Social Media Election Policies: The Good, the Bad and the Misinformed

Published: 02.23.2023

False and misleading information threatens the integrity of our elections even when those elections are months or years away. When politicians and other public figures <u>promote conspiracy</u> theories about "stolen" elections, they <u>undermine public trust in democratic processes</u>, <u>normalize baseless challenges to election outcomes</u>, stoke <u>harassment of elections officials</u>, and <u>incite violence</u> against political opponents. With the 2024 U.S. Presidential election cycle on the horizon, election deniers and spreaders of misinformation will continue to exploit social media platforms.

The ADL Center for Technology & Society (CTS) evaluated five major social media platforms' election misinformation policies and enforcement after the 2022 midterms concluded. Before the most recent U.S. election cycle, <u>ADL</u> found most major social media platforms had inadequate policies for preventing election misinformation.

We monitored accounts on Facebook, Instagram, TikTok, Twitter, and YouTube operated by candidates, political media personalities, and news outlets between October 18 and November 18, 2022. We reported false and misleading non-advertising content to these platforms, and recorded examples of platforms either taking action or declining to do so (see the Appendix). While platforms' policies and enforcement affect elections around the world, our analysis focuses on their impact for U.S. elections. Our findings and recommendations for platforms and policymakers to consider ahead of future elections are:

- Policy enforcement was inconsistent across and within platforms. Platforms allowed some
 content that violated one or more election misinformation policies to remain online. In other
 cases, similar or even identical pieces of content were treated differently without any clear
 justification.
- Platforms' policies are inadequate for preventing election misinformation. Vulnerabilities and loopholes persist, including accounts that post screenshots containing misinformation from fringe platforms and links that direct others to false and misleading content.
- Post-election misinformation remains concerning. Election denial narratives grow and develop between elections. The January 6th House Select Committee emphasized in its final report that far-right extremists who claim the 2020 election was stolen are an <u>ongoing threat</u>. Platforms need to be proactive in countering those narratives.

Although the 2022 elections were largely undisturbed, platforms cannot be complacent about combatting election misinformation; more than <u>170 election deniers</u> won their races nationwide. All but two of the 123 Congressional incumbents who voted against certifying the 2020 presidential election results <u>won re-election</u>.[1] Many election-denying candidates lost by only razor-thin margins, and a <u>few refused to concede</u>.

With Facebook's decision to reinstate Trump's <u>account</u>, he once again has access to the world's largest social media platform to amplify false claims that his re-election in 2020 was stolen or incite violence like he did prior to the U.S. Capitol insurrection on January 6, 2021. At the same time, <u>platforms are laying off staff charged with combating misinformation</u>, leaving them more vulnerable to manipulation

in the future. As we clearly saw in the <u>lead up</u> to the January 6th insurrection, election misinformation policies matter even after voting ends.

Election Misinformation Enforcement by Platform

Twitter

The Good:

- Prevents users from liking, sharing, or commenting on tweets that have a "misleading" label
- Links to authoritative sources for flagged content

The Bad:

- Inconsistent enforcement
- No restrictions on problematic search terms
- Users can share false and misleading claims in screenshots
- Labels have unofficial expiration dates

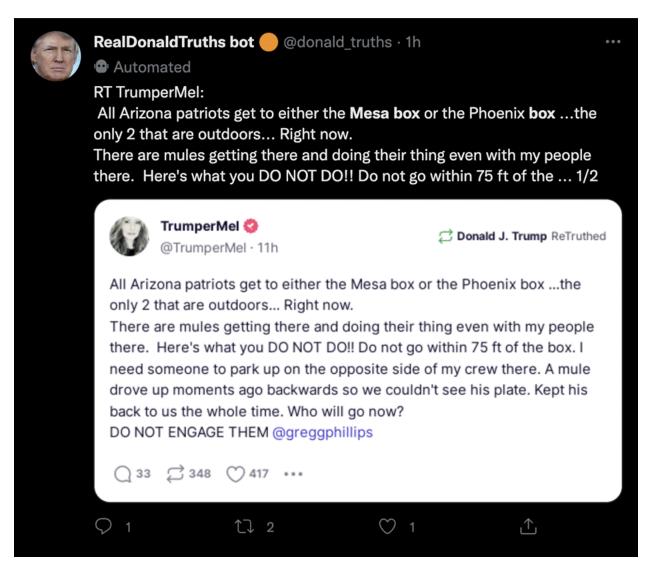
Ever since Elon Musk acquired Twitter in late October 2022, the platform's misinformation policies have been in flux. Twitter laid off 15% of its Trust and Safety team—which is responsible for content moderation and policy enforcement—in the week after Musk's acquisition. Since then, it has dissolved the Trust and Safety Council and made more cuts to its global content moderation workforce. Musk himself has promoted misleading narratives about platforms colluding with the U.S. government to interfere in the 2020 elections. Twitter has also announced that it will relax its policy on political advertising, which could open the door to more amplifying of misinformation.

Currently, Twitter's election misinformation policy is covered by its <u>Civic Integrity Policy</u>, which prohibits:

- False and misleading claims about how to participate in elections (where to vote, when to vote, how to cast a ballot, etc.).
- Claims that undermine public confidence in elections.
- Attempts to dissuade or suppress voters.

Twitter alerted users to trusted election resources when they searched for specific terms. For example, searches for "election corruption," "stolen election," and "voter fraud" generated a "Know the Facts" message, including links to the U.S. Election Commission Assistance's website and official Twitter account.[2] Twitter also promoted its election information hub at the top of users' timelines one week prior to Election Day. Both of these mechanisms prompted users but did not force them to review election information or prevent them from accessing search results.

Twitter is especially vulnerable to election misinformation in the form of screenshots from other social media platforms, primarily Truth Social. For example, Twitter suspended an automated account that reposted screenshots from Trump's Truth Social after it posted a screenshot of a <u>well-known election</u> <u>denier's</u> call to monitor drop box locations in Arizona.



An automated account that retweeted former President Trump's Truth Social posts. Twitter suspended the account. ADL took this screenshot on October 18, 2022.

However, accounts posting other Truth Social screenshots went unmoderated, like one in which Trump claimed that the Arizona governor's election had been stolen from Republican candidate Kari Lake.



Daily Wire senior writer Ryan Saavedra reposted this screenshot from former President Trump's Truth Social account. ADL took this screenshot on November 14, 2022.

At the time, Trump was permanently suspended from Twitter, but Twitter users could amplify his false claims by exploiting the loophole of posting screenshots, which may be harder to moderate.

Twitter was reluctant to remove content, even if it violated their policies, and preferred to moderate by applying labels to tweets. Our research indicates that Twitter used three types of labels for election misinformation corresponding to different levels of harm:

- The "misleading" label (see below) includes a link to Twitter's election information hub and relevant fact-checking resources. Tweets with a "misleading" label cannot be shared, liked, or replied to, thereby limiting their distribution and users' ability to engage with them.
- The "stay informed" label directs users to Twitter's election information hub and fact checks, but does not limit users from engaging with the labeled tweet.

But according to ADL's investigation, these labels can disappear with no reason given. Disappearing labels are one example of how Twitter does not enforce election misinformation policies year round.



Example of a tweet that received a "stay informed label." ADL took this screenshot on November 8, 2022. As of February 6, 2023, the label no longer appears.

• Twitter's <u>Community Notes</u> feature (previously known as Birdwatch) is its method of crowdsourcing content moderation. Community Notes are suggested by volunteer contributors rather than official content moderators. If enough users rate a note as "helpful," then the note is added to the <u>tweet in question.[3]</u> No additional restrictions are placed on the tweet.

While Twitter relied on labeling to enforce its election misinformation policies, labels were applied inconsistently. For instance, an accusation that the Republican party was "cheating" in Texas's elections was labeled, while another was not.

• • •

Just got home from voting and the "calibration is off" on the machines again. Marked my ballot for Beto and it showed I marked Abbott! I switched that shit back so fast! Happened on every single statewide Dem candidate. TX GOP is up to the same dirty tricks!!





Learn why election experts say election technology in the US is safe and secure. Find out more

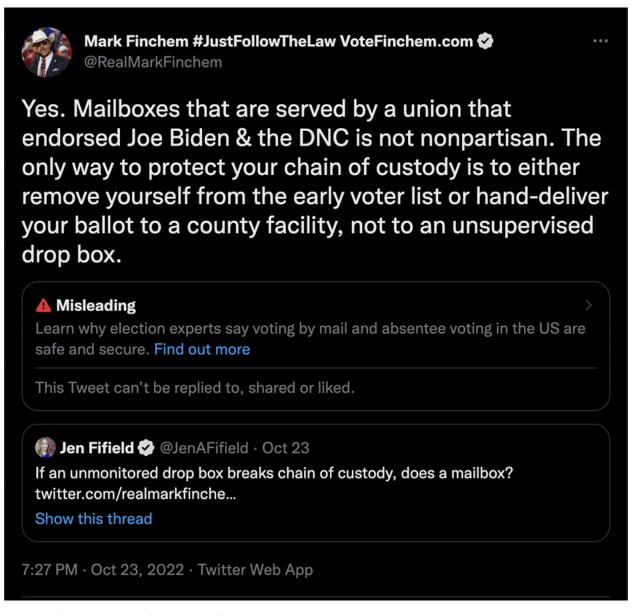
This Tweet can't be replied to, shared or liked.

4:24 PM · Oct 24, 2022



Tweets alleging that the Republican party in Texas was "cheating" in the 2022 elections, one with a label and one without. ADL took these screenshots on December 1, 2022. As of February 6, 2023, the label on the first tweet no longer appears.

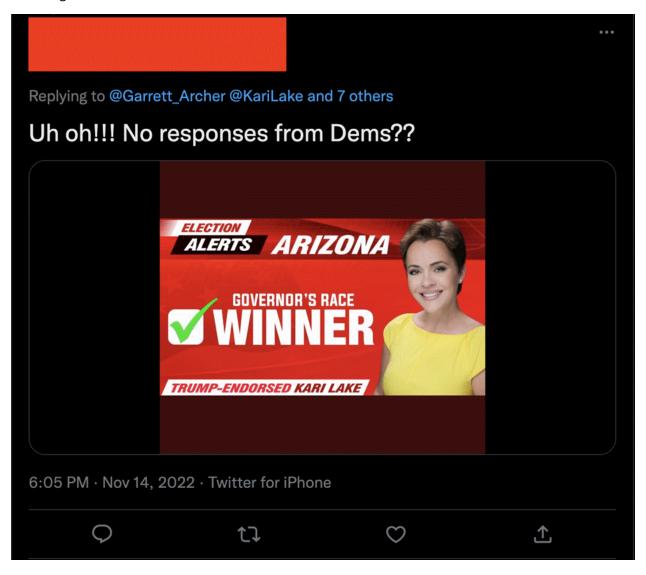
We found that labels were sometimes applied and then later removed. A tweet from Arizona Secretary of State candidate Mark Finchem arguing mailboxes could not be trusted because the American Postal Workers Union endorsed President Biden was initially given a "misleading" label, but the label <u>disappeared</u> some time between October 27 and October 31.[4]



Finchem's tweet with a "misleading" label. ADL took this screenshot on October 27, 2022.

In other cases, Twitter did not take any action on content we identified as violating one or more of its misinformation policies. For example, Twitter prohibits the sharing of "synthetic, manipulated, or out-of-context media that may deceive or confuse people." However, it allowed an image to remain on the

platform fabricated to look like an official cable news election alert declaring Kari Lake the winner in the Arizona governor's race.



A fabricated image falsely declaring Kari Lake the winner of the governor's race in Arizona. ADL took this screenshot on November 14, 2022.

While the poster's intent is impossible to determine, the image promoted a false claim that could deceive people. The tweet's context is also important: Associated Press called the election for Lake's opponent, Arizona Secretary of State Katie Hobbs, <u>less than 90 minutes</u> after the image was posted. At a time when the election results were still unknown and observers were eagerly awaiting confirmation, the image had significant potential to mislead and fuel false narratives.

Facebook (Meta)

The Good:

Restricts search results for problematic search terms

- Links to authoritative sources for flagged content
- Blurs images and links in posts with content that fact checkers rated false
- Labels appear to be permanent

The Bad:

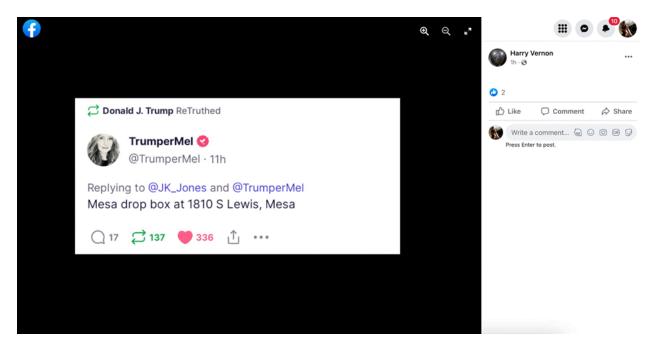
- Inconsistent enforcement
- No restrictions on comments, reactions, or shares for flagged content
- · Users can share false and misleading claims in screenshots

Facebook's <u>election misinformation policy</u> since 2019:

- Prohibits any content intended to "interfere with or suppress voting."
- Labels misleading content and relies on its <u>fact-checking partners</u> such as Factcheck.org, Politifact, and Associated Press to vet information.
- Similar to Twitter, Facebook prompts users to visit its <u>Voting Information Center</u> when they search for terms like "election fraud" or "stolen election."

Unlike Twitter, Facebook fully restricts users' ability to view search results for certain hashtags. For example, following the January 6th insurrection, Facebook banned search results for "#stopthesteal," a slogan that election deniers used to coordinate in the lead up to January 6th. It had previously banned searches for "#wwg1wga," which stands for "where we go one we go all," and similar hashtags popularized by QAnon conspiracy theory adherents. Searches for these hashtags trigger a message alerting users that "some content in those posts goes against [Facebook's] Community Standards."

Like Twitter, Facebook is also vulnerable to posts with screenshots from other platforms. We observed one account removal for a user who posted a screenshot from former President Trump's Truth Social account. The screenshot was itself a repost from an <u>election denier</u> who shared the street address of a ballot drop box in Arizona:



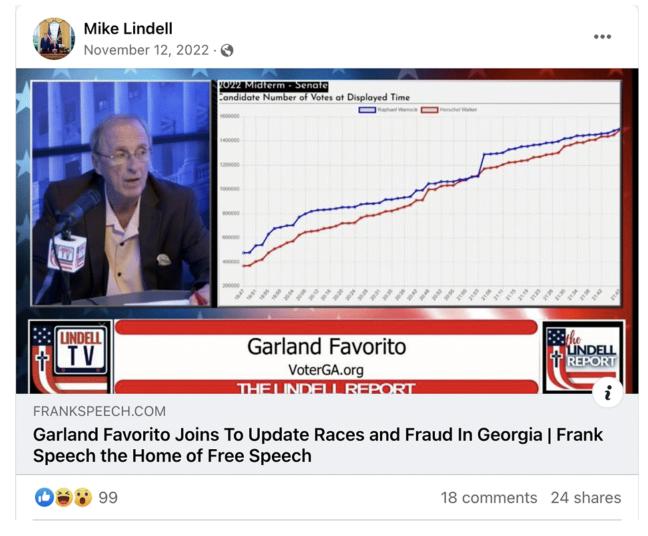
Truth Social screenshot posted to Facebook. ADL took this screenshot on October 18, 2022.

We are unable to determine whether Facebook suspended the account or its owner decided to delete it.

Facebook's primary method of policy enforcement for election misinformation is labeling. We found two types of labels that appear to be based on the quality of available fact checks:

- <u>False information</u>: Content that independent fact checkers have rated false. Facebook blurs out any images or links, and provides users with options to view the post and links to credible sources. Users may engage with the content as they would any post.
- <u>Missing context</u>: Content that is not entirely false but is misleading. Images and links remain intact and there are no obstacles to engagement.

We observed several inconsistencies with Facebook's labeling. MyPillow CEO Mike Lindell, who has held <u>events</u> presenting what he claims is "proof" of fraud in the 2020 U.S. election, posted multiple graphs of vote tallies from different races on Election Night. Lindell claimed that these graphs showed evidence of election fraud due to large "vote spikes" for Democratic candidates. Fact-checkers quickly debunked Lindell's claims. Facebook added a "false information" label to <u>some</u> of his posts, but took no action on others. One post that was left unlabeled contained the same false claims of election fraud as the labeled post, included the same image of a graph, and linked to the same source, Lindell's own "Frank Speech" website. It is unclear why there was a discrepancy in how Facebook applied these labels.



Misleading graph of vote counts posted on Mike Lindell's Facebook page. Facebook added labels to similar images on Lindell's page, but not this one. ADL took this screenshot on November 14, 2022.

Instagram (Meta)

The Good:

- Links to authoritative sources for flagged content
- Blurs images and links in posts with content that fact checkers rated false
- Labels appear to be permanent

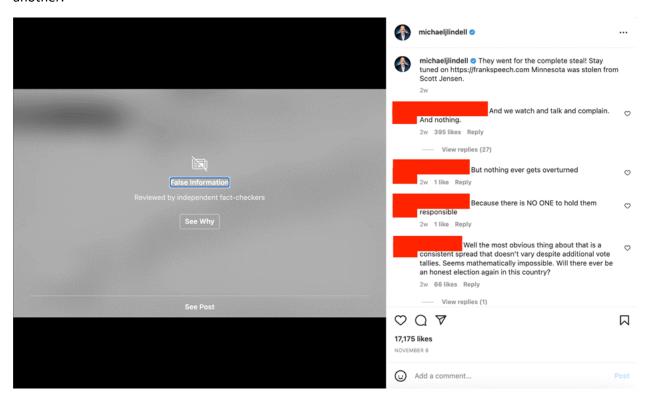
The Bad:

- Inconsistent enforcement
- No restrictions on comments, reactions, or shares for flagged content
- No restrictions on problematic search terms

Instagram has nearly the same election misinformation policies and resources as Facebook since both platforms are owned by the same parent company, Meta. There are a few differences, however. Unlike Facebook, Instagram places no restrictions on search terms or hashtags related to election denial, and does not show users any messages encouraging them to visit its Voting Information Center. [5] As of December 20, 2022, the Center was only accessible via Facebook.

Like Facebook and Twitter, Instagram enforces its misinformation policies primarily via labeling, and uses the same process as Facebook: "False information" and "missing context" for misleading but not entirely false content.

We found the same inconsistencies with labeling on Instagram as on Facebook regarding Mike Lindell's Election Night graphs and claims of fraud: "false information" for some and "missing context" for another.



Example of a Mike Lindell post that Instagram labeled as false information. ADL took this screenshot on November 9, 2022.



Example of a Mike Lindell post that Instagram labeled as missing context. ADL took this screenshot on November 9, 2022.

The only differences among these cases were the results of the races mentioned in each post. The posts that Instagram labeled as false information were called on Election Night while the one labeled missing context—the Georgia Senate race between Senator Raphael Warnock and Herschel Walker—went to a run-off the next month. Instagram provided specific fact-checks about each of the races in the posts labeled as false information, clarifying that candidates had conceded and the election had not been stolen. For the post labeled as missing context, the fact-checkers noted Lindell's graphs were not proof of stolen elections. It is unclear why the posts received different labels despite having similar false claims of election fraud.

There were also 14 cases where we reported a violation and Instagram took no action. Meta claimed that it would remove "misinformation about the dates, locations, times, and methods for voting." On Election Day, tabulation machine malfunctions at a handful of locations in Maricopa County caused delays for some Arizona voters. Kari Lake posted a message to her Instagram followers advising them not to change voting locations if they had already checked in with an election official. Lake's claim was missing important context: If a voter checked out of their current location with a poll worker, they would be able to safely and securely vote at any other location in Maricopa County. While there may be no reason to believe that Lake intended to mislead anyone, her advice was misinformed about a method for voting. According to Meta's policy, her post should have been removed or at the very least received a missing context label. In a confusing situation like this, voters often turn to social media—and to voices like Lake—for information. Her inaccurate advice in this case had the potential to create more confusion and undermine confidence in the election results.

YouTube (Google)

The Good:

- Links to authoritative sources for flagged content
- Flags false and misleading claims about past elections
- Some problematic search terms trigger information panels

The Bad:

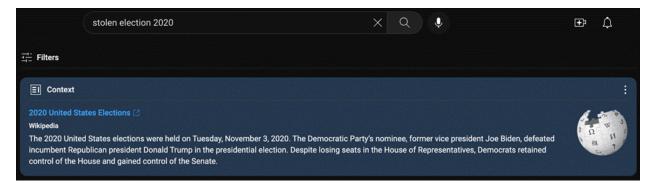
- Inconsistent enforcement
- No restrictions on comments or likes for flagged videos
- Some problematic search terms are unrestricted

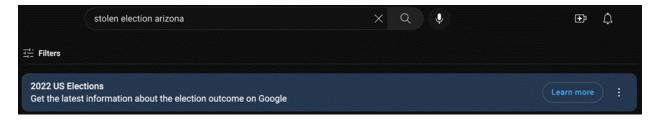
YouTube's election misinformation policy prohibits:

- Content that misleads or discourages voters.
- Spreading false information about candidate eligibility.
- Interfering with election processes.
- Distributing hacked materials.
- False claims about widespread fraud in the process or outcomes of specific past elections including "any past U.S. Presidential election."

The last point about past elections distinguishes YouTube from the other platforms we analyzed, none of which have a similar policy. It also means that YouTube moderates false and misleading claims about the 2020 election that other platforms do not.

Like Twitter, YouTube does not restrict users from viewing search results for specific words or phrases like "stolen election" or "election fraud." We cannot independently confirm false or misleading results for queries like these are demoted, but it is likely given the top results are news segments from reliable sources and not conspiracy theory videos. Search terms like "stolen election 2020" or "stolen election Arizona" trigger messages at the top of the results: for the first example, a link to a Wikipedia article about the 2020 U.S. elections; for the second, a link to a Google[6] search for 2022 U.S. election results.



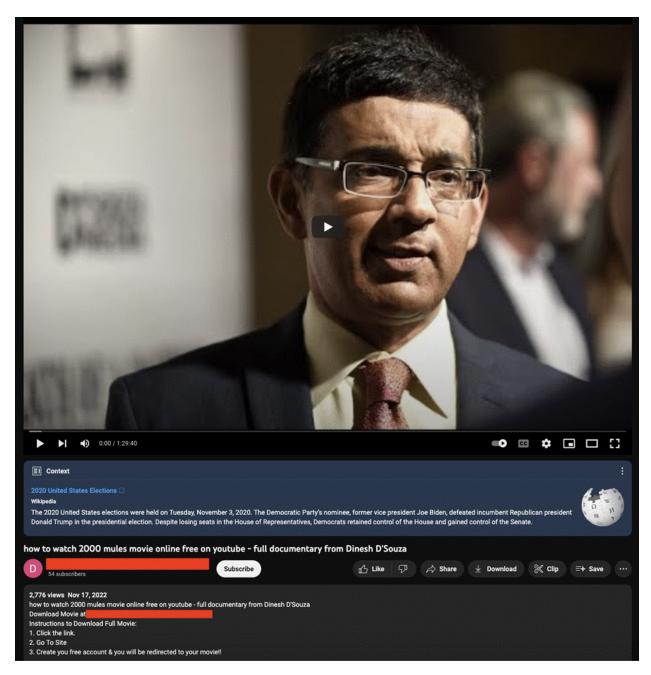


Examples of YouTube searches with information panels directing users to credible sources. ADL took these screenshots on January 23, 2023.

However, searching for "<u>fix2020</u>," a slogan that is <u>popular among election deniers</u>, does not trigger any such warnings. Some of the top results for "fix2020" were videos promoting claims that the 2020 U.S. Presidential election was stolen.

Our sample of candidates' YouTube channels posted campaign videos, clips of their media appearances, and clips of their speeches in the Senate or House of Representatives to their YouTube channels. Election misinformation on YouTube came almost exclusively from media personalities and news outlets, not individual candidates. For example, conservative political commentator Dinesh D'Souza promoted his conspiracy film about the 2020 elections, "2000 Mules," on his channel. Fact-checkers have thoroughly debunked the film's claims. However, since its release in May 2022, election deniers—including former President Trump—have often cited "2000 Mules" as proof that the 2020 election was stolen. YouTube did not host the video and attached information panels to videos that discussed it or showed clips from it with links to Wikipedia's article about the 2020 elections, including on D'Souza's own channel. D'Souza's video discussing "2000 Mules" has nearly 80,000 views and includes links to his Rumble account where it can be watched in its entirety. On Rumble, the film has over 3.2 million views.

YouTube did not attach these information panels consistently, however. One America News Network (OAN) posted the "2000 Mules" trailer along with instructions about how to watch it, attracting over 36,000 views. YouTube did not add an information panel with a Wikipedia link to OAN's video but did for other news networks that posted segments referencing the film. YouTube also did not immediately prevent channels from posting outbound links to sites that hosted the film, a loophole that election misinformation spreaders exploit.



Examples of a video with an outbound link to "2000 Mules." ADL took this screenshot on November 30, 2022. YouTube <u>removed</u> the video some time before January 23, 2023 for violating its policy on spam, deceptive practices, and scams.

TikTok

The Good:

- Labels all election-related content on mobile app
- Restricts problematic search terms and hashtags

The Bad:

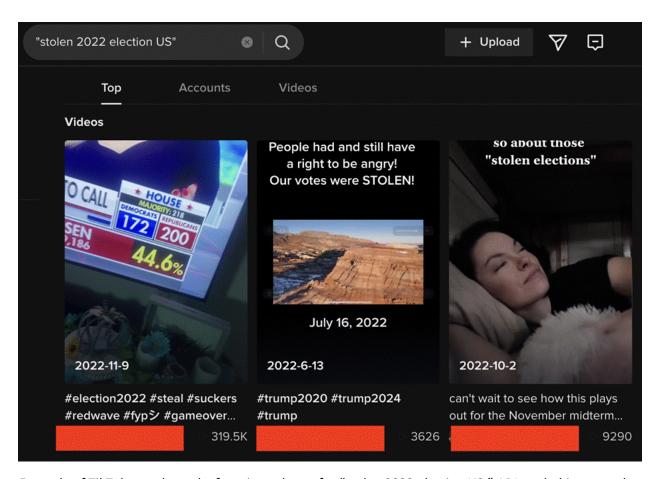
- Inconsistent enforcement
- Inconsistencies between mobile and web app
- Users can find false and misleading content by making small changes to search terms
- Users can share false and misleading claims in screenshots
- Easy to evade account bans by creating new accounts

Few of the candidates, media personalities, or news outlets that we identified in our initial sample have TikTok accounts, and only a handful consistently post content. As we learned, election misinformation on TikTok spreads primarily via ordinary users (see the Appendix for details about how we selected TikTok accounts). As a result, TikTok was unique in how we chose which accounts to monitor. TikTok also had a more noticeably different user experience between its web and mobile apps than other platforms, specifically regarding its Elections Center page: Content identified as election-related received a "get info on the U.S. midterm elections" label that linked to the Center on the mobile app, while the same content did not have a label on the web app.



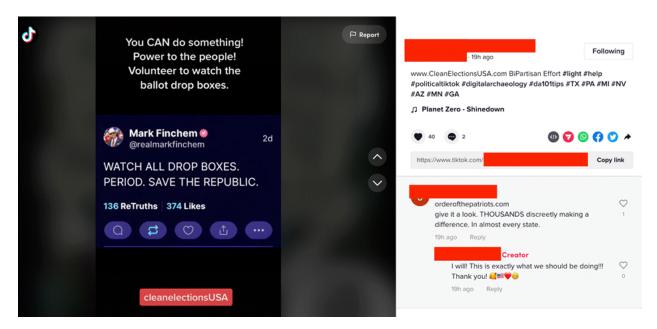
TikTok's "get info" label, screenshot taken from the mobile app on November 11, 2022. This same video does not have a label on the web app.

Like Facebook, TikTok fully restricts search terms and hashtags for content that violates its Community Guidelines, such as "stolen election," "election fraud," or "#wwg1wga." However, by changing the search terms slightly, e.g., "stolen 2022 election US," users can access search results that include election misinformation, thereby exploiting this moderation loophole.



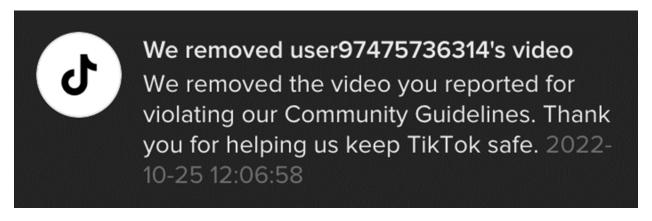
Example of TikTok search results from its web app for "stolen 2022 election US." ADL took this screenshot on January 23, 2023.

Besides the "get info" label, which TikTok applied to all election-related content on the mobile app, misinformation or not, TikTok's primary enforcement mechanisms were content removal and account suspension. We documented multiple removals and suspensions in response to our reports, but enforcement was inconsistent. For example, TikTok prohibits "attempts to intimidate voters or suppress voting." We reported a video with a screenshot of a Mark Finchem Truth Social post asking viewers to watch ballot drop boxes in Arizona on the grounds that the post encouraged voter intimidation.

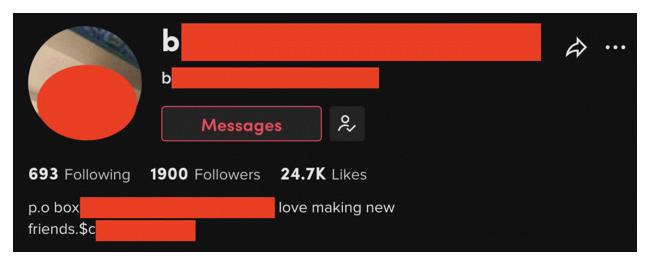


TikTok video, screenshot taken from web app on October 20, 2022.

TikTok responded to our report within minutes, saying that it had not found any violations. However, TikTok updated its ruling five days later and notified us that the video had been removed. The account that posted the video was subsequently deleted, but the user was able to create another using a nearly identical account name and similar profile picture. TikTok did not give any explanation for its reversal, did not specify which of its Community Guidelines the video had violated, and has not removed the second account as of January 23, 2023.



TikTok's update to our report, screenshot taken from web app on October 25, 2022.



New account created by the same user who posted a screenshot of Mark Finchem's Truth Social post. ADL took this screenshot on January 23, 2023.

TikTok also did not act in cases where our team identified violations such as "false claims that seek to erode trust in public institutions," including claims of stolen elections or voters being paid and others voting in their place. By comparing these cases of inaction with content that was removed, we infer that TikTok applies a narrow definition of "false claims," leaving room to deny election results and cast doubt on the integrity of voting processes.

Recommendations

ADL noted in our pre-Election Day analysis of platforms' policies that being reactive to misinformation rather than proactive sets platforms up for a continuous "game of content moderation whack-a-mole." The threat of false narratives around elections remains substantial, thus platforms must adopt more comprehensive policies and consistently enforce them. Until they do, bad faith actors will continue to enjoy a safe harbor on social media for spreading claims that erode democratic processes and institutions.

We recommend platforms:

- Enforce consistently on both mobile and web interfaces. By labeling or restricting posts on mobile but not on web, or vice versa, misinformation can more easily reach audiences.
- Prohibit election misinformation year-round, not just during election cycles. Platforms should make their election hubs and resource pages permanent and accessible. They should also regularly update these resources as false and misleading election narratives evolve.
- Close loopholes to problematic external links and media. Platforms should make it more difficult
 to post or view external links and media from known sources of misinformation. Users should be
 prohibited from sharing screenshots of social media posts from accounts that are banned on the
 platform or that contain information that violates the platforms' policies
- Be explicit about enforcement decisions for different types of misinformation. Erroneous claims about where and when to vote, incitement to voter intimidation or suppression, explicit threats, and false claims about election outcomes must be removed immediately before they can

spread. Misleading claims that are less urgent should be moderated via labeling, adding context, demoting, and redirecting users to trusted information sources. Platforms should be clear about why they took a specific action.

- Hold politicians and candidates for public office to higher standards. Social media users may
 perceive politicians as more credible than other sources, especially if that politician can claim
 expertise about election processes because of their position, e.g., a Secretary of State whose job
 is overseeing elections. Platforms should prioritize adding context and labels to politicians' and
 candidates' false and misleading claims about election processes and results rather
 than exempting them from moderation efforts.
- Share moderation data with independent researchers. Independent researchers in academia
 and civil society can only understand and vet platforms' policies if they can access moderated
 content. Platforms must provide researchers the necessary access to such data, while respecting
 privacy concerns. Platforms should also provide the public with reports about their policies and
 enforcement records after national elections and allow researchers to verify the methods and
 data in those reports.

Self-regulation, however, has proven to be insufficient. Platforms have instituted often reactive policies that allow them to better moderate false and misleading content, but enforcement of these policies is inconsistent and not to scale. When it comes to protecting the integrity of elections, and social media's role in combating election misinformation, we recommend that U.S. legislators:

- Require platforms share more information about misinformation content policies and
 enforcement: In line with consumer protection regulations for other industries, social media
 platforms must be more transparent about their election misinformation policies, how they are
 enforced, and the process for changing them. Transparency reports should include clear,
 consistent, and easily accessible details about the thresholds for labeling election
 misinformation content, as well as the fact-checking process. Additionally, platforms should be
 required to produce a public-facing audit of their election misinformation policy enforcement
 metrics after every national election in the U.S.
- Mandate archiving of problematic content: As part of helping to combat threats to election
 integrity and democracy, social media platforms should maintain an internal repository of all
 flagged images, videos, content, and incident reports for election-related misinformation.
 Platforms should make their repositories available to third-party researchers, upon request,
 provided they have passed a verification process.
- Conduct pre-election assessments: Policymakers should engage social media platforms prior to
 national elections in the U.S. to understand social media platforms' fact-checking processes,
 policies for content labeling—including thresholds—and what steps platforms have taken to
 ensure that users are not exposed to election-related misinformation. The government can use
 existing election security mandates, such as those of the Cybersecurity and Infrastructure
 Security Agency and the Department of Justice, to inform companies on how to identify and
 address threats. Relatedly, platforms should produce regular risk assessments about their
 policies' effectiveness.

Note: We did not include recommendations about combating misinformation in paid content or advertising because it is beyond the scope of this particular report.

Appendix: Methodology

We analyzed five social media platforms between October 18 and November 18, 2022: Facebook, Instagram, TikTok, Twitter and YouTube. We chose these platforms based on four criteria:

- Number of users: Each has over 300 million monthly active users.
- Dedicated election policies: Each has specific election misinformation policies.
- User moderation tools: Each gives users the ability to report problematic content.
- Platform design: Each platform's design features enable misinformation to spread widely.

We created a list of accounts to track on each platform based on convenience sampling, a non-probability sampling method that draws from easily-accessible sources. Because public figures play an outsized role in spreading misinformation on social media, we focused on three types of accounts: candidates running for statewide office in 2022, political media personalities, and news outlets. For each platform, we selected accounts to track based on the following criteria:

- Election denial content: candidates who <u>fully denied</u> the results of the 2020 U.S. presidential election and were running in the five states where the <u>2020 popular vote margin</u> for president was within 2%: Arizona, Georgia, North Carolina, Pennsylvania, and Wisconsin.
- Close races: election-denying candidates' opponents in races within a 5% polling margin two weeks before Election Day.[8]
- Promoting election fraud conspiracy theories: political media personalities and news outlets
 that have claimed the 2020 U.S. Presidential election was "stolen," including <u>Dinesh</u>
 <u>D'Souza</u>, <u>Tom Fitton</u>, <u>The Gateway Pundit</u>, and <u>Mike Lindell</u>, among others.

TikTok was an outlier among the five platforms because few candidates, media personalities, or news outlets have TikTok accounts or post content to them regularly. To select which TikTok accounts to follow, we used keyword searches like "election 2022 fraud," "voter 2022 fraud," and "fix2020" to find potential election misinformation content. We reviewed the results manually and followed accounts that posted false or misleading election-related content. This method did not allow us to compare and contrast how specific accounts were treated across all five platforms, but it did give us enough material to evaluate TikTok's policy enforcement.

We ultimately selected 498 accounts across the five chosen platforms, most of which were duplicates across these platforms, e.g., Dinesh D'Souza has Twitter, Facebook, Instagram, YouTube, and TikTok accounts. We hypothesized that these accounts were likely to post election misinformation because of their past histories and that races with slim margins would be fodder for misinformation. We also reasoned that accounts in our initial sample might repost misinformation content from other accounts that we could add to our list via snowball sampling, i.e., identifying accounts linked to accounts in our initial sample. As our goal was to assess how well each platform enforced its policies, we did not report every piece of election misinformation on each platform we tracked nor did we quantify the

percentage of reports platforms acted on (as we have done here). Instead, we focused on reporting violative content and observing where, when, and how platforms took action.

We reviewed organic posts—i.e., non-advertising content—from each account in our sample daily between October 18 (three weeks before Election Day) and November 18 (the second Friday following Election Day). We reported any content we identified as violating that platform's published election misinformation policy. [9] Since each platform's policies were different—with the exception of the two Meta products, Facebook and Instagram—prohibited content on one site might be allowed on another.

This study has several limitations. Because we only reported content from a selection of accounts, our data does not represent the sum total of election misinformation and policy enforcement on these five platforms. Additionally, all five platforms make exceptions for content that is deemed to be "newsworthy" or "in the public interest." Meta makes special allowances for political candidates. [10] By focusing on public figures and news outlets, it is possible we observed accounts that platforms were less likely to moderate. We were also unable to detect or evaluate some enforcement methods such as demoting posts, reducing their visibility, or otherwise changing how they were recommended to users on the platform.

Endnotes

- [1] 147 Congresspeople voted against certifying the election on January 6, 2021. Of those 147, nine retired before the 2022 election, three died, and four lost their 2022 primaries. None of the eight Senators who voted against certification were on the 2022 ballot, leaving 123 incumbent candidates. 119 won re-election to the House of Representatives, two were elected to the Senate, and two lost. We arrived at these numbers by cross-referencing the list in the Vox article linked above with Associated Press's 2022 election results and Wikipedia articles for politicians who were not on the 2022 ballot.
- [2] The "Know the Facts" message only appeared close to Election Day. As of mid-December 2022, these search terms and similar ones returned results just like any other Twitter search.
- [3] <u>Critics</u> have noted the Community Notes feature does little to stop the spread of misinformation. Twitter is also still <u>adding updates</u> as it learns more about the program's vulnerabilities to manipulation.
- [4] We were unable to pinpoint the exact date of removal.
- [5] This was true for Instagram's web interface. We did not test its mobile interface.
- [6] YouTube is a Google subsidiary.
- [7] For example, Reddit has more active users than Twitter, but did not have a platform-wide policy for election misinformation. Snapchat has more active users than both Reddit or Twitter, but its platform architecture prioritizes private, peer-to-peer communication, making it more difficult for misinformation to spread widely and for researchers to observe and compare activity. https://en.wikipedia.org/wiki/List of social platforms with at least 100 million active users
- [8] Where polling was available via FiveThirtyEight.com or Real Clear Politics.
- [9] See ADL's analysis of individual platform policies for details: https://www.adl.org/resources/blog/major-platforms-midterm-election-policies-are-they-enough.

[10] Twitter's public-interest exception: https://help.twitter.com/en/rules-and-policies/public-interest; Meta's approach to newsworthy content: https://transparency.fb.com/features/approach-to-newsworthy-content/; YouTube's educational, scientific, documentary, or artistic content exception: https://www.youtube.com/howyoutubeworks/policies/community-guidelines/#allowing-edsa-content; TikTok's community guidelines: https://www.tiktok.com/community-guidelines?lang=en; Meta's fact-checking

policies: https://www.facebook.com/business/help/315131736305613?id=673052479947730.