu.edu/experience-stern/faculty-research/who-moderates-social-media-giants-call-end-outsourcing

27 https://about.fb.com/news/2018/11/myanmar-hria/

28 https://www.tandfonline.com/doi/full/10.1080/23808985.2023.2280824; https://bridge.georgetown.edu/research/factsheet-soldiers-of-odin/

29 https://www.facebook.com/help/135275340210354

30 https://www.washingtonpost.com/technology/2022/01/04/facebook-election-misinformation-capitol-riot/

31 https://www.tandfonline.com/doi/full/10.1080/1057610X.20201862895; https://www.bellingcat.com/news/2020/05/27/the-boogaloo-movement-is-not-what-you-think/

32 https://www.routledge.com/Researching-Far-Right-Move-ments-Ethics-Methodologies-and-Qualitative-Inquiries/Toscano/p/book/9780367581930

33 https://www.hachettebookgroup.com/titles/renee-diresta/invisible-rulers/9781541703377/?lens=publicaffairs; https://www.theguardian.com/us-news/2020/oct/15/qanon-vio-lence-crimes-timeline

34 https://www.nbcnews.com/tech/tech-news/six-months-after-christchurch-shootings-videos-attack-are-still-face-book-n1056691

35 https://link.springer.com/article/10.1007/s10551-020-04635-w

36 https://www.statista.com/statistics/272014/global-social-net-works-ranked-by-number-of-users/

37 https://www.justice.gov/opa/pr/proud-boys-leader-sentenced-22-years-prison-seditious-conspiracy-and-other-charges-related

38 https://www.tandfonline.com/doi/full/10.1080/17439884.2018.1544149

39 https://journals.sagepub.com/doi/abs/10.1177/1461444818825133

40 https://www.globalmediainsight.com/blog/youtube-users-statistics/: https://www.pewresearch.org/shortreads/2024/02/02/5-facts-about-how-americans-use-facebook-two-de-cades-after-its-launch/

41 https://www.wired.com/story/christchurch-shooter-you-tube-radicalization-extremism/; https://christchurchattack.royalcommission.nz/the-report/

42 https://www.nytimes.com/2022/06/01/technology/fringe-mainstream-social-media.html; https://ag.ny.gov/sites/default/files/buffaloshooting-onlineplatformsreport.pdf

43 https://www.degruyter.com/document/doi/10.1515/opis-2020-0007/html

44 https://www.science.org/doi/10.1126/sciadv.add8080; see also a study led by researchers at the University of Penn-sylvania, entitled, "Examining the consumption of radical content on YouTube," https://www.pnas.org/doi/10.1073/pnas.2313377121

45 https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

46 https://www.brookings.edu/articles/misunderstood-mechanics-how-ai-tiktok-and-the-liars-dividend-might-affect-the-2024-elections/

47 https://www.theverge.com/2022/6/15/23168887/face-book-discovery-engine-redesign-tiktok

48 https://www.vice.com/en/article/yw74gy/tiktok-neo-nazis-white-supremacy; https://www.wsj.com/articles/islam-ic-states-tiktok-posts-include-beheading-videos-11571855833

49 https://www.isdglobal.org/digital_dispatches/naztok-an-organized-neo-nazi-tiktok-network-is-getting-millions-of-views/

50 https://gnet-research.org/2024/03/04/why-not-bring-weapons-to-school-how-tiktoks-algorithms-contribute-to-a-culture-of-violence-in-brazilian-schools/

51 https://www.isdglobal.org/isd-publications/hatescape-an-in-depth-analysis-of-extremism-and-hate-speech-on-tiktok/; see also: https://gnet-research.org/2024/04/08/the-creeping-influ-ence-of-the-extreme-rights-meme-subculture-in-southeast-asias-tiktok-community/

52 https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/; https://www.pew-research.org/short-reads/2024/02/02/5-facts-about-how-americans-use-facebook-two-decades-after-its-launch/

53 https://www.hachettebookgroup.com/titles/renee-diresta/invisible-rulers/9781541703377/?lens=publicaffairs

54 https://www.cambridge.org/core/journals/american-politi-cal-science-review/article/abs/from-isolation-to-radicaliza-tion-antimuslim-hostility-and-support-for-isis-in-the-west/C11A754C706DB9F9CAD86D1486A9B97A

55 https://journals.sagepub.com/doi/abs/10.1177/1461444820945317

56 https://ijoc.org/index.php/ijoc/article/view/10076

57 bit.ly/3yd8lFc

58 https://journals.sagepub.com/doi/abs/10.1177/1461444820958123

59 https://www.nytimes.com/2019/06/26/us/politics/reddit-donald-trump-quarantined.html; https://www.npr.org/2020/06/29/884819923/reddit-bans-the_donald-forum-of-nearly-800-000-trump-fans-over-abusive-posts

60 https://www.washingtonpost.com/technology/2023/01/17/jan6-committee-report-social-media/; https://www.washing-tonpost.com/technology/2021/02/05/why-thedonald-moderator-left/

61 https://www.tandfonline.com/doi/full/10.1080/09546553.2022.2161373#.Y9DznWgNMEM.twitter

62 https://telegram.org/press; https://www.statista.com/statis-tics/1306022/whatsapp-global-unique-users/

63 https://gnet-research.org/2022/09/12/analysing-terrorgram-publications-a-new-digital-zine/

64 https://www.taylorfrancis.com/chapters/edit/
10.4324/9781003105251-16/praise-saints-graham-macklin

65 https://icct.nl/publication/ted-kaczynski-anti-technology-
radicalism-and-eco-fascism/

66 https://www.lse.ac.uk/media-and-communications/assets/
documents/research/projects/WhatsApp-Misinformation-
Report.pdf

67 https://www.researchgate.net/publication/331768800_On_
whatsapp_rumours_and_lynchings

68 https://www.degruyter.com/document/doi/1
0.1515/9783839446706/html?lang=en

69 We did not find sufficiently illuminating research on Donald
Trump's Truth Social to warrant a section on that platform,
which launched in 2022.

70 https://www.tandfonline.com/doi/full/10.1080/105761
0X.2020.1862895; https://dl.acm.org/doi/10.1145/3587094

71 https://www.degruyter.com/document/doi/10.1515/
9783839446706/html?lang=en; see also Greta Jassen's
chapter in https://rowman.com/ISBN/9781786614926/
Rise-of-the-Far-Right-Technologies-of-Recruitment-
and-Mobilization

72 https://nyupress.org/9780814706497/the-legitimization-
of-violence/

73 https://firstmonday.org/ojs/index.php/fm/article/
view/11574/11249

74 https://www.jstor.org/stable/26984798

75 https://onlinelibrary.wiley.com/doi/full/10.1111/nana.12780

76 https://journals.sagepub.com/doi/
full/10.1177/14614448231172963

77 https://ieeexplore.ieee.org/document/9771124

78 https://polarisproject.org/wp-content/uploads/2021/02/
Polaris-Report-Countering-QAnon.pdf

79 https://www.frontiersin.org/journals/sociology/arti-
cles/10.3389/fsoc.2022.876070/full

80 https://ojs.aaai.org/index.php/ICWSM/article/
view/31409/33569

81 https://ojs.aaai.org/index.php/ICWSM/article/view/18117

82 https://jmiw.uitm.edu.my/images/Journal/Vol14No2/6-A-
Study-on-Donald-Trump-Twitter-Remark-A-Case-Study-on-
the-Attack-of-Capital-Hill.pdf

83 https://www.taylorfrancis.com/chapters/edit/
10.4324/9781003409427-1/2020-presidential-election-
social-media-trump-john-allen-hendricks-dan-schill

84 https://psycnet.apa.org/record/2022-58538-001

85 https://www.aiac.org.au/journals/index.php/alls/article/
view/7151

86 https://www.washingtonpost.com/world/2021/10/27/po-
land-facebook-algorithm/

87 https://uscneelycenter.substack.com/p/introducing-the
-neely-center-design

88 https://d3qi0qp55mx5f5.cloudfront.net/cpost/i/docs/
Pape_AmericanInsurrectionistMovement_2022-01-02.pdf?
mtime=1641247264

89 https://escholarship.org/uc/item/9x20f636

90 https://www.proquest.com/openview/8f00b09a809faeaa2d-
60d0c5c11761e1/1

91 https://carnegieendowment.org/research/2023/09/po-
larization-democracy-and-political-violence-in-the-unit-
ed-states-what-the-research-says?lang=en

92 https://humanrights.fb.com/wp-content/uploads/
2023/09/2022-Meta-Human-Rights-Report.pdf

93 https://transparency.meta.com/policies/community-
standards/

94 https://about.fb.com/news/2023/11/how-meta-is-planning-
for-elections-in-2024/

95 https://www.washingtonpost.com/technology/2024/02/10/
politics-meta-threads-instagram/

96 https://support.google.com/youtube/answer/9288567?s-
jid=301535242468677810-NA

97 https://support.google.com/youtube/answer/
10047535?hl=en-GB&sjid=11929519135582130843-NC

98 https://health.ucdavis.edu/news/headlines/new-study-
looks-at-attitudes-towards-political-violence/2023/09

99 https://www.pewresearch.org/short-reads/2024/01/31/most-
americans-say-elected-officials-should-avoid-heated-or-
aggressive-speech/

100 https://www.tiktok.com/community-guidelines/en/
safety-civility

101 https://www.tiktok.com/community-guidelines/en/
fyf-standards

102 https://www.redditinc.com/policies/content-policy

103 https://support.reddithelp.com/hc/en-us/articles/360045715951-Promoting-Hate-Based-on-Identity-or-Vulnerability; https://support.reddithelp.com/hc/en-us/articles/360043513151-Do-not-post-violent-content

104 https://news.osu.edu/true-stories-can-win-out-on-social-media-study-finds/

105 https://help.x.com/en/rules-and-policies/enforcement-options

106 https://gab.com/about/tos

107 https://help.parler.com/hc/90853520/40/community-guidelines-user-commitment-and-content-standards?category_id=33

108 https://bhr.stern.nyu.edu/publication/digital-risks-to-the-2024-elections-safeguarding-democracy-in-the-era-of-disinformation/; https://bhr.stern.nyu.edu/publication/spreading-the-big-lie-how-social-media-sites-have-amplified-false-claims-of-u-s-election-fraud/; see also https://assets.pubpub.org/l2w7s191/Preventing_Tech-Fueled_Political_Violence-01716308264401.pdf

109 https://www.techpolicy.press/the-unbearably-high-cost-of-cutting-trust-safety-corners/

110 Renèe DiResta's new book, _Invisible Rulers: The People Who Turn Lies Into Reality_, summarizes a number of worthy design improvements in a chapter entitled, "The Path Forward." See also the Conclusion and Recommendations section of our Center's February 2024 report, _Digital Risks to the 2024 Elections_.

111 https://www.washingtonpost.com/technology/2024/07/14/trump-allies-disarm-misinformation-researchers-ahead-election/; https://www.science.org/doi/10.1126/science.adr3820; https://www.nytimes.com/2024/06/25/opinion/stanford-disinformation-election-jordan-twitter.html

112 https://www.justsecurity.org/97268/murthy-missouri-social-media-disinformation/

113 https://www.brennancenter.org/our-work/analysis-opinion/poll-election-officials-finds-concerns-about-safety-political; https://www.citizen.org/article/tracker-state-legislation-to-protect-election-officials/

114 Our earlier proposal: https://bhr.stern.nyu.edu/ftc-whitepaper. The two Senate bills: https://www.bennet.senate.gov/public/index.cfm/2023/5/bennet-welch-reintroduce-landmark-legislation-to-establish-federal-commission-to-oversee-digital-platforms; and https://www.warren.senate.gov/newsroom/press-releases/warren-graham-unveil-bipartisan-bill-to-rein-in-big-tech.

115 https://www.techpolicy.press/child-online-safety-law-clears-the-us-senate-but-its-fate-in-the-house-remains-unclear/

116 https://www.coons.senate.gov/news/press-releases/senator-coons-colleagues-introduce-legislation-to-provide-public-with-transparency-of-social-media-platforms; https://trahan.house.gov/news/documentsingle.aspx?DocumentID=2389; https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en

117 https://www.justsecurity.org/97660/netchoice-content-moderation/