July 2022

Misleading Information and the Midterms

How Platforms are Addressing Misinformation and Disinformation Ahead of the 2022 U.S. Elections

Spandana Singh & Quinn Anex-Ries

Last edited on July 18, 2022 at 7:20 p.m. EDT
About the Author(s)

Spandana Singh is a Policy Analyst at New America’s Open Technology Institute, where she researches and reports on policies and practices related to content moderation, disinformation, algorithmic accountability, government surveillance, and privacy.

Quinn Anex-Ries is a Legal/Public Policy intern with New America’s Open Technology Institute, working with the platform accountability and privacy teams.

About New America

We are dedicated to renewing the promise of America by continuing the quest to realize our nation’s highest ideals, honestly confronting the challenges caused by rapid technological and social change, and seizing the opportunities those changes create.

About Open Technology Institute

OTI works at the intersection of technology and policy to ensure that every community has equitable access to digital technology and its benefits. We promote universal access to communications technologies that are both open and secure, using a multidisciplinary approach that brings together advocates, researchers, organizers, and innovators.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>4</td>
</tr>
<tr>
<td>Sharing Authoritative Information and Promoting Informed User Decision-Making</td>
<td>7</td>
</tr>
<tr>
<td>Moderating and Curating Misleading Information</td>
<td>9</td>
</tr>
<tr>
<td>Tackling Misleading Advertising</td>
<td>12</td>
</tr>
<tr>
<td>Providing Meaningful Transparency and Accountability</td>
<td>14</td>
</tr>
<tr>
<td>Conclusion</td>
<td>16</td>
</tr>
</tbody>
</table>
Introduction

Since 2016, online platforms have been at the heart of discussions around misinformation and disinformation. Over the past two years, many platforms significantly expanded their efforts to identify and combat misleading information, particularly related to the COVID-19 pandemic and the 2020 U.S. presidential election.¹

Before the 2020 presidential election, the Open Technology Institute (OTI) published a report outlining how 11 internet platforms were addressing the spread of election-related misinformation and disinformation on their services.² The report issued recommendations on how these platforms can improve their efforts to share authoritative information and empower users to make informed decisions, moderate and curate misleading information, tackle misleading advertising, and provide meaningful transparency and accountability around these efforts. It also proposed actions policymakers could take to support these efforts.

But, while internet platforms ramped up attempts to combat election misinformation and disinformation in 2020, many have pulled back on these efforts, noting that they were temporary measures.³ In the meantime, conspiracy theories and misleading information about the 2020 election continue to circulate and contribute to ongoing voter suppression, which particularly impacts communities of color. Ongoing misinformation and disinformation also contributed to the January 6 insurrection at the U.S. Capitol.⁴ As the 2022 U.S. midterm elections draw near, it is important to examine what steps internet platforms are taking to curtail the spread of election-related misinformation. National security agencies and independent experts have already warned that foreign and domestic actors are likely to continue disseminating misinformation and disinformation ahead of the midterms.⁵ Social media platforms, in particular, can facilitate the spread of posts and advertisements with false information about candidates, election results, the overall voting process, and more. The continued spread of this content threatens to suppress voting, undermine public trust in elections, and erode the health of our democracy.

This scorecard evaluates how major internet platforms are combating election and voter suppression-related misinformation and disinformation ahead of the midterm elections. Using this data, we demonstrate which platforms have made the most progress towards tackling misleading election information, which platforms are falling behind, and where companies need to invest more resources. The scorecard measures platforms against a selection of recommendations included in our 2020 report, which we consider baseline policies and practices that companies should implement. These recommendations are broken into four categories:
Sharing Authoritative Information and Promoting Informed User Decision-Making

- Partner with reputable fact-checking organizations and entities to promote, verify, or refute information circulated through organic content and search results.

- Partner with reputable government or civil society entities to promote, verify, or refute information circulated through organic content and search results.

- Notify users who are or have been engaging with misleading election-related content and/or direct them to authoritative sources of information.

- Conduct regular impact assessments of algorithmic curation tools so they do not direct users to or surface misleading content when they search for election-related topics.

Moderating and Curating Misleading Information

- Create a comprehensive set of content policies to address the spread of election-related misinformation and disinformation.

- Institute a dedicated reporting feature that enables users to flag misinformation and disinformation to the company.

- Remove, reduce the spread of, or label content that has been fact-checked and/or deemed to contain election-related misinformation.

Tackling Misleading Advertising

- Create and implement comprehensive policies for the content and targeting of ads that prohibit election-related misinformation and disinformation in advertisements.

- Establish a comprehensive review process for election-related ads and ad targeting categories that includes fact-checking.

- Create policies that prevent users and entities from being able to monetize and advertise on the platform if they repeatedly spread misinformation and disinformation.
Providing Meaningful Transparency and Accountability

- Publish data related to the moderation, curation, and labeling of election-related misinformation and disinformation in their regular transparency reports.

- Create a publicly available online database of all ads in categories related to elections and social and political issues that a company has run on its platform.

- Publish data on the company’s ad enforcement efforts.

The scorecard includes data on ten of the internet platforms discussed in the 2020 report—Facebook/Instagram, Google, Pinterest, Reddit, Snap, TikTok, Twitter, WhatsApp, and YouTube. While the original report included Amazon, we chose to omit the company from the scorecard due to its unique features as an e-commerce platform, which made comparative analysis challenging. For the purposes of this scorecard, the Facebook and Instagram platforms are grouped together, as parent company Meta typically applies similar content and advertising policies to both services. We evaluate WhatsApp separately as it offers different services and therefore has different policies. We distinguish between Google and YouTube because at times the services have differing policies and practices. Reddit employs a decentralized approach to content moderation, allowing individual users to serve as moderators of subreddits. For the purposes of this scorecard, we focus on policies and practices that Reddit applies across the platform, rather than individual policies and practices deployed by user moderators. The data in the charts is based on publicly available information about platforms’ misinformation and disinformation efforts as of June 2022. We focused our research and analysis on each company’s primary platform. For example, we focused on Google’s search product and Snap Inc.’s Snapchat product.

Editorial disclosure: This report discusses policies by Google (including YouTube), Facebook (including WhatsApp and Instagram), and Twitter, all of which are funders of work at New America but did not contribute funds directly to the research or writing of this report. New America is guided by the principles of full transparency, independence, and accessibility in all its activities and partnerships. New America does not engage in research or educational activities directed or influenced in any way by financial supporters. View our full list of donors at www.newamerica.org/our-funding.
Sharing Authoritative Information and Promoting Informed User Decision-Making

Internet users often come to internet platforms for information and discussion on candidates, electoral processes, and more. As a result, many platforms have instituted efforts to surface and promote reliable information and ensure users can make informed decisions about voting. Ahead of the 2020 presidential election, many internet platforms partnered with fact-checking entities to promote, verify, or refute information circulating on their services. Based on our findings, there appears to be no significant change in platforms’ fact-checking efforts.

Platforms such as Facebook/Instagram and TikTok that had partnerships with fact-checking organizations in 2020 continue to do so, while those that did not have not initiated new partnerships. The only exception to this is Twitter, which formed a partnership with the Associated Press in August 2021 to surface reliable information in Twitter Moments, Trends, and other surfaces on the platform. Additionally, in January 2021, Twitter introduced Birdwatch, which allows Twitter users to add “notes” to tweets in order to flag misleading information and add contextual information. Twitter has partnered with reputable fact-checking entities to evaluate the nature of the tweets. However, for the most part, all of the fact-checking is done by Twitter users, raising concerns about the credibility of the fact-checks. Additionally, the Birdwatch program is still in its pilot stage and is therefore not operating at full scale. The Birdwatch program does, however, offer a unique perspective on how platforms can adopt a decentralized approach to managing misleading information on their services.

---

1 Pinterest recently announced that it is taking steps to direct users to official sources of information and to include search advisories in advance of elections in Australia, France, Brazil, and the Philippines, but the company has not clarified if these measures will be adopted in the U.S.

2 When users try to share content containing unreliable information, TikTok presents them with a warning. However, they do not notify users if they view the content or direct users to more information.

3 WhatsApp offers end-to-end encrypted messaging services. As a result, it cannot provide notices to users who engage with election-related disinformation. However, WhatsApp provides users with search tools to independently verify information, such as a feature that allow users to directly chat with independent fact-checkers and an in-app function that allows users to double-check the content of a message on the web.

NEW AMERICA
Prior to the 2020 presidential election, several platforms partnered with reputable government and civil society entities to promote, verify, or refute organic content circulating on their services. For many platforms, it is unclear whether these partnerships will continue into the midterm election season.

When evaluating how platforms notify users who engage with misleading election-related content and/or direct them to authoritative sources of information, we wanted to account for the fact that platforms may use different mechanisms to notify users that they’re engaging with misleading or unverified content. For example, TikTok presents a warning label to users who attempt to share unverifiable content. Similarly, WhatsApp notifies users that a piece of content has been shared several times. But, WhatsApp cannot comment on the details of the content itself, as WhatsApp offers end-to-end encrypted messaging services. We awarded partial credit to platforms if they either provide notification to users or direct users to authoritative services, but not both.

Very few platforms notify users if they have engaged with or are engaging with content that is misleading or unverified. Additionally, very few platforms connect users with authoritative information on elections as they navigate their services. Only three platforms—Facebook, Google, and Instagram—do both. Given that users often come to internet platforms to access information, this is a critical gap in platforms’ efforts to combat misleading information online.

OTI and numerous other organizations have written extensively about how internet platforms’ algorithmic curation tools can amplify harmful content, including misinformation and disinformation. In the run-up to the 2020 election, some internet platforms, such as Twitter, instituted changes to their algorithmic curation and content recommendation systems to prevent the amplification of misleading content. At that time, we pushed platforms to conduct regular impact assessments of their algorithmic curation tools to ensure they were not directing users to or surfacing misleading content. However, we were unable to find any public reporting indicating that platforms adopted this recommendation. Some platforms conduct impact assessments internally. However, if they do not provide any public transparency around this, there is no mechanism to evaluate their processes and hold them accountable for their policies and procedures.

One additional recommendation we evaluated platforms against is “providing vetted researchers with access to tools and datasets that could enable them to better evaluate company efforts to combat election-related misinformation and disinformation.” We believe it is critical that independent researchers have access to platform data, as independent evaluation helps hold platforms accountable. However, we were unable to find reliable information from many platforms and, as a result, omitted it from the scorecard.
Moderating and Curating Misleading Information

<table>
<thead>
<tr>
<th>Platform</th>
<th>Create a comprehensive set of content policies</th>
<th>Dedicated reporting feature</th>
<th>Remove, reduce the spread of, or label election-related misinformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook/Instagram</td>
<td>☑ Yes</td>
<td>☑ Yes</td>
<td>☑ Yes</td>
</tr>
<tr>
<td>Google</td>
<td>☑ Yes</td>
<td>No</td>
<td>☑ Yes</td>
</tr>
<tr>
<td>Pinterest</td>
<td>☑ Yes</td>
<td>Yes</td>
<td>☑ Yes</td>
</tr>
<tr>
<td>Reddit</td>
<td>☑ Yes</td>
<td>Yes</td>
<td>☑ Yes</td>
</tr>
<tr>
<td>Snap</td>
<td>☑ Yes</td>
<td>Yes</td>
<td>☑ Yes</td>
</tr>
<tr>
<td>TikTok</td>
<td>☑ Yes</td>
<td>Yes</td>
<td>☑ Yes</td>
</tr>
<tr>
<td>Twitter</td>
<td>☑ Yes</td>
<td>Yes</td>
<td>☑ Yes</td>
</tr>
<tr>
<td>WhatsApp</td>
<td>☑ Yes</td>
<td>Yes</td>
<td>☑ Yes</td>
</tr>
<tr>
<td>YouTube</td>
<td>☑ Yes</td>
<td>Yes</td>
<td>☑ Yes</td>
</tr>
</tbody>
</table>

Employing content moderation and algorithmic content curation to tackle misleading election content can yield promising results. All of the platforms we evaluated have a comprehensive set of content policies to address the spread of election-related misinformation and disinformation. There is variation in how platforms outline these policies, however.

For example, Reddit has an impersonation policy prohibiting content that “impersonates individuals or entities in a misleading or deceptive manner” and “deepfakes or other manipulated content presented to mislead, or falsely attributed to an individual or entity.” Twitter, on the other hand, has a Civic Integrity Policy, that prohibits the use of Twitter’s services “for the purpose of manipulating or interfering in elections or other civic processes” including “posting or sharing content that may suppress participation or mislead people about when, where, or how to participate in a civic process.” WhatsApp’s policies also prohibit “publishing falsehoods, misrepresentations, or misleading statements.” Other platforms have broader policies on false information, misinformation, or misleading information, which include election-related content. If a platform had a content policy category that could be applied to election misinformation and disinformation, we gave them full credit.

One limitation of existing efforts in this area is that not all platforms specifically discuss voter suppression in their policy. Given that online services are often used to misinform users about when, where, and how they can vote and deter voting, platforms must provide more clarity around their policies ahead of the midterms. Additionally, companies who do not have a clear stance on voter suppression should engage in broad, multi-stakeholder engagement to develop robust policies around voter suppression.
A major improvement that we have seen across platforms over the past several years is that most allow users to flag content as misinformation and disinformation. The major exception to this is Google, which allows users to flag search results, but not for misleading content. In discussions with certain companies, we learned there was initially resistance to providing this flagging feature as some companies observed that users were often flagging content that they disagreed with, rather than content that was truly misleading in nature. However, most platforms now offer users the ability to flag misleading content, which is an important way of providing users with greater agency to surface harmful content they are seeing and ensuring the content moderation process is not one-sided.

We provided companies with full credit if they had a reporting feature for misinformation or misleading information broadly, rather than a dedicated feature for election misinformation and disinformation. It is important to note that some platforms combine misinformation reporting with other content categories. For example, YouTube users can flag content under the category “spam and misleading.” This may be because of the high number of false misinformation reports companies receive. The limitations of this approach will be discussed in the last section on transparency.

Our evaluation of WhatsApp was slightly different for this category because the platform does not engage in traditional content moderation and curation due to the end-to-end encrypted nature of its messaging services. End-to-end encryption is a critical feature for privacy and security. While WhatsApp cannot scan the content of user messages, the company has created several features designed to limit the spread of misinformation, including forwarding limits for viral messages, spam detection to ban mass messages, and enabling users to block and report messages containing misinformation. WhatsApp also allows users to directly interact with fact-checkers and to independently verify information on the web.16

Before the 2020 election, there was significant variation in terms of how platforms approached misleading election information. This remains true. Facebook and Instagram, for example, rely on a “Remove, Reduce, Inform” approach, in which the services remove content that violates their Community Standards, algorithmically reduce the spread of misleading or harmful content that does not violate their Community Standards in their News Feeds, and inform users with additional context, including through labels.17 On the other hand, Twitter uses a range of approaches including tweet deletions, labeling, temporary and permanent account locks, and suspensions. The company determines which enforcement actions to use based on the severity of a tweets’ content, past account history, and a predetermined strike system.18

Platforms should be able to use a broad range of enforcement mechanisms to tackle misleading content and accounts. However, few platforms have released
granular data outlining the effectiveness of their efforts. This makes it difficult to identify which mechanisms are most effective and hold platforms accountable. This will be discussed in greater detail in the last section.
Tackling Misleading Advertising

As demonstrated during the 2016 presidential election, online advertising can be used to spread misleading information and foment social divides. We have written extensively about how platforms’ ad-targeting tools allow advertisers to precisely target users based on sensitive characteristics, political affiliation, and personal interests, as well as how advertising delivery algorithms can generate discriminatory and harmful results.

Internet platforms have adopted a variety of approaches when it comes to tackling misleading election information in advertisements. Some companies have clear policies on misinformation in advertising, which can help prevent the spread of election misinformation and disinformation. Other companies, including TikTok and Twitter, have banned political advertisements as a conduit for avoiding election information. While bans may help reduce election misinformation and disinformation in advertising, they are not a foolproof solution. It is often very difficult to draw clear lines around what is political content. TikTok, for example, notes that advertisements must not “reference, promote, or oppose a candidate for public office, current or former political leader, political party, or political organization. They must not contain content that advocates a stance (for or against) on a local, state, or federal issue of public importance in order to influence a political outcome.” This is a very specific definition of political advertising, but it does not account for a broad range of social issues that are often very politicized, including climate change and immigration. Facebook, on the other hand, broadly categorizes ads related to social issues, elections, and politics together, noting that these touch on topics including candidates for public office, political action committees, ongoing elections, referendums, or ballot initiatives, and social issues such as civil and...
Because it is difficult to draw clear lines around political content, we expect companies to have policies related to misinformation and disinformation in advertising, regardless of whether or not they ban political ads. If a platform has an advertising policy category that could be applied to election misinformation and disinformation, we gave them full credit.

We did not give Google or YouTube credit for this category because their advertising policies are vague. For example, the policy notes that the companies support “responsible political advertising.”24 Given that Google operates the largest digital advertising platform and there is a breadth of evidence demonstrating the relationship between online advertising tools and misleading information, the companies should have stronger policies against misleading election information. We did not evaluate WhatsApp against the recommendations in this chart because WhatsApp does not allow any advertisements on its services.

While many platforms have policies related to misinformation and disinformation in advertising, fewer have comprehensive processes for reviewing and fact-checking advertisements. Only two companies—Reddit and Snap—have both. If platforms do not have both, we did not give them full credit. Without a comprehensive review and fact-checking process, it is likely that companies will fail to consistently enforce their advertising policies, allowing misleading information to slip through the cracks, including on services that have banned political ads.25 Currently, few platforms discuss how much their advertising review processes rely on automated tools compared to human reviewers. This is an area where we need significantly more transparency to ensure that automated tools are being trained and deployed effectively, and to ensure that humans are kept in the loop in situations where there are lower levels of confidence in automated decision-making.26

Lastly, only three platforms have policies in place that prevent repeat spreaders of disinformation from monetizing their content. This is a major concern. Most internet platforms rely on a targeted advertising business model. This means that platforms have a strong incentive to increase user engagement and time spent on their services, as they can deliver more ads to users. Because of this structure, platforms often fail to remove harmful or misleading content, as it is often more engaging. These business models incentivize some of the largest spreaders of misleading information to generate more of such content.27 This is a major gap in how platforms are addressing election misinformation and disinformation that must be filled.
Providing Meaningful Transparency and Accountability

Disappointingly, while most platforms we evaluated have comprehensive content policies that cover election misinformation and disinformation, platforms have made little progress in providing adequate transparency around how they apply these policies and how this has impacted content and user accounts online.

In addition, companies that do report data often fail to disaggregate data between categories of content. For example, YouTube reports on spam and misleading information together, making it difficult to understand how much misleading content the company took action against. Additionally, none of the companies provide a comprehensive overview of how all of their enforcement actions have impacted content. For example, Twitter primarily reports on the number of accounts actioned, suspended, and the amount of content it removed under its Civic Integrity Policy. It does not discuss its use of labels and other enforcement mechanisms. Similarly, TikTok reports on how much election misinformation, disinformation, or manipulated media content they removed and how many times banners were added to election-related videos in the “For You” feed, but not the broader set of enforcement actions it takes on its platform against misleading information. Without this data, it is difficult to understand the effectiveness of platforms’ efforts to combat misleading information, and to hold them accountable for falling short. We encourage platforms to expand their reporting. Additionally, while we encourage some degree of standardization in reporting, we believe that platforms should report on metrics that allow researchers to understand the unique nature of their services and their unique

newamerica.org/oti/reports/misleading-information-and-the-midterms-how-platforms-are-addressing-disinformation/
efforts to combat misleading information (e.g., as TikTok has done with the “For You” label metric).\textsuperscript{31}

Similarly, only two platforms publish comprehensive ad databases or libraries of all ads in categories related to elections and social and political issues, while three offer limited libraries. For example, Google’s U.S. political advertising library is limited to four categories of ads, which include ads related to a “federal or state level political party” and “a current office holder or candidate for an elected federal office.”\textsuperscript{32} Once again, it is challenging to draw clear lines around political and non-political content. As a result, we expect and urge companies that have instituted political ad bans to still establish comprehensive ad libraries for ads that are tangentially related but that do not fall within their political advertising definition. Some of the existing libraries, such as Snap’s Political Ads Library, are also difficult to access and navigate, decreasing the utility of the feature.\textsuperscript{33}

This information is critical for understanding what kinds of advertisements run on these services, which advertisers benefit from certain policies, and where platform advertisement moderation efforts are falling short. Publishing data around ad enforcement is particularly critical, as this facilitates evaluation of platform advertising policies and practices and allows independent assessment of where and how advertisements—including advertisements with misleading information—can slip through the cracks. As of now, only two platforms—Reddit and TikTok—publish any such data, although this data is not directly related to election misinformation or election and issue advertising. Given that there is a dearth of data available in this area right now, we awarded these two companies credit.
Conclusion

With the 2022 midterm elections on the horizon, it is all the more urgent that internet platforms take concrete steps to combat misinformation and disinformation related to election and voter suppression. Since the 2020 elections, we have seen promising improvements in the creation of platform policies to address the spread of misleading election content and in the addition of features that allow users to flag misleading content. However, there is still significant room for platforms to improve their efforts to address misinformation and promote civic engagement.

Many platforms still lack specific policies on voter suppression and do not have mechanisms for notifying users who engage with content that is misleading or unverified. Additionally, despite significant pressure, most platforms have not yet addressed changes tied to their core business models, such as reviewing and fact-checking advertisements or preventing repeat spreaders of misinformation from monetizing or advertising on a platform.

Most concerningly, there is an overall lack of transparency and accountability across all efforts to tackle election-related misinformation and disinformation. Going forward, companies should make this information available so that the effectiveness of current efforts can be accurately evaluated.

As people continue to turn to internet platforms to discuss, debate, and find information on elections, companies must enhance their efforts to identify and combat misleading information, and ensure they are providing adequate transparency and accountability around these efforts.
Notes


2 Singh and Blase, Protecting the Vote, https://www.newamerica.org/oti/reports/protecting-vote/.


12 Singh and Blase, Protecting the Vote, https://www.newamerica.org/oti/reports/protecting-vote/.


18 “Civic Integrity Policy,” Twitter Help Center.


Singh and Blase, Protecting the Vote, https://www.newamerica.org/oti/reports/protecting-vote/.


32 “Political Advertising on Google,” Google, https://adstransparency.google.com/political?political=®ion=US.


34 Singh and Blase.
This report carries a Creative Commons Attribution 4.0 International license, which permits re-use of New America content when proper attribution is provided. This means you are free to share and adapt New America’s work, or include our content in derivative works, under the following conditions:

- **Attribution.** You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

For the full legal code of this Creative Commons license, please visit [creativecommons.org](http://creativecommons.org).

If you have any questions about citing or reusing New America content, please visit [www.newamerica.org](http://www.newamerica.org).

All photos in this report are supplied by, and licensed to, [shutterstock.com](http://shutterstock.com) unless otherwise stated. Photos from federal government sources are used under section 105 of the Copyright Act.