

ELECTION INTEGRITY ROADMAP FOR SOCIAL MEDIA PLATFORMS

HIGHLIGHTS

The U.S. is entering an election season fraught with unprecedented risks: a global pandemic, widespread online voting disinformation, leaders attempting to sow distrust in the integrity of the election, and the potential for online communities to foment real-world violence in the wake of election results being called. While every institution and citizen has a role to play to curb these threats, social media giants have an outsized influence on our modern information ecosystem -- and an added responsibility to address these challenges. Companies are taking important steps, but significant gaps and vulnerabilities remain.

To that end, this Election Integrity Roadmap was designed in consultation with technologists, civil rights leaders, and disinformation experts to help social media companies responsibly navigate the 2020 election season, from early voting through the official certification of results. Recognizing the difficulty of implementing changes at scale, the Roadmap is grounded in platforms' existing policies, tools, and technologies – and is limited in scope to address election-related threats only, not political misinformation issues more broadly. As Americans begin to cast their ballots in this unprecedented election, the integrity of democracy is at stake.

Social media companies have a responsibility to act and meet this moment, including by:

- > **As early voting begins: Implementing a clearly defined Election Integrity Strike System.** To deter the worst vectors of disinformation at the Account, Page, Group and website level, the Election Integrity Strike System offers specific tiered restrictions and penalties. It would progressively limit the reach of serial disinformers and provide context to help inform and inoculate other users, defanging the worst actors before the most volatile period of the election season.
- > **As Election Day nears: Implementing a Platform Poll Watchers program.** Just as election observers are deployed to ensure the legitimacy of democratic processes and report problems, Platform Poll Watchers would help ensure the integrity of the online information sphere. Platforms should create specialized verification labels for state Election Directors and nonpartisan civil society groups with expertise in voting rights, election law, or online disinformation, allowing them to promote credible information, flag specific pieces of misleading content, and counter-message false narratives in real-time. As in the context of COVID-19, these trusted accounts would serve as the first line of defense against misinformation.
- > **As polls close: Implementing Violence Prevention Preclearance.** Just as the Voting Rights Act required certain states to preclear new voting laws, social media platforms should implement Violence Prevention Preclearance for election-related posts from highly influential accounts. These posts would be subject to proactive detection and rapid human review, preempting content that would violate incitement of violence or civic integrity policies. Posts that haven't been checked within 10 minutes would be published with a flag noting review is pending.

ELECTION INTEGRITY ROADMAP PROPOSED TIMELINE

All proposals would be phased in at the listed dates and remain in place until election results are certified.

AS EARLY VOTING BEGINS

- > NOW - Eliminate [fact-checking exemptions for politicians](#) on posts about voting, election integrity, or results.
- > **SEPTEMBER 14 – Implement a clearly defined *Election Integrity Strike System* aimed at deterring and penalizing the worst vectors of disinformation at the account, page, group, and website level.**
- > SEPTEMBER 14 – Start correcting the record with specific [fact-checking treatments both in users’ feeds and via retroactive alerts](#) to users previously exposed.
- > SEPTEMBER 14 – Begin expanding context and friction to educate voters and slow the spread of viral lies, including by appending authoritative information to all voting-related posts, proactively surfacing key information about the origins of content and actors disseminating it, and [prompting users to think critically about the accuracy of election-related information before they share content](#).
- > SEPTEMBER 14 – Release [weekly transparency reports](#) on civic integrity enforcement.

AS ELECTION DAY NEARS

- > **OCTOBER 1 – Implement a *Platform Poll Watchers* program, empowering state Election Directors and NGOs specializing in voting rights, election law or disinformation to promote credible information, flag specific pieces of misleading content, and counter-message false narratives in real-time.**
- > OCTOBER 1 – Promote [voter assistance tools](#) from the nonpartisan Election Protection coalition, including the [866-Our-Vote](#) hotline and [ProtectTheVote.net](#) volunteer program.
- > OCTOBER 1 - Introduce additional context, friction, and transparency, including by [appending flags to all new accounts/groups and those that change names](#) and implementing virality circuit breakers if possible.
- > OCTOBER 1 - Escalate proactive information distribution, including by [beginning voter education about the likely timeline](#) for election results, providing free advertising credits to state election officials, and sending users push alerts about how to vote early or by mail.
- > OCTOBER 1 – Detail steps being taken to train all content moderators on election-related policies and all plans to surge moderation capacity in the lead-up to November 3 and following days.
- > OCTOBER 20 - Tighten advertising standards to prevent foul play, including by [applying the highest transparency standards to all ads, not just “political” ones](#), and placing a freeze on new advertisers until post-election.
- > OCTOBER 20 - Limit algorithmic amplification, including by temporarily turning off algorithmic curation in default settings, turning off YouTube’s autoplay and “Up Next” engine, and [halting Facebook Group recommendations and Twitter Trending Topics](#).

AFTER POLLS CLOSE

- > **NOVEMBER 3 – Implement *Violence Prevention Preclearance* for election-related posts from the highest-reach accounts for proactive detection and rapid human review, preempting content that would violate incitement of violence or civic integrity policies before they can do harm.**
- > NOVEMBER 3 – Strictly [adhere to Associated Press or National Election Pool results and standards for partial reporting](#), providing as much context to users as possible about the current circumstances and what to expect moving forward.