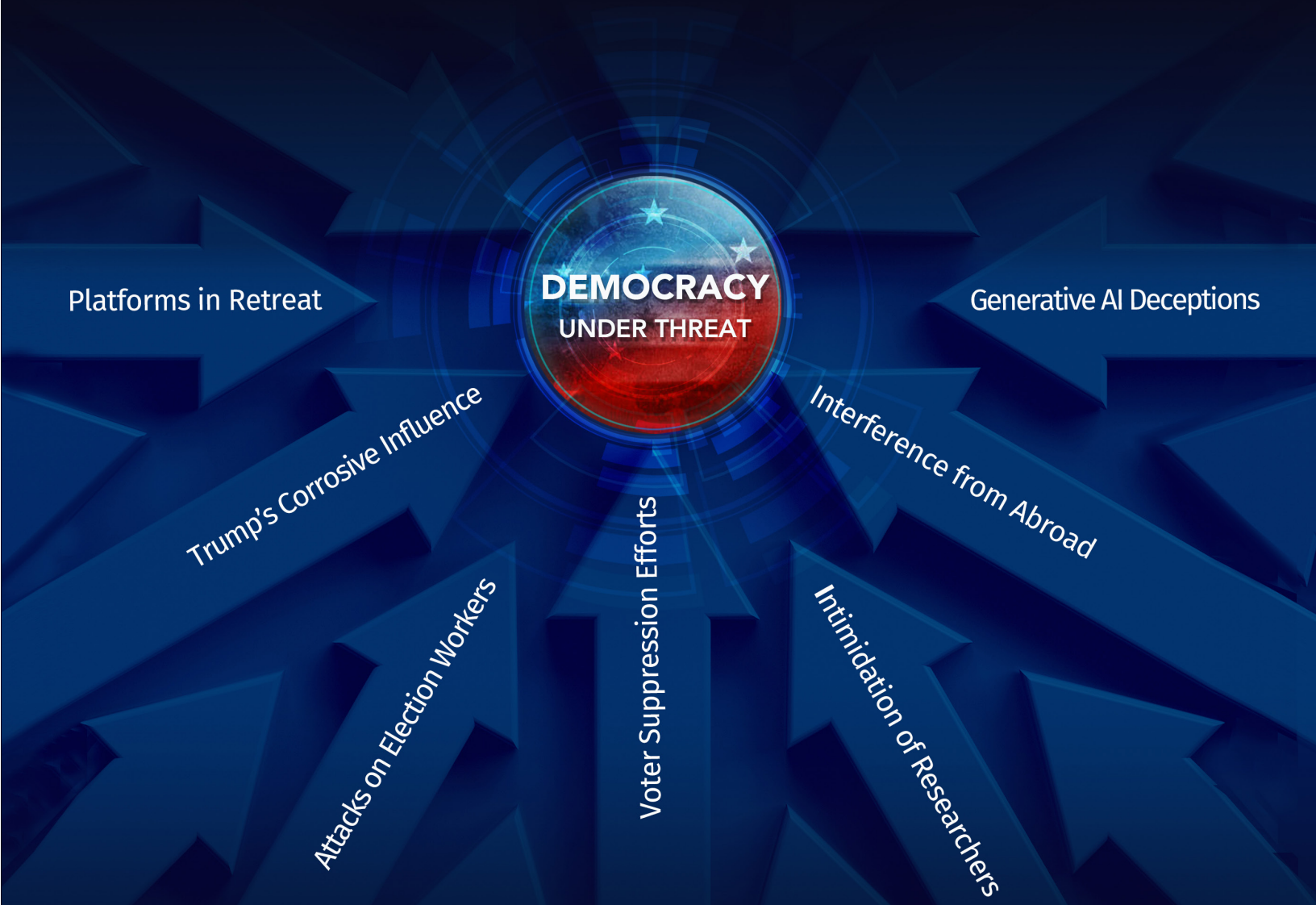
Digital Risks to the 2024 Elections:

Safeguarding Democracy in an Era of Disinformation

PAUL M. BARRETT, JUSTIN HENDRIX, AND CECELY RICHARD-CARVAJAL



NYU STERN
Center for Business and Human Rights

February 2024

Contents

Authors

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

Justin Hendrix is an associate research scientist and adjunct professor at the NYU Tandon School of Engineering and the CEO and editor of Tech Policy Press, a nonprofit media venture concerned with the intersection of technology and democracy.

Cecely Richard-Carvajal is a Masiyiwa-Bernstein Fellow at the NYU Stern Center.

Prithvi Iyer, a program manager with Tech Policy Press, and Smita Samanta, a graduate student research assistant with the NYU Stern Center, did valuable research.

Acknowledgments

Our thanks go to Craig Newmark Philanthropies and the Open Society Foundations for their continued support of our work on technology and democracy.

Executive Summary

A record two billion people are expected to vote in elections in more than 50 countries in 2024. But this impressive display of democracy faces a number of potential disruptions related to technology.

> **The leading tech-related threat to this year's elections stems not from the *creation* of content with AI but from a more familiar source: the *distribution* of false, hateful, and violent content via social media platforms.**

Since late 2022, popular and corporate attention has focused on advances in generative artificial intelligence and apps such as OpenAI's ChatGPT, which are built with large language models. New forms of AI create risks to elections, including the more efficient production of misinformation and "deepfake" imagery.

The leading tech-related threat to this year's elections, however, stems not from the *creation* of content with AI but from a more familiar source: the *distribution* of false, hateful, and violent content via social media platforms. Despite the disruptions and violence that roiled the U.S. presidential election in 2020 and Brazil's election in 2022, major platform companies like Meta (parent of Facebook, Instagram, and WhatsApp); Google (YouTube); and X, formerly known as Twitter, have retreated from some of their past commitments to promote election integrity. TikTok raises distinctive concerns, in part because its parent company, ByteDance, is Chinese and in part because of the design of its potent content-recommendation algorithm.

In the U.S., the degradation of political discourse—a process spurred primarily by former President Donald Trump, who is the likely 2024 Republican nominee, and some of his supporters—has created an online environment in which other deleterious phenomena unfold. These include harassment and threats aimed at public officials, including state and local election workers; conservative attempts to discourage registered Democratic voters from going to the polls; and Republican attacks on university and civil society researchers who study misinformation and played a constructive role in 2020 by identifying falsehoods and conspiracy theories.

Other digital threats come from abroad: China, Russia, and Iran are likely to attempt to take advantage of political polarization and sow confusion as Americans prepare to go to the polls.

Beyond the upcoming U.S. elections, this report looks at contests in a number of other places, including South Korea, Mexico, and the European Union. In India, Hindu nationalists are exploiting the WhatsApp messaging platform to incite violent hostility against Muslims and rally support for the ruling Bharatiya Janata Party in anticipation of national elections in April and May 2024.

The report concludes with practical recommendations aimed at technology companies and governments. We summarize the recommendations here:

Recommendations to Technology Companies

1. **Add many more humans to the content moderation loop.** For years, social media companies have done content moderation on the cheap, outsourcing this crucial corporate function to low-cost vendors operating in low-wage locations. To improve effectiveness, platform companies need to bring human moderation in-house and hire many more reviewers.

2. **Fund more outside fact-checkers to prowl for falsehoods.** While companies hire more in-house moderators to enforce platform policies, they should also fund more outside fact-checking organizations capable of separating reality from fiction.

3. **Blunt election delegitimization and prepare for a crisis.** As the U.S. presidential campaign unfolds, Donald Trump is likely to step up efforts to undercut the legitimacy of the 2024 elections by claiming that his loss in 2020 was "rigged." Platforms should label such claims as false or remove them altogether and direct users to authoritative, nonpartisan information.

4. **Make platform design and policy adjustments for election season, and beyond.** Social media companies need to scrutinize whether rampant resharing of content tends to promote misinformation, institute "circuit breakers" to slow the spread of certain "viral" posts, impose "interstitial" warnings about problematic content, and remove provably false material from feeds.

5. **Mitigate the risks of generative artificial intelligence.** Companies that design and market generative AI systems should label all AI content, seek more effective methods to detect deceptive AI content, and create a clearinghouse for collaborative identification of misleading AI video, audio, and text.

Recommendations to Government

6. **Enforce existing laws as they apply to digital industries.** In the U.S., executive branch agencies need to use their full authority to enforce existing laws against election fraud, voter suppression, cyberattacks, and other offenses potentially relevant to protecting elections.

7. **Strengthen legal protections for election workers.** Governments should raise the stakes for those who seek to intimidate election workers by hardening existing penalties and introducing new ones that take into account the coordinated, networked disinformation campaigns that have become the norm.

8. **Enhance federal authority to oversee digital industries.** Longer term, Congress should enhance the consumer protection authority of the Federal Trade Commission to regulate digital industries in a more systematic fashion. This idea does not relate specifically to elections, but it would create incentives for improved conduct, benefiting democracy.

9. **Mandate more transparency.** Greater disclosure of how digital businesses make decisions will provide more insight into why technology sometimes harms democracy and how to extend useful government oversight.

10. **Bolster public sector and academic research capacity.** Lawmakers should shore up public sector and academic capacity to study technology's effects on democracy by appropriating additional funds and providing authority to federal agencies to foster robust research on the effects of technology on elections.

1. Introduction

> The social media industry has retreated from safeguarding elections from the manipulation, hatred, and conspiracy mongering that all too often now characterize online political discourse.

Looking ahead to the digital risks facing the 2024 elections in the United States and around the world, it is tempting to fixate on generative artificial intelligence. Since late 2022, the release of ChatGPT and other AI systems that can generate text, imagery, and audio based on simple written prompts has captured enormous popular and media attention. Significant potential dangers accompany the rise of generative AI. It could, for example, make political disinformation, including "deepfake" imagery, easier to produce and more believable.[1]

But the leading technology-related risk facing the coming elections stems not from new ways that bad actors may use AI to *create* harmful online content. It comes instead from well-established methods of *distributing* deleterious digital material—namely, via major social media platforms like Facebook, Instagram, YouTube, TikTok, and X (formerly known as Twitter).

In the U.S., one reasonably would expect that the 2020 election crisis, culminating in the Capitol riot on January 6, 2021, would prompt social media companies to redouble their efforts to protect the democratic process in 2024. After all, Republicans led by former President Donald Trump exploited the platforms to undermine popular trust in the 2020 election and then incite the violent attempt to stop Congress from certifying President Biden's victory. But the opposite has occurred. The social media industry has retreated from safeguarding elections from the manipulation, hatred, and conspiracy mongering that all too often now characterize online political discourse.

The reasons for this retreat vary from company to company and range from economics to ideology to a form of collective fatalism, according to interviews with current and former industry executives and employees. The main factor, by far, is that after years of rapid growth and over-hiring, Silicon Valley experienced an advertising revenue decline leading to severe layoffs in 2022 and 2023; at some companies, the firings are continuing in 2024. Corporate teams responsible for "trust and safety," including those devoted to "election integrity," have suffered.[2]

X presents a special case. Under Elon Musk's volatile ownership, the financially unstable company has

slashed its ranks more deeply than any other, eliminating much of its elections team. Musk has justified these moves, in part, as reflecting his devotion to free speech.[3] The practical result has been a resurgence of racism, antisemitism, and other forms of hatred. Without going nearly as far as Musk, other companies have used the deterioration at X to rationalize diminished vigilance, Yoel Roth, the former head of trust and safety at X, said in an interview. "Why was it a strategically simple choice, not just at [X] but at every one of these companies?" he added. "It's because there is no air cover for them to maintain this [elections] work. The only thing they experienced is the sustained pressure to cut it back."

Seven digital risks

This report examines seven technology-related risks to elections in 2024. We start with the two mentioned above: **platform backsliding** and **nascent threats posed by generative AI**.

A third risk, different in its nature, is the continued **degradation of the larger online environment**—a trend driven primarily, although not exclusively, by former President Trump and some of his supporters. With heightening threats of "retribution" against political foes, scorn for the rule of law and elections, and ominous echoes of authoritarian rhetoric, Trump and his backers are setting the stage for a political clash that could devolve into chaos and violence. While there are extreme voices on the political left, the sustained viciousness from the right is asymmetric in its extent and potential for harm.

The corrosiveness of the digital information environment enables other malign tendencies. A fourth risk is that since 2020, public officials generally and **state and local election workers in particular have become targets** of harassment and violent threats. Election workers have resigned in droves, leading to a loss of institutional knowledge and a greater likelihood of problems at the polls, which, in turn, could feed more conspiracy theories.

An adjacent fifth risk with roots in the digital fever swamp is the Republican **campaign to suppress the Democratic vote** by means of purging voters as fraudulently registered. These efforts by Trump-affiliated conservative activists are already underway.

Yet another outgrowth of right-wing information distortion, and our sixth risk, is an **ongoing attack on academic and civil society researchers** who played a constructive role in 2020 in identifying online misinformation narratives. One line of assault has come from Republican members of Congress led by Representative Jim Jordan of Ohio, who has orchestrated an investigation based on the baseless idea that a collection of researchers, social media company officials, and Democratic operatives conspired to censor right-leaning views. On a related front, Republican attorneys general from Missouri and Louisiana have sued the Biden administration for its role in this supposed plot. Known as Murthy v. Missouri, the high-profile case is expected to be decided by the Supreme Court just as the presidential campaign heats up in late spring.

We also assess the **risks posed to the U.S. elections by foreign nation states**—in particular, China, Russia, and Iran, each of which has a stake in who controls the White House and Congress beginning in January 2025.

In a separate section, we look beyond the United States to identify tech-related problems in India, Mexico, South Korea, and the European Union. As is our practice, we conclude the report with recommendations to technology companies and governments.

"Election workers have resigned in droves, leading to a loss of institutional knowledge and a greater likelihood of problems at the polls, which, in turn, could feed more conspiracy theories."

News Online: True Or False?

Results from an NYU Stern Center survey of likely U.S. voters

Voters get *most* of their news about politics and elections from these sources.

Social media	33.8%
Cable TV	21.1%
Broadcast TV	19.5%
Newspapers	6%
Websites, podcasts & other	19.6%

Most voters believe they can identify false news.

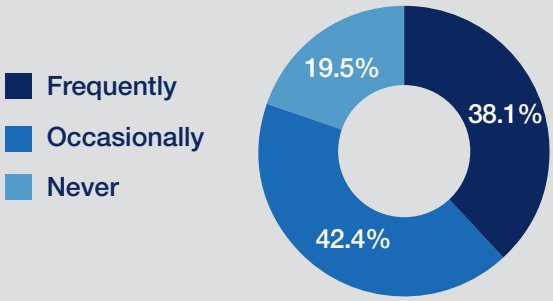
On traditional media

89.2% are "somewhat confident" or "very confident" they can pick out false news

On social media

87.7% are "somewhat confident" or "very confident" they can detect false news

How often do voters try to verify the truth of news hosted by social media?



Where do voters check the truth of this news?

Search engines	76.8%
Fact-checking organizations	32.2%
Messaging services	12.5%
Generative AI apps	10.5%

Voters believe that both traditional and social media increase divisiveness.

How does traditional media affect political divisiveness?

Increases it	Decreases it	Has no impact
68%	11.4%	20.6%

How does social media affect political divisiveness?

Increases it	Decreases it	Has no impact
78%	10%	11.9%

Source: NYU Stern Center for Business and Human Rights survey conducted by Dynata in December 2023 of 1,104 likely U.S. voters, 41% Democratic or lean Democratic, 41% Republican or lean Republican, and 18% independent

2. Seven Digital Risks to U.S. Elections in 2024

> **'We're seeing in many ways teams that were built up following public outrage in 2017 [over Russian interference] now get pruned back. We clawed these roles into existence because they needed to be there, because it was impossible to deliver on election-security work without each and every one of these roles…. The elimination of these roles topples load-bearing infrastructure within these companies.'**
>
> —Yoel Roth, former head of trust and safety at Twitter (now X)

Election year 2024 constitutes an historic global challenge, as some two billion people are expected to vote in more than 50 countries around the world.

In the U.S., the presidential campaign will unfold eight years after Russian operatives exploited Facebook, Instagram, YouTube, and Twitter (now X) to heighten divisiveness over race, religion and immigration to try to boost the fortunes of then-candidate Donald Trump. Since then, domestically generated misinformation and manipulation have come to play a more prominent role than foreign interference. But the need for Silicon Valley companies to protect against online distortion and hatred remains every bit as urgent.

Rather than stepping up their efforts to protect elections, however, major social media companies have pulled back since 2020-2021, shrinking teams devoted to civic and election integrity and making subtle policy shifts that signal diminished attention to these concerns.

1. Platforms in retreat

Any discussion of platform retreat necessarily begins with X, which tech mogul Elon Musk acquired in October 2022. By going to extremes, Musk made more modest retrenchment at rival platforms seem reasonable, or even inevitable, when, in fact, the industry should have been building up its capacity for promoting trust in elections.

By spring 2023, Musk had fired 6,000, or 80%, of X's employees.[4] The company's election integrity team shrank drastically. In September 2023, Musk said of the elections unit, "Yeah, they're gone."[5] The new owner also scaled back content moderation more broadly, contributing to a surge of racist and antisemitic expression that prompted advertisers to quit X in droves.[6] Musk says he is acting on principle: "Free speech is the bedrock of democracy," he has posted on X.[7] In December 2023, the European Union opened a formal investigation of X, the first under a newly enacted E.U. law called the Digital Services Act. The probe focuses, in part, on how X deals with hate speech, disinformation, and terrorist content related to the Israel-Gaza war.[8]

Yoel Roth resigned as X's head of trust and safety in the fall of 2022, undergoing the bizarre and frightening experience of having the new owner, Musk, publicly accuse him of having tried to silence conservative voices. What Roth actually had done was oversee a cautious response to then-President Trump's baseless allegations of a "rigged" election, mainly by labeling some of Trump's incendiary tweets and preventing them from being

liked or retweeted. When Trump incited the January 6, 2021, attack on the Capitol and then used social media to praise the mob while it was still marauding through the halls of Congress, Roth participated in the decision to block the then-president's Twitter account. Musk's unsubstantiated smear of Roth, combined with separate attacks by Trump, resulted in the former executive receiving death threats that prompted him to move his family from their home.[9]

Today, Roth advises other social media companies as a consultant. He said in an interview that he has observed first-hand the degradation of election integrity efforts spread from X to rival platforms. "We're seeing in many ways teams that were built up following public outrage in 2017 [over Russian interference] now get pruned back," he said. "We clawed these roles into existence because they needed to be there, because it was impossible to deliver on election-security work without each and every one of these roles." He added: "The elimination of these roles topples load-bearing infrastructure within these companies."

Musk's attempt to discredit Roth was part of a broader campaign that also featured the strategic release of internal company communications that came to be called the "Twitter Files." In February 2023, Republicans in Congress, echoing Musk, held public hearings on the Twitter Files and the unsupported claim of a Silicon Valley plot to censor conservatives.[10] This partisan activity coincided with economically driven layoffs of tens of thousands of tech workers across the industry in response to deteriorating advertising revenue. The confluence of these developments led most social media companies to roll back election integrity efforts as they tried simultaneously to slash spending and avoid right-wing political ire.

Anika Collier Navaroli agreed with her former Twitter colleague Roth: "I see the sort of preparedness of the digital information ecosystem [for the 2024 elections] essentially being pre-2016, before anybody staffed up, before anybody believed bad things could happen," Navaroli told us in an interview. Navaroli worked as a senior content policy expert at Twitter from 2019 through 2021 and then at the Amazon-owned live-streaming platform Twitch until 2022. "We're going into literally the most historic election year in the world, and the team that does that thing no longer exists after finally being able to staff up, finally being able to maybe do half capacity," said Navaroli, who is now a senior fellow at the Tow Center for Digital Journalism at Columbia Journalism School. "And again, I see the sort of ripple effect throughout the industry."

Separately, a senior executive still at X who worked there for years before Musk's arrival acknowledged in an interview that the platform has stepped back from policing election-related misinformation. "We don't want to make those decisions," the executive said, adding that other platforms likely would follow suit in 2024. At the same time, this person said that X adequately polices harmful content of all sorts by relying on a "crowdsourcing" program called Community Notes, which allows certain X users to comment on posts that might be misleading. Musk has said that Community Notes "has incredible potential for improving information accuracy."[11] X has said that its in-house research shows that people who see Community Notes are less likely to reshare misinformation.[12] But outside analysts have pointed out that the program makes only a small percentage of user comments public, limiting its effectiveness, and that Community Notes itself is vulnerable to manipulation by partisan participants.[13]

In the run-up to, and immediate aftermath of, the 2020 election, all of the major platforms responded to efforts by then-President Trump and some of his supporters to undermine the results. Most dramatically, Facebook, YouTube, and Twitter suspended Trump's accounts in January 2021, citing the danger that he would use the platforms to incite more violence and block the peaceful transition of presidential power. At Meta, founder and chief executive Mark Zuckerberg had built an elections team of more than 300 employees and met regularly with its leadership. But by the November 2022 U.S. mid-term elections, Meta had reversed course. Amid company-wide layoffs totaling 21,000 employees, or nearly a quarter of Meta's workforce, the elections team was reduced to 60 people, and Zuckerberg stopped regularly meeting with its leaders. Today, the company maintains that many employees who don't have "elections" in their job title nevertheless do important work on the topic, and top management receives regular updates.[14]

> **"By mid-2023, whatever progress had been made on elections had unraveled, 'so much so that the current state of platform content moderation is more like 2016 than 2020.'**
>
> —Media scholars Daniel Kreiss and Bridget Barrett

Media scholars Daniel Kreiss and Bridget Barrett observed recently that by mid-2023, whatever progress had been made on elections had unraveled, "so much so that the current state of platform content moderation is more like 2016 than 2020."[15] Kreiss, an associate professor at the University of North Carolina, and Barrett, an assistant professor at the University of Colorado, noted in Tech Policy Press that Meta, YouTube, and X had all reinstated Trump's accounts even though "the former president continues to spout lies about the

2020 election and is actively working to undermine confidence in the next one." Kreiss and Barrett pointed to research published in 2021 that found that "deplatforming" by a major social media company can limit the reach of disinformation and extreme speech, even if some problematic content migrates to fringe platforms.[16]

Beyond reinstating Trump, Meta has said that, in the interest of promoting free speech, it will loosen its rules to allow political ads on Facebook and Instagram to question the legitimacy of the 2020 presidential election. Sure enough, in August 2023, Trump ran a campaign spot on Facebook declaring: "We won in 2016. We had a rigged election in 2020 but got more votes than any sitting president."[17] Meta also continues to exempt politicians from its extensive outside fact-checking operation, opening the door for influential public officials and candidates to lie about election fraud and other topics. YouTube, meanwhile, has announced that it is rescinding a policy under

TikTok's Potential Misinformation Problem

TikTok is different from the other major U.S.-based social media platforms. Its parent, ByteDance, is a giant Chinese company, which has raised questions about the autocratic Chinese government's potential influence over TikTok (see discussion on page 15). In addition, some of TikTok's characteristics are different in ways that could result in election misinformation reaching a more receptive audience, according to some experts.

TikTok's famously effective recommendation algorithm—the software that selects what content to present to users—is distinct from that of rival platforms. X, Facebook, and Instagram are structured around a "social graph," which selects content for users based on what is shared in their personal network: the people whom they follow and who follow them.

TikTok, by contrast, selects short videos for its "For You" page based on algorithmic recommendations of content from outside of their social network, as well. This difference may help explain why TikTok is so successful at serving up videos that users find novel and compelling. But it could present a danger during election season, according to researchers with NYU's Center for Social Media and Politics (CSMaP): "With generative AI making fabricated videos easier to produce," the researchers wrote in January 2024, "we could see political misinformation reaching users on TikTok that it wouldn't reach on other social graph-based platforms."[1] What's more, TikTok users are younger, and studies show that young people are more likely to believe misinformation.[2] Without calling it a certainty, the CSMaP team observed that this combination of platform characteristics could "potentially make misinformation more effective on this platform."

Another difference is that, at least according to one investigation, TikTok is markedly less effective in detecting false U.S. election advertising. Global Witness, a progressive nonprofit advocacy group, and Cybersecurity for Democracy, a research team at NYU's Tandon School of Engineering, reported in October 2022 that TikTok fared much worse than Facebook and YouTube when investigators presented the platforms with misinformation about the November 2022 midterms that could stop people from voting. The misinformation included incorrect guidance on when and where to vote, encouragement to vote twice, and messages designed to undermine voting by mail. Even though TikTok has an announced policy of not accepting any political advertising, the platform approved 18 out of 20—or 90%—of the blatantly false ads, according to the study sponsors (which said they immediately deleted the dummy ads so that they never actually received public attention). Facebook approved seven out of 20, or 35%; YouTube rejected all of the false ads and banned the dummy YouTube channel set up to host the ads.[3]

Asked for comment, a TikTok spokesperson said: "The CSMaP comment about misinformation being more effective on our platform is entirely speculative and betrays a poor understanding of how our platform works, how we enforce our rules, and the steps we take to reduce the reach of misinformation." The political ads study from 2022 "is outdated," the spokesperson said. "We regularly review and improve our policies and processes in order to further strengthen our systems." To identify and remove prohibited political advertising, the company has "added hundreds of new detection keywords" and "invested in dedicated teams" that monitor for this content, the spokesperson added.

1 https://www.brookings.edu/articles/misunderstood-mechanics-how-ai-tiktok-and-the-liars-dividend-might-affect-the-2024-elections/
2 https://www.kateto.net/covid19/COVID19 CONSORTIUM REPORT 14 MISINFO SEP 2020.pdf; https://www.theguardian.com/us-news/2023/aug/16/teens-online-conspiracies-study
3 https://www.globalwitness.org/en/campaigns/digital-threats/tiktok-and-facebook-fail-detect-election-disinformation-us-while-youtube-succeeds/

which it removed tens of thousands of videos claiming that the 2020 election was illegitimate.[18] (In a contradictory signal, the company continues to "demonetize" content making this false claim, meaning that creators cannot earn ad revenue from such videos.)

"Discrediting previous elections sets the stage for discrediting future ones," Kreiss and Barrett argued. "Trump's election denialism, coupled with his rhetoric that his political opponents are [as he said in a recent Facebook post] out to 'destroy America,' is a clear and present threat to U.S. democracy and a likely catalyst for violence."

2. Generative AI deceptions

With the capacity to produce human-sounding prose, as well as realistic imagery and audio—all based on simple written prompts—generative AI systems offer potent tools to people, groups, or countries seeking to disrupt elections. Generative AI technology has made it easier, for example, to create convincing "deepfakes," which are videos or photographs that make subjects appear to do or say things they never did or said.

The former presidential campaign of Florida Governor Ron DeSantis disseminated phony images in June 2023 of Donald Trump embracing Dr. Anthony Fauci, the former U.S. official who steered COVID-19 policy and is reviled by many Trump backers. Two months earlier, the Republican National Committee promoted a YouTube video it disclosed had been generated by AI that showed an imagined dystopia following a successful Joe Biden re-election campaign. And a consortium of responsible-AI advocates garnered some attention for posting a fake Trump-Biden debate on Twitch, the video-game-streaming platform owned by Amazon.[19]

Some observers have been underwhelmed. In August 2023, *The Economist* published a long article arguing that American voters are too savvy, or maybe too jaded, to be fooled by AI fakery. Brendan Nyhan, a political scientist at Dartmouth College, told the magazine: "We still have not one convincing case of a deepfake making any difference whatsoever in politics."[20]

Of course, there hasn't been a national election in the U.S. since the Silicon Valley start-up OpenAI publicly released its generative AI app, ChatGPT, in late November 2022. The incentive to deploy the technology in a disruptive fashion will be far greater in 2024.

Skeptics "fail to understand the trend of the technology [by] looking at where we are today, not where we are going," Hany Farid, a computer scientist at the University of California at Berkeley who specializes in digital forensics, said in an interview. Farid pointed to the wave of "viral fake images that have emerged from the conflict in the Middle East, something we did not see during the Ukraine conflict just two years ago." As for the notion that social media users can readily identify deepfakes, Farid's recent empirical research has found that AI-synthesized faces are now often indistinguishable from real faces—and strike many viewers as particularly "trustworthy."[21]

Generative AI designers and marketers are implicitly conceding that their products can be used to undermine elections—and, to their credit, are beginning to impose restrictions. OpenAI, maker of the ChatGPT chatbot and DALL-E image generator, announced in January 2024 that it won't allow use of its technology to build apps for political campaigns or lobbying, to discourage people from voting, or to spread falsehoods about the voting process. Shortly thereafter, OpenAI suspended the account of a developer that had created a chatbot imitating Dean Phillips meant to boost his long-shot Democratic presidential aspirations.[22] In December, Google said it would restrict the answers its AI products give to questions about elections.[23] Meta requires political advertisers on its platforms to disclose if they use AI.[24]

Tricking voters about a particular politician or voting procedure is not the only way generative AI could undermine elections. The technology may already be contributing to a general erosion of trust in information about politics and public affairs, which diminishes the legitimacy of election results and may influence people not to vote at all. Donald Trump has tried to discredit ads attacking him with old footage of his gaffes by falsely claiming that the clips were generated by AI. Amazon's Alexa AI voice assistant reportedly has falsely declared that the 2020 presidential election was stolen—a problem Amazon said it has fixed.[25]

"Tricking voters about a particular politician or voting procedure is not the only way generative AI could undermine elections. The technology may already be contributing to a general erosion of trust in information about politics and public affairs."

Between May and December 2023, the number of websites hosting AI-created false articles increased by more than 1,000%, to more than 600 from 49, according to NewsGuard, which tracks misinformation. Presented as news sites, these outlets—with names like iBusiness Day and Ireland Top News—often juxtapose real reports with AI-generated fakes, sometimes seeking to make political points and other times to use polarizing content to draw clicks and advertising revenue.[26]

Lawrence Lessig, an expert on the internet and democracy who teaches at Harvard Law School, has used medical imagery when describing the threat of misinformation generated by AI: "They will just spread pathogens throughout the system, and we have no antibodies for them."[27]

In Their Own Words

Social media companies describe their plans for protecting the 2024 elections.

X

Enforcement via labeling

The company bans "posting or sharing content that may suppress participation, mislead people about when, where, or how to participate in a civic process, or lead to offline violence during an election. Any attempt to undermine the integrity of civic participation undermines our core tenets of freedom of expression and as a result, we will apply labels to violative posts informing users that the content is misleading."[1]

Meta

Remaining 'broadly consistent' with past practices

"We continually review and update our election-related policies, and take action if content violates our Community Standards, including our policies on election and voter interference, hate speech, coordinating harm and publicizing crime and bullying and harassment. We remove this content whether it was created by a person or AI." Starting in 2024, advertisers will have to "disclose when they use AI or other digital techniques to create or alter a political or social issue ad in certain cases."[2]

YouTube

Recommending 'authoritative sources'

"We will stop removing content that advances false claims that widespread fraud, errors, or glitches occurred in the 2020 and other past U.S. presidential elections….[But] we quickly remove content that incites violence, encourages hatred, promotes harmful conspiracy theories, or threatens election workers. At the same time, our systems recommend election news and information from authoritative sources and display information panels at the top of search results and below videos to provide even more context….[On AI:] Our misinformation policies prohibit technically manipulated content that misleads users and could pose a serious risk of egregious harm. And for election ads, we require advertisers to disclose when their ads include digitally altered or generated materials."[3]

TikTok

Collaborating with outside experts

"We invest in media literacy as a counter-misinformation strategy as well as technology and people to fight misinformation at scale. This includes specialized misinformation moderators with enhanced tools and training, and teams on the ground who partner with experts to prioritize local context and nuance. We partner with 17 global fact-checking organizations, who assess the accuracy of content in over 50 languages so that our moderators can apply our misinformation policies accordingly. We added three new global fact-checking partners in 2023, and will continue to expand our fact-checking program this year."[4]

[1] *https://help.twitter.com/en/rules-and-policies/election-integrity-policy*
[2] *https://about.fb.com/news/2023/11/how-meta-is-planning-for-elections-in-2024/*
[3] *https://blog.youtube/inside-youtube/us-election-misinformation-update-2023/; https://blog.youtube/inside-youtube/supporting-2024-united-states-election/*
[4] *https://newsroom.tiktok.com/en-us/protecting-election-integrity-in-2024*

3. Trump's corrosive influence

Since his first run for president began in earnest in 2015, Donald Trump has influenced the tone and content of digital political discourse in the U.S. more than any other single person. By lying habitually about topics large (e.g. legitimacy of the 2020 election) and small (square footage of his Trump Tower apartment), he has blurred the line between truth and falsehood with singular efficacy.

Research published in September 2023 in the academic journal *Perspectives on Politics* concluded that "Donald Trump's 'big lie'" about election fraud "is pervasive and sticky: the number of Republicans and independents saying that they believe the election was fraudulent is substantial, and this proportion did not change appreciably over time or shift after important political developments."[28] Journalists and fact-checking organizations struggle to keep up with the volume of Trump's falsehoods, as predicted by his former adviser Steve Bannon, now a prominent right-wing podcaster, who in 2018 famously told the writer Michael Lewis that the way to thwart inquisitive reporters is to "flood the zone with shit."[29]

Lately, Trump's rhetoric online and elsewhere has veered increasingly toward that of authoritarian leaders who dehumanized their foes by comparing them to detested animals. "We pledge to you that we will root out the communists, Marxists, fascists, and the radical left thugs that live like vermin within the confines of our country," Trump told a New Hampshire crowd in November 2023.[30] In an interview with the far-right website The National Pulse, he called immigration "a very sad thing for our country; it's poisoning the blood of our country."[31] Hitler, in his manifesto, Mein Kampf, warned that "all the great civilizations of the past became decadent because the originally creative race died out, as a result of contamination of the blood." Trump has denied borrowing Hitler's turn of phrase.[32] One of Trump's favorite epithets for the media—"enemy of the people"—was used by a litany of 20th century totalitarians, including Soviet dictator Joseph Stalin and Chinese Communist leader Mao Zedong.[33]

People listen. On the day after Christmas 2023, Trump himself shared an image on his social media platform, Truth Social, that showed the words voters most commonly associate with his potential second term; they included "revenge" and "dictatorship."[34]

On social media and in speeches, Trump legitimizes political violence by relentlessly describing the January 6 rioters—about 900 of whom have been convicted or pleaded guilty to criminal charges—as innocent political prisoners or "hostages" whom he plans to pardon and see released if he is reelected. "The cops should be charged and the protesters should be freed," declared a post that Trump re-posted on Truth Social.[35] Again, people listen: A Washington Post/University of Maryland poll released in January 2024 found that 25% of Americans say it is "probably" or "definitely" true that the Federal Bureau of Investigation instigated the January 6 attack.[36]

Political developments involving Trump routinely elicit violent threats and hate speech online. On December 19, 2023, the Colorado Supreme Court rejected Trump's presidential eligibility under the 14th Amendment to the U.S. Constitution, which bans insurrectionists from running for office. In the 24 hours after the Colorado ruling, hundreds of menacing posts targeting members of the state court, as well as Democrats and other "leftists," appeared on Trump's Truth Social and other right-wing venues such as Gab, 4chan, and GETTR. More than 150 such posts appeared on Truth Social alone, including calls for members of the Colorado court to "be arrested" or "hang for sedition."[37]

"Lately, Donald Trump's rhetoric online and elsewhere has veered increasingly toward that of authoritarian leaders who dehumanized their foes by comparing them to detested animals."

This is not the place for a full accounting of Trump's baleful influence on political discourse in the U.S. The point here is that, along with other influential voices on the right, the likely Republican nominee in 2024 has helped create a digital environment that has spawned other phenomena that threaten the orderly conduct of the November elections. The next several sections examine three of these phenomena: attacks on election workers, voter suppression schemes, and intimidation of academic and civil society misinformation researchers.

4. Attacks on election workers

Members of the Colorado Supreme Court were only one group of public officials to face violent intimidation after being castigated by Donald Trump on social media and in speeches. Shenna Bellows, Maine's secretary of state, was targeted with death threats after she announced in late December 2023 that Trump should be excluded from that state's ballot because of the 14th Amendment. Tanya S. Chutkan, the federal judge in Washington, D.C., presiding over Trump's criminal prosecution for election subversion, has had her home "swatted," meaning that someone called in a false emergency to police, resulting in armed law enforcement descending on the residence. Jack Smith, the special

> **'The atmosphere is that a mistake [by election workers] that would get no press, no attention in 2016 suddenly gets front page news, and then that cascades [into] a series of threats.'**
>
> — Colorado Secretary of State Jena Griswold

counsel prosecuting the election subversion case, has also been swatted, as have some Trump allies, such as Rep. Marjorie Taylor Greene (R., Ga.).[38]

This spreading aura of intimidation has the potential to disrupt elections, especially because many of the targets have been election workers, who typically do not enjoy the bodyguards and security protocols afforded officials higher in the government pecking order. In response, many election employees are quitting, taking with them years of institutional knowledge and practical experience.

One context in which election workers have been attacked is in the wake of routine snafus at polling places or other temporary mixups. These events receive hysterical attention on social media platforms as evidence of a "stolen election," Jena Griswold, Colorado's secretary of state, told us in an interview. Now "the atmosphere is that a mistake that would get no press, no attention in 2016 suddenly gets front page news and then that cascades [into] a series of threats," Griswold said.

Some harassment of election workers evolves from seemingly random hostility. Tonya Wichman, director of elections in Defiance County, Ohio, recounted a situation in her locale where a voter in the 2022 midterm election loudly proclaimed that voting machines were rigged and harassed volunteer poll workers. This individual has since turned his attention to Wichman herself. "He will sit and stare at me when I go to my nephew's ball games," she said. "It is mentally and physically exhausting."

Griswold of Colorado became the target of an intensive harassment campaign as a result of her involvement in the litigation initiated in September 2023 over Trump's eligibility for the presidency. "Within three months, I had received 64 death threats and over 950 non-lethal, non-death threats," she told us.

After upholding the results of the 2020 presidential election in Maricopa County, Bill Gates, the county supervisor, received a barrage of threats that caused him to move his family out of their home for a time. Gates has since spoken openly about the mental health struggles he continues to experience as a result of this harassment.[39]

Some interference with election workers takes the form of people calling in or emailing with questions about conspiracy theories they have read about on social media or blogs. The high volume of such inquiries has become a serious distraction, according to Gates. "There's only so many hours in the day," he told us.

Residents purporting to be concerned about election fraud are also inundating local government offices with formal public records requests for evidence to substantiate these theories —evidence that typically does not exist. Officials often are obliged by law to respond to such requests. In Harris County, Texas, which includes Houston, from 2020 to 2022, records requests more than doubled, to 1,200, according to Votebeat, a nonprofit news organization. In Maricopa County, requests increased twelvefold to 1,193 between 2019 and 2022.[40] Wichman, the Ohio elections official, said the "nonstop" records requests she received during the 2022 midterms—many of which were identically worded—seemed like another effort to subvert the democratic process: "It was like a copy-and-paste, where they were just trying to overwhelm the Board of Elections."

In April 2023, the Brennan Center for Justice at NYU School of Law published a survey of local election officials across the country showing that one in three had experienced harassment or threats because of their jobs. Forty-one percent of the respondents had received threats via social media; more than half, in person. In response, many are leaving their jobs. More than one in five election officials will be serving in their first presidential election in 2024, the Brennan Center found.[41] This turnover means that in many places, relatively inexperienced employees will help administer elections, increasing the chances of mistakes. The miscues, in turn, are likely to fuel baseless claims of election fraud and further erosion of trust in elections.

5. Voter suppression efforts

The corrosive digital environment fueling harassment of election workers also propels efforts to suppress the vote. In recent years, conservative groups and individuals have challenged voter registrations and alternatives to in-person voting, such as absentee ballots and ballot drop boxes. While the claim of widespread Democratic election fraud has become an article of faith on the political right, in fact, such fraud—by members of either party—is rare in the U.S.[42]

In one hotly contested recent episode, True the Vote, a Texas-based right-wing group, spearheaded a registration challenge in Georgia in the lead-up to a pair of pivotal Senate run-off elections in January 2021. Working with members of the Republican Party, True the Vote challenged the eligibility of some 360,000 voters, based on their residency. Georgia election officials ended up

rejecting only a few dozen ballots, and Democrats defeated their Republican foes by tens of thousands of votes, securing control of the Senate.[43]

In the subsequent run-up to the 2022 midterms, a number of Republican activists and candidates across the country used YouTube and other social media platforms to promote "2000 Mules," a documentary film designed to discredit the 2020 presidential election. The film contended that thousands of Democratic-paid "mules" illegally stuffed ballots into drop boxes in five swing states. Directed by conservative firebrand Dinesh D'Souza based on research by True the Vote, the movie was applauded by Donald Trump but de-bunked by PolitiFact, the Associated Press, and even Bill Barr, Trump's penultimate attorney general.[44]

Battles over voter registration for the 2024 election are already underway. One began, in a roundabout way, with the Electronic Registration Information Center (ERIC), which maintains a database that helps states keep accurate voter rolls and identify illegal practices. ERIC had existed without controversy since 2012, but in January 2022, the far-right news outlet Gateway Pundit described it as a "left-wing voter registration drive disguised as voter roll clean-up."[45] The accusation gained traction on social media, and at least nine Republican-led states, including Louisiana, Missouri, and Florida have pulled out of ERIC, leaving 24 states as members.[46]

One of ERIC's most prominent critics is Cleta Mitchell, an influential Republican attorney who worked with Donald Trump to overturn the 2020 election results. Mitchell was on the telephone line with Trump on January 2, 2021, when he asked Georgia election officials to "find" 11,780 votes to swing the state in his favor.[47] Mitchell leads the Election Integrity Network (EIN), a coalition of activists devoted to "making it easy to vote and hard to cheat," according to its website.[48] The network is a project of the Conservative Partnership Institute, a think tank with close ties to the Trump political operation.[49] Mitchell has used EIN to attack ERIC in terms similar to those used by Gateway Pundit. The EIN website calls ERIC "a taxpayer funded left-wing get-out-the-vote organization pretending to clean up voter rolls while exploiting private citizen data in violation of law." The site adds that ERIC "hides behind its private sector status to avoid public information requests, and discourages its government members from honoring such requests."[50]

Mitchell's EIN has taken an interest in a Georgia-based project called EagleAI Network, a digital platform that can be used to gather voter registration data "from governmental and other highly reliable sources" and displays the information "in a convenient, efficient manner," according to John Richards, a medical doctor who developed EagleAI.[51] Dr. Richards said in an interview that his software can be used by individuals interested in voter roll accuracy, as well as by government officials. He said that EIN had invited him to join Zoom calls to provide demonstrations and updates on EagleAI. Mitchell attended one of the online demonstrations and "made nice statements about what we are doing," Dr. Richards added. He also said that "there are many people who volunteer with EIN who use EagleAI, support EagleAI, and like EagleAI." But he declined to provide more specifics on users. He said that EagleAI is not partisan in nature and should not be tied to "any political or controversial figure." In an email, Dr. Richards said that "EagleAI Network is a tool to *prevent* disenfranchisement."[52]

In a separate email, Mitchell said: "You cannot find a single instance, EVER, when I or any of my colleagues have tried to 'suppress the vote.'… What we do and have done is to study the election systems and procedures, and we identify the many ways in which there are votes cast and counted by persons whose votes were illegal, violated state law, or otherwise were not cast in accordance with applicable law. So if you want to say that we are dedicated to suppressing ILLEGAL votes, that would be correct."

Only one jurisdiction in the country so far has signed on to use EagleAI: heavily Republican Columbia County, Georgia, near Augusta.[53] And according to documents posted by the investigative group Documented, the chairman of the Georgia State Elections Board has sent a letter to the Columbia County Board of Elections warning against the adoption of the system because of voter-privacy concerns.[54]

6. The campaign against misinformation research and mitigation

A third outgrowth of the poisonous online ecosystem is the ongoing campaign by right-wing partisans to target efforts to identify and mitigate election mis- and disinformation by academic and civil society experts, government officials, and certain employees at social media platforms. While the Republican politicians and conservative activists leading this charge have failed to substantiate their allegations of a leftist conspiracy to marginalize right-leaning viewpoints, the intimidation is likely to diminish the role these experts and their organizations have played protecting the integrity of past elections.

The campaign—including a barrage of record requests, congressional subpoenas, and lawsuits—seeks to characterize a range of activities aimed at flagging false election claims to the platforms for potential moderation as what Rep. Jim Jordan (R., Ohio) calls a "censorship industrial complex."[55] Jordan has helped lead hearings and directed a legislative investigation in

> **Elon Musk's X is pursuing its own litigation strategy, suing nonprofit organizations that research misinformation and have published reports critical of the platform.**

his role as chairman of the House Judiciary Committee and its Select Subcommittee on the Weaponization of the Federal Government.[56]

Despite an absence of persuasive evidence, these efforts to impugn the work of researchers have succeeded in chilling collaboration among the experts, government officials, and platform employees. One casualty has been dialogue between platforms and government agencies about anti-U.S. disinformation generated by Russia, China, and Iran. In a report issued in November 2023, Meta noted that "threat sharing by the federal government in the U.S. related to foreign election interference has been paused since July."[57]

Kate Starbird, a University of Washington computer scientist and leading misinformation researcher, has faced voluminous demands by Rep. Jordan's subcommittee for her email and other communications.[58] Being publicly identified as a target of the investigation has led to sustained online harassment, including threats on her life. "It has become a lot harder to do our work," Starbird said in an interview, pointing to recent hesitancy on the part of some funding organizations and reluctance on the part of younger, non-tenured researchers to specialize in the field. Adding fuel to the fire has been the selective disclosure of the so-called Twitter Files, which X owner Elon Musk provided to a group of Substack newsletter writers with the goal of portraying prior management of the company as involved in precisely the kind of liberal plotting that Jordan and his congressional allies have alleged. A House hearing in February 2023 centered on the Twitter Files produced testimony about favoritism, but former company employees said that before Musk took over, the platform went out of its way to favor *conservatives*, not liberals.[59]

While Jordan's congressional investigators have sought information from dozens of universities, including NYU (although not from the NYU Stern Center[60]), the probe has focused heavily on the Election Integrity Partnership (EIP), a collaboration among the University of Washington's Center for an Informed Public, which Starbird heads; the Stanford Internet Observatory and Program on Democracy and the Internet; the social media analytics firm Graphika; and the Digital Forensic Research Lab, an arm of the Atlantic Council, a Washington, D.C., think tank. Established during the 2020 election cycle, the EIP described its purpose as enabling "real-time information exchange between election officials, government agencies, civil society organizations, social media platforms, the media, and the research community."[61] It coordinated with federal agencies like the Department of Homeland Security and its Cybersecurity and Infrastructure Security Agency to identify disinformation generated domestically and by U.S. rivals. An interim report released in November 2023 by the Jordan subcommittee claims that "the EIP provided a way for the federal government to launder its censorship activities in hopes of bypassing both the First Amendment and public scrutiny."[62]

By contrast, a fair-minded reading of the partnership's work reveals meticulous description and analysis of the spread of election falsehoods, with an eye toward reducing their distortion of the democratic process.[63]

Litigation is another weapon conservatives have deployed against anti-misinformation efforts. The leading lawsuit, initiated by the Republican attorneys general of Missouri and Louisiana and now pending before the U.S. Supreme Court, alleges that major social media platforms censored content criticizing federal Covid-19 policies at the behest of the Biden administration. In September 2023, the New Orleans-based U.S. Court of Appeals for the Fifth Circuit ruled largely in favor of the state officials, finding that in some instances, administration officials violated the First Amendment by effectively "coercing" platform employees to take down posts disputing government vaccination policies, among other topics.[64]

In its appeal to the Supreme Court, the Biden administration countered that the communication at issue represented legitimate expression of government policy.

Elon Musk's X is pursuing its own litigation strategy, suing nonprofit organizations that research misinformation and have published reports critical of the platform. In separate lawsuits filed in 2023, X accused the Center for Countering Digital Hate and the liberal advocacy group Media Matters for America of inaccurately describing harmful content on X to alienate advertisers from the platform—allegations the two organizations have denied.[65]

Yet another misinformation authority who has been targeted in a separate lawsuit filed by conservatives, Alex Stamos, told Congress in testimony in December 2023 that lawmakers, not the courts, should set rules for how government interacts with social media companies. Stamos, the former head of the Stanford Internet Observatory and prior to that, Facebook's chief security officer, made a worthwhile distinction: "Instead of this being a five-year fight in the courts, I think Congress needs to act and say, 'These are the things that the government

is not allowed to say; this is what the administration cannot do with social media companies,' he testified. "But if the FBI knows that this IP address is being used by the Iranians to create fake accounts, they can contact Facebook."[66]

In this contentious environment, many people trying to blunt the effects of election mis- and disinformation are switching strategies, Emma Steiner, information accountability project manager with the nonpartisan group Common Cause, said in an interview. "We are focused less on 'Whac-a-Mole' and more on promoting positive inoculation content—[which is] still a contested method—but we have reason to believe it may be a more fruitful use of resources," Steiner added. The shift reflects necessity, she said. "If there is an emergency 'break-the-glass' issue, we simply don't know who to contact" at platforms where election integrity teams have been depleted.

7. Interference from abroad

The lethal wildfires that afflicted the Hawaiian island of Maui in August 2023 provided an unlikely preview of the sort of disinformation that China and Russia may aim at the 2024 U.S. election. An apparently coordinated Russian campaign by false accounts on X, which began a day after the fires ignited, promoted the idea that the U.S. government should spend more money to help its own citizens rather than provide military aid to Ukraine. These messages spread from the phony Russian X accounts to conservative U.S. outlets like Breitbart and ultimately to Russian state media organizations. Meanwhile, the Chinese government used multiple major social media platforms to disseminate the false conspiracy theory that the U.S. government had caused the Maui fires with a mysterious "weather weapon."[67]

The Russian disinformation about Maui resembled past Kremlin attempts to stir discord in the U.S. by means of inauthentic social media accounts and hack-and-leak operations—most notably around the 2016 presidential election. For China, seeking to sow social and political division in the U.S. is a relatively new strategy; previously, Beijing has focused its influence operations primarily on supporting its policies toward Hong Kong, Taiwan, and other Asian countries.

China's apparent emulation of Russia, as well as increasing digital influence operations by Iran, prompted Microsoft's Threat Analysis Center to predict that "Election 2024 may be the first presidential election during which multiple authoritarian actors simultaneously attempt to interfere with and influence an election outcome."[68]

China

U.S. intelligence agencies believe that China increasingly seeks to use information operations to portray its main rival, the U.S., as riven by social conflict—in contrast to the official Chinese self-image of a unified society loyal to its autocratic leadership. In a joint report declassified and released in December 2023, a consortium of American intelligence agencies noted that during the 2022 U.S. midterm elections, China had demonstrated greater interest than in the past in trying to shape American politics. In a handful of midterm races, Chinese operatives covertly posted material on social media platforms meant to undermine politicians seen as opposed to China's interests and boost those perceived as more friendly. The heavily redacted published version of the report did not identify specific American politicians.[69]

On some occasions, major platforms have identified and removed networks of foreign accounts targeting the U.S. In November 2023, Meta published a report describing its takedown of a Chinese network of 4,789 accounts for violating its policy against "coordinated inauthentic behavior." Meta said the accounts posted in English about U.S. politics and U.S.-China relations, often "criticiz[ing] both sides of the U.S. political spectrum by using what appears to be copy-pasted partisan content from people on X." In December, YouTube said it terminated 1,953 video channels and 52 blogs as part of an ongoing investigation into "coordinated influence operations" linked to China that upload content in Chinese and English about China and U.S. foreign affairs.[70]

"A distinctive concern about Chinese influence has swirled around TikTok, the short-video platform used by 150 million Americans (and an estimated 1.5 billion people worldwide), whose parent company is China-based tech giant ByteDance."

A distinctive concern about Chinese influence has swirled around TikTok, the short-video platform used by 150 million Americans (and an estimated 1.5 billion people worldwide), whose parent company is China-based tech giant ByteDance. Though originally seen as a venue for trading dance moves and celebrity gossip, TikTok increasingly hosts political content, much of it presented by, and seemingly targeted at, young adults.

The debate about the platform focuses on allegations—so far, unproven—that ByteDance collects sensitive user data and censors content at the behest of the Chinese government. For example, the Network Contagion Research Institute, an organization affiliated with Rutgers University, issued a report in December 2023 finding "a strong possibility that TikTok systematically

promotes or demotes content on the basis of whether it is aligned with or opposed to the interests of the Chinese government." The institute based this conclusion on a comparison of the frequency that hashtags referring to issues of concern to Beijing—such as Taiwan, Hong Kong, and Tiananmen Square—appeared on TikTok, as compared to on Instagram. Hashtags relating to these hot-button issues appeared far less frequently on TikTok than Instagram, the institute found.[71]

But the institute's conclusion seems debatable. For one thing, users holding anti-China views might avoid TikTok because of its association with China. For its part, TikTok said the institute used "a flawed methodology to reach a predetermined, false conclusion." After the report's publication, TikTok restricted access to the transparency tool that had allowed the institute to search for hashtag data on the platform.[72]

TikTok has emphatically denied that the Chinese government has any influence over it or ByteDance. The platform says that it routes its users' personal data and traffic to a secure system in the U.S. overseen by the American tech company Oracle, with no access provided to Beijing. And TikTok points out that it periodically removes "covert influence networks that are operating from China or amplifying pro-China narratives."[73]

The multi-agency Committee on Foreign Investment in the United States, which is led by the Treasury Department, demanded in March 2023 that ByteDance spin off TikTok as a separate U.S. business or face a possible ban. But since then, the committee has failed to follow through, and legislation introduced to give the executive branch the authority to impose such a ban has not advanced in Congress.[74]

Russia

The Kremlin began accelerating its online information activity in late 2023, according to Microsoft's Threat Analysis Center. Russian state news outlets and covert Russia-affiliated social media networks aligned their messages, the company's researchers reported, "focusing their propaganda and disinformation on Western military aid to Ukraine and messaging against candidates committed to it."[75] The report didn't name specific candidates.

Clint Watts, who heads the Microsoft center, said in an interview that as of December 2023, the Russians hadn't yet stepped up information activity targeting the U.S. elections. "Everything is on pause" until the general election, he explained. The Russians assume that President Biden will once again face Donald Trump. "That doesn't mean that by next spring [2024] they won't be raring to go," Watts added.

Russian operatives have been experimenting with generative AI, primarily by "spoofing legitimate media coverage of fake content espousing Kremlin-preferred narratives delegitimizing Ukraine and casting blame for the current Israel and Gaza conflict on the U.S. and Ukraine," the Microsoft center noted. Watts told us that if the Russians attempt to disrupt the U.S. election cycle with an "October surprise" deepfake video, it's likely that they will use the very latest version of generative AI—technology that probably is not even on the market yet.

Iran

Tehran hasn't yet asserted itself in the 2024 U.S. election cycle and has fewer resources for sophisticated information operations, according to U.S. intelligence analysts. But Iranian leaders have the ambition of fueling distrust in U.S. political institutions and increasing social tensions, in large part to undermine American support for Israel.

In 2022, Iranian-controlled Twitter accounts masqueraded as left-leaning Americans supporting progressive Democratic candidates in mid-term elections, apparently because those politicians were likely to be skeptical of military aid to Israel.[76]

3. Digital Risks in Elections Outside of the U.S.

> India's ruling Hindu nationalist Bharatiya Janata Party has built an extraordinary digital army to push its political messages—including disinformation smearing Muslims—while at the same time, the Indian government aggressively censors social media content posted by dissenters.

Countries around the world holding elections in 2024 are also likely to experience digital disruptions of various kinds. Slovakia, where voters went to the polls in September 2023, provided a preview. In the months before the Slovakian election, disinformation surged, often with the apparent aim of undermining the country's membership in the North Atlantic Treaty Organization (NATO) and support for Ukraine in its war with Russia.[77]

In one prominent example, an audio recording posted on Facebook just two days before the election appeared to capture the leader of the pro-NATO, pro-Ukraine Progressive Slovakia Party discussing with a journalist how to rig the results by buying votes. The supposed participants immediately denounced the recording as fake, and the fact-checking department of the Agence France-Presse said it showed signs of manipulation using artificial intelligence.[78] But the audio post came during a pre-election media moratorium, limiting the degree to which it could be debunked in Slovakia. And as an audio file, the post exploited a loophole in Facebook's rules, which ban certain manipulated video but not audio.[79] Meanwhile, Robert Fico, the prime minister candidate from the opposing party, known as Smer, reportedly used Facebook to spread false claims that the war in Ukraine started with Ukrainian fascists murdering Russians. In the election, Smer defeated Progressive Slovakia, Fico became prime minister, and his government promptly cut off military aid to Ukraine.[80]

Here are snapshots of digital risks in countries holding elections in 2024:

South Korea (April 10)

In South Korea, the prospect of disinformation has become a weapon used against the press. President Yoon Suk Yeol and some of his supporters have tried to intimidate press critics by calling their coverage "fake news." Yoon, a former prosecutor, and his backers are using threats, lawsuits, and criminal investigations to attack journalism outlets publishing articles critical of his government.[81] These allegations could mute skeptical coverage and influence legislative elections scheduled for April 10. In a December 2023 statement to the Voice of America, Yoon's office said that his administration "holds freedom of the press in the highest regard and exerts its utmost to protect it as it is the core value of a robust democracy."[82]

Separately, researchers at South Korea's Gachon University reported in December 2023 that they had found more than 50 false accounts on Naver, a popular Korean web portal, that they

> **Mis- and disinformation spread on social media platforms often originate with political elites, including lame duck Mexican President Andrés Manuel López Obrador and his party.**

surmise are linked to the Chinese Communist Party (CCP). These accounts generated some 30,000 comments critical of friendly Korean relations with the U.S. and Japan—an attempt to exacerbate internal conflicts in South Korea, according to the researchers. In November 2023, Korea's National Intelligence Service found that suspect media companies serving as fronts for the CCP had set up 38 websites imitating local Korean news outlets. The websites had attempted to spread false information and incite pro-Communist and anti-U.S. sentiment, the intelligence service said.[83]

India (April - May)

India's ruling Hindu nationalist Bharatiya Janata Party (BJP) has built an extraordinary digital army to push its political messages—including disinformation smearing Muslims—while at the same time, the Indian government aggressively censors social media content posted by dissenters. And knowledgeable observers warn that a third problem, the distribution of deepfake video using generative artificial intelligence, could disrupt the general elections scheduled to take place over several weeks this spring, In India, "the effective use of AI and other digital tools by powerful political actors can help nullify the efforts of weaker political actors," Amber Sinha, a senior fellow with the Mozilla Foundation and Tech Policy Press fellow, said in an interview.

Led by Prime Minister Narendra Modi, the BJP has assembled an "information technology cell" of some 150,000 workers devoted to promoting the party line, primarily via Meta's WhatsApp messaging platform. In India, the world's biggest electoral democracy, with more than 1.4 billion people, WhatsApp has half a billion users, the app's largest market. The BJP's enormous online workforce pumps out praise for the party's accomplishments but also seeks to inspire loyalty among the majority Hindu population by inciting hatred of the country's 14% Muslim minority. An exhaustive investigation published in autumn 2023 by *The Washington Post* found that the BJP barrages voters' phones with false accounts of Muslims murdering Hindus and enticing Hindu girls to convert to Islam and marry Muslims, all to persuade Hindus that they will be safe only if the BJP remains in control.[84]

WhatsApp has responded to this kind of mass messaging by placing restrictions on the number of phone numbers to which messages can be forwarded at one time. But the sheer scale of the BJP's IT cell allows the party to send countless incendiary messages to huge swaths of the Indian electorate.

Along with its exploitation of WhatsApp, the BJP has enacted legislation giving itself sweeping power over online content moderation. Through amendments to India's technology laws, the government has established grievance committees with veto power over content moderation decisions.[85] The BJP aggressively demands that major platforms like Meta take down content that the party contends undermines national security in what skeptics describe as blatant censorship of political foes. X, which in its former days as Twitter, enjoyed a reputation for resisting such government take-down requests, now routinely complies. In January 2023, Twitter and YouTube bowed to Indian government orders to remove links in India to a BBC documentary that blamed Modi for tolerating riots in 2002 that left more than 1,000 people dead, most of them Muslims, in Gujarat state, where he was chief minister at the time.[86]

Last fall, 14 leaders of Indian opposition parties jointly wrote a letter to Meta's Mark Zuckerberg and his counterpart at Google, Sundar Pichai. The letter called on the CEOs to cease "aiding the communal hatred campaign of the ruling BJP" ahead of the 2024 elections.[87]

Experts on artificial intelligence are warning that the danger of generative AI deepfakes is acute in India, where 71% of the adult population owns a smart phone, data rates are relatively low, and usage is high.[88] For instance, in the run-up to state assembly elections in Telangana in November 2023, the incumbent BJP lodged a complaint against the rival Indian National Congress Party for using deepfake video and audio to target BJP president K Chandrashekar Rao and other BJP leaders and candidates.[89] The BJP, too, has deployed deepfake technology during elections. In February 2020, during state elections in Delhi, the BJP IT cell reportedly partnered with a political communications firm to use deepfake videos to reach voters from different linguistic backgrounds.[90]

Karen Rebelo, deputy director of Boom, a fact-checking service in India, has noted a substantial rise in the use of deepfakes. "2023 was the year of generative AI," she said in an interview. Rebelo pointed to one example where

the campaigns for opposing candidates from BJP and National Congress in a Madhya Pradesh assembly race both apparently used AI voice-cloning technology to spread disinformation.[91] "Generative AI and deepfakes are so potentially scary because they offer sophistication and scale," she added. "We are not talking about some hypothetical end-of-the-world Terminator type situation; we are already seeing the havoc that they are wreaking."

Mexico (June 2)

Mis- and disinformation spread on social media platforms often originate with political elites, including lame duck President Andrés Manuel López Obrador and his party.[92] Inauthentic accounts linked to López Obrador were discovered in the 2022 election cycle, targeting Mexico's electoral authority. Headed into 2024, the sitting President's "blatant disregard for the National Electoral Institute itself and its ability to set electoral rules…is an ominous sign of institutional deterioration and the lack of adherence to written norms and practices ahead of the presidential election," according to analysts at the Center for Strategic & International Studies.[93] His willingness to deploy falsehoods and challenge electoral norms may benefit his political protégée, presidential front-runner Claudia Sheinbaum, a former mayor of Mexico City.

To prepare for elections in Mexico and other Spanish-speaking countries, social media platforms need to address persistent discrepancies between the moderation of English and Spanish content on multiple social media platforms, Roberta Braga, founder and executive director of the Digital Democracy Institute of the Americas, said in an interview. "Social media platforms need to improve machine learning models to better detect inauthentic coordination and networks spreading Spanish-language disinformation, but they also need to hire more culturally fluent Spanish-language moderators, and make sure they have effective and efficient integrity systems in place for national and local Mexican elections." Braga recommended that Meta, in particular, partner with organizations studying Mexican and Mexican-American communities on WhatsApp. "This is vital for understanding how conversations evolve and how encrypted content can be amplified or go viral," she said.

European Union (June 6 - 9)

Across the 27 member states of the European Union, the most common theme of online disinformation related to the 2024 E.U. Parliamentary elections is expected to be attempts by Russia to undermine aid to Ukraine and loyalty to NATO and the United States. Populist and nationalist political parties are also focusing on the demonization of immigrants, including Ukrainian refugees, as a way to sow discord within European countries. The European Digital Media Observatory, which is affiliated with the European Commission, has noted that false narratives regarding election fraud, popularized in the U.S. by Donald Trump, are also being deployed in the EU to cast doubt on the workings of democracy.[94]

In November 2023, the EU DisinfoLab, an independent nonprofit, reported on a Russia-based influence network that has been operating in Europe since at least May 2022. Dubbed "Doppelgänger" by analysts, the Russian campaign uses multiple "clones" of authentic media outlets—including *Bild*, *The Guardian*, and 20minutes—to spread fake articles, videos, and polls. A fake *Der Spiegel* article, circulated online in August 2023, claimed that Germans were being increasingly pushed out of the country by migrants. The same month, a video made to appear as if it came from the German broadcaster DW started circulating on X and TikTok. The video claimed to report on an anti-Ukrainian "flash mob" in Warsaw and that Poles had renamed wi-fi networks with phrases like "Ukraine is hell," and "Murderers from Ukraine." The two major German outlets denied having ever produced this content.[95]

The European Policy Centre, a Brussels think tank, has noted that Bulgaria has likewise been targeted by pro-Russian information distorters seeking to erode the country's support for Ukraine and connections to NATO. False posts on social networks have claimed that NATO ordered the mobilization of 30,000 Bulgarian soldiers. A separate Facebook post widely shared in November 2023 falsely stated that NATO troops were en route to Bulgaria Related TikTok videos claimed without evidence that Bulgarian territory had been donated to NATO and that U.S. military bases were going to be built in the Bulgarian city of Yambol. In fact, while NATO was using some Bulgarian facilities, there was no gift of land to the alliance for base construction.[96]

4. Conclusions and Recommendations

> All off the platforms must put in place contingency plans for a crisis on the order of the January 6, 2021, assault on the U.S. Capitol or the January 8, 2023, attack on Brazil's federal buildings by supporters of then-president Jair Bolsonaro, who had been defeated in his country's 2022 election.

Mainstream technology companies cannot be expected on their own to insulate elections in the United States or anywhere else from all of the disruptions caused by conspiracy mongers, bigots, and other anti-democratic forces. Political leaders, as well as hyper-partisan cable television outlets, fringe social media sites, podcasters and millions of voters who for a variety of reasons are drawn to extremist, anti-democratic views all bear responsibility for the precarious state of politics in the U.S. and countries around the world.

But we *can* expect and demand that major tech companies not retreat from the modest attempts they have made in the past to protect the integrity of the election processes in the countries where they operate. They should be stepping up their efforts in this regard, not backsliding.

What follows are practical recommendations to the industry and government: not solutions that would miraculously erase the damaging side effects of technology, let alone the larger alienation and distrust that currently clouds politics, but modest steps, which, if pursued diligently and in combination, could mitigate some of the harm.

We began proposing some of these ideas years ago.[97] For the most part, progress has been limited to a shift in the terms of debate, with precious little tangible improvement. But the only hope for progress is to keep pushing an achievable agenda so that people of good will—inside companies and government and among the public at large—have a map pointing toward change.

Recommendations to Technology Companies

1 Add many more humans to the content moderation loop.

Elections raise content questions that require human judgment, ranging from whether a claim of voting fraud is provably false to whether a dubious advertisement is "political" in nature. For years, social media companies have done content moderation on the cheap by outsourcing this crucial corporate function to low-cost vendors operating in low-wage locations. This strategy limits Silicon Valley payrolls but undercuts effectiveness, as platform companies shirk responsibility for hiring, training, and directly overseeing the workers evaluating content. During the recent wave of tech layoffs, the companies even reduced the modest ranks of moderators who work in-house. All of this needs to change, and what better reason to do so than protecting elections? Social media companies should double or even triple the number of people doing content moderation, properly resourcing teams with cultural expertise and language skills necessary in each market where they operate, and make them all direct employees who receive the oversight and compensation commensurate with the importance of their task. Expensive? Yes. It's the cost of doing business responsibly.

2 Fund more outside fact-checkers to prowl for falsehoods.

While companies beef up human moderation to enforce platform policies, they should also fund more outside fact-checking organizations capable of separating reality from fiction. Fact-checking groups employ reporters who use interviews, document reviews, and other journalistic techniques to dig deeper than content moderators typically can. Research has shown that fact-checking has a positive effect on people's ability to distinguish truth from lies.[98] But there aren't enough fact-checkers, a deficit that is especially obvious during a global profusion of elections. Meta, to its credit, funds a network of more than 90 fact-checking groups in the U.S. and around the world. But the other major platform companies have invested far less; some, like X, do nothing on this front. Even Meta undercuts its fact-checking investment by maintaining an exemption for politicians, who may lie on its platforms without penalty—a policy mistake that needs to be reversed right away.

3 Blunt election delegitimization and prepare for a crisis.

As the U.S. presidential campaign unfolds, Donald Trump is likely to step up his efforts to undercut the legitimacy of elections. False claims about the 2020 election are already serving as the basis for false claims about the legitimacy of the 2024 cycle. Platforms should label these claims as baseless or remove them altogether and direct users to authoritative, nonpartisan information about the country's generally sound election record. Meta should reverse its ill-advised decision to allow political ads that question the legitimacy of the 2020 election, and YouTube needs to undo its policy of allowing videos that propagate falsehoods about past election fraud. And all of the platforms must put in place contingency plans for a crisis on the order of the January 6, 2021, assault on the U.S. Capitol or the January 8, 2023, attack on Brazil's federal buildings by supporters of then-president Jair Bolsonaro, who had been defeated in his country's 2022 election. Contingency plans should include the capacity to add personnel—from top executives to policy experts and rank-and-file moderators—and to adjust recommendation algorithms so that they prioritize reliably gathered news rather than conspiracy theories. The safety of other features, such as the ability to form groups and readily recruit new group members, should also be assessed with these concerns in mind.

Recommendations to Technology Companies

4 Make platform design and policy adjustments for election season, and beyond.

The following changes, some of which have been proposed by Accountable Tech and others, would make digital disruptions of elections less likely.[99] They also would improve the tone and content of online communication beyond election years. Rather than institute them as temporary measures for elections, platform companies should make them permanent.

Scrutinize resharing.

Algorithms that promote content based on whether it will prompt user engagement are not the sole reason sensational and false posts "go viral." Users are part of the equation, often clicking to share—and reshare—content based on a cursory glance at whether it reinforces their ideological or cultural predilections. According to leaked Meta internal research, "when a user sees a reshare of a reshare…they are 4 times more likely to be seeing misinfo compared to when they see links or photos on News Feed in general."[100] So, during election season, social media companies should monitor resharing patterns, with an eye toward possibly restricting the feature. And if such interventions have a salutary effect, they should impose the limits all of the time. Notably, in the world of messaging, Meta in 2019 restricted the number of WhatsApp groups to which a user can forward a post with a single click. That change seemed to help slow the spread of misinformation.[101] WhatsApp should tighten the limits further to prevent viral sharing of falsehoods in India, Mexico, Brazil, and other countries where the app is particularly popular.

Adopt 'circuit breakers.'

Following the example of financial markets that impose automatic "circuit breakers" that pause trading in response to panic about a stock or a sector, platform companies should impose circuit breakers that briefly halt algorithmic amplification of certain rapidly spreading content. "This would buy time for a quick check to determine what the content is, whether it's reputable or malicious and—for certain narrow categories in which false information has high potential to cause harm—if it is accurate," Renée DiResta, the technical research manager at the Stanford Internet Observatory, has written.[102]

Impose 'interstitial' warnings.

Studies have found that simply slowing down the use of social media by imposing "friction" can improve the quality of online discourse. For all of its missteps, X commendably continues to ask readers in an automatic pop-up whether they want to read an article before sharing it. But "not all warnings are created equal," according to a group of researchers from Princeton and Cornell. They found that users rarely notice, let alone respond to, "contextual" warnings that rely on small icons and discreet placement of text below the content in question. But readers do notice and sometimes change their behavior in response to bolder "interstitial" warnings, like the X suggestion to first read what you're trying to share or labels pointing out that fact-checkers have found a claim to lack a basis in fact. To get at the content in question, users have to pause and click through the warning—a source of friction that may cause some to think more carefully about what they are doing.[103]

Recommendations to Technology Companies

Remove provably false content.

For the relatively small subset of mis- and disinformation that is provably false—for example, claims that Joe Biden was not legitimately elected president in 2020—platforms should remove the content altogether from user feeds, rather than merely down-rank or label it. A record copy of this material should be made available only in a cordoned-off part of a platform, marked with a prominent diagonal red stripe or other badge of falsehood, and made ineligible for sharing, liking, or commenting. In this manner, the false material is not entirely censored, but it will not circulate on that platform.

5 Mitigate the risks of generative artificial intelligence.

A number of major technology firms recently have announced their intention to assure that AI-generated imagery and audio content are readily identifiable as such—a welcome development that could reduce risks that political mis- and disinformation will disrupt elections.[104] Now, it is vital that these companies, which include Microsoft, Google, Meta, and OpenAI, swiftly follow through on their commitments and also respond to the inevitable attempts by bad actors to circumvent these protections. Hardware manufacturers need to pitch in as well by designing their devices to apply "content credentials" whenever an image or video is created. Here are specific steps the industry should take:

Label all AI-generated content.

Developers and marketers of AI systems need to ensure that content is designated as created by artificial intelligence, whether text, audio, or imagery. Better tools must be invented to detect deceptive, potentially harmful uses of AI-generated content. Meta's promise to apply labels automatically to AI-generated images is the kind of response that is needed.

Embrace industry-wide standards.

Google recently joined Microsoft, Meta, Adobe, and other companies in the Coalition for Content Provenance and Authenticity, signaling that the big technology firms agree on the adoption of common standards for labeling media. All firms that operate social media platforms or permit user-generated content should follow suit. These standards allow the tracking of image- or video-creation history, source credibility, and editing processes.

Prioritize collaboration and information sharing.

The industry should form a clearinghouse to facilitate labeling or removal of deceptive AI content, using as a model the Global Internet Forum to Counter Terrorism, which currently allows for coordination on the removal of content promoting terrorism. Standardized information-sharing protocols are essential for effectiveness.

Recommendations to Governments

6 Enforce existing laws as they apply to digital industries.

In the U.S., a dysfunctional Congress is not going to pass meaningful legislation related to digital industries before Election Day in November. In the meantime, though, executive branch agencies—primarily the Federal Trade Commission, Federal Election Commission, Justice Department, Consumer Financial Protection Bureau, and their state counterparts—need to use their full authority to enforce existing laws against election fraud, voter suppression, cyberattacks, and other offenses potentially relevant to protecting elections. Consumer protection and privacy laws could also come into play when regulators assess alleged misuse of generative AI in connection with elections.

7 Strengthen legal protections for election workers.

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. In the U.S., a legislative tracker from the nonprofit watchdog group Public Citizen finds that 15 states have passed new laws to protect election workers, and 15 more have introduced legislation that would do so.[105] Legislators should press ahead with this effort. In addition to measures against violence, harassment, and intimidation, many of these laws address phenomena such as doxxing—revealing personal and contact information online—or introduce reporting mechanisms to assess the scale of the problem. Given that real world violence often begins with online threats, prosecutions under such laws may ultimately serve as a deterrent to behavior offline.

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 protecting elections, but it would create incentives for social media companies, AI designers and marketers, and other digital enterprises to conduct themselves in a more constructive fashion. This would have positive ripple effects on politics and other aspects of society.

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.[106]

Recommendations to Governments

9 Mandate more transparency.

All serious proposals for more vigorous regulation of digital companies begin with the need for greater disclosure of how these businesses make decisions. More transparency—related to social media platforms, generative AI systems, and other digital technology—will provide better insight into why the technology sometimes goes awry and how to extend useful government oversight. These revelations will benefit the election process, as well as every other aspect of democracy affected by digital activity. 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.[107] The European Union's Digital Services Act also contains transparency provisions worth considering in any U.S. legislation.[108]

10 Bolster public sector and academic research capacity.

Congress needs to shore up public sector and academic capacity to study technology and its effects on democracy by appropriating additional funding and providing authority to the National Science Foundation and other agencies to foster robust research on the interaction of technology and elections, among other relevant topics. Beyond bolstering research on the spread of mis- and disinformation, such an initiative should address the disparity in computing power between the private sector, on the one hand, and the public sector and academia, on the other. This imbalance is particularly striking in connection with artificial intelligence. Building and testing generative AI models requires enormous computing capacity; private companies possess it, but the government and universities generally do not. Congress needs to address this disparity so strengthened public and academic computing infrastructure can be used to evaluate what Silicon Valley produces.

Endnotes

1 https://bhr.stern.nyu.edu/tech-generativeai

2 https://www.nytimes.com/2024/01/11/technology/discord-tech-layoffs.html; https://www.nbcnews.com/tech/tech-news/tech-layoffs-hit-trust-safety-teams-raising-fears-backsliding-efforts-rcna69111; https://www.cnbc.com/2023/05/26/tech-companies-are-laying-off-their-ethics-and-safety-teams-.html?utm_source=newsletter&utm_medium=email&utm_campaign=newsletter_axiosam&stream=top

3 https://www.theinformation.com/articles/musks-x-cuts-half-of-election-integrity-team-after-promising-to-expand-it; https://twitter.com/elonmusk/status/1707147926789554422

4 https://www.cnn.com/2023/04/12/tech/elon-musk-bbc-interview-twitter-intl-hnk/index.html

5 https://www.theinformation.com/articles/musks-x-cuts-half-of-election-integrity-team-after-promising-to-expand-it; https://twitter.com/elonmusk/status/1707147926789554422

6 https://www.nytimes.com/2023/11/24/business/x-elon-musk-advertisers.html

7 https://twitter.com/elonmusk/status/1737488430202851389

8 https://www.washingtonpost.com/technology/2023/12/18/european-union-x-probe-elon-musk/

9 https://www.nytimes.com/2023/09/18/opinion/trump-elon-musk-twitter.html

10 https://thehill.com/opinion/technology/3852594-the-twitter-bias-hearings-point-to-favoritism-but-not-for-liberals/; https://bhr.stern.nyu.edu/bias-report-release-page

11 https://twitter.com/elonmusk/status/1588933974470332418

12 https://arxiv.org/pdf/2210.15723.pdf

13 https://www.poynter.org/fact-checking/2023/why-twitters-community-notes-feature-mostly-fails-to-combat-misinformation/; https://www.wired.com/story/x-community-notes-disinformation/; https://mashable.com/article/twitter-x-community-notes-misinformation-views-investigation

14 https://www.nytimes.com/2022/06/23/technology/mark-zuckerberg-meta-midterm-elections.html

15 https://www.techpolicy.press/platforms-are-abandoning-u-s-democracy/

16 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3867818

17 https://www.wsj.com/tech/meta-allows-ads-claiming-rigged-2020-election-on-facebook-instagram-309b678d?st=rqlu-09a42lhue3p

18 https://blog.youtube/inside-youtube/us-election-misinformation-update-2023/; https://support.google.com/youtube/answer/6162278?hl=en#Harmful_dangerous&zippy=,policy-detail

19 https://www.semafor.com/article/06/08/2023/desantis-campaign-shares-fake-trumpfauci-images-prompting-new-ai-fears; https://www.theverge.com/2023/4/25/23697328/biden-reelection-rnc-ai-generated-attack-ad-deepfake; https://www.politico.com/newsletters/digital-future-daily/2023/06/21/biden-vs-trump-and-the-future-of-debate-00103000

20 https://www.economist.com/united-states/2023/08/31/ai-will-change-american-elections-but-not-in-the-obvious-way

21 https://www.pnas.org/doi/10.1073/pnas.2120481119

22 https://openai.com/blog/how-openai-is-approaching-2024-worldwide-elections#OpenAI; https://www.washingtonpost.com/technology/2024/01/20/openai-dean-phillips-ban-chatgpt/

23 https://blog.google/outreach-initiatives/civics/how-were-approaching-the-2024-us-elections/

24 https://www.washingtonpost.com/technology/2023/11/08/meta-artificial-intelligence-political-ads/

25 https://www.washingtonpost.com/technology/2024/01/22/ai-deepfake-elections-politicians/; https://www.washingtonpost.com/technology/2023/10/07/amazon-alexa-news-2020-election-misinformation/

26 https://www.newsguardtech.com/special-reports/ai-tracking-center/

27 https://www.theverge.com/23929233/lawrence-lessig-free-speech-first-amendment-ai-content-moderation-decoder-interview?ref=everythinginmoderation.co

28 https://www.cambridge.org/core/journals/perspectives-on-politics/article/donald-trump-and-the-lie/A438DF5A45FE78CB-2BC887859EFAB587

29 https://www.bloomberg.com/view/articles/2018-02-09/has-anyone-seen-the-president

30 https://www.npr.org/2023/11/17/1213746885/trump-vermin-hitler-immigration-authoritarian-republican-primary

31 https://www.cnn.com/2023/10/06/politics/trump-anti-immigrant-comments/index.html

32 https://www.nytimes.com/2023/12/19/us/politics/trump-immigrants-hitler-mein-kampf.html

33 https://www.theguardian.com/us-news/2018/aug/03/trump-enemy-of-the-people-meaning-history

34 https://ny1.com/nyc/all-boroughs/politics/2023/12/27/trump-poll-dictatorship-revenge-second-term

35 https://truthsocial.com/@realDonaldTrump/posts/111500556888904384; https://www.npr.org/2024/01/04/1218672628/the-trump-campaign-embraces-jan-6-rioters-with-money-and-pardon-promises

36 https://www.washingtonpost.com/dc-md-va/2024/01/04/fbi-conspiracy-jan-6-attack-misinformation/

37 https://globalextremism.org/post/colorado-court-decision/

38 https://www.washingtonpost.com/politics/2024/01/09/public-officials-death-threats-swatting-surge/

39 https://www.washingtonpost.com/politics/2023/05/06/bill-gates-maricopa-county-arizona-ptsd/

40 https://www.votebeat.org/arizona/2023/9/15/23874134/celia-nabor-maricopa-county-arizona-election-fraud-signature-verification-public-records/

41 https://www.brennancenter.org/our-work/research-reports/local-election-officials-survey-april-2023

42 https://www.washingtonpost.com/politics/2022/11/01/truth-about-election-fraud-its-rare

43 https://www.ajc.com/politics/eligibility-of-364000-georgia-voters-challenged-before-senate-runoff/3UIMDOVRFVERX-OJ3IBHYWZBWYI/; https://www.ajc.com/politics/georgia-voter-challenges-fall-short-with-few-ballots-thrown-out/SN-PHXD4YXVB7LMIL5N5L3RZPLA/; https://www.nytimes.com/live/2021/01/06/us/georgia-election-results

44 https://bhr.stern.nyu.edu/tech-big-lie; https://www.politifact.com/article/2022/may/04/faulty-premise-2000-mules-trailer-about-voting-mai/; https://apnews.com/article/2022-midterm-elections-covid-technology-health-arizona-e1b49d2311bf900f44fa5c6dac406762; https://www.youtube.com/watch?v=Nz6smxo-MkE

45 https://www.thegatewaypundit.com/2022/01/cleaning-voter-rolls-soros-founded-funded-eric-now-used-31-states/

46 https://www.npr.org/2023/10/20/1207142433/eric-investigation-follow-up-voter-data-election-integrity

47 https://www.washingtonpost.com/politics/trump-raffensperger-call-transcript-georgia-vote/2021/01/03/2768e0cc-4ddd-11eb-83e3-322644d82356_story.html

48 https://whoscounting.us/

49 https://www.nytimes.com/2022/05/30/us/politics/republican-poll-monitors-election-activists.html

50 https://whoscounting.us/wp-content/uploads/2023/11/EIN-12-Reasons-to-Leave-ERIC-2023-The-Dirty-Dozen.pdf

51 The investigative journalism organization, Documented, has posted documents and analysis about EIN and EagleAI Network: https://documented.net/investigations/meet-eagle-ai-the-cleta-mitchell-backed-project-for-maga-activists-to-file-mass-voter-challenges

52 An EagleAI "Capabilities Summary" dated May 1, 2023 and posted by Documented explains that the platform can be used by county election officials and also by "Individuals interested in voter roll accuracy and integrity" seeking "to ascertain whether voter registrations meet their state criteria for 'eligible voter.'" https://documented.net/media/eagle-ai-network-capabilities-study

53 https://www.nytimes.com/2023/12/01/us/politics/georgia-county-election-deniers-trump.html

54 https://documented.net/media/georgia-state-election-board-may-11-2023-letter-to-columbia-county

55 https://twitter.com/Jim_Jordan/status/1721674461408006431

56 https://www.congress.gov/event/118th-congress/house-event/115611/text?s=1&r=58

57 https://transparency.fb.com/metasecurity/threat-reporting

58 https://www.washingtonpost.com/technology/2023/06/06/disinformation-researchers-congress-jim-jordan/

59 https://www.washingtonpost.com/technology/2023/02/08/house-republicans-twitter-files-collusion/; https://thehill.com/opinion/technology/3852594-the-twitter-bias-hearings-point-to-favoritism-but-not-for-liberals/

60 Congressional Republicans reportedly have sought information from NYU's Center for Social Media and Politics and Tandon School of Engineering. Co-author Justin Hendrix is an associate research scientist and adjunct professor at Tandon. See https://www.washingtonpost.com/technology/2023/06/06/disinformation-researchers-congress-jim-jordan/

61 https://www.eipartnership.net/report

62 https://judiciary.house.gov/sites/evo-subsites/republicans-judiciary.house.gov/files/evo-media-document/EIP_Jira-Ticket-Staff-Report-11-7-23-Clean.pdf

63 https://www.eipartnership.net/

64 https://law.justia.com/cases/federal/appellate-courts/ca5/23-30445/23-30445-2023-09-08.html

65 https://www.npr.org/2023/08/01/1191318468/elon-musk-sues-disinformation-researchers-claiming-they-are-driving-away-adverti; https://apnews.com/article/elon-musk-media-matters-lawsuit-advertising-neonazi-1fe499daa600f513af27ffa68d2e8b91

66 https://www.techpolicy.press/transcript-house-hearing-on-dhs-and-cisas-role-in-securing-ai/

67 https://go.recordedfuture.com/hubfs/reports/ta-2023-0830.pdf; see also New York Times coverage: https://bit.ly/41K24kh

68 https://blogs.microsoft.com/wp-content/uploads/prod/sites/5/2023/11/MTAC-Report-2024-Election-Threat-Assessment-11082023-2-1.pdf

69 https://www.dni.gov/files/ODNI/documents/assessments/NIC-Declassified-ICA-Foreign-Threats-to-the-2022-US-Elections-Dec2023.pdf; https://www.nytimes.com/2023/12/24/us/politics/china-russia-election-interference.html

70 https://transparency.fb.com/integrity-reports-q2-2023/; https://blog.google/threat-analysis-group/tag-bulletin-q4-2023/

71 https://networkcontagion.us/wp-content/uploads/A-Tik-Tok-ing-Timebomb_12.21.23.pdf

72 https://www.nytimes.com/2024/01/08/business/media/tiktok-data-tool-israel-hamas-war.html

73 https://newsroom.tiktok.com/en-us/delivering-on-our-us-data-governance; https://newsroom.tiktok.com/en-au/the-truth-about-tiktok; https://www.tiktok.com/transparency/en-us/community-guidelines-enforcement-2023-3/

74 https://www.theguardian.com/technology/2023/dec/31/us-tik-tok-ban-federal

75 https://blogs.microsoft.com/wp-content/uploads/prod/sites/5/2023/11/MTAC-Report-2024-Election-Threat-Assessment-11082023-2-1.pdf

76 https://www.dni.gov/files/ODNI/documents/assessments/NIC-Declassified-ICA-Foreign-Threats-to-the-2022-US-Elections-Dec2023.pdf

77 https://www.nytimes.com/2023/09/27/technology/disinformation-law-european-union.html

78 https://fakty.afp.com/doc.afp.com.33WY9LF

79 https://www.wired.com/story/slovakias-election-deepfakes-show-ai-is-a-danger-to-democracy/; https://transparency.fb.com/en-gb/policies/community-standards/manipulated-media/

80 https://www.bloomberg.com/news/articles/2023-09-29/eu-s-ability-to-fight-disinfo-gets-fact-checked-in-slovakia?sref=LW7poGYk; https://www.reuters.com/world/europe/new-slovak-government-rejects-final-military-aid-package-ukraine-2023-11-08/

81 https://www.nytimes.com/2023/11/10/world/asia/south-korea-fake-news-disinformation.html

82 https://www.voanews.com/a/under-yoon-south-korea-defamation-cases-against-media-rise-/7388864.html

83 https://www.ntd.com/ccp-develops-disinformation-campaign-to-interfere-in-south-koreas-election_962277.html

84 https://www.washingtonpost.com/world/2023/09/26/hindu-nationalist-social-media-hate-campaign/

85 https://www.reuters.com/world/india/india-sets-up-govt-panel-hear-social-media-content-moderation-complaints-2022-10-28/

86 https://www.washingtonpost.com/world/2023/11/08/india-twitter-online-censorship/

87 https://www.hindustantimes.com/india-news/indialeaders-write-to-google-facebook-on-neutrality-during-2024-elections-101697137511412.html

88 https://www.statista.com/statistics/1229799/india-smartphone-penetration-rate/#:~:text=In 2023, the penetration rate,end of 2020 was Xiaomi

89 https://timesofindia.indiatimes.com/india/brs-complains-to-ec-on-congress-alleged-use-of-deepfake-technology-in-telangana-poll-campaign/articleshow/105621669.cms?from=mdr

90 https://www.vice.com/en/article/jgedjb/the-first-use-of-deepfakes-in-indian-election-by-bjp

91 https://www.boomlive.in/decode/madhya-pradesh-elections-polls-assembly-shivraj-singh-chouhan-kamal-nath-bjp-congress-ai-voice-clones-deepfakes-elevenlabs-24147

92 https://apnews.com/article/mexico-fake-news-social-media-elections-2024-487943383b8f57f5eb0cc9ae1324fcc1

93 https://www.csis.org/analysis/presidential-elections-and-fragmenting-political-landscape-mexico

94 https://edmo.eu/wp-content/uploads/2023/10/EDMO-TF-Elections-disinformation-narratives-2023.pdf

95 https://www.disinfo.eu/doppelganger/

96 https://www.epc.eu/content/PDF/2023/Disinformation_DP_-_Eiw_and_EMD.pdf

97 See, e.g., https://bhr.stern.nyu.edu/tech-content-moderation-june-2020?_ga=2.51005679.794379451.1705328068-727471783.1685995516; https://bhr.stern.nyu.edu/polarization-report-page; https://bhr.stern.nyu.edu/youtube-report; https://bhr.stern.nyu.edu/tech-big-lie

98 https://www.tandfonline.com/doi/abs/10.1080/10584609.2019.1668894

99 https://accountabletech.org/wp-content/uploads/Democracy-By-Design.pdf

100 https://facebookpapers.com/wp-content/uploads/2021/11/Insurrection_Redacted.pdf

101 https://www.technologyreview.com/2019/09/26/434/whatsapp-disinformation-message-forwarding-politics-technology-brazil-india-election/

102 https://www.wsj.com/articles/how-to-fix-social-media-11635526928#renee-diresta-circuit-breakers-to-encourage-1d04f057; https://accountabletech.org/wp-content/uploads/Democracy-By-Design.pdf

103 https://www.lawfaremedia.org/article/warnings-work-combating-misinformation-without-deplatforming; https://www.usenix.org/conference/usenixsecurity21/presentation/kaiser

104 https://www.nytimes.com/2024/02/08/business/media/google-aai.html; https://www.nytimes.com/2024/02/06/technology/meta-ai-standards-labels.html.

105 https://www.citizen.org/article/tracker-state-legislation-to-protect-election-officials/

106 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.

107 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

108 https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en

NYU Stern Center for Business and Human Rights
Leonard N. Stern School of Business
44 West 4th Street, Suite 800
New York, NY 10012
+1 212-998-0261
bhr@stern.nyu.edu
bhr.stern.nyu.edu

Center for Business
and Human Rights