PREVENTING TECH-FUELED VIOLENCE:
What online platforms can do to ensure they do not contribute to election-related violence

Yael Eisenstat, Senior Policy Fellow, Cybersecurity for Democracy
Justin Hendrix, CEO and Editor, Tech Policy Press
Daniel Kreiss, Principal Researcher, UNC Center for Information, Technology, and Public Life
On the heels of reporting that far right extremist militias are once again organizing on Facebook in advance of the 2024 U.S. presidential election, it is urgent for platforms to assess – and immediately act on – threats to the peaceful conduct of elections and the holding and transfer of power.

On March 26th, 2024, a working group of experts met to discuss the relationship between online platforms and election-related political violence. The goal was to provide realistic and effective recommendations to platforms on steps they can take to ensure their products do not contribute to the potential for political violence, particularly in the lead-up and aftermath of the U.S. general election in November, but with implications for states around the world.

Relying on online platforms to “do the right thing” without regulatory and business incentives that reinforce pro-democratic conduct may seem increasingly futile, but we believe there remains a critical role for independent experts to play in both shaping the public conversation and shining a light on where we believe these companies can act more responsibly.

Here are seven recommendations in brief:

1. **Prepare for threats and violence:** Platforms must develop robust standards for threat assessment and crisis planning, and ensure they have adequate, multidisciplinary expertise, and collaboration across teams. They should prepare for potential violence by engaging in scenario planning, crisis training, and engagement with external stakeholders, with as much transparency as possible.

2. **Develop and enforce policies to meet the threat:** Platforms should enforce clear and actionable content moderation policies that address election integrity year-round, proactively addressing election denialism and potential threats against election workers and systems while being transparent about enforcement and sensitive to local contexts.

3. **End exceptions for high-value users:** Politicians and other political influencers should not receive exemptions from content policies or special treatment from the platforms. Platforms should enforce their rules uniformly, and be especially attentive to preventing such users from monetizing harmful content and promoting election mis- and disinformation.

4. **Resource adequately:** Platforms should scale up teams focused on election integrity, content moderation, and countering violent extremism while ensuring that outsourcing to third parties in the growing trust & safety vendor ecosystem doesn’t impede expertise and rapid response to emerging threats and crisis events.

5. **Transparency on content moderation decisions:** Platforms must clearly explain important content moderation decisions during election periods, ensuring transparency especially when it comes to the moderation of high profile accounts. Platforms should establish a crisis communication strategy, and be ready to explain cooperation with fact-checkers and other sources that inform enforcement actions.

6. **Collaborate with researchers, civil society and government where appropriate:** Platforms should collaborate with independent researchers, civil society groups, and government officials to enhance the study and mitigation of election misinformation, within the context of applicable laws. Platforms should maximize data access, and proactively counter false claims about such collaborations.

7. **Develop industry standards:** Platforms should establish industry standards that respect diverse approaches to moderating speech but prioritize protecting elections. An industry body, similar to the Global Internet Forum to Counter Terrorism (GIFCT) could help develop clear threat assessment capabilities, enforce consistent policies, and facilitate collaboration to defend against democratic threats and political violence.
INTRODUCTION

Across the United States, election officials are preparing for the worst. In Arizona, officials are distributing tourniquets, conducting active-shooter drills, and training staff how to defend themselves if they are attacked. In Maine, election officials are being trained to conduct threat analysis and reporting. And, in multiple states, such as New Mexico, officials are racing to put in place rules that would limit the presence of firearms at polling sites.

On March 26th, 2024, a working group of experts on social media, election integrity, extremism, and political violence (see “contributors” below) met to discuss the relationship between online platforms and election-related political violence. The goal was to provide realistic and effective recommendations to platforms on steps they can take to ensure their products do not contribute to the potential for political violence, particularly in the lead-up and aftermath of the U.S. general election in November, but with implications for states around the world.

While perpetrators of violence are responsible for their own actions, online platforms are tools for distributing messages and narratives intended to incite political violence, and for recruiting and organizing groups that may carry out violent acts. Whether platforms are prepared for another election cycle that may end in bloodshed – in the U.S. or any other of the numerous countries conducting democratic elections in 2024 – is a significant concern.

Scope: For this paper, we use the term “online platforms” to refer to social media companies and a variety of messaging apps. We did not focus on AI developers or generative AI companies, which also have serious challenges that merit equal consideration with regards to the potential fueling of political violence. While there are dozens of elections happening around the globe this year, we used the U.S. general election in November, and related threats of political violence, as key reference points.
THE CURRENT LANDSCAPE

Democracies around the globe are increasingly under threat and even assault by anti-democratic leaders, parties, and movements. Concerted, strategic campaigns to delegitimize elections by questioning – without evidence – the safety, security, and process of votes are tools in the arsenal of those seeking to illegitimately gain or hold onto power. Another tool is the strategic targeting, including threats of violence, of election administrators. And, in extreme cases, anti-democratic actors use violence to prevent the peaceful transfer or holding of legitimate, democratically-elected power.

Platforms are not the only forms of communication that can weaken democratic processes and lead to violence, such as the speeches and statements of political leaders covered in cable and local news broadcasts. But, platforms are often central forums for anti-democratic leaders, parties, and movements to circulate their delegitimizing rhetoric and sabotage elections and their administration. Social media companies are powerful precisely because their algorithms boost certain content, and they provide users opportunities to amplify these claims and ways to inspire and organize election-related violence. Technology itself does not cause political leaders to seek to illegitimately gain or hold onto power, and in fact it can be used by both pro- and anti-democracy forces. However, online platforms provide a set of tools both for would-be authoritarians to subvert the will of the people and extreme voices to influence political leaders, which makes them worthy of urgent consideration and reform.

The role of online platforms during the January 6, 2021 attempted coup at the U.S. Capitol and the January 8, 2023 sacking of government offices in Brazil are cases in point about the imperative of addressing these issues now. There is substantial legislative and regulatory action in the European Union and Brazil, as well as in countries such as the United Kingdom that have passed online safety legislation. The United States still largely relies on platforms to self-regulate, but practices, to date, have not adequately met the challenges anti-democratic forces pose to democratic functions.

Despite these violent events and the ongoing threats to free and fair elections in democracies around the world, online platforms have rolled back a number of election protections.

In the wake of the January 6 insurrection, online platforms took comparatively bold actions to protect democracy. The major social media companies deplatformed former President Trump, after years of election lies. Less than two years later, however, these same online platforms failed to protect Brazil from a similar situation. Brazil’s former president, Jair Bolsonaro, spent months undermining the country’s election and claiming that it would be stolen from his supporters. False claims about the election propagated widely on social media platforms ahead of the events in Brasilia, and unlike Trump, Bolsonaro was never deplatformed.

Despite these violent events and the ongoing threats to free and fair elections in democracies around the world, online platforms have rolled back a number of election protections in the United States and elsewhere – including the dozens of countries and regional bodies that are holding elections this calendar year. What’s more, in the U.S. four years later we have no more transparency into what, if anything, they are doing. The process, the design choices at play, and the trade-offs remain shrouded in secrecy.
The experts gathered in this working group shared a number of baseline assumptions about the role of platforms in election violence. First, broadly, it is clear that democracy is under assault by various illiberal and anti-democratic movements around the world. This has been well-documented by various research bodies and includes everything from the rise of nationalist parties to violence directed at racial, ethnic, and religious minorities. Second, more relevant for our arguments here, is the rise of anti-democratic political leaders and members of parties that engage in concerted campaigns to delegitimize elections and sabotage election administration. These political elites have weaponized the rhetoric of election safety and security to undermine election integrity and, ultimately, people’s ability to choose at the ballot box.

Lessons from January 6th insurrection

The results of multiple investigations, including that of the House Select Committee to Investigate the January 6th Attack on the United States Capitol, offer important lessons about the connection between social media and political violence related to elections. A draft report produced by one of the Committee’s investigative teams contains a number of such lessons, including for platform planning, policy, and operations.

With regard to planning, the report makes clear that in 2020, the platforms were unprepared for the possibility of violence in the post-election and even post-inauguration period, having focused primarily on threats to the voting process and other concerns about mis- and disinformation before the election. After the election, platforms like Facebook and Twitter relaxed policies that could have mitigated the spread of violent incitement, demonstrating a gap in their understanding of the threat and why the post-election period was particularly ripe for unrest, despite numerous warnings from outside experts.

When it comes to policy, the Committee’s report points to specific failures across major social media platforms, such as Twitter’s refusal to implement policies against coded incitement to violence and Facebook’s inadequate policing of disinformation and violent content within groups associated with the "Stop the Steal" movement. The lack of consideration for coded incitement, in particular, is an ongoing problem. Groups whose speech may not technically run afoul of platform policies may still be engaged in speech that is intended to build capacity toward future acts of incitement, or that is heard as a call to violence even if that call is not explicit.

And when it comes to operations, the Committee’s draft social media report noted a number of failures among the platforms, including technical ones. For instance, it recounts how a technical failure meant that many groups and pages on Facebook were not penalized for hate speech for months.

Another major learning which the Committee addressed and which is well studied by academic researchers is the relationship between major platforms and fringe sites where more extreme speech, organizing, and planning takes place. Signals issued on Twitter (“Be there, will be wild!”) are regarded as direct instructions by users on sites such as Patriots.win. The dynamic of this relationship, a kind of cross-platform call and response, deserves more attention from researchers and the trust and safety field.
WHAT ONLINE PLATFORMS CAN, AND SHOULD, DO

We understand that online platforms are not responsible for individual or group actions and cannot solve society’s social and political problems. Drawing on our collective experience as academics, policy advocates, and former social media employees, our group came to agreement on a number of important steps that online platforms could, however, take to ensure they do not contribute to election-related political violence. Relying on online platforms to “do the right thing” without the proper regulatory and business incentives in place may seem increasingly futile, but we believe there remains a critical role for independent experts to play in both shaping the public conversation and shining a light on where we believe these companies can act more responsibly.

While the January 6th U.S. insurrection was an anchoring event for our discussion, these recommendations are broad enough that they could be adopted by a variety of social media and messaging platforms and apply to elections and political violence globally.

Most companies have publicized their 2024 election integrity policies (see appendix), but we remain concerned that they are falling short on their responsibilities to ensure they are not exacerbating tensions that could spill over into violence. A number of recommendations that experts have made over the years remain unheeded, and we have seen backsliding on various policies, trust and safety resourcing, and collaboration with academics and civil society.

This paper does not address the larger media ecosystem, legislative shortcomings, or systemic business and design issues of online platforms. But there are concrete, actionable steps they can take right now to help ensure they do not, again, contribute to political violence. In addition to the election integrity policies various platforms have already announced, we want to help crystalize the arguments around key things that can be done or improved upon to protect against political violence. This list is by no means exhaustive, but they are the most salient points that our group agreed can and should be done.
RECOMMENDATIONS
Before we can discuss policy and enforcement recommendations, it is critical to understand how companies assess and prepare for potential violence. While there has been increasing public attention on transparency requirements for content decisions, we have not seen enough discussion of how companies develop their policy and enforcement decisions and prepare for crisis moments.

We believe more robust discussion must occur around these questions:

- **How do companies assess threats?** Are their assessments grounded in social science and/or best practices from industries that regularly conduct threat assessments?

- **Who is included in conducting threat assessments?** Do they include a multi-disciplinary set of expertise and knowledge from people with real-world experience in national security, political violence, human rights, conflict mediation, communications, etc? Do they include expertise from people who are the most vulnerable to or targeted by violent threats?

- **How do those assessments factor into platform policy and enforcement decisions?** Is a suite of potential policies developed based on the results of threat assessments? Are the assessment teams siloed, or do they work seamlessly across the content, public policy, legal, growth and enforcement teams?

- **When do platforms prepare for crises?** Do platforms use threat assessments to engage in scenario planning, to prepare in advance for how to react to any number of potential violent or volatile situations?

- **How often do platforms update their threat assessments and ensure they factor into policy decisions?** For example, ahead of the January 6th insurrection, Twitter was focusing on formed militias as the only groups that could incite violence. The company was not nimble enough to update that assumption as it became apparent that smaller groups and individuals were inciting followers to violence as well.

Without an understanding of the above, the public is left to take companies at their word when they explain their policies, or lack thereof. For example, in justifying a 2023 reversal of its 2020 election denialism policies, **YouTube wrote**: “In the current environment, we find that while removing this content does curb some misinformation, it could also have the unintended effect of curtailing political speech without meaningfully reducing the risk of violence or other real-world harm [emphasis added].” **Facebook basically said the same thing**. There was no explanation of how YouTube assessed the “current environment” or the “risk of violence.”

While there are internal tensions over how much leadership might want to actually engage in this level of research and preparation, any company that claims to want to protect the integrity of elections and not help fuel violence must do so. We recommend that online platforms develop adequate standards for risk assessments and scenario/infrastructure planning. They must proactively prepare for a range of future threats, which require not only assessing these threats, but engaging in crisis training, scenario planning, collaboration across various workstreams, and engagement with relevant external stakeholders. For particularly high-stakes threats, there should be an all-hands-on-deck playbook, and teams must study, practice, and continue to update the playbook. And these companies must be prepared to resource adequate levels of enforcement and support for any of the potential crisis scenarios.

**DEVELOP AND ENFORCE POLICIES TO MEET THE THREAT**

Content moderation remains one of the most important, albeit imperfect, tools for protecting against the spread of false information about elections and incitements to violence. Whether it’s conspiracies such as “The Great Replacement theory” or lies about elections being “stolen”, having clear, actionable content policies that are enforced year-round is critical.
In addition to the various existing election integrity policies, we recommend:

- Preventing online incitement to political violence must be a year-round job, not just something that is only considered and resourced in the weeks leading up to a heated election. In past years, social media companies have disbanded their election “war rooms” or stopped prioritizing election integrity work following the end of voting, missing the critical period leading up to the violence we saw in the U.S. on January 6th, 2021 or in Brazil on January 8th, 2023. There are potential violent touch points throughout the year, directly tied to elections, and monitoring for those threats and enforcing election and political violence policies year-round must become the norm.

- All lies about elections should be covered, including lies about the 2020 U.S. election; companies should not pick and choose which elections merit enforcement. In June, 2023, YouTube announced it was reversing its content policies regarding lies about the 2020 election, and in August Meta followed suit, allowing political ads disputing past election results and carrying lies about fraud or election rigging to resume. These companies joined Twitter, which reversed its policy in early 2022. We already see rhetoric about the 2020 election being used to foment anger, and lies about those results will surely be used to prime those who believe the 2020 election was “stolen” to act if the 2024 results do not go their way. Election deniers will continue to build networks and audiences for groups and pages that will eventually convert their 2020 claims to target 2024. These lies directly led to violence on January 6th, and there is no reason to believe that threat no longer exists.

- Consider the local/national context when developing policies. Not every policy has to be “scalable” to the entire world. While each country’s elections will have lessons and dynamics that apply globally, rejecting potential content policies in one country because they would not scale to the rest of the world continues to be problematic.

- Proactively and clearly categorize what will constitute or contribute to “incitement to violence.” Content policies that only focus on the intent of the speaker risk missing the important question of how that speaker’s message is being interpreted, and acted upon, by their followers. Dog whistles and coded language must be understood and considered, especially when they inspire or incite followers of influential voices. A playbook exists: after January 6th, Twitter began implementing an internal policy around “coded incitement,” although they later rolled it back. Platforms should continue developing these policies, in coordination with outside researchers and experts, and enforce them year-round.

- Develop clear policies for threats against election administrators, judges, and other public actors. There has been a substantial rise in violent threats against public officials, including through doxing and mobilization of harassment toward specific election officials and agencies. Threats against those who work to protect and administer elections should be treated as forms of violence and enforced accordingly. These should have a particularly low bar for enforcement.

- Enforce policies against all types of content, text, video and audio, including artificially-generated content. Watermarking and labeling are not enough, and all election and violence-related policies must apply.

- Develop “break glass” measures to deploy if violence appears to be imminent or does erupt. As mentioned above, preparing for potential violence requires understanding and assessing the threat landscape, engaging with outside experts, and preparing for a variety of potential scenarios. That preparatory work should include designing playbooks of policies and enforcement mechanisms for the worst-case scenarios, like we saw on January 6th, and proactive approaches to content moderation when those situations erupt.
Content policies, however, are only as good as their enforcement. Companies often point to their written policies when challenged about election denialism or violent threats on their platforms, but there is a wide body of evidence that policies have often not been enforced equally or at scale.

In addition to continuing to push companies to enforce their own policies, we recommend:

- **Broaden enforcement to include threats that have demonstrated potential to catalyze violence in the future.** One example is a narrative such as the “great replacement theory,” that even if not used to call for violence in the moment has led to violence. The same applies to narratives that have already proved to incite violence in the past, such as 2020 election denialism.

- **Enforce these policies in a timely manner:** as long as content moderation remains a key mechanism, timeliness is a critical component. Taking down content days or weeks after a violent event, or after it has already been amplified and spread to millions, is not only too late, but it mostly serves only to make it more difficult for researchers and journalists after the event already occurred.

- **Continually test and check that automated systems and classifiers are functioning properly.** The January 6 Committee’s final report demonstrated how a technical failure at Meta resulted in groups and pages not being penalized for violating hate speech rules for months. Upon addressing the technical failure, Facebook's actions resulted in approximately 10,000 groups receiving strikes for policies related to hate speech and misinformation. Additionally, about 500 groups associated with the "Stop the Steal" movement were immediately removed for policy violations.

- **Err on the side of caution during a violent event, but be transparent after the fact:** During a violent event, the potential to “over moderate” is high. There is no perfect solution for this, but companies must be prepared to strictly enforce the rules and make tough decisions while the event is unfolding. Where over-enforcement occurs, be transparent after the fact, in as timely manner as possible, about those mistakes.

---

### END EXCEPTIONS FOR HIGH-VALUE USERS

Historically, most social media platforms have treated ‘high-value’ users – loosely defined as those with large accounts or of public importance – as special exemptions to the rules governing other users. In the political domain, this has meant that politicians, for instance, could do things that would get other users suspended or even barred, such as posting vaccine misinformation. The justification often offered by platforms is that people deserve to see the posts of those vying to represent them – or, those that are already politically, religiously, culturally, or socially important. Of course, not taking action against these users likely also has economic or political advantages for technology companies (such as not deplatforming popular figures and being open to charges of censorship, or protecting those who afford the monetization of content.)

Regardless of these justifications, this is the wrong approach. Political elites are uniquely influential over their followers, who often share a set of partisan, political, and other identities and interests with them. As we know from decades of research, what political elites say and do shapes what their supporters say and do. Look no further than how the lie that the election was stolen took root throughout the MAGA movement during the 2020 U.S. presidential election. Or, how supporters of Jair Bolsonaro were called to violent action by the loser of the presidential election.
Platforms should expect even more of their high-value users, not less. They have unique responsibilities because they are uniquely capable of influencing their supporters. Even more, political elites exploit platform speech carve outs in the attempt to undermine the voices of citizens at the ballot box – which was directly at issue in the U.S. and Brazilian cases. This includes influencers and other high-reach users who might not be running for office. The influence they potentially have over their networks increases the responsibilities they have to use platforms in accord with civic integrity policies.

Platforms should consider a range of options to address this reality:

- They must uniformly apply and fairly enforce their existing rules for all users.

- High-value users (along with other users)—whether in the domain of politics or not—should be barred from monetizing harmful content. This would remove that incentive from speech that undermines democratic processes or plays a role in potential violence.

- High-value users should be barred from paid promotions of content, such as targeted advertising, that undermines democratic processes such as election administration, preventing these users from using commercial means of promoting content likely to play a contributing factor in political violence.

- In times of heightened political contestation, the content of direct parties to political contests should be reviewed prior to allowing the amplification or monetization of content (similar to Snap’s model of prior review).

We realize that determining who is a politician and where we draw the line around a high-value user is difficult – such as political influencers. Our stance is that companies should draw bright lines around election disinformation and content that potentially will lead to election violence, and enforce their policies uniformly. If there are exemptions, they should be well justified and transparent, and should be in line with not only threat assessments, but also industry standards.

**4 RESOURCE ADEQUATELY**

Platforms must ensure they are appropriately resourced to handle the serious risk of election-related political violence. The posture of the major social media platforms, at the time of writing, does not suggest this is the case. Rather than reducing the number of trust and safety professionals on staff, platforms need to scale up teams focused on election integrity, content moderation, and countering violent extremism. This includes being ready to do sufficient monitoring, rapid response, and proactive detection of emerging risks. While there is a move at some platforms to outsource these functions to third parties in the growing trust and safety vendor ecosystem, platforms should be certain that any such arrangements do not impede the speed at which they can respond to threats, and that organizationally the lines of communication are not blocked by layers of client-customer hierarchy and account management.

Ultimately, the public needs to understand what degree of safety spending is appropriate as a percentage of revenue for major social media platforms. While this number may be difficult to determine across platforms, current levels of investment would appear to be too low given the possible societal impacts of failure, particularly with regard to elections.
While this may be an imperfect way to think about the resources necessary to ensure platforms are doing enough, an increase in these figures should signal that concerns about possible political violence are a high priority at the platforms.

In general, this recommendation is in line with the draft communication “on the mitigation of systemic risks for electoral processes pursuant to the Digital Services Act” issued to online platforms by the European Commission in March 2024. The Commission encourages platforms to adequately resource teams and operations related to mitigating such risks, and to calculate what is adequate based on an assessment of the threat for each specific election context. That process requires working with outside experts.

5  TRANSPARENCE ON CONTENT MODERATION DECISIONS

Social media platforms should be prepared to clearly explain content moderation decisions during the election period. They must be transparent about enforcing policies, particularly for high-profile accounts, and acknowledge errors in judgment and move to rectify mistakes swiftly. The controversy over the moderation of a New York Post story about Hunter Biden’s laptop in 2020 demonstrates the need for readiness in future similar situations.

Platforms should:

- Ensure all moderation actions—whether reducing virality, removing content, or suspending accounts—are explained clearly and promptly. Notices of such actions should reference specific policies and the behavior that led to the action, and in certain cases, the decision not to act. Initial communications are critical and should be detailed to avoid perceptions of partisanship.

- Identify a suitable spokesperson and internal structures to handle communications in a crisis posture. A PR executive’s tweet or a blog post on a corporate site alone is insufficient; substantive evidence should accompany the announcement and executives should be prepared to promptly answer media inquiries. The approach should mimic the readiness of public officials in emergency scenarios.

- Be prepared to explain protocols for cooperation with third party fact checkers and sources of other threat information and analysis. Diagrams for the workflow should be available for public scrutiny to ensure transparency and efficiency.

6  COLLABORATE WITH RESEARCHERS, CIVIL SOCIETY, AND GOVERNMENT WHERE APPROPRIATE

The past three years have seen a concerted effort to attack and disrupt the work of independent researchers and government officials engaged in efforts to study and mitigate false claims related to elections. But such collaboration is crucial, not only for real time insight into the emergence and propagation of false claims, but for the study of such phenomena intended to produce insights that may help ameliorate the problem in the future.

Platforms should:

- Reassess protocols for communication and collaboration with independent researchers, civil society groups, and government officials in order to maximize the benefit of their efforts while minimizing claims that such collaboration is itself anti-democratic.

- Make as much data access for independent research available as possible during the election period in order to support analysis in as close to real-time as possible.

- Be prepared to counter false claims about such collaboration by being as transparent as possible about the mechanics and protocols in place.
DEVELOP INDUSTRY STANDARDS

There should be a set of industry standards that respects a diversity of approaches to speech and expression, but has a strong baseline of protecting civic processes. There are good reasons for a diversity of approaches to expressive content on platforms. Having multiple approaches means many different types of forums for speech (content on Reddit, for example, is different from content on Facebook, which is different from content on Wikipedia). It also means differing rules and enforcement mechanisms (such as algorithmically determined content moderation versus human moderation.)

That said, there should be a strong baseline of defending against democratic threats and political violence. There are potential models for how to do this, including the creation of a body similar to the Global Internet Forum to Counter Terrorism, an industry-led content moderation entity that collaborates with governments to identify terrorist-related content. Or the creation of a body that helps multiple platforms collaborate with election administrators and law enforcement might help identify and protect against political violence.

At the very least, we believe a standards body could help the development and implementation across the platform industry of a few of the recommendations that we detail above:

- **Develop clear threat assessment capabilities** and provide a degree of transparency into them for researchers, journalists, and civil society stakeholders.

- **Move from a narrow focus on ‘elections’** to ongoing monitoring of threats to democratic processes and the potential for political violence.

- **Extend policies and enforcement to all election disinformation**, including illegitimate claims of electoral fraud in past elections that are used to undermine future ones.

- **Develop country-specific expertise** with a clear understanding of the local context.

- **Clearly delineate policies** that consistently define, and enforce against, incitements to violence and threats against election administrators.

- **Have clear action plans for when violence does occur.**

- **Have policies cover all users**, but with special scrutiny over high-value and publicly important users.
CONCLUSION

The complex interaction between political leaders, media, social media, and the public in election periods requires diligence on the part of all stakeholders, particularly in contexts where the potential for political violence is high.

Platforms, which create incentives that affect the behaviors of all stakeholders, have a particular duty to ensure they are prepared for helping to protect legitimate and fair democratic elections and the peaceful transfer or continued holding of power in their aftermath.

Protecting democratic processes must be an urgent priority, particularly in the democracy that created many of the conditions for the industry’s success.
CONTRIBUTORS

Scott Althaus, Merriam Professor of Political Science and Professor of Communication, University of Illinois Urbana-Champaign

Susan Benesch, Executive Director at Dangerous Speech Project and Faculty Associate at Berkman Klein Center for Internet and Society

Laura Edelson, Assistant Professor at Northeastern University; Co-Director at Cybersecurity for Democracy

Jacob Glick, Georgetown Institute for Constitutional Advocacy and Protection; Former Investigative Counsel, January 6th Select Committee

Ellen Goodman, Distinguished Professor at Rutgers Law School and Co-Director, Rutgers Institute for Information Policy & Law (RIIPL)

Dean Jackson, Principal, Public Circle LLC; Former January 6 Investigative Analyst

Jaime Longoria, Manager of Research and Training, Disinfo Defense League

Anika Collier Navaroli, Senior Fellow, Tow Center for Journalism; former senior policy official at Twitter and Twitch

Lorcan Neill, University of North Carolina at Chapel Hill; Knight Fellow at The Center for Information, Technology, & Public Life

Spencer Overton, Professor of Law and Multiracial Democracy Project Director, George Washington University

Katie Paul, Director, Tech Transparency Project

Kate Starbird, Associate Professor at University of Washington; Director of Center for an Informed Public
## APPENDIX

<table>
<thead>
<tr>
<th>Current Platform Policies Related to Politics and Elections</th>
<th>Facebook</th>
<th>Instagram</th>
<th>X</th>
<th>YouTube</th>
<th>TikTok</th>
</tr>
</thead>
<tbody>
<tr>
<td>Includes civic integrity or content moderation policies that:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specifically prohibit content that promotes political or election-related violence</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Prohibit false claims related to past elections</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Deplatform anti-democratic political leaders</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
</tr>
<tr>
<td>Requires labels for AI-generated content</td>
<td>Yes*</td>
<td>Yes*</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*Based on a proposed Meta AI policy released in April 2024. Meta has yet to confirm that labeling efforts for AI-generated content have been enacted.