



September 2020

Protecting the Vote

How Internet Platforms Are Addressing
Election and Voter Suppression-Related
Misinformation and Disinformation

Spandana Singh & Margerite Blase

Disclaimer

Given that the elections are quickly approaching, the landscape of platform policies and practices is constantly changing. In this report, we have aimed to capture as much accurate and relevant information related to how internet platforms are responding to the spread of election-related misinformation and disinformation as possible up until the date at which we went to press.

About the Author(s)

Spandana Singh is a policy analyst with New America's Open Technology Institute.

Margerite Blase 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

Executive Summary	5
Introduction	11
Amazon	13
Facebook/Instagram	15
Google	22
Pinterest	27
Reddit	29
Snapchat	33
TikTok	36
Twitter	39
WhatsApp	44
YouTube	46
Recommendations	49
Recommendations for Internet Platforms	49
Recommendations for Policymakers	55
Conclusion	57

Executive Summary

The 2016 U.S. presidential election demonstrated how internet platforms can be used to spread false and misleading information and suppress voting. This can be done through various means, including through posts and advertisements that spread inaccurate information about dates, locations, and voting procedures, as well as content that threatens or intimidates particular communities, particularly communities of color, into not voting. Our research dives into what several popular internet platforms—Facebook and Instagram, Google, Pinterest, Reddit, Snapchat, TikTok, Twitter, WhatsApp and YouTube—are doing to combat false and misleading election information, including potential voter suppression content around the 2020 U.S. presidential election and beyond. It also lays out recommendations for how platforms and policymakers can better protect the public from such content.

These platforms, and the policies and practices they deploy, can have a strong influence on the strength and nature of democracy and discourse, both in the United States and around the world. As the 2020 U.S. presidential election approaches, and as the ongoing COVID-19 pandemic shapes how and where people vote, strong practices to combat election misinformation and disinformation, including voter suppression material, including voter suppression material, are critical

Key Findings

- Many internet platforms are establishing online hubs to house information related to voter registration, voting processes, and more ahead of the 2020 U.S. election. However, not all of these hubs are easily accessible by users.
- A number of platforms we researched do not house content and advertising policies in one central location and do not have comprehensive policies that outline their approach. This makes it difficult to understand the parameters of these policies, and puts the onus on users and researchers to dig through numerous web pages and documents to figure out how and when these policies apply.
- Platforms are addressing misleading information in political ads in a myriad of ways, from banning political ads completely, to instituting restrictions on political ad targeting and delivery. However, there is no consensus on which approach is optimal, as they all present flaws and limitations.

- There is a discrepancy between the election-related policies that companies are creating and how they are being enforced. Companies fail to provide adequate transparency and accountability around the scope and scale of these policy enforcement and moderation efforts, and where they fall short. This raises some serious concerns around whether these policies and practices are effectively addressing the spread of election-related misinformation and disinformation online.
- Companies are relying on middle-ground moderation and curation efforts, such as labels and downranking, to handle election-related content. However, these policies and practices are applied inconsistently, and there is little transparency and accountability around the policies that guide their use.
- There is a fundamental lack of transparency and accountability around what platforms are doing to handle election misinformation and disinformation online, and if platforms' efforts are effective. Most internet platforms covered in this report publish transparency reports which include some data on the scope and scale of their content moderation efforts. However, only one company publishes data directly related to the moderation of election-related content and very few publish data related to the moderation of misleading information. This information is vital to understand where companies are taking action, what influence these actions are having, and where these efforts are falling short.
- There is contention over whether or not internet platforms should fact-check content and ads on their services. While some internet platforms fact-check user-generated content, fewer fact-check ads. This is concerning because entities and politicians could precisely target specific audiences with potentially false information.
- Some platforms are adopting labels and identity verification standards to provide transparency around which businesses—such as foreign media outlets and advertisers—are sharing information online. These efforts are largely seen as a response to the 2016 U.S. presidential election, where foreign actors used U.S. internet platforms to spread misleading information among U.S. voters without having to disclose their identity or location.
- Publishing an ad transparency library provides insight into the kinds of political ads that are run on a platform, but this practice is not widely adopted (and some platforms have banned political ads completely). In addition, there is a serious lack of quantitative and qualitative transparency around how platforms moderate ads based on their

advertising content policies, and what impact these efforts have on the ads available on their platforms.

- Algorithmic curation and amplification processes can significantly boost or undermine the reach of a piece of content or an ad. Some companies have recalibrated their algorithmic ranking and recommendation systems to prevent the amplification of election-related misinformation and disinformation. However, there is still a significant lack of transparency around how these tools are trained and used, what impact these disclosed changes have had on the spread of misleading and false election information, and how or if humans are kept in the loop.

Recommendations

Going forward, internet platforms and policymakers should consider the following set of recommendations prior to the 2020 U.S. presidential election as well as in the long-term to address future elections. The section below includes an excerpt of our recommendations on how companies can improve their efforts to connect users to, and lift up, authoritative information; address the spread of misleading information through content moderation and curation; tackle misleading ads; and provide meaningful transparency and accountability around these efforts. This section also includes recommendations for how U.S. policymakers can encourage greater accountability and integrity from internet platforms, although they are limited in the extent to which they can direct how internet platforms decide what content to permit on their sites.

Recommendations for Internet Platforms

Sharing and Lifting Up Authoritative Information and Empowering Informed User Decision-Making

- Partner with reputable fact-checking organizations and entities, as well as local and state election bodies to verify or refute information circulated through organic content and ads.
- Notify users who have engaged with misleading election-related content and direct them to authoritative sources of information.
- Institute a public interest exception policy that permits companies to leave content posted by world leaders, candidates for political office, and other government officials on their services, even if the content has been fact-checked and contains misleading information. In instances where the company determines that the content posted by officials could result in

imminent harm, this public interest exception policy should not be applied and the content should be removed.

- Conduct regular impact assessments and audits of algorithmic curation tools (e.g. ranking and recommendation systems), and recalibrate them as necessary so they do not direct users to or surface misleading content when they search for election-related topics and do not algorithmically amplify such content in trending topics and recommendations.
- Label organic content and ads that have been produced by state-controlled media outlets to inform users of the content's origins.

Moderating and Curating Misleading Information

- Create a comprehensive set of content policies to address the spread of election-related misinformation and disinformation with specific considerations for voter-suppressive content. Companies should house these policies in one location, provide public notice if their policies change, and include an archive of past policies.
- Companies should clarify to what extent election-related policies interface with content policies related to hate speech, deepfakes, bots, coordinated inauthentic behavior, etc. While manipulated media may be a part of user expression on social media and therefore permissible for user-generated content, platforms should consider banning the use of such manipulation technologies for political advertising.
- Institute a dedicated reporting feature which enables users to flag election-related misinformation and disinformation to the company.
- Remove, reduce the spread of, or label content that has been fact-checked and 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 ads. The policies should include specific considerations for voter-suppressive ad content and should clarify to what extent these policies interface with advertising policies related to hate speech, bots, deepfakes, etc.
- Establish a comprehensive review process for election-related ads and ad targeting categories. Companies should require all election-related ads to

be fact-checked and reviewed by a human reviewer before they are permitted to run on a platform. Companies should publicly disclose high-level information on what this review process consists of and to what extent it relies on automated tools and human reviewers.

- Create a comprehensive vetting process for advertisers which requires them to verify their identity and which country they are based in before running ads.
- Append “paid for” disclosures to all paid political, social, and issue ads and ensure labels are maintained even if ad campaigns end or if ads are organically shared online.
- Create policies that prevent users and entities from being able to monetize and advertise on the platform if they repeatedly spread misinformation and disinformation.

Provide Meaningful Transparency and Accountability

- Explain to users how and to what extent content that is flagged for violating election-related misinformation and disinformation policies is reviewed, moderated, and curated by human reviewers and by automated tools. Users should be notified of any significant updates to these processes.
- Preserve data on election-related content and advertising removals. Vetted researchers should have access to this data so they can identify where these content and advertising moderation policies and practices fell short and make recommendations on how they can be improved.
- 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 election-related ad content and targeting policy enforcement efforts.

Recommendations for Policymakers

- Policymakers should enact rules to require greater transparency from online platforms, including regular reporting regarding their content moderation, curation, labeling, and ad targeting and delivery efforts.
- Authoritative election authorities such as the Federal Elections Commission (FEC), state election boards, and other state and local authorities should partner with internet platforms to provide and promote verified and legitimate information related to the election on their platforms. These entities should also help debunk misleading claims and information using their own online accounts.
- Policymakers should clarify that the Voting Rights Act, which prohibits suppressing voting through intimidation, applies in the digital environment. Further, Congress should amend the Act or pass new legislation to prohibit suppression of voting through deception, which is the primary means of vote suppression online.

Introduction

Following the 2016 U.S. presidential election, internet platforms have come under increased scrutiny for how they handle the spread of misinformation and disinformation on their services. Since that election, numerous researchers have found evidence that social media platforms served as hotbeds for the spread of election-related misleading content, including content designed to suppress voting. These efforts fanned existing societal tensions around race and socioeconomics, and they disproportionately impacted communities of color and other marginalized groups. For example, during the 2016 elections, Russian operatives fraudulently posed as Black Americans to actively dissuade the Black community from voting.¹

Social media platforms can and have been used to spread false information and suppress voting in a number of ways. These include posts and advertisements that spread inaccurate information about dates, locations, and voting procedures, as well as content that threatens or intimidates particular communities into not voting. These types of content can undermine trust in the electoral process and discourage voters from participating at all.

Since 2016, internet platforms have instituted a range of policies and practices that seek to identify and curb the spread of election-related misinformation and disinformation. However, experts and users have little confidence in the efficacy of these measures.² According to a 2018 national survey conducted by the Brookings Institution, 57 percent of those surveyed felt that they had seen fake news or misleading information during the 2018 U.S. midterm elections, and 19 percent believed that this information had influenced how they planned to vote.³ In addition, a January 2020 Pew Research Center study found that just 25 percent of U.S. adults felt confident that tech companies would be able to prevent the misuse of their platforms during the upcoming elections. This was a decrease from 33 percent prior to the 2018 midterm elections.⁴

As the 2020 U.S. presidential election draws near, experts are concerned that social media platforms will be used by both foreign and domestic actors to suppress votes and spread misleading information.⁵ In particular, many experts fear that because individuals are relying more on digital resources to learn about voting procedures and policies, they will be especially susceptible to misinformation and disinformation.⁶ In addition, watchdog organizations have also expressed concerns that these platforms will be used to suppress voting by exploiting users' fears around COVID-19 in order to encourage them to avoid polling places, which could particularly affect participation among older voters.⁷ Further, there are also concerns that entities seeking to suppress voting could use the ongoing protests related to racial justice in the United States to push out messaging that voters from certain communities should protest racial injustice by

not participating in the electoral process.⁸ Thus far, internet platforms have responded to concerns of election and voting-related misinformation and disinformation in a number of ways. Major tech companies, including Facebook, Google, Pinterest, Reddit, and Twitter have announced that they plan to meet regularly with each other and government agencies to discuss ongoing trends and ways to protect information around the 2020 election.⁹ Many platforms have created new policies or expanded existing ones to cover these categories of content, as well as related forms of content that could impact elections such as hate speech, fake accounts, and inauthentic behavior. In addition, many internet platforms have begun examining the role political advertising can play in fostering a false information ecosystem on their services. Some companies, such as Amazon and Twitter, have banned political advertising altogether. Other platforms, such as Google and Snapchat, have instead introduced guidelines for political ads. However, there is still a significant lack of transparency and accountability around how these platforms are creating and implementing these policies, sparking concerns that these policies are not being implemented consistently and are ineffective. In addition, this has also raised concerns that platforms may be prioritizing profit over the safeguarding of user rights and the electoral process.¹⁰ Internet platforms, used by millions of people in the United States every day, have assumed a central role as gatekeepers of speech in society. Given that there are no clear laws that address the spread of election-related misinformation and disinformation online, these platforms are also the de facto “legislative, judicial, and executive branches” in terms of preventing online voter suppression.¹¹ As a result, these platforms, and the policies and practices they deploy, can have a strong influence and impact on the strength and nature of democracy and discourse, both in the United States and around the world.

This report will provide an overview of how various internet platforms are addressing the rapid spread of election-related misinformation and disinformation, and particularly content that promotes voter suppression. The report concludes by offering recommendations on how platforms can improve the efficacy of their efforts and provide greater transparency for their users and the public. The report also includes recommendations on how U.S. policymakers can encourage further accountability and support efforts to combat the spread of misinformation and disinformation around voting.

Editorial disclosure: This report discusses policies by Facebook (including Instagram and WhatsApp) and Google (including YouTube), both of which are funders of work at New America, but neither of which contributed 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.

Amazon

Amazon is one of the most popular online platforms in the United States, with an average of 200 million unique visitors per month.¹² The platform has also played an important role during the COVID-19 pandemic as more consumers rely on online shopping.¹³ Between January and March of 2020, the company made \$75.5 billion in revenue, a 26 percent increase over the prior year's first quarter. Amazon, an e-commerce and advertising platform, can be a source of misleading information and voter suppression content through both its products and advertisements. Amazon sells products, such as t-shirts, books, and other merchandise that are related to politics, voting, and the election. The platform does not place specific requirements on these types of products. Rather, it utilizes the same general rules that apply to all of Amazon's products, which includes acting "fairly and honestly on Amazon to ensure a safe buying and selling experience."¹⁴ However, there is little transparency around how these policies are enforced or whether the company takes any steps to address voter suppression content within its e-commerce operations.

In 2018, Amazon accrued over \$10 billion in revenue from its advertising platform, making it the third-largest ad platform in the United States behind Facebook and Google.¹⁵ To help avoid the issue of voter and election misinformation, Amazon banned political ads. The platform's policy prohibits ad campaigns that are for or against a politician or political party or that are related to an election or political issue.¹⁶ Still, it is unclear how the platform intends to enforce this policy. An Amazon representative speaking with CNBC at the beginning of this year stated that Amazon uses automated tools and review teams to monitor and remove policy-violating ads.¹⁷ However, CNBC claims that it was able to easily locate ads for political products sold on Amazon, including "Talk Bernie to Me Sanders 2020" and "Trump 2020 The Sequel Make Liberals Cry Again" t-shirts. Going forward, Amazon should clarify whether its policy prohibiting political ads pertains to products sold on the platform that promote a particular candidate or party. The company should also share information around how it enforces its political ads ban, including information on how its automated review tools and moderation teams are trained and deployed to moderate political advertisements, and how effective these approaches are.

Furthermore, Amazon owns Alexa, an automated home assistant and cloud-based voice service that is available on millions of devices. Alexa can function as a search engine for voting and election information, such as the latest polling information or a candidate's stance on a certain topic.¹⁸ In order to stay up-to-date on general election and voting information, such as when voting polls are open, Alexa pulls information in from authoritative sources including Associated Press, Ballotpedia, RealClearPolitics, and Factba.se.¹⁹ However, Alexa is often unable to provide more contextual and regulatory information, such as whether

voting machines are going to be used in a certain area or whether a specific jurisdiction requires a voter ID.²⁰ According to Amazon, it decided to focus on areas where the platform sees the most amount of customer interest and need.²¹ However, because Alexa promotes itself as a service where individuals can obtain election and voting information, the platform should expand its efforts to connect users with more relevant, contextual, and accurate information related to the elections.

While Amazon has disclosed some of the information sources Alexa uses for voting and election information, the company has not publicly published information outlining how it addresses election-related misinformation, and it is unclear how, or if, the company vets the information it receives from its sources for accuracy. Going forward, the company should provide greater transparency on its policies for misinformation and disinformation on Alexa and whether it takes any steps to fact-check the information Alexa shares with its users.

Furthermore, Amazon should provide greater clarity on its advertising policies for Alexa. Advertising is generally not allowed on Alexa except for a few exceptions, such as promotional offers or deals in response to specific requests from customers.²² It is unclear whether Alexa has political advertising-specific policies and whether Amazon's ban on political advertising on its platform applies to Alexa as well.

Amazon receives millions of monthly unique users and is therefore at risk of playing a significant role in spreading misinformation. As elections occur across the country, the company should provide greater transparency and accountability around its policies and practices related to misinformation and disinformation across all of its products. In addition, the company should provide greater transparency around its ad policy enforcement efforts, especially as they relate to political ads. This should include information around how Amazon uses automated and human review components to enforce its ad policies, how effective these policies are, and how many policy-violating political ads the company has removed after erroneously permitting them to run on the service.

Facebook/Instagram

Facebook is the largest social media platform in the world, with over 2.4 billion active users.²³

Since the 2016 U.S. presidential election, numerous researchers concluded that Facebook was a prominent site for election-related misinformation and disinformation, including voter suppression content. Since then, Facebook has taken steps to improve its policies and practices related to elections and misleading information. Facebook has also adopted many similar policies and practices for Instagram, a photo and video sharing social media platform that it owns. As a result, this section also outlines some efforts implemented by Instagram to combat election and voter suppression misinformation and disinformation.

Prior to the 2018 U.S. midterm elections, the company expanded its policies addressing voter suppression and intimidation to explicitly ban misrepresentation of the dates, locations, times and ways that voting or voter registration can take place; misleading information about who is qualified to vote, whether a vote will be counted, and other parts of the voter process; and threats of violence related to voting, voter registration, or the outcome of an election.²⁴ According to Facebook, the company removes content in these categories regardless of who posted it.²⁵ The company also said that prior to the midterm elections its Elections Operations Center removed over 45,000 pieces of content that violated these policies, out of which 90 percent was proactively detected by Facebook's automated systems.²⁶ Further, in a June 2020 blog post, Zuckerberg stated that Facebook will remove any misleading claims that aim to discourage voting, and noted that politicians will be subject to these policies as well.²⁷

Facebook also introduced a reporting feature which enables users to flag potentially incorrect voting information. Additionally, the company established dedicated reporting channels that state election authorities can use to flag potentially false voting information as well.²⁸ Further, Facebook has established partnerships with over 30 voting rights and election protection groups, enabling these groups to monitor and flag election-related content that potentially violates the platforms' content policies for review.²⁹ In September 2020, the company announced that it would expand its partnership with state election authorities to address misleading information about polling conditions.

Facebook's Election Operations Center, which is responsible for enforcing voter and election-related policies, will be specifically tasked with addressing false claims about polling conditions. Facebook initially aimed to focus the Center's work on addressing misleading information about polling conditions on the 72 hours prior to election day,³⁰ when content volumes and flags are typically higher.

³¹ However, Facebook began instituting these efforts in September in response to the large number of early voters expected due to COVID-19.³²

Facebook partners with independent third-party fact-checking organizations such as the Associated Press, Reuters Fact Check, and PolitiFact to review content on the platform that is suspected to be misleading.³³ Once these fact-checking partners have debunked a piece of content, Facebook reduces the distribution, or downrank, this content in the Facebook News Feed. Facebook may also apply a warning label to the debunked content.³⁴ Facebook appends labels to photos and videos, as well as Instagram Stories,³⁵ with the intent of allowing users to decide whether or not they'd like to view the content. Each label contains a link to the fact-checkers' evaluation of the content at hand.³⁶ Facebook may also feature "Related Articles" written by fact-checkers alongside debunked content in order to add context to the debunked post.³⁷ When a user tries to share this debunked content on either Facebook or Instagram, they see a pop-up warning them that the content has been proven to be inaccurate.³⁸ Further, if Pages, domains, and Groups repeatedly post misleading content on the service, Facebook down ranks them and restricts Page owners from advertising and monetizing.³⁹ On Instagram, this content, as well as content posted by accounts that continuously share misleading information, is omitted from the Explore and hashtag pages.⁴⁰ Coordinated inauthentic behaviors and campaigns were a hallmark of the disinformation campaigns which sought to suppress voting and sow discord during the 2016 U.S. presidential election. As a result, the company has also updated its policies and practices in this regard. In particular, Facebook updated its policy regarding inauthentic behavior to better explain how the company responds to foreign, domestic, state, and non-state led deceptive efforts.⁴¹ Further, in order to prevent Page owners from masking their identity, Facebook requires that all Pages, including those that are election-focused, have a confirmed Page owner and provide verified information such as the organization's legal name and its website, among other things.⁴² Advertising also played a prominent role in spreading misleading information during the 2016 presidential election. As a result, Facebook introduced a series of metrics and features to provide greater transparency around its advertising operations. First, the company introduced a tracker which enables users to see how much money U.S. presidential candidates have spent on ads. This ad spending information can be broken down at the state or regional level to demonstrate what specific geographies candidates are focusing their ad spend on. The company is also making efforts to clarify whether an ad ran on Facebook, Instagram, Messenger, or on Facebook's Audience Network.⁴³ Facebook Audience Network allows advertisers to extend their Facebook and Instagram campaigns across the internet.⁴⁴ The company also introduced new features including API filters that allow journalists, researchers, and others to access and download ad creatives as well as a collection of frequently used API scripts.⁴⁵ Further, the company instituted new rules which require advertisers to assign a verified Page Owner to

their pages in order for them to run issue, electoral, or political ads in the United States.⁴⁶

Now, closer to the 2020 presidential election, Facebook has instituted additional policies and procedures that aim to tackle the spread of election and voter-suppression related misinformation and disinformation. This includes the launch of the Voting Information Center in August 2020, which has been dubbed “the largest voting information effort in U.S. history.” Through the Center, Facebook aims to increase participation in the election by helping 4 million Americans register to vote across the Facebook, Instagram, and Messenger products⁴⁷ and also seeks to promote accurate and authoritative information about elections in order to counter misinformation and disinformation.⁴⁸ The Center includes guidance on how to register to vote and how to vote (including information on both mail and in-person voting), as well as election results.⁴⁹ The Center is also a hub for updates from local election authorities regarding any changes to the voting process. The information in the Center is drawn from state election officials and other nonpartisan civic organizations.⁵⁰ However critics have expressed concerns that the Center is difficult to locate on the platform, as it requires users to navigate multiple drop-down menus.⁵¹ In September 2020, Facebook CEO Mark Zuckerberg announced that the company will place information from the Center at the top of Facebook and Instagram in the days running up to the election.⁵²

In addition, in order to provide greater transparency around what governments and other entities are behind news posts on Facebook and Instagram, Facebook has instituted a labeling policy for media outlets that are partially or entirely state-controlled.⁵³ This is particularly important given the role foreign governments played in pushing content labeled as news during the 2016 presidential election. Facebook also said it would begin labeling ads from such publishers later in 2020, although a concrete launch date has not been announced.⁵⁴ Advocates and experts have raised concerns, however, about whether the effectiveness of these labeling efforts will be undermined by a failure to implement them consistently.⁵⁵ In response to growing concerns that Facebook’s advertising platform can and likely will be used to promote election-related disinformation including voter suppression content, the platform is permitting users to opt-out of all social issue, electoral, and political ads across all Facebook products.⁵⁶ Users who choose to view political ads will be able to see who paid for these ads even after they have been shared by other users.⁵⁷ However, this approach falls short in several ways. First, this new policy puts the onus on users to explicitly opt-out of viewing political ads. Further, the policy does not address pre-existing concerns around Facebook’s flawed advertising policies and policy enforcement process,⁵⁸ especially concerns that the company does not fact-check political ads. Facebook CEO Mark Zuckerberg has long argued that no social media company, including Facebook, should be the arbiter of truth.⁵⁹ As a result, he has pushed back on calls for the company to remove

false claims, particularly those made by politicians, on the platform.⁶⁰ While Facebook should not be the arbiter of truth, it can and should do more to ensure it is not amplifying and enabling the spread of harmful misinformation and disinformation on its platform. One way of doing this could be by fact-checking political ads and subsequently notifying users when content in political ads has been debunked.⁶¹ Finally, this new policy does not address broader concerns related to access to microtargeting tools, which enable advertisers to precisely target users based on a range of personal data points, and can be used to target specific groups of users with misleading information.⁶²

In September 2020, Facebook also announced that it would ban new political advertisements on the platform during the week preceding the November 3 election.⁶³ Campaigns can, however, continue to promote ads that they placed on or before October 27, as long as the ads were viewed by at least one Facebook user.⁶⁴ Although this policy change prevents new advertisements from being introduced immediately prior to the election, it does not address growing concerns that the company does not fact-check content in its advertisements. As a result, false information can still circulate through ads as long as they are posted on or before October 27.⁶⁵

In addition, Facebook said that since 2016, the company has tripled its workforce that focuses on “security and safety issues” and is responsible in part for content moderation on the platform.⁶⁶ The company has also stated that it uses machine-learning to rapidly identify and remove inaccurate voting information, and that these efforts have become more effective over time.⁶⁷ Due to the unprecedented COVID-19 pandemic, Facebook, and many other internet platforms, had to adjust its content moderation operations as its content moderator workforce could not initially work remotely.⁶⁸ As a result, Facebook relies more heavily on automated tools to detect and flag certain categories of content. Although Facebook has since been able to readjust its content moderation operations to an extent, the company warned that users should expect more mistakes.⁶⁹ Given that a significant amount of election-related content moderation is occurring during the pandemic, Facebook should preserve data related to election-related content removals during this period so that researchers can evaluate these efforts later on.⁷⁰ This is a good best practice in general, as it allows researchers to assess where moderation efforts fell short and can be improved.

There is currently little transparency around how the company enforces its voter suppression and election-related content policies. In Facebook’s Community Standards Enforcement Report (CSER), the company outlines how it enforces some of its content policies, and how often it receives and takes actions based on appeals. However, the CSER does not include any data related to the enforcement of misinformation or voter suppression-related policies and related appeals. The report includes data on categories of content that could intersect with the company’s election-related misinformation policies, such as fake

accounts and hate speech. However, this data is not sufficient to fully understand the nature of election and voter suppression-related misinformation and disinformation on the platform.⁷¹ In addition, there is little transparency around how Facebook's machine-learning and automated tools are trained, refined, and used, and how effective they are.⁷²

Further, currently, content that users report to the platform as voter interference is not immediately sent for review by human review teams. Rather these flags are considered “user feedback” and are used to evaluate aggregate trends. If Facebook receives a large volume of user reports for a piece of content, then that content will be reviewed by its policy and operational teams. According to Facebook, the company relies on this process because during the 2018 midterm elections, a low number of user reports of voter interference involved content that actually violated the platform’s policies. The majority of flagged content instead were posts that expressed differing political opinions from the flagger. Instead, Facebook said that during the midterm elections, over 90 percent of the content it removed for violating its voter suppression policy was detected proactively using its automated tools.

However, both online efforts to suppress voting and Facebook’s voter suppression policies have changed since 2018, and civil rights experts have raised concerns around whether Facebook should subsequently change its moderation practices to route user-flagged content for human review. In addition, if Facebook does not review content, users cannot appeal moderation decisions. The decision to not review user flags for voter interference therefore denies users a right to appeal and redress.⁷³

In July 2020, Facebook introduced a new election-related labeling policy in response to push back from a range of organizations, including those who led the #StopHateForProfit campaign,⁷⁴ who argued that Facebook does little to address misinformation and hate speech in its content moderation and advertising practices.⁷⁵ The new policy allows Facebook to append “Get Voting Information” labels⁷⁶ to content that mentions voting to provide users with relevant information about the voting process.⁷⁷ However, although Zuckerberg stated this policy will also apply to politicians, researchers have expressed concerns that the labels will not be applied consistently and that they will fail to have a meaningful impact.⁷⁸ These concerns have been underscored by the fact that Facebook seems to be appending these labels to posts that refer in any way to voting, rather than posts that are inaccurate or misleading. For example, the company not only attached a “Get Voting Information” label to a July 21 post by President Trump which states mail-in ballots would result in the “most CORRUPT ELECTION” in the history of the United States⁷⁹ but also to a straightforward post by Kimberly Klacik, a Congressional candidate for Maryland’s District 7, which states “Please vote KIM KLACIK on November 3rd. We are getting all of our ducks in a row. On Day 1 you will see you made a great

choice.”⁸⁰ Civil rights groups are concerned that Facebook’s decision to apply a label to any voting-related post, regardless of the content, creates no distinction between accurate and misleading content. Additionally, these groups have stated that the use of such broad labeling procedures could reduce the company’s sense of urgency around removing false election-related information, since the content will have a label directing users to the Voting Information Center.⁸¹ In addition, some watchdog groups have pressed the company to go one step further than labeling content and notify users who have viewed or engaged with misleading election-content while on the platform.⁸²

The platform also shared it would broaden its existing prohibition⁸³ on posting content that misleads individuals on how they can vote. As a result, claims such as Immigration and Customs Enforcement (ICE) agents will be reviewing immigration papers at polling stations or individualized threats of voter interference would be banned on the platform.⁸⁴ These updated policies also ban threats of coordinated interference that could intimidate or discourage individuals from voting.⁸⁵ Facebook has also said that when posts aim to delegitimize the outcome of the election or undermine the legitimacy of voting methods, the company will add information labels to these posts that include links to authoritative information.⁸⁶

Further, Zuckerberg announced that politicians will be subject to all these new policies, although if there is public interest or newsworthy value in some of the content, it will be left up and labeled.⁸⁷ This sparked concern given recent events where political figures spread misleading content about the election process⁸⁸ that some experts say amounts to voter suppression.⁸⁹ Going forward, when companies deem there is a public interest value in leaving such content up, any labels that they use should provide sufficient contextual information that explains the content is misleading and is being left up for awareness purposes. The company should also create a central policy that guides such cases, rather than use a disparate series of ad hoc statements and policies to make these determinations. In September 2020, the company also updated its policies to prohibit using the COVID-19 pandemic to discourage voting in both content and advertising. The company also stated it would include a link to authoritative information around COVID-19 in such posts.⁹⁰

In 2018, Facebook committed to participating in an independent civil rights audit of the impact of its policies and practices on communities of color and other underrepresented groups.⁹¹ The final civil rights audit report, released in July 2020, outlined how Facebook has broadened its voter suppression and intimidation policies over the past two years to cover a more expansive set of threats and scenarios. However, the report also stated that in order for these policies to be effective, the company needs to interpret them in a more comprehensive and consistent manner.⁹² The overall audit outlined that the company’s lack of a strong civil rights foundation has resulted in numerous

concerning outcomes, including the creation of opaque policies and practices related to elections and voting-related content, as well as the inconsistent and incomplete enforcement of these policies. By contrast, the platform responded more proactively and aggressively to the rapid spread of COVID-19 misinformation and disinformation. Some have suggested that this shows that when Facebook is committed to addressing a category of harmful misinformation, it has greater capabilities than it has demonstrated in the context of election-related misinformation.⁹³ In furtherance of a commitment made by Facebook in the civil rights audit report released in June 2019,⁹⁴ the company implemented a new policy banning paid ads that state that voting is meaningless or discourages people from voting.

Google

Google, one of the world's largest technology companies, sees roughly 3.5 billion searches per day.⁹⁶ As the 2020 election approaches, people are likely going to increase their voting and election-related searches online, and it will be critical for Google to combat misleading information to ensure that its users can successfully participate in the electoral process.⁹⁷

Google does not have content guidelines or restrictions for the websites that appear within its search engine. According to Google, its goal is to identify relevant information on the web based on users' queries, not to decide whether the material is "in some sense wrongful."⁹⁸ However, due to recent concerns around election misinformation, Google announced on September 10, 2020 that it updated its search function to no longer suggest autocomplete search phrases for users looking up information on candidates or voting.⁹⁹ Searches for informational claims on a candidate, how to cast a ballot, or the legitimacy of the electoral process will need to be typed out fully. The search engine will still suggest phrases for questions around voting or the election, i.e. how to vote, but it will not autocomplete claims about voting or the election, i.e. you can not vote by mail. Google stated that it made this change to help prevent bad information from appearing in autocomplete suggested phrases.¹⁰⁰

Apart from its search function, Google places requirements on advertising that runs on its platforms. According to Google's ad policies, all ads must be "clear and honest, and provide the information that users need to make informed decisions."¹⁰¹ The company prohibits ads that deceive users by including misleading information about products, services, or businesses. This includes the use of false claims or "deceptively doctoring media related to politics, social issues, or matters of public concern."¹⁰² As of September 2020, Google's misrepresentation policy for ads also prohibits accounts from coordinating with other sites or accounts to conceal or misrepresent their identities or other material details if the content of the account relates to politics, social issues, or matters of public concern.¹⁰³ Although these general policies can help Google combat election and voting-related misinformation in advertising, the company does not have policies that specifically address ads aimed at suppressing voting. To ensure that these types of ads would not be allowed to run on its platform, Google's policies should be updated to specifically address voter suppression content as well as election-related misinformation, especially leading up to the 2020 election.

Google states that accounts that violate its updated misrepresentation policy will be suspended "upon detection and without prior warning" and will not be allowed to advertise on Google's platform in the future.¹⁰⁴ It is unclear if Google's standard appeals process for advertisers would apply to these accounts.

Typically, if an advertiser feels that their ad was unfairly removed, they can appeal the decision directly in Google Ads, Google's online advertising platform, by hovering over a disapproved or limited ad and clicking on an "appeal" link.¹⁰⁵

Due to the COVID-19 pandemic, Google reduced its moderation and review workforce capacity because the number of people going into offices decreased and some content moderation tasks dealing with sensitive customer data or graphic content must be done onsite for security and wellness purposes.¹⁰⁶ The company warned that this may result in more content moderation mistakes as well as slower appeals review processes.¹⁰⁷ Given that a significant amount of election-related content moderation is occurring during the pandemic, Google should preserve data related to election-related content removals during this period so that researchers can evaluate these efforts later on. Additionally, the company should ensure that users and advertisers have access to a robust and timely appeals process. It is important for the company to remove *misleading* ads, but accurate voting and election-related ads on Google can be beneficial to encourage active participation and keep users informed, and advertisers who have their ads erroneously flagged should have the opportunity to appeal these decisions.

Google's general ad targeting and delivery practices have raised concerns in the past for infringing on users' privacy and for generating discriminatory outcomes.¹⁰⁸ In an effort to increase accountability for political ads and help combat misleading information on the platform, Google updated its political ad policies in November 2019.¹⁰⁹ Under Google's updated policies, political ads must "comply with local legal requirements, including campaign and election laws for any geographic areas they target."¹¹⁰ This covers any ad that is for a political organization, political party, political advocacy or fundraising event, or any ad pertaining to an individual candidate or politician.¹¹¹

While Google claims that it has never allowed "granular microtargeting" for election ads, under the new policies, election-related campaigns have additional targeting limitations compared to non-election campaigns.¹¹² For example, election ads may only target users based on geographic location (i.e. state or zip code), age, gender, or contextual targeting such as by topic or keywords against sites. Election ads may not utilize Google's other advertising options such as remarketing, customer matches, geographic radius targeting, or third-party audiences. Limitations on microtargeting and other granular targeting, like radius targeting, are important because these techniques can be used to divide the public into very small groups,¹¹³ allowing misinformation or misleading content to avoid scrutiny from the broader public and spread more easily.¹¹⁴ While having a ban on geographic radius targeting can have a positive impact on ad visibility, Google does allow political advertisers to target their ads based on other geographic information, including zip code. Zip code targeting can be problematic when it comes to ads for housing, employment, and credit

opportunities, because zip codes can be used as a proxy for race.¹¹⁵ While zip code targeting could be beneficial for election related ads on local elections or for a presidential candidate attempting to increase their visibility in a certain region, it is worth examining whether advertisers have been able to use zip code targeting to promote voter suppression tactics for the presidential election.¹¹⁶

Along with targeting restrictions, Google requires identity verification for advertisers running federal or state election ads.¹¹⁷ There is a specific U.S. election ads verification form that requires an organization or individual to provide a Google Ads Customer ID and email.¹¹⁸ Google reviews the application and provides an email notification within three to five business days. If the platform discovers that an account has violated election-related advertising policies or provided false information during the verification process, the account's verification is revoked. Political ads are also required to have an in-ad disclosure that displays the name of the party or individual paying. For most ad formats, the "paid for by" disclosure is automatically generated from the account's information. Otherwise, it is the advertiser's responsibility to add the disclosure.¹¹⁹

Google publishes a transparency report to keep the public informed about political advertising on Google, YouTube, and partner properties.¹²⁰ The report includes information such as the number and the amount of money spent on political ads since May 31, 2018, top advertisers by total ad spend, and ad spend by state. It also allows people to view or search for ads. Although these data points are a good starting point, the report does not break down certain information, such as ad type (i.e. search versus display), or provide granular details such as the type or amount of policy-violating ads that have been removed. In addition, the report only includes ads that feature a "current officeholder or candidate for an elected federal or state office, federal or state political party, or state ballot measure, initiative, or proposition that qualifies for the ballot in a state." As a result, the report does not provide a comprehensive overview of all political ads that are run on the platform, and it creates a significant gap in terms of which ads are available for public scrutiny.¹²¹ In addition, Google does not currently publish a comprehensive transparency report outlining its content policy enforcement for its web search product (it does publish one for YouTube).¹²² Going forward, the company should expand its transparency reporting efforts to include information on policy-violating content, and ensure that the report includes data on the enforcement of its election-related misinformation and disinformation policies. Further, in the run up to 2020 elections, the company should provide periodic updates on the enforcement of its election-related misinformation content policies and ad policies, including how much user content and how many accounts have been removed, and how many ads have been rejected and removed for violating election-specific policies or supporting voter suppression. Following the 2020 presidential election, the company should issue a comprehensive transparency

report detailing its moderation efforts that includes data on voter suppression and election related content.

Outside of advertising, Google also offers Google News, an aggregator that helps connect users to different news and media outlets.¹²³ The platform utilizes an algorithm to determine how news is ranked, which relies on factors such as relevance of content, prominence, location, and language.¹²⁴ Google does not accept payments to expedite or increase a site's visibility within the news aggregator.¹²⁵ To appear in Google News, a publisher does not need to submit their site for approval, they simply need to "produce high-quality content and comply with Google News content policies."¹²⁶ However, it is unclear how Google decides what type of content qualifies as high-quality. Under Google News' community policies for publishing partners, content may not contain deceptive practices or attempt to mislead users.¹²⁷ The platform also does not allow sites or accounts to impersonate any person or organization or misrepresent their primary purpose. However, Google News' policies do not specifically address election or voter suppression content.¹²⁸ Apart from providing information on how its algorithm ranks news sources, Google News should clarify how it first selects which news sources to include. The platform should also clarify if it takes steps to monitor the news sources for potential voter suppression and election misinformation content.

According to Google, in order to combat the spread of misinformation on its platforms, its Trust and Safety and Threat Analysis Group (TAG) teams monitor and combat inauthentic activity, disinformation campaigns, and other forms of abuse across Google's platforms on a 24/7 basis.¹²⁹ Google utilizes Jigsaw, a unit within the company that forecasts and fights against emerging threats, to help combat harassment and misinformation across a variety of topics, including voting and election information.¹³⁰ Google explains that Jigsaw's technology enables "deepfake" detection, therefore allowing the company to protect user accounts and campaigns that are targets of hacks or phishing attacks.¹³¹ Additionally, Google works with other technology companies and government agencies, such as the FBI's Foreign Influence Task Force, to help combat harmful interference online.¹³²

While Google has taken steps to combat voter and election related misinformation, the measures have not always proven effective. According to a Tech Transparency Project (TTP) analysis, Google failed to remove ads linking to websites that charged fees, in some cases fees up to \$129, to register to vote, although eligible voters can typically register directly with their states or territories for free.¹³³ From the analysis, it is unclear whether Google failed to remove these ads because they were overlooked or because they did not explicitly violate Google's policies. These types of ads can suppress voting because people may not be able to afford the supposed registration fees, or may not be willing to pay to register to vote. Also, if online users looking for

information about the election are sent to confusing websites or feel they are being manipulated, they may give up on trying to find important voting-related information. Furthermore, users may be directed to websites that are trying to extract personal information or receive fraudulent donations, which can have negative repercussions beyond voter suppression.

On August 13, Google released a blog post with additional steps they are taking to combat misinformation around the 2020 U.S. election.¹³⁴ The platform released two new features in Google Search that allow users to search for detailed information on how to register and when to vote by state, by searching “how to vote.” Google collects this information from Democracy Works, a third-party non-partisan organization that collects data directly from state and county election administrators.¹³⁵ The post also states that Google is reaffirming its investment in the company’s TAG and Trust and Safety teams, which are assigned to work together to prevent government-backed attacks and phishing campaigns on Google users. TAG tracks over 270 targeted or government-backed groups from over 50 countries, whose goals include “intelligence collection, stealing intellectual property, targeting dissidents and activists, destructive cyber attacks, or spreading coordinated disinformation.”¹³⁶ Google, along with other major tech companies including Facebook, Pinterest, Reddit, and Twitter, also participates in meetings with government agencies that are responsible for election integrity to discuss trends in the spread of misinformation.

While Google has been vocal about taking measures to combat voter and election-related misinformation and disinformation, the company should go further in providing transparency around how effective these efforts have been. For example, the company should include statistics in its transparency report on how many ads have been removed for violating ad policies and on how many government-backed attacks it stopped leading up to the 2020 election.

Pinterest

Pinterest, an image sharing and social media platform, allows users to discover new ideas or projects and share their own content with other users. As of 2019, the platform has more than 200 billion “pins” and more than 320 million monthly active users.¹³⁷ Given Pinterest’s role in allowing users to share and discover new ideas across a wide variety of topics, it could play an important role in elections and other forms of civic engagement. The company recently joined a cohort of other major tech companies, including Facebook, Google, and Twitter, and government agencies to discuss ongoing trends in deceptive online behavior and ways to combat misinformation.¹³⁸ In order to help avoid issues with voter misinformation and disinformation on its platform, Pinterest has banned all political advertising.¹³⁹ The platform’s ad policy does not allow advertising for:

- The election or defeat of political candidates running for public office, including fundraising for political candidates or parties
- Political parties or action committees
- Political issues with the intent to influence an election
- Legislation, including referendums or ballot initiatives
- Merchandise related to political candidates, parties, or elections

However, similar to other platforms that have banned political advertising, like Amazon, it is unclear how Pinterest enforces these policies on political ads. The platform does not provide details around the advertising approval process and whether potential ads are reviewed and approved by human-led teams or whether the platform utilizes algorithms.

Along with banning political advertising, Pinterest’s general Community Guidelines state that the platform is not a place for “misinformation, disinformation, mal-information or the individuals or groups spreading or creating it.”¹⁴⁰ The platform states that it removes any content it deems to fall into these categories. Under its misinformation section, Pinterest explicitly prohibits “false or misleading content that impedes an election’s integrity or an individual’s or group’s civic participation, including registering to vote.”¹⁴¹ Users can report content they believe violates Pinterest’s policies, and the platform states that it uses these reports to learn and improve its standards. The platform ensures that content meets its Community Guidelines through automated processes and human review,¹⁴² but it does not clarify if the company plans to use its standard review process for this or if it will take additional steps to combat

voter misinformation.¹⁴³ Given that there is little transparency around these review processes, going forward Pinterest should clarify what type of automated processes it has in place, how its tools are trained, updated, and used, and how effective they are. Further, in a January 2020 blog post Pinterest stated “we’ll take down content that misleads people about where, when or how to vote.” The company also shared that for accounts that have been suspended due to a single or repeat violation, the account owner will receive a suspension notice when they try to log in to Pinterest. There is a form on the platform’s website that allows users to appeal their suspension, but the guidelines do not provide a timeline for how long the review process can be.¹⁴⁴ A timely and robust appeals process is a vital method of remedy and redress for users. Going forward, the company should ensure users have access to such a process and provide clarity around the timeline for these processes.

While Pinterest has taken some steps to combat misinformation and disinformation, it should be more transparent on how they are actively combating voting and election misinformation. For example, the platform should share what steps it is taking to detect and remove voting-related ads or voting-related misinformation on users’ Pinterest accounts. For other areas of misinformation, such as anti-vaccination content, Pinterest has been more vocal about its efforts, which include conducting internal content moderation sweeps and utilizing automated tools to block URLs that frequently share anti-vaccination content.¹⁴⁵ Pinterest should confirm if they utilize similar steps for voting and election-related misinformation.

Currently, Pinterest’s annual transparency report only includes information on U.S. law enforcement requests for user information, such as the number of subpoenas or court orders the platform receives. The company currently does not provide data on its own content moderation efforts for ads or user-generated content. This lack of transparency makes it difficult to gauge the scope of misinformation on the platform and what potential impacts it may have on voter suppression. In order to increase visibility and accountability, Pinterest should expand its transparency reporting practices to include comprehensive content policy enforcement data. The report should include data related to the enforcement of misinformation and voter suppression-related policies and related appeals.

Reddit

Reddit is a popular social media platform that dubs itself “the front page of the internet.” Although the platform has a small user base (approximately 330 million monthly active users)¹⁴⁶ compared to platforms such as Facebook and YouTube, it is recognized as one of the online services on which viral content is frequently created and spread.¹⁴⁷ A significant amount of the conversation around the spread of misinformation and disinformation during the 2016 U.S. presidential election focused on larger internet platforms. However, research indicates that Reddit also played an instrumental role in this false information ecosystem.¹⁴⁸ The platform has not been a focal point of ongoing policy conversations, potentially because only approximately 4 percent of Americans use Reddit, and the majority of its users are U.S.-based.¹⁴⁹

Despite this lack of attention, Reddit’s unique format fosters an environment in which misinformation and disinformation related to voter suppression can easily spread. Reddit relies on a decentralized model of content moderation, in which the majority of content policy development and subsequent content moderation is carried out by volunteer user moderators, colloquially referred to as “Mods,” who are responsible for specific subreddits. Reddit has high-level content guidelines for the service, which are enforced by a team of employee moderators, known as Admins.¹⁵⁰ This approach to content moderation has allowed for niche, localized communities and norms to prosper on the platform. However, these individualized approaches have also enabled misinformation and disinformation to spread among subreddits.¹⁵¹ As a result, some of the most well-known conspiracy theories and misinformation-laden stories, such as the 2016 Pizzagate conspiracy¹⁵² and the QAnon conspiracy theory,¹⁵³ have gone viral on the platform. Reddit is also home to misleading election-related information which seeks to suppress voting. Reddit, however, does not have specific policies that cover voter suppression content, although the company says that its existing policies, such as those on impersonation, could cover such content.¹⁵⁴ These policies, introduced in January 2020, prohibit the impersonation of an individual or entity “in a misleading or deceptive manner.” The policy applies to instances including when a Reddit account is being used to impersonate someone, when a domain is being used to mimic others, and when deepfakes or other manipulated media are used to mislead users or are misleadingly attributed to an entity or person. Deepfake technology allows for the creation of falsified and manipulated content that could be used to spread misinformation by making it appear as if an individual is doing or saying something they did not actually do.¹⁵⁵ Reddit’s policy creates exceptions for parody and satire, and the company says it will take context into consideration when applying the policy.¹⁵⁶

According to Reddit, the company introduced its policy on impersonation to protect against elements that the platform had not yet seen numerous instances

of, but could in the future.¹⁵⁷ According to the company's latest transparency report, impersonation accounted for 0.6 percent of content removed and 1.4 percent of accounts removed or suspended by Admins for content policy violations.¹⁵⁸ Reddit says it could also use these policies in certain instances to clamp down on misinformation campaigns. In 2018, the company said it identified 944 "suspicious accounts" it associated with the Internet Research Agency (IRA),¹⁵⁹ a Russian-backed professional troll-farm and online influence operations company that has carried out campaigns to support Russian business and political aims.¹⁶⁰ Expert analysis after the 2016 U.S. presidential election found that the IRA was responsible for numerous voter suppression campaigns on social media platforms, including Reddit, which targeted Hillary Clinton voters, particularly voters of color.¹⁶¹

Although Reddit utilizes a more decentralized approach to content moderation, the company takes a more active role in moderating advertising on the platform. This includes steps to address the potential spread of misleading content related to voter suppression and elections in its advertising. The company bans "deceptive, untrue, or misleading advertising" on the platform, including in political ads. Additionally, Reddit manually reviews and approves the messaging and creative content of each ad that is run on the platform. Reddit's political ads policies apply generally to ads that relate to campaigns or elections, solicit political donations, encourage voting or voter registration, and issue or advocacy ads that relate to topics of legislative or political importance, among other things. The company only permits ads from candidates and advertisers who are inside the United States, and who are running ads at the federal level. The company also explicitly states that discouraging voting or voter registration through its advertising services is prohibited.¹⁶² Further, Reddit says that all political ads must feature "paid for by" disclosures within the ads themselves, must be in alignment with all relevant laws and regulations, and must align with Reddit's content policies.¹⁶³

In order to provide transparency around its political ads operations and enforcement mechanisms, Reddit launched a subreddit dedicated to political ads that the company itself moderates. The subreddit includes data on all political ad campaigns that ran on the platform after January 1, 2019, as well as data on individual advertisers, their targeting selections, the impressions ads receive, and instances in which Reddit mistakenly approves ads.¹⁶⁴ This data on advertisements, as well as Reddit's errors during the enforcement process, is valuable for understanding how the company enforces its policies and how these practices shape the political ads landscape on the platform, as well as the misleading information ecosystem within it. Going forward, the platform should share further granular engagement data, such as the number of upvotes, downvotes, and comments political ads receive.¹⁶⁵ Reddit has responded to misinformation and disinformation on its platform by introducing "misinformation" as a category that Mods can choose to flag posts and

comments under. This reporting flow surfaces this content to Admins.¹⁶⁶ According to the platform, misinformation can be understood as “malicious and coordinated attempts to spread false information,” as well as users inadvertently spreading false information.¹⁶⁷ In the context of COVID-19 related misinformation and disinformation, Reddit says that unless a subreddit is specifically dedicated to spreading misleading information, the company will always aim to educate and cooperate with subreddits to address these forms of content, and will only use enforcement actions such as banning subreddits or “quarantining” if these cooperative efforts fail. When a community is quarantined, it does not appear in search results. Additionally, if a user tries to visit the quarantined community, they will be notified that the subreddit may contain misleading content, and they must explicitly opt in to viewing the content.¹⁶⁸ However, it is unclear whether the same policies apply in the context of other categories of misinformation and disinformation.¹⁶⁹ Along these lines, the company is also monitoring for content manipulation efforts, particularly ahead of the 2020 U.S. presidential election. One of the most common avenues for content manipulation on the platform takes advantage of the content voting system. On Reddit, users can rate each piece of content by “upvoting” or “downvoting” it. Reddit’s algorithms use these votes to assign each post a score and rank them in the news feed.¹⁷⁰ In this way, some of the content on the platform is community-curated and trends often emerge as a result of this democratic process. However, as experts have outlined, this system can be gamed by users who create several accounts to downvote or upvote a post, by coordinated attacks on certain forms of content,¹⁷¹ and through the use of existing features on Reddit such as “gilding” which is similar to a “super-vote” and is a mechanism that is generally available to users who have a Gold Reddit subscription or who purchase Reddit coins.¹⁷² In alignment with the company’s desires to better understand and track content manipulation efforts on the platform, the company shared some public information about the kinds of coordinated influence campaigns they have detected, such as one led by a Russian-connected group known as Secondary Infektion on the r/redditsecurity subreddit.¹⁷³ This is a valuable form of transparency that enables users to comment and ask questions about how these types of content are spreading on the platform and what the company is doing to address these issues. Further, the company began issuing a security report in Q4 of 2019 that focuses on efforts to keep the platform and user accounts safe, and includes content manipulation related data such as the number of reports, Admin removals, Admin account sanctions, and Admin subreddit sanctions. Prior to introducing its new hate speech policy, the company also expanded the data in the June 2020 report to include figures related to Admin account sanctions and Admin subreddit sanctions for abuse.¹⁷⁴ In addition, the company outlined in its security report that it is working on detecting bots on the platform and providing clear guidance around the use of bots. These policies will aim to address the use of malicious bots that can spread spam and abusive content at scale, manipulate the

amplification of content on the service by gaming the voting system, and more. These policies will not look to prohibit bots such as those used by Mods for content moderation purposes.¹⁷⁵ Although Reddit has a relatively comprehensive set of policies and transparency practices to address the spread of potential voter suppression misinformation and disinformation on its platform, the company can do more. In particular, given that the platform was an active hotspot for the spread of voter suppression content during the 2016 elections, the company should expand its content policies to explicitly address the voter suppression content, as it has done in its ads policy. Reddit's existing policies, including its impersonation policy, and its new hate speech policy, could address these issues. However, the lack of one central policy that lays out the company's stance and practices around this form of content could result in gaps and prove extremely problematic ahead of the 2020 elections. To this end, the company should also provide greater transparency around how much content has been removed under a new centralized policy through its transparency report. The company should also alert users who have come into contact with election-related misinformation and disinformation campaigns, particularly content that aims to suppress voting, and the company should clarify what the legitimate parameters around voting and voter registration are. Finally, given that content moderation efforts do not always yield entirely accurate results, the company should notify affected users with information related to the policies they violated, and provide them with the opportunity to appeal the decision.

Snapchat

Snapchat, a multimedia messaging app owned by Snap Inc., saw roughly 210 million daily active users in Q3 2019; active Snapchat users opened the app up to 30 times per day.¹⁷⁶

Snapchat's main service is allowing users to upload photos and video messages that disappear after they have been viewed. The platform has affirmatively sought to engage in the electoral process through both its advertising services and user offerings. In 2018, for example, the app displayed a link to register to vote on the profile page of every user who was 18 years old or older, and allowed users to register directly within the app through a service called TurboVote.¹⁷⁷ This helped register over 400,000 voters, 57 percent of which were later confirmed to have cast a ballot, demonstrating the influence Snapchat can have when it comes to elections.¹⁷⁸ On August 6, 2020, Snapchat also announced it was planning to release new features in the coming months to encourage users to register to vote in the 2020 Presidential election.¹⁷⁹ These features include a voter checklist card as well as voter-related "Minis," which are miniature applications made by third-parties that run inside Snapchat.¹⁸⁰ The platform is planning to release a "Before You Vote" mini that lets users know where and how to vote and whether voting by mail is available in their state.¹⁸¹ These digital avenues for voter registration are particularly important in the run-up to the 2020 presidential election, as the COVID-19 pandemic has made it harder for people to register to vote in person at libraries or at local DMV offices. As a result, more people may be likely to rely on online platforms to register and participate in the election process.¹⁸²

Snap Inc.'s CEO Evan Spiegel has stated that the platform fact-checks all ads from political candidates and about the voting process and does not allow misinformation in these types of ads.¹⁸³ Spiegel explained that Snapchat wants to allow political advertising to encourage users, particularly young people and first-time voters, to engage in the political process.¹⁸⁴ Under its advertising policies, Snapchat prohibits ads that "are false or misleading, including deceptive claims, offers, functionality, or business practices" within its content guidelines.¹⁸⁵

In addition, as per Snapchat's political and advocacy advertising policies,¹⁸⁶ political ads must adhere to all applicable laws and regulations, including national election, copyright, and defamation laws, as well as (where applicable) Federal Election Commission (FEC) regulations.¹⁸⁷ The policies state that it is the responsibility of the advertiser to comply with all laws and regulations.¹⁸⁸ It is therefore unclear if Snapchat itself takes steps to ensure ads are compliant with state and federal laws. Furthermore, political ads must include a "paid for by" disclaimer that contains the name of the paying person or entity.¹⁸⁹ This applies to any ad with political messaging or any ad that links to political content.

Election related ads must also state whether or not an ad was authorized by a candidate or an organization.¹⁹⁰ If the ad is not paid for by the candidate, it must include contact information for the sponsoring organization. While Snapchat's general ad policies ban harassment or threats, the platform does not explicitly ban attack ads, (ads that express disagreement with candidates or political parties), as long as they do not violate any other guidelines.¹⁹¹ Snapchat does not currently provide information on what tools and processes it uses to review and approve ads. Going forward, the company should confirm to what extent it uses automated tools and human reviewers to review and approve ads and how these tools and individuals are trained.

Currently, Snapchat maintains a Political and Advocacy Ads Library to provide transparency around political advertising on Snapchat.¹⁹² The library includes information, such as the amount of ad spend and the identity of paying entities, on all political and advocacy advertising that has run on the platform from 2018 to 2020, and it is updated daily. Although this report is a helpful first step towards providing transparency around political advertising on the service, the reports are not readily accessible, as ad information must be downloaded from the website and viewed in Excel. Going forward, the platform should improve the user-friendliness and accessibility of the report by creating a web version of the report. The company should also expand the report to include information such as how many ads were removed for violating Snap's policies, particularly around voting and election misinformation.

When it comes to user-generated content, Snap prohibits hateful content, violent content, impersonation, spam, and illegal speech and activity.¹⁹³ The platform does not explicitly prohibit political or election-related misinformation—or any categories of misinformation—from standard users (i.e. non media partners), and has no restrictions on false or misleading content that may support voter suppression. The platform should expand its guidelines and clarify what type of political content is permitted on the service, and how, or if, it intends to address election and voter suppression misinformation. In addition, Snapchat's Community Guidelines state that, in certain cases, the platform “won't take action against content when it is newsworthy and relates to a matter of political, social, or other general concern to our community.”¹⁹⁴ It is unclear in what circumstances this policy would be applied, and whether it could potentially apply to political content broadly, as this content could be of general public interest.

Snapchat's Community Guidelines also have additional requirements for media partners, whose content is displayed in the Discover feed. Snapchat's Discover section is a space for publishers, like the New York Times, BuzzFeed, and Mashable, to post stories. These media partner-specific guidelines are not applicable to standard users, and they are similar to Snap's advertising policies in that they require that content is fact-checked and accurate.¹⁹⁵ Because Snapchat partners with these accounts and actively promotes their content, it is an

important step that there are additional requirements on these accounts. However, Snap should provide greater transparency around how they review these accounts and ensure that their content is accurate and fact-checked.

If Snapchat does, or plans to, prohibit voter and election related misinformation in user-generated content, the platform should ensure that it provides adequate notice to parties who have had their content removed for violating the platform's policies. The company should also allow these parties to access a robust and timely appeals process. The Community Guidelines state that Snapchat reviews reports of policy-violating content, and if an account violates the company's policies Snapchat "may remove the offending content, terminate the account, and/or notify law enforcement."¹⁹⁶ The Community Guidelines also state that if an account is terminated for violating Snapchat's content policies, the account holder may not use Snapchat again. However, it is unclear if the company currently offers users an appeals process and the Community Guidelines do not outline what steps a user can take if they feel their content was removed in error.

Unlike other platforms like Facebook and Twitter, Snap has not taken an active stance against deepfake videos.¹⁹⁷ While manipulated media such as deepfakes can exacerbate voter suppression by misleading users on election or candidate information, there is likely less risk that user-generated images manipulated with Snapchat's filters (and that disappear after viewing) will be a source of election misinformation. However, there is still a risk that political advertising on the platform could utilize its technology to create misleading deepfake videos. , Therefore, the company should develop explicit guidelines for this type of technology around political advertising. Snapchat's "paid for by" policy for political ads is one way users can understand who is creating certain videos and decide if they trust its content . However, the platform should go further in providing transparency around the allowed or restricted uses of deepfakes technology in political ads.¹⁹⁸

TikTok

TikTok is a relatively new internet platform compared to platforms such as YouTube and Facebook.¹⁹⁹ However, over the past several years the company has seen rapid growth around the world, with approximately 800 million active users globally today.²⁰⁰ The company's popularity makes its platform a focal point for misleading information. Further, experts have expressed concerns that it could particularly become home to election-related misinformation and disinformation, including voter suppression-related content, given that the company is nascent and has less robust content moderation practices.²⁰¹

In January 2020, in response to concerns that the platforms' rapid growth had not been met with tandem efforts to create responsive policies that safeguard users,²⁰² TikTok released a more comprehensive version of its Community Guidelines.²⁰³ These updated policies include a section on misleading information which states "we do not permit misinformation that could cause harm to our community or the larger public." This policy includes a prohibition on content that is meant "to incite fear, hate, or prejudice," "hoaxes, phishing attempts, or manipulated content meant to cause harm," content that misleads users "about elections or other civic processes," and "content distributed by disinformation campaigns."²⁰⁴ Under the company's integrity and authenticity content policies, it also bans spam-related activity, including "coordinated attempts to manufacture inauthentic activity" and operating accounts under false pretenses.²⁰⁵ In August 2020, TikTok broadened and clarified its Community Guidelines to address the spread of misinformation, disinformation, and related content that seeks to undermine the 2020 elections. As part of these efforts, the company updated its general policy on misleading content to include a clear prohibition on manipulated media such as deepfakes.²⁰⁶ In addition, the company clarified that it does not permit coordinated inauthentic behavior.²⁰⁷ Although these policies can apply to election-related content, the Community Guidelines do not include specific voter suppression-related content policies. In an August 2020 announcement, TikTok stated that it was expanding its partnerships with fact-checking organizations, such as PolitiFact and Lead Stories, to help review content and identify and debunk misleading election-related information. Further, the company shared that it would soon introduce an in-app feature permitting users to report content or accounts for election misinformation. The company will also establish an election information center to connect users to authoritative information surrounding the 2020 elections.²⁰⁸

TikTok's mission is "to inspire creativity and bring joy." According to the company, politics is not a topic that it views as bringing joy to its users. As a result, the company has actively discouraged the use of its services for political means. To this end, in 2019, the company banned political advertising.²⁰⁹ The ban prohibits any paid advertising that references, promotes, or opposes a

political candidate, current or former elected official, or political party or group. The ban also includes any content that advocates for a particular position on a local, state, or federal issue of public relevance that could influence political outcomes.²¹⁰ It is unclear how effective policy enforcement around political advertising is, however, as the company does not share any data related to enforcement of its political ads policy.

Some research indicates that the company has discouraged political content by suppressing its promotion and recommendation during election periods.²¹¹ The company states it does not remove political content, although it has not explicitly addressed the topic of algorithmic amplification and suppression.²¹² Although TikTok says it does not view politics and political content as creating joy for its users, and it therefore has instituted a political ads ban, some experts suggest the real reason for the ban is that the company is a small and nascent one that lacks the capacity to moderate and engage with such content in a scalable manner.²¹³ However, the majority of TikTok users in the United States are between the ages of 18 and 24,²¹⁴ and as a result, avoiding politics has been challenging given that this demographic often uses social media to engage in social and political discussions.²¹⁵ In addition, 70 percent of TikTok users are of voting age. TikTok videos with political content have been found to generate a significant amount of traffic on the platform.²¹⁶ In addition, despite the prohibition on paid political advertising, politicians and political groups have partnered with TikTok influencers to promote their ideas and gain popularity.²¹⁷ For example, progressive nonprofit ACRONYM has worked with influencers to encourage voter registration. Similarly, a Trump campaign manager stated he is exploring solidifying deals with TikTok influencers.²¹⁸

TikTok has also introduced a range of programs which could help address the spread of voter suppression misinformation and disinformation on its service. For example, in March 2020, the company established the TikTok Content Advisory Council, a group of external technology and safety experts tasked with providing the company guidance on its content moderation policies, including its misinformation and hate speech related policies. Further, in July 2020, the company introduced a media literacy and safety video series titled “Be Informed,” which features popular creators on the platform encouraging users to be cognizant of the spread of false information on the service. The video includes guidance on how users can evaluate content and sources on the platform, use in-app features to protect against the spread of misleading content, and recognize facts versus opinions.²¹⁹ According to senior officials from the U.S. Department of Homeland Security (DHS) Cybersecurity and Infrastructure Security Agency (CISA), TikTok has an open line of communication and collaboration with DHS to combat election-related disinformation.²²⁰ However, TikTok is owned by a Chinese company called ByteDance; as a result, the expectations and outcomes surrounding this arrangement may be different. This is compounded by the fact that TikTok is under heavy scrutiny from U.S. lawmakers for its ties with China,

²²¹ prompting many U.S. agencies and branches of the armed forces to ban the use of the app on employee phones, citing national security concerns.²²² Most recently, President Trump called for a ban on the app.²²³ It is difficult to assess how these factors might impact the effectiveness of a collaboration between TikTok and CISA.

TikTok needs to implement clearer and more direct policies to address voter suppression misinformation and disinformation. In addition, the company should provide greater transparency and accountability around how it enforces these policies. In its July 2020 transparency report, which for the first time featured data on how the company enforces its own Community Guidelines, the platform states that less than 1 percent of content actioned violated the company's policies on hate speech, integrity and authenticity, and dangerous individuals and organizations. The majority of content the company actioned violated its policies on adult nudity and sexual activities, minor safety, and illegal activities and regulated goods. The company does not, however, include data related to misinformation, although the transparency report does state that the company often proactively removes harmful misinformation.²²⁴ This lack of transparency makes it difficult to understand the scope of voter suppression or election-related misinformation and disinformation on the platform, and how the company addresses these forms of content. Transparency around the effectiveness of the company's enforcement actions is also important, as despite the company's ban on disinformation campaigns, researchers have found examples of disinformation related to topics such as the COVID-19 pandemic circulating on the service.²²⁵ Similar transparency is needed around the enforcement of the company's political ads policies. The company also should outline what kinds of enforcement actions it takes against misleading content, and provide disaggregated data which outlines how often the company removes such content compared to how often the company employs another enforcement action such as algorithmically downranking content or appending a label to such content.

Twitter

Twitter is a microblogging and social media platform that has approximately 152 million daily active users.²²⁶ Twitter became a focal point of conversations on misleading information and the electoral process following the 2016 U.S. presidential election.

According to Twitter’s civic integrity policy, individuals “may not use Twitter’s services for the purpose of manipulating or interfering in elections or other civic processes. This includes posting or sharing content that may suppress participation or misleading people about when, where, or how to participate in a civic process.”²²⁷ The company defines civic processes as “events or procedures mandated, organized, and conducted by the governing and/or electoral body of a country, state, region, district, or municipality to address a matter of common concern through public participation.”²²⁸ For example, in September 2020, the company took action against a Tweet posted by Democratic House candidate Elizabeth Hernandez, which suggested that Republicans should vote on November 4, for violating its rules against voter suppression. Twitter required the campaign account to remove the Tweet, which Hernandez’ team said was posted as a joke, before it could regain access to its account.²²⁹ Under its civic integrity policy, Twitter prohibits three categories of manipulative content and actions:

1. Misleading information about how an individual can participate in an election or civic process: This category includes the posting or promotion of misleading information about when a civic process such as an election is taking place as well as misleading information on how to participate, such as false claims that individuals can vote by Tweet or text messages.²³⁰
2. Content that aims to suppress, intimidate, or discourage individuals from participating in an election or other civic process: This category includes false claims that polling places are closed or experiencing equipment problems, misleading claims about law enforcement activity related to voting in an election, misleading statements related to process procedures which could discourage voting, and threats related to voting locations.²³¹ Twitter’s violent threats policy is also applicable to instances of violent threats that are not covered by the civic integrity policy.²³²
3. False or misleading affiliations: This category prohibits the creation of fake accounts which misrepresent their affiliation or share and promote content that misrepresents an affiliation with a candidate, elected official, government entity, and so on.²³³

In September 2020, Twitter announced that it is updating its civic integrity policy to allow the company to label or remove false or misleading information that aims to undermine public confidence in an election or other civic processes. This policy change will apply to content that creates confusion around the laws related to a civic process, shares unverified claims related to election rigging and other procedures that could undermine faith in the civic process, or promotes misleading claims that call for interference in or that relate to the results of civic process.²³⁴ The company clarified that the civic integrity policy does not apply to inaccurate statements about an elected or appointed official; election or political-related content that is polarizing, controversial, or hyper partisan in nature; and high-level statements about the integrity of civic processes.²³⁵ In addition, if an elected or appointed official shares content that violates this policy but has a public interest value, the company may leave the content up under its public interest notice policy (discussed further below).²³⁶ Twitter's civic integrity policy also includes information on when and how users can report content they believe violates this policy. Reporting is available to users in relevant jurisdictions prior to the first officially-sanctioned event associated with major civic processes, and users can report content both via the Twitter app and desktop website.²³⁷ In addition, Twitter works with several government and civil society partners around the world who flag policy-violating content and receive expedited review on their flags.²³⁸ Further, the civic integrity policy details that consequences for violations vary depending on the nature of the violation and the user's history with the platform. In situations where a user violates the civic integrity policy for the first time, the platform blocks the user's ability to publish new tweets and mandates that the user deletes the violating tweet or content from their profile before they can regain full access. If a user violates the policy again after receiving an initial warning, their account will be permanently suspended. The company offers users whose accounts have been flagged for violating the policy access to an appeals process, although it is unclear whether this also applies to tweets.²³⁹ In August 2020, the company announced it is expanding its misinformation policies related to mail-in ballots and early voting.²⁴⁰

As the 2020 U.S. presidential election draws near, Twitter is under a significant amount of pressure to improve its efforts to detect and curb the spread of election-related misinformation and disinformation, particularly voter suppression-related content, and to connect users with reliable information on voting. In September 2020, Twitter debuted its election hub by adding a "US Elections" tab in the Explore menu on the platform. The hub will feature Twitter-selected election-related news in English and Spanish, debate live streams, state-related voting information and resources, and candidate information. Twitter has stated that the hub will also include public service announcements that aim to inform voters about important election-related topics such as voter registration, how to obtain a mail-in ballot, and guidance for safe voting during the pandemic.²⁴¹ The company has also banned deepfakes²⁴² and in January 2020, a few days before the Iowa caucuses, Twitter permitted users in the United States to report

misleading content related to the elections as well as instances of voter intimidation or suppression.²⁴³

According to Twitter, it designed the reporting tool to empower users in the United States to flag content that could harm the electoral process. However, experts have raised concerns that the tool could be abused by trolls seeking to attack candidates or undermine individuals they disagree with. In addition, there is little transparency around how effective this tool is. It was originally rolled out in India,²⁴⁴ during the general elections in April 2019, and in the European Union ahead of its May 2019 elections, but Twitter has not released information about whether the tool curbed the spread of voter suppressive-related misinformation and disinformation.²⁴⁵ Although Twitter publishes a transparency report which outlines how it enforces its content policies, the company only recently included data on how it enforces its civic integrity policy in the report.²⁴⁶ Currently, the report includes data on the unique number of accounts reported and actioned and the amount of content actioned for hateful conduct, impersonation, and violent threats.²⁴⁷ However, the metrics and data offered in the report do not provide any specific election-related data, and the data provided is not granular enough to understand the scope and scale of voter suppression-related misinformation and disinformation on the service, and how Twitter aims to combat such content.²⁴⁸ In addition, Twitter has continuously been criticized for inconsistently enforcing its policies.²⁴⁹ This raises further questions around whether Twitter's civic integrity policy and this new reporting tool will have any positive impact on preventing voter suppression on the platform.

Twitter has also tried to address the spread of election and voter suppression misinformation and disinformation in advertising. In August 2019, the company banned all advertising from state-backed media.²⁵⁰ The company defines state-controlled media as entities that are financially or editorially controlled by a state. It does not include entities that receive some taxpayer funding but are otherwise independent—such as independent public broadcasters—in this definition. The company worked with academic and civil society leaders to curate its list of state-controlled media organizations.²⁵¹ In addition, in October 2019, the company banned political ads on the platform.²⁵² Further, Twitter introduced ad targeting limitations to prevent targeting of cause-based ads using an individual's age, race, or location.²⁵³ However, there is little transparency around how this ban and the limitations on cause-based ads are being enforced, and how effective these enforcement mechanisms have been to-date. In June 2019, Twitter introduced a public interest notice policy. The policy details that in certain instances in which a government official, individual who is running for public office, individual who is being considered for a government position, or user who is verified or has over 100,000 followers violates the company's policies, the company may leave the content up as it believes the content has a public interest value, but will append a notice over the tweet. The notice, which appears over these tweets in both the news feed and search results, informs users

the content has been deemed as violating, details which of the content policies it violated, and explains that the tweet has been left up because it has a public interest value. However, Twitter removes content that features direct threats of violence or calls to commit violence against an individual.²⁵⁴

The company has deployed these public interest notices in numerous instances. For example, in May 2020, President Donald Trump and his campaign shared unsubstantiated content that claimed vote-by-mail programs are efforts to commit voter fraud. The company responded by fact-checking the tweets and appending a warning label to two of President Trump's tweets that featured false claims related to mail-in voting. The notices also included a link where users could learn more about mail-in ballots.²⁵⁵ However, Twitter has not taken action on similar content, sparking concerns among civil society and civil rights groups that these policies are not applied consistently or transparently,²⁵⁶ which can undermine their effectiveness.²⁵⁷

In August 2020, the company also announced it will label Twitter accounts belonging to senior government officials and entities (e.g. foreign ministers, institutional entities, diplomatic leaders, etc.) and accounts belonging to state-affiliated media entities as well as their editors-in chief and their senior staff, from the five permanent members of the UN Security Council (China, France, Russia, the United Kingdom, and the United States).²⁵⁸ The labels will contain information such as "Russia state-affiliated media" to provide greater transparency around who is sharing content on the platform. Twitter will not apply these labels to heads of states' personal accounts. Twitter also announced that state-affiliated media accounts and their Tweets will no longer be amplified through the platform's recommendation systems. This will impact their visibility on the home timeline, in notifications, and in search.²⁵⁹ Users who click on these new labels on account pages will be redirected to an article which explains this new policy²⁶⁰ as well as to the Twitter Transparency Report to provide further information.²⁶¹

Going forward, the company should provide greater transparency and accountability around how it uses methods such as labels to address voter suppression-related misinformation and disinformation on the platform. The company could do this by publishing data on how many times it has used such labels for different categories of content and different types of officials or entities in its transparency report. This is particularly important given that advocates have expressed concerns that such labeling processes could be ineffective if they are implemented inconsistently or are not implemented in a comprehensive manner.²⁶² Although the platform can still improve its efforts to prevent the spread of voter suppression-related content on its service, its efforts are notable compared to other similar platforms such as Facebook, who have taken a more hands-off approach on these issues.

On July 21, 2020, Twitter announced a range of enforcement actions against accounts related to far-right conspiracy group QAnon, one of the entities identified as responsible for promoting misinformation in the 2016 U.S. presidential election and promoting voter suppression.²⁶³ These enforcement actions include preventing QAnon-related content and accounts from appearing in the algorithmically-curated trending topics and recommendations, preventing QAnon-related links from being shared on the platform, and attempting to prevent the algorithmic amplification of QAnon-related content in search and conversation threads. These are important efforts to prevent the spread of voter suppression-related misinformation and disinformation on the platform. However, as mentioned above, in order for the impact of these efforts to be quantified, the company should provide greater transparency around the effectiveness and results of these enforcement actions.

Twitter should also expand its transparency reporting to include data related to misinformation, disinformation, and voter suppression, as well as data outlining how effective their enforcement of their political ads policies are. In addition, the company should also provide greater transparency around how users' news feeds and recommendations are algorithmically curated, and how this could result in the promotion of related misleading information. Users should also have access to controls that allow them to determine whether and how their data is used to enable these algorithmic curation processes, and to institute preferences around the types of content users see online.²⁶⁴ Finally, the company should preserve data related to election-related content removals and provide researchers with access to this data following elections so they can assess where the company's moderation policies fell short.

WhatsApp

WhatsApp is the world's largest messaging app, with over 2 billion users across the globe.²⁶⁵

The company is owned by Facebook and it offers end-to-end encrypted messaging services. Because WhatsApp offers encrypted messaging services, which are critical for privacy and security, the company is not able to view or review the content that users share. As a result, the company has adopted other approaches to address the spread of misleading information. This includes identifying indicators of problematic content at large, introducing mechanisms to limit the spread of content, and introducing features that enable users to fact check content that they receive.

Generally, WhatsApp aims to reduce the virality of misleading information on the service. In particular, in April 2020, the company instituted a new policy limiting the number of times a forwarded message can be shared to five. If a message has already been forwarded five times, the receiving user can only share it to other chats one at a time.²⁶⁶ According to Facebook, this approach has proven effective at preventing the spread of misinformation in many different countries and locations across the world, especially during elections. As a result, in September 2020, Facebook introduced similar forwarding limits on Facebook Messenger, requiring users to forward messages one at a time.²⁶⁷

The company also began labeling messages that are forwarded five or more times with a double arrow to indicate that they were not written by the sender.²⁶⁸ Further, the platform has a feature enabling users to prevent unknown numbers and contacts from communicating with them and adding them to groups through the app.²⁶⁹ WhatsApp also recently piloted a feature that allows users to upload forwarded messages into their browser to see if online sources support the information in the message. Users can access this feature without having to reveal the message to WhatsApp, thus maintaining their privacy and security. They can also use this feature to fact-check information in the messages they receive.²⁷⁰ Some researchers also suggest that the company institute features enabling device-side hashing and comparison of images against a pre-distributed on-device hash list of known disinformation images.²⁷¹ However, such an approach would not be consistent with offering fully end-to-end encrypted messaging services, and could undermine the privacy and security benefits that strong encryption provides. It would also raise some freedom of expression concerns in that it involves screening user content before it is uploaded and shared.²⁷²

In the context of the upcoming U.S. presidential election, WhatsApp is partnering with the International Fact-Checking Network and its member fact-checking

organizations to establish WhatsApp tip lines. These fact-checking organizations can use these tip lines to engage with users around misleading content and debunk and verify content. During elections around the world, WhatsApp accounts are often used to distribute messages at scale.²⁷³ This can result in the rapid spread of election and voter suppression misinformation and disinformation. As a result, the company works to identify and remove accounts that engage in automated or spam-like behaviors, which do not reflect the behaviors of human users.²⁷⁴ To do this, WhatsApp developed machine-learning systems to detect suspicious accounts at multiple stages of the product use cycle, including registration, during messaging, and in response to user feedback such as user reports and blocks. These systems then calculate a spam score for the accounts in question based on a range of indicators,²⁷⁵ and subsequently ban accounts that are found to be engaging in automated or spam-like behaviors.²⁷⁶ According to the company, it removes over two million accounts through this process every month, and over 75 percent of these removals take place without a user report flagging an account in the first place.²⁷⁷ However, aside from these figures there is little transparency around the scope and scale of these moderation efforts. The company offers users the right to appeal these decisions, and has a team of reviewers who manage these appeal requests.²⁷⁸ This is important given that appeals are a vital mechanism for providing accountability and redress, and given that these enforcement actions are largely taking place in an opaque setting. Going forward, the platform should publish data explaining these enforcement actions, and where possible break down this data by potential relevance to the elections.

YouTube

YouTube, one of Google's subsidiaries, is the most popular video platform in the market with approximately 2 billion users worldwide.²⁷⁹ YouTube is a major source of online information and advertising, and is poised to play a substantial role in the 2020 presidential election. Similar to other online platforms, YouTube has received scrutiny for potential election misinformation and disinformation, including voter suppression content, on its site.²⁸⁰ YouTube follows the same policies for ads and political content as Google, which ban misleading information in general. YouTube also maintains its own Community Guidelines that include policies that prohibit false or misleading content.²⁸¹

Under its Community Guidelines, YouTube prohibits content that contains “spam, scams, or other deceptive practices.”²⁸² The policies specifically call out voter suppression content by prohibiting content “aiming to mislead voters about the time, place, means or eligibility requirements for voting.”²⁸³ The policies also address issues like deepfake videos or fake content by prohibiting malicious manipulated media and stating that the company will terminate channels that attempt to impersonate others. The platform says that it is able to remove policy-violating content by investing in new technologies and tactics for identifying malicious actors. In 2018, for example, YouTube formed an Intelligence Desk to help detect new trends in inappropriate content and behavior. YouTube also partners closely with TAG, Google's Threat Analysis Group, to combat foreign and domestic entities trying to interfere with the electoral process. General users and Trusted Flaggers, which consists of individual users, government agencies, and non-governmental organizations, may also flag content for violating YouTube policies.²⁸⁴ Flagged content is then later reviewed by the platform's content moderators and either removed or kept online.²⁸⁵ In order to demonstrate accountability around the spread of misinformation on its service, YouTube should notify users who see or engage with content that has been flagged as misinformation and provide them with additional contextual information to understand why the post was misleading or false.²⁸⁶

YouTube also announced that it is working to raise authoritative voices on its platform to help reduce misinformation.²⁸⁷ In 2017, it started prioritizing known sources it deems authoritative—such as CNN, Fox News, and the Guardian—for news and information in search results and “watch next” panels. Although YouTube does not provide information on how it determines whether a source is “authoritative.” The platform continues to expand its Top News and Breaking News sections to highlight videos from news sources and display breaking news events directly on its homepage. YouTube also announced on April 28, 2020 that it was expanding its work on fact-check information panels, which connect users to authoritative information based on their search queries.²⁸⁸ One way YouTube uses information panels is to show whether a channel is owned by a news

publisher that is publicly funded or funded by a government.²⁸⁹ Increasing the visibility of reputable sources for news and keeping users informed about the sources of the content they view on YouTube can be a helpful tool to combat voter misinformation and voter suppression tactics.

Over the past several years, researchers have outlined the ways YouTube's algorithmic recommendation system contributes to the spread of misinformation.²⁹⁰ In response to these criticisms, the company instituted a number of changes.²⁹¹ YouTube's recommendation process typically ranks and recommends videos to users based on a range of signals, including likes, dislikes, watch history, and data from user feedback surveys.²⁹² Since January of 2019, the platform has worked to reduce the recommendation of borderline content that comes close to violating its Community Guidelines, but does not merit removal.²⁹³ As a result, YouTube limits recommendations for videos that, for example, promote a miracle cure for a serious illness or claim the Earth is flat. Users may also turn off recommendations to have more controls over the content they see.²⁹⁴ Although these tactics could be helpful for preventing the rapid spread of election misinformation and voter suppression content, the platform has not published information about how it tackles this type of content in its recommendations. In addition, YouTube does not provide users with a comprehensive set of controls for determining why YouTube recommends certain content to them. Going forward, the company should provide greater transparency around how it addresses misleading election content in its recommendation system. Further, it should empower users to decide how their data is used to shape the recommendations they receive.

YouTube uses both automated tools and human reviewers to moderate content and enforce its Community Guidelines.²⁹⁵ However, there is little transparency around how the platform's automated tools are trained, updated, or used, and how effective they are at combating misinformation. In addition, as previously discussed, Google (and YouTube) are increasingly relying on automated tools to review potentially violating content during the COVID-19 pandemic. The company has not provided adequate transparency around what categories of content this new process applies to, and what the consequences of this shift are.

YouTube also warns that this increased reliance may result in a higher number of removals for videos, some of which "may not violate policies."²⁹⁶ The company states that it won't issue strikes on content removed by automated systems without human review, unless it has a high confidence that the content actually violates policies. Strikes are typically issued when a user's content is removed for violating YouTube's Community Guidelines and, if a user receives three strikes within a 90-day period, their channel may be permanently removed from YouTube.²⁹⁷ If an account holder believes that their content was improperly removed, they can appeal the decision. Appeals are an important mechanism for remedy and redress, however, the company stated that appeals processes may

take longer than usual due to the pandemic. This raises concerns that a potentially higher number of mistaken removals combined with a slower appeals process could negatively affect election and voting content prior to the election. While it is important to remove and block misleading content, accurate voting and election-related content, which can be beneficial for potential voters, could also be taken down accidentally. The platform should therefore invest more in ensuring election content gets priority review in the run up to the election. In addition, because a significant amount of election-related content moderation is occurring during the pandemic, YouTube should preserve data related to election-related content removals during this period so that researchers can evaluate these efforts later on.

Google publishes a political advertising transparency report, which features data on Google, YouTube, and partner properties.²⁹⁸ However, the consolidated report does not break out reporting for each platform, and it does not provide granular information for YouTube specifically. The report also does not include the number of political ads that were flagged or removed for violating Google or YouTube's advertising policies, making it difficult to understand how effective Google's ad enforcement practices are. In addition, Google publishes a Community Guidelines Enforcement Report for YouTube, which outlines how the platform enforces its content policies.²⁹⁹ While the report shows the volume of videos, channels, and comments removed for being spam, misleading, or scams, it does not specifically break out the volume of videos, channels, or comments removed for attempting to mislead users about elections or voting information.

Recommendations

The 2016 U.S. presidential election illustrated the alarming levels of misinformation and disinformation that could spread and potentially influence an electorate. These misinformation and disinformation campaigns, many of which were designed to suppress voting on a large scale, particularly impacted communities of color. As the 2020 U.S. presidential election draws near, internet platforms can play an important role in promoting civic engagement. However, these platforms can also be manipulated and can cause serious harm to the electoral process. It is therefore critical that internet companies institute comprehensive policies and practices to respond to the continuous and rapid spread of election-related misinformation and disinformation, while distributing accurate voting and election information.

The section below includes short-term recommendations that internet platforms and policymakers should implement prior to the 2020 U.S. presidential election as well as long-term recommendations that should be used to help address future elections. The recommendations outline how companies can improve their efforts to connect users to, and lift up, authoritative information; address the spread of misleading information through content moderation and curation; tackle misleading advertisements; and provide meaningful transparency and accountability around these efforts. This section also includes recommendations for how U.S. policymakers can encourage greater accountability and integrity from internet platforms.

Recommendations for Internet Platforms

Sharing and Lifting Up Authoritative Information and Empowering Informed User Decision-Making

- Partner with reputable fact-checking organizations and entities, as well as local and state election bodies to verify or refute information circulated through organic content and advertisements.
- Partner with reputable organizations to launch media literacy efforts which aim to educate users on how to identify and evaluate misleading election-related content they may engage with online. These campaigns should also explain how users can report this content.
- Fund and partner with vetted fact-checking organizations to ensure that fact-checking efforts can adequately tackle the growing volume of election-related misinformation and disinformation.

- Educate users about potential attacks and scams related to elections that may appear on the platform and on methods to avoid becoming a victim of such efforts.
- Notify users who have engaged with misleading election-related content and direct them to authoritative sources of information.
- Institute a public interest exception policy that permits companies to leave content posted by world leaders, candidates for political office, and other government officials on their services, even if the content has been fact-checked and contains misleading information. In these cases, the company should label the content and provide additional context to users which explains that the content has been debunked but there is a public interest value in creating public awareness that political and government officials posted such content. Companies should also include links to authoritative information sources in the labels. In instances where the company determines that the content posted by officials could result in imminent harm, this public interest exception policy should not be applied. Rather, the companies should remove the content as they would with any other user.
- Conduct regular impact assessments and audits of algorithmic curation tools (e.g. ranking and recommendation systems), and recalibrate them as necessary so they do not direct users to or surface misleading content when they search for election-related topics and do not algorithmically amplify such content in trending topics and recommendations.
- Label organic content and advertisements that have been produced by state-controlled media outlets to inform users of the content's origins.
- Educate users on how their personal data is being collected and to what extent this data is being used to curate the content and ads that users are seeing online. Companies should also provide users with controls which allow them to determine how their data is collected, shared, and used to shape their content and ad experiences, especially as it relates to political advertising and election-related content.
- Provide vetted researchers with access to tools and datasets that could enable them to better evaluate company efforts to combat election-related misinformation and disinformation.

Moderating and Curating Misleading Information

- Create a comprehensive set of content policies to address the spread of election-related misinformation and disinformation with specific considerations for voter-suppressive content. Guidelines should include examples of how these policies are enforced and what kinds of content the policies do not apply to. Companies should house these policies in one location, provide public notice if their policies change, and include an archive of past policies.
- Companies should clarify to what extent election-related policies interface with content policies related to hate speech, deepfakes, bots, coordinated inauthentic behavior, etc. While manipulated media may be a part of user expression on social media and therefore permissible for user-generated content, platforms should consider banning the use of such manipulation technologies for political advertising.
- Institute a dedicated reporting feature which enables users to flag election-related misinformation and disinformation to the company.
- Remove, reduce the spread of, or label content that has been fact-checked and deemed to contain election-related misinformation.
- Label content that has been fact-checked and deemed to contain misinformation but does not qualify for removal. Labels should direct users viewing such content to authoritative sources of information. Companies should also provide adequate notice to users explaining what specific policies the user has violated and include information on how the user can appeal this decision.
- Establish a Trusted Flaggers program which allows vetted and reputable civil rights organizations, civil society groups, and individuals to flag election-related misinformation and disinformation at scale and receive priority review for these flags. Companies should publicly disclose how this program works, how entities and individuals can apply, and other relevant information.
- Collaborate with other internet platforms to share information on and strategies for addressing trending misinformation and disinformation campaigns, fraudulent accounts, coordinated inauthentic behavior, and debunked content. Any collaborations should be publicly disclosed and should be respectful of users' privacy and comply with antitrust laws.

Tackling Misleading Advertising

- Create and implement comprehensive policies for the content and targeting of ads that prohibit election-related misinformation and disinformation in advertisements. The policies should include specific considerations for addressing voter suppressive ad content and should clarify that advertisers must adhere to all applicable laws and regulations. Companies should include examples of how these ad policies are enforced and what kind of content does not fall under these policies. If these policies change, companies should provide public notice of these changes and share an archive of past policies. Companies should also clarify to what extent these policies interface with ad content and targeting policies related to hate speech, bots, deepfakes, coordinated inauthentic behavior, etc.
- Establish a comprehensive review process for election-related ads and ad targeting categories. Companies should require all election-related ads to be fact-checked and reviewed by a human reviewer before they are permitted to run on a platform. Companies should publicly disclose high-level information on what this review process consists of and to what extent it relies on automated tools and human reviewers.
- Explain to users to what extent advertisements that are flagged for violating election-related ad policies are reviewed, moderated, and curated by human reviewers and by automated tools. Users should be notified of any significant updates to these processes.
- Create a comprehensive vetting process for advertisers which requires them to verify their identity and which country they are based in before running ads.
- Provide adequate notice to advertisers who have had their ads removed, algorithmically curated (e.g. downranked), or labeled. This notice should explain what specific policies the advertiser violated and include information on how the advertiser can appeal this decision.
- Give political advertisers the opportunity to appeal ad moderation decisions. This appeals process should be timely and enable advertisers to provide additional information on the case and have their case reviewed by a new reviewer or group of reviewers.
- Append “paid for” disclosures to all paid political, social, and issue ads and ensure labels are maintained even if ad campaigns end or if ads are organically shared online.

- 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

- Explain to users how and to what extent content that is flagged for violating election-related misinformation and disinformation policies is reviewed, moderated, and curated by human reviewers and by automated tools. Users should be notified of any significant updates to these processes.
- Provide adequate notice to users who have had their content removed, algorithmically curated (e.g. downranked), or labeled. This notice should explain what specific policies the user has violated and include information on how the user can appeal this decision.
- Give users the opportunity to appeal moderation decisions. This appeals process should be timely and enable users to provide additional information on the case and have their case reviewed by a new reviewer or group of reviewers. Users who flag content and accounts should also have access to an appeals process.
- Preserve data on election-related content and advertising removals. Vetted researchers should have access to this data so they can identify where these content and advertising moderation policies and practices fell short and make recommendations on how they can be improved.
- Publish data related to the moderation, curation, and labeling of election-related misinformation and disinformation in their regular transparency reports. At a minimum, this data should include:
 - The number of accounts flagged, the number of accounts suspended, and the number of accounts removed for violating these policies
 - The number of pieces of content that were flagged, removed, downranked, and labeled as a result of policy violations
 - How much of the content and accounts that were removed, suspended, downranked, and labeled were identified proactively using automated tools and how much of the content and accounts were identified through human flags (e.g. from users, Trusted Flaggers, etc.)

- A breakdown of content and accounts that were removed, suspended, downranked, or labeled by product (e.g. Facebook, Instagram, or WhatsApp)
 - A breakdown of content and accounts that were removed, suspended, downranked, or labeled by format (e.g. video, text, image)
 - A breakdown of content and accounts that were removed, suspended, downranked, or labeled by category of misinformation/disinformation (e.g. voter suppression, impersonation, etc.)
 - The number of appeals received for action taken against content and accounts in this category
 - The number of pieces of content restored and the number of accounts restored as a result of appeals in this category
 - The number of pieces of content restored and the number of accounts restored as a result of proactive recognition of errors by the company
- 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. This database should include search functionality. In order to protect privacy, the information in this database should not permit the identification of specific users who received the ads. At a minimum, this database should disclose the following information about each of the ads:
 - The format of the ad (e.g. text, video, etc.)
 - The name of the advertiser
 - What state the ad was run in
 - How much the ad spend was
 - The time period during which an ad was active
 - Granular engagement and interaction information such as how many users saw the ad and the number of likes, shares, and views an ad received

- What targeting parameters the advertiser selected
 - What categories of users the ad was eventually delivered to (i.e. what targeting parameters did the ad delivery system eventually select and optimize for)
 - Whether the ad was delivered to a custom set of users or ones generated by an automated system
- Publish data on the company’s election-related ad content and targeting policy enforcement efforts. This should include:
 - The total number of ads and advertiser accounts removed for violating the platform’s election-related ad content and targeting policies
 - A breakdown of ads and advertiser accounts removed based on which policy they violated
 - A breakdown of ads and advertiser accounts removed based on the format of the ad (e.g. text, audio, image, etc.)
 - A breakdown of ads and advertiser accounts removed based on the country of the advertiser
 - A breakdown of ads and advertiser accounts removed based on the product or service on which the ad was run
 - The detection method used (e.g. user flag, automated tools, etc.). This data should not reveal the identity of individual flaggers
 - Provide periodic updates on content and advertising moderation, curation, and labeling efforts in the run up to the 2020 U.S. presidential election.
 - Following major elections, publish an election-specific transparency report that summarizes the scope and scale of content and advertising moderation, curation, and labeling efforts surrounding the elections.

Recommendations for Policymakers

Although the U.S. government is limited in the extent to which it can direct platforms how to decide what content to permit on their sites, policymakers can

do more to encourage greater transparency and accountability from internet platforms around how they are addressing the rapid spread of election-related misinformation and disinformation on their services.

- Policymakers should enact rules to require greater transparency from online platforms, including regular reporting regarding their content moderation, curation, labeling, and ad targeting and delivery efforts.
- Government agencies and representatives should ensure that when they post online they are only disseminating verified information related to the elections and are not spreading unproven or debunked information.
- Authoritative election authorities such as the Federal Elections Commission (FEC), state election boards, and other state and local authorities should partner with internet platforms to provide and promote verified and legitimate information related to the election on their platforms. These entities should also help debunk misleading claims and information using their own online accounts.
- Policymakers should clarify that the Voting Rights Act, which prohibits suppressing voting through intimidation, applies in the digital environment. Further, Congress should amend the Act or pass new legislation to prohibit suppression of voting through deception, which is the primary means of vote suppression online.
- Policymakers should fund vetted fact-checking organizations around the world to ensure that fact-checking efforts can adequately tackle the growing volume of election-related misinformation and disinformation.
- Policymakers should update campaign finance laws to address gaps and ensure that federal laws and regulations comprehensively cover digital political advertising.

Conclusion

Over the past several years, it has become clear that internet platforms can play an influential role in providing a platform for civic conversations and the distribution of accurate election-related information. It has also become apparent that many of these platforms are hotbeds for the spread of election-related misinformation and disinformation, which can suppress voter participation and particularly impact communities of color. Going forward, internet companies should demonstrate greater commitment to improving their efforts to tackle misleading election-related information, while preserving the important role their platforms can play in providing a space for civic engagement and access to legitimate election information. They should also ensure that they provide greater transparency and accountability around these efforts. Given the significant impact these platforms can have during electoral processes, these companies should continuously work to improve these approaches, for the upcoming 2020 U.S. presidential election, and beyond.

Notes

- 1 Select Committee on Intelligence, (U)Report of the Select Committee on Intelligence United States Senate On Russian Active Measures Campaigns and Interference in the 2016 U.S. Election Volume 2: Russia's Use of Social Media With Additional Views, S. Rep. No. 116th, 1st (Oct. 8, 2019). https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf.
- 2 Cat Zakrzewski and Tonya Riley, "The Technology 202: Social networks haven't done enough to prevent voter manipulation, tech leaders say," *The Washington Post*, March 10, 2020, <https://www.washingtonpost.com/news/powerpost/paloma/the-technology-202/2020/03/10/the-technology-202-social-networks-haven-t-done-enough-to-prevent-voter-manipulation-tech-leaders-say/5e667771602ff10d49ac68f4/>.
- 3 Darrell M. West, "Brookings Survey Finds 57 Percent Say They Have Seen Fake News During 2018 Elections and 19 Percent Believe It Has Influenced Their Vote," Brookings Institution, last modified October 23, 2018, <https://www.brookings.edu/blog/techtank/2018/10/23/brookings-survey-finds-57-percent-say-they-have-seen-fake-news-during-2018-elections-and-19-percent-believe-it-has-influenced-their-vote/>.
- 4 Ted Van Green, "Few Americans Are Confident In Tech Companies To Prevent Misuse Of Their Platforms in the 2020 Election," Pew Research Center, last modified September 9, 2020, <https://www.pewresearch.org/fact-tank/2020/09/09/few-americans-are-confident-in-tech-companies-to-prevent-misuse-of-their-platforms-in-the-2020-election/>.
- 5 Jessica Guynn, "Facebook civil rights audit warns of Trump, voter suppression ahead of presidential election," *USA Today*, July 8, 2020, <https://www.usatoday.com/story/tech/2020/07/08/trump-facebook-audit-hate-speech-racism-civil-rights/5395533002/>
- 6 Miles Parks, "Social Media Usage Is At An All-Time High. That Could Mean A Nightmare For Democracy," *NPR*, May 27, 2020, <https://www.npr.org/2020/05/27/860369744/social-media-usage-is-at-an-all-time-high-that-could-mean-a-nightmare-for-democr>.
- 7 Carrie Levine, "Online Misinformation During the Primaries: A Preview Of What's To Come?," *The Center for Public Integrity*, last modified March 11, 2020, <https://publicintegrity.org/politics/elections/online-misinformation-during-the-primaries-a-preview-of-whats-to-come/>.
- 8 Naomi Nix and Kurt Wagner, "Social Media Braces for a Deluge of Voter Misinformation," *Bloomberg Businessweek*, July 24, 2020, <https://www.bloomberg.com/news/articles/2020-07-24/social-media-civil-rights-groups-brace-for-voter-disinformation>.
- 9 Mike Issac, Kate Conger, "Google, Facebook and Others Broaden Group to Secure U.S. Election," *NYTimes*, August 12, 2020, <https://www.nytimes.com/2020/08/12/technology/google-facebook-coalition-us-election.html>
- 10 Nathalie Maréchal and Ellery Roberts Biddle, *It's Not Just the Content, It's the Business Model: Democracy's Online Speech Challenge*, March 17, 2020, <https://www.newamerica.org/oti/reports/its-not-just-content-its-business-model/>.
- 11 Nix and Wagner, "Social Media".
- 12 Jillian Hufford, "Amazon Statistics: Need To Know Numbers about Amazon [Infographic]," *nChannel*, February 12, 2020, <https://www.nchannel.com/blog/amazon-statistics/>
- 13 Dana Mattioli, Sebastian Herrera, "Amazon's Sales Jump as Coronavirus Prompts Surge in Online Shopping," *The Wall Street Journal*, April 30, 2020, <https://www.wsj.com/articles/amazons-sales-jump-as-coronavirus-prompts-surge-in-online-shopping-11588278740>

- 14 Amazon, "Selling Policies and Seller Code of Conduct," Amazon Seller Central, https://sellercentral.amazon.com/gp/help/external/G1801?language=en_US
- 15 Garrett Sloane, "Amazon Makes Major Leap In Ad Industry With \$10 Billion Year," AdAge, January 31, 2019, <https://adage.com/article/digital/amazon-makes-quick-work-ad-industry-10-billion-year/316468>
- 16 Amazon, "Posts creative acceptance policies," Amazon Advertising, <https://advertising.amazon.com/resources/ad-policy/posts>
- 17 Megan Graham, Annie Palmer, "Amazon bans sponsored products related to political campaigns, but you can still find them on the site," CNBC, January 7, 2020, <https://www.cnbc.com/2020/01/07/amazon-struggles-to-enforce-its-own-rules-for-political-ads.html>
- 18 "Alexa, tell me about the election," Amazon Blog, September 18, 2019, <https://blog.aboutamazon.com/devices/alexa-tell-me-about-the-election>
- 19 Denny Arar, "Alexa, where do I vote? Who's running? Amazon assistant provides voter help," Digital Trends, November 5, 2018, <https://www.digitaltrends.com/home/alexa-midterms/>
- 20 Pete Pachal, "Amazon is beefing up Alexa with real-time election data," Mashable, October 31, 2018, <https://mashable.com/article/amazon-election-data-midterms/>
- 21 Pete Pachal, "Amazon is beefing up," Mashable
- 22 "Policy Testing," Amazon Alexa, <https://developer.amazon.com/en-US/docs/alexa/custom-skills/policy-testing-for-an-alexa-skill.html>
- 23 J. Clement, "Most Popular Social Networks Worldwide as of July 2020, Ranked by Number of Active Users (In Millions)," Statista, last modified August 21, 2020, <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>
- 24 Guy Rosen et al., "Helping to Protect the 2020 US Elections," Facebook Newsroom, last modified October 21, 2019, <https://about.fb.com/news/2019/10/update-on-election-integrity-efforts/>
- 25 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 26 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 27 Harrison Mantas and Susan Benkelman, "How Users See Facebook's Labels Will Determine Their Effectiveness," Poynter, last modified July 23, 2020, <https://www.poynter.org/fact-checking/2020/how-users-see-facebooks-labels-will-determine-their-effectiveness/>. Mark Zuckerberg, "Three weeks ago, I committed to reviewing our policies ahead of the 2020 elections. That work is ongoing, but today I want to share some new policies to connect people with authoritative information about voting, crack down on voter suppression, and fight hate speech.," Facebook, June 26, 2020, 2:25 PM, <https://www.facebook.com/zuck/posts/10112048980882521>.
- 28 Jessica Leinwand, "Expanding Our Policies on Voter Suppression," Facebook Newsroom, last modified October 15, 2018, <https://about.fb.com/news/2018/10/voter-suppression-policies/>.
- 29 *Facebook's Civil Rights Audit – Final Report*, July 8, 2020, <https://about.fb.com/wp-content/uploads/2020/07/Civil-Rights-Audit-Final-Report.pdf>.
- 30 Andrew Whalen, "Facebook to Ban Posts Meant to Suppress Votes, Including ICE Agent Threats," *Newsweek*, June 26, 2020, <https://www.newsweek.com/facebook-political-ads-policy-misinformation-mark-zuckerberg-2020-1513789>.
- 31 *Facebook's Civil*.

- 32 Facebook, "New Steps to Protect the US Elections," Facebook Newsroom, last modified September 3, 2020, <https://about.fb.com/news/2020/09/additional-steps-to-protect-the-us-elections/>.
- 33 Jared Newman, "Tech Platforms Screwed Up The Last Election. Here's How They're Prepping for 2020," *Fast Company*, March 4, 2020, <https://www.fastcompany.com/90467733/tech-platforms-screwed-up-the-last-election-heres-how-theyre-prepping-for-2020>.
- 34 Newman, "Tech Platforms".
- 35 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 36 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 37 Leinwand, "Expanding Our Policies," Facebook Newsroom.
- 38 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 39 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 40 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 41 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 42 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 43 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 44 "About Audience Network," Facebook for Business, <https://www.facebook.com/business/help/788333711222886?id=571563249872422>.
- 45 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 46 Rosen et al., "Helping to Protect," Facebook Newsroom.
- 47 Alex Hern and Julia Carrie, "Facebook Plans Voter Turnout Push – But Will Not Bar False Claims From Trump," *The Guardian*, June 17, 2020, <https://www.theguardian.com/technology/2020/jun/17/facebook-plans-to-boost-voter-turnout-and-transparency-in-us-elections>.
- 48 Hern and Carrie, "Facebook Plans".
- 49 Facebook, "New Steps," Facebook Newsroom.
- 50 Hern and Carrie, "Facebook Plans".
- 51 Alexandra S. Levine, "Could Uber and Lyft Drive Business Off the Road?," *Politico Morning Tech*, August 20, 2020, <https://www.politico.com/newsletters/morning-tech/2020/08/20/could-uber-and-lyft-drive-business-off-the-road-790054>
- 52 Facebook, "New Steps," Facebook Newsroom.
- 53 Newman, "Tech Platforms".
- 54 Nathaniel Gleicher, "Labeling State-Controlled Media On Facebook," Facebook Newsroom, last modified June 4, 2020, <https://about.fb.com/news/2020/06/labeling-state-controlled-media/>.
- 55 Courtney C. Radsch, "Tech Platforms Struggle to Label State-Controlled Media," Committee to Protect Journalists, last modified August 12, 2020, <https://cpj.org/2020/08/tech-platforms-struggle-to-label-state-controlled-media/>.
- 56 Hern and Carrie, "Facebook Plans".
- 57 Hern and Carrie, "Facebook Plans".
- 58 Cecilia Kang and Mike Isaac, "Defiant Zuckerberg Says Facebook Won't Police Political Speech," *The New York Times*, October 21, 2019,

<https://www.nytimes.com/2019/10/17/business/zuckerberg-facebook-free-speech.html>.

59 Kang and Isaac, "Defiant Zuckerberg".

60 Kang and Isaac, "Defiant Zuckerberg".

61 Craig Timberg and Andrew Ba Tran, "Facebook's Fact-Checkers Have Ruled Claims in Trump Ads Are False — But No One Is Telling Facebook's Users," *The Washington Post*, August 5, 2020, <https://www.washingtonpost.com/technology/2020/08/05/trump-facebook-ads-false/>.

62 Spandana Singh, *Special Delivery: How Internet Platforms Use Artificial Intelligence to Target and Deliver Ads*, February 18, 2020, <https://www.newamerica.org/oti/reports/special-delivery/>.

63 Facebook, "New Steps," Facebook Newsroom.

64 Thomas Germain, "Facebook to Halt New Political Ads in Week Before November Election," Consumer Reports, last modified September 3, 2020, <https://www.consumerreports.org/social-media/facebook-to-halt-new-political-ads-in-week-before-election/>.

65 Donie O'Sullivan and Brian Fung, "Facebook Will Limit Some Advertising in the Week Before the US Election -- But It Will Let Politicians Run Ads with Lies," *CNN Business*, September 3, 2020, <https://www.cnn.com/2020/09/03/tech/facebook-political-ads-election/index.html>.

66 Nick Clegg, "Facebook Is Preparing For an Election Like No Other," *The Telegraph*, June 17, 2020, <https://www.telegraph.co.uk/news/2020/06/17/facebook-preparing-election-like-no/>.

67 Rosen et al., "Helping to Protect," Facebook Newsroom.

68 Spandana Singh, "AI Proves It's a Poor Substitute for Human Content Checkers During Lockdown," *VentureBeat*, May 23, 2020, <https://venturebeat.com/>

[2020/05/23/ai-proves-its-a-poor-substitute-for-human-content-checkers-during-lockdown/](https://venturebeat.com/2020/05/23/ai-proves-its-a-poor-substitute-for-human-content-checkers-during-lockdown/).

69 Kang-Xing Jin, "Keeping People Safe and Informed About the Coronavirus," Facebook Newsroom, last modified August 19, 2020, <https://about.fb.com/news/2020/08/coronavirus/>.

70 Ian Vandewalker, *Digital Disinformation and Vote Suppression*, September 2, 2020, <https://www.brennancenter.org/our-work/research-reports/digital-disinformation-and-vote-suppression>. Letter by Emma Llansó, "COVID-19 Content Moderation Research Letter – in English, Spanish, & Arabic," April 22, 2020, <https://cdt.org/insights/covid-19-content-moderation-research-letter/>.

71 Facebook, *Community Standards Enforcement Report*, <https://transparency.facebook.com/community-standards-enforcement>.

72 Spandana Singh, *Everything in Moderation: An Analysis of How Internet Platforms Are Using Artificial Intelligence to Moderate User-Generated Content*, July 22, 2019, <https://www.newamerica.org/oti/reports/everything-moderation-analysis-how-internet-platforms-are-using-artificial-intelligence-moderate-user-generated-content/>.

73 *Facebook's Civil*.

74 Whalen, "Facebook to Ban Posts".

75 Stop Hate for Profit, last modified 2020, <https://www.stophateforprofit.org/>.

76 Mantas and Benkelman, "How Users," Poynter.

77 Mantas and Benkelman, "How Users," Poynter.

78 Mantas and Benkelman, "How Users," Poynter.

79 Donald J. Trump, "Mail-In Voting, unless changed by the courts, will lead to the most CORRUPT ELECTION in our Nation's History! #RIGGEDELECTION," Facebook, July 21, 2020, 8:39,

<https://www.facebook.com/DonaldTrump/posts/10165125044025725>.

80 Kimberly Klacik, "Please vote KIM KLACIK on November 3rd. We are getting all of our ducks in a row. On Day 1 you will see you made a great choice. Maryland's District 7 KimKForCongress.com," Facebook, August 28, 2020, 12:42, <https://www.facebook.com/KimberlyNKlacik/posts/3299662563450413>.

81 *Facebook's Civil*.

82 Accountable Tech, *Election Integrity Roadmap for Social Media Platforms*, September 2020, <https://accountabletech.org/wp-content/uploads/2020/09/Election-Integrity-Roadmap-for-Social-Media-Platforms.pdf>.

83 Whalen, "Facebook to Ban Posts".

84 Jonathan Shieber, "As Advertisers Revolt, Facebook Commits to Flagging 'Newsworthy' Political Speech That Violates Policy," *TechCrunch*, June 26, 2020, <https://techcrunch.com/2020/06/26/as-advertisers-revolt-facebook-commits-to-flagging-newsworthy-political-speech-that-violates-policy/>.

85 Whalen, "Facebook to Ban Posts".

86 Facebook, "New Steps," Facebook Newsroom.

87 Whalen, "Facebook to Ban Posts".

88 Aaron Holmes, "Trump Made False Claims About Vote-By-Mail On Facebook and Twitter. Here's Why the Tech Companies Won't Ban Him or Take Down the Posts.," *Business Insider*, May 22, 2020, <https://www.businessinsider.com/why-facebook-twitter-wont-ban-trump-vote-by-mail-2020-5>.

89 Sam Levine, "'It Could Have a Chilling Effect': Why Trump is Ramping Up Attacks on Mail-In Voting," *The Guardian*, June 1, 2020, <https://www.theguardian.com/us-news/2020/jun/01/trump-mail-in-voting-election-2020-election-attacks>.

90 Facebook, "New Steps," Facebook Newsroom.

91 Sara Fischer, "Exclusive: Facebook Commits to Civil Rights Audit, Political Bias Review," *Axios*, May 2, 2018, <https://www.axios.com/scoop-facebook-committing-to-internal-pobias-audit-1525187977-160aaa3a-3d10-4b28-a4bb-b81947bd03e4.html>.

92 *Facebook's Civil*.

93 Spandana Singh and Koustubh "K.J." Bagchi, *How Internet Platforms Are Combating Disinformation and Misinformation in the Age of COVID-19*, June 1, 2020, <https://www.newamerica.org/oti/reports/how-internet-platforms-are-combating-disinformation-and-misinformation-age-covid-19/>.

94 Sheryl Sandberg, "A Second Update on Our Civil Rights Audit," Facebook Newsroom, last modified June 30, 2019, <https://about.fb.com/2019/06/second-update-civil-rights-audit/>.

95 Rosen et al., "Helping to Protect," Facebook Newsroom.

96 "Google Search Statistics," Internet Live Stats, <https://www.internetlivestats.com/google-search-statistics/>

97 Scott Spencer, "An update on our political ads policy," Google, November 20, 2019, <https://blog.google/technology/ads/update-our-political-ads-policy/>

98 "Remove information from Google," Google, <https://support.google.com/webmasters/answer/6332384?hl=en>

99 Steven Overly, "Google makes changes to search to curb misleading info about U.S. election," Politico, September 9, 2020, <https://www.politico.com/news/2020/09/10/google-search-to-curb-election-misinformation-411882>

100 Overly, "Google makes changes," Politico

- 101 “Google Ads Policies,” Google, <https://support.google.com/adspolicy/answer/6008942?hl=en>.
- 102 “Misrepresentation,” Google, <https://support.google.com/adspolicy/answer/6020955?hl=en>
- 103 “Update to Misrepresentation policy (September 2020),” Google, last modified July 2020, <https://support.google.com/adspolicy/answer/9991401?hl=en#:~:text=In%20September%202020%2C%20the%20Google,or%20matters%20of%20public%20concern.>
- 104 “Update to Misrepresentation policy,” Google
- 105 “Easier resubmissions and appeals of ad policy issues,” Google, March 19, 2020, <https://support.google.com/google-ads/answer/9790008?hl=en>
- 106 “Actions to reduce the need for people to come into our offices,” Google, March 16, 2020, <https://blog.google/inside-google/company-announcements/update-extended-workforce-covid-19/>
- 107 “Actions to reduce,” Google
- 108 Singh, *Special Delivery*. Krista Bradford, “Google Shows Men Ads for Better Jobs,” The Good Search, December 1, 2019, <https://tgsus.com/diversity/google-shows-men-ads-for-better-jobs/>
- 109 Spencer, “An update on,” Google.
- 110 “Political content,” Google, <https://support.google.com/adspolicy/answer/6014595?hl=en>
- 111 “Political content,” Google
- 112 Spencer, “An update on,” Google
- 113 Dr Siddhartha Bandyopadhyay, “Less is better: guarding your privacy in a world of micro-targeting,” Birmingham Business School Blog, March 29, 2018, <https://blog.bham.ac.uk/business-school/2018/03/29/2499/>
- 114 Dipayan Ghosh, “What is microtargeting and what is it doing in our politics?,” Mozilla, October 4, 2018, <https://blog.mozilla.org/internetcitizen/2018/10/04/microtargeting-dipayan-ghosh/>
- 115 Pares Dave, “Google’s new rules clamp down on discriminatory housing, job ads,” Reuters, June 11, 2020, <https://www.reuters.com/article/us-alphabet-google-ads/googles-new-rules-clamp-down-on-discriminatory-housing-job-ads-idUSKBN23I2YK>
- 116 “Google Ads Policies,” Google
- 117 “About verification for election advertising in the United States,” Google, <https://support.google.com/adspolicy/answer/9002729?hl=en>
- 118 “U.S. election ads verification,” Google, https://support.google.com/adspolicy/contact/us_election_ads
- 119 “About verification for election,” Google
- 120 Google, *Political Advertising in the United States, Google Transparency Report*, <https://transparencyreport.google.com/political-ads/region/US>.
- 121 Google, *Political Advertising*.
- 122 “YouTube Community Guidelines enforcement,” Google Transparency Report, <https://transparencyreport.google.com/youtube-policy/removals?hl=en>
- 123 Richard Gingras, “A look at how news at Google works,” The Keyword, May 6, 2019 <https://blog.google/products/news/look-how-news-google-works/>

- 124 “Ranking within Google News,” Google, <https://support.google.com/news/publisher-center/answer/9606702?hl=en>
- 125 “Ranking With,” Google
- 126 “Appear in Google News,” Google, https://support.google.com/news/publisher-center/answer/9607025?hl=en&ref_topic=9603441
- 127 “Google News content policies,” Google, <https://support.google.com/news/publisher-center/answer/6204050?hl=en>
- 128 “Google News content,” Google
- 129 Kristen Canegallo, “Supporting the 2020 U.S. election,” Google, February 3, 2020, <https://blog.google/technology/safety-security/supporting-2020-us-election/>
- 130 Jigsaw, <https://jigsaw.google.com/>
- 131 Canegallo, “Supporting the 2020,” Google
- 132 Canegallo, “Supporting the 2020,” Google
- 133 “Google Pushing Scam Ads on Americans Searching for How to Vote,” Tech Transparency Project, June 29, 2020, <https://www.techtransparencyproject.org/articles/google-pushing-scam-ads-americans-searching-how-vote>
- 134 Amanda Storey, “An update on our 2020 U.S. election efforts,” Google, August 13, 2020, <https://blog.google/outreach-initiatives/civics/update-our-2020-us-election-efforts/>
- 135 Democracy Works, <https://www.democracy.works/>
- 136 Shane Huntley, “Protecting users from government-backed hacking and disinformation,” Google, November 26, 2019, <https://blog.google/threat-analysis-group/protecting-users-government-backed-hacking-and-disinformation/>
- #:-:text=Google's%20Threat%20Analysis%20Group%20(TAG,against%20Google%20and%20our%20users.&text=TAG%20tracks%20more%20than%20270,from%20more%20than%2050%20countries.
- 137 Lauren Cover, “11 Pinterest facts (and 30 stats) marketers must know in 2020,” Sprout Social, January 28, 2020, <https://sproutsocial.com/insights/pinterest-statistics/>
- 138 Isaac, Conger, “Google, Facebook and Others Broaden,” NYTimes.
- 139 “Advertising Guidelines,” Pinterest, <https://policy.pinterest.com/en/advertising-guidelines#subsection-political-campaigning>
- 140 “Community Guidelines,” Pinterest, <https://policy.pinterest.com/en/community-guidelines>
- 141 “Community Guidelines,” Pinterest
- 142 “Account Suspension,” Pinterest, <https://help.pinterest.com/en/article/account-suspension>
- 143 Aerica Shimizu Banks, “Safeguarding civic participation and the U.S. 2020 census,” Pinterest Newsroom, January 29, 2020, <https://newsroom.pinterest.com/en/post/safeguarding-civic-participation-and-the-us-2020-census>
- 144 “Get more help,” Pinterest, https://help.pinterest.com/en/contact?current_page=about_you_page&appeals=account_suspension
- 145 Mark Wilson, “The tech giant fighting anti-vaxxers isn’t Twitter or Facebook. It’s Pinterest,” Fast Company, February 26, 2019, <https://www.fastcompany.com/90310970/the-tech-giant-fighting-anti-vaxxers-isnt-twitter-or-facebook-its-pinterest>
- 146 Lauren Feiner, “Reddit Users Are the Least Valuable of Any Social Network,” CNBC, August 14, 2019, <https://www.cnbc.com/2019/02/11/reddit->

users-are-the-least-valuable-of-any-social-network.html.

147 Laura Hautala, "Reddit Was a Misinformation Hotspot in 2016 Election, Study Says," *CNET*, December 19, 2017, <https://www.cnet.com/news/reddit-election-misinformation-2016-research/>.

148 Hautala, "Reddit Was a Misinformation".

149 Hautala, "Reddit Was a Misinformation".

150 Singh, *Everything in Moderation*.

151 Hautala, "Reddit Was a Misinformation".

152 BBC Trending, "The Saga of 'Pizzagate': The Fake Story That Shows How Conspiracy Theories Spread," *BBC*, December 2, 2016, <https://www.bbc.com/news/blogs-trending-38156985>.

153 Adrienne LaFrance, "The Prophecies of Q," *The Atlantic*, June 2020, <https://www.theatlantic.com/magazine/archive/2020/06/qanon-nothing-can-stop-what-is-coming/610567/>.

154 Newman, "Tech Platforms".

155 Lisa Kaplan, "How Campaigns Can Protect Themselves From Deepfakes, Disinformation, and Social Media Manipulation," Brookings Institution, last modified January 10, 2019, <https://www.brookings.edu/blog/techtank/2019/01/10/how-campaigns-can-protect-themselves-from-deepfakes-disinformation-and-social-media-manipulation/>.

156 U/LastBluejay, "Updates to Our Policy Around Impersonation," *r/redditsecurity*, last modified January 9, 2020, https://www.reddit.com/r/redditsecurity/comments/emd7yx/updates_to_our_policy_around_impersonation/

157 U/LastBluejay, "Updates to Our Policy," *r/redditsecurity*.

158 Reddit, *Transparency Report 2019*, <https://www.redditinc.com/policies/transparency-report-2019>.

159 Lucas Matney, "Reddit Has Banned 944 Accounts Linked to the IRUA Russian Troll Farm," *TechCrunch*, April 18, 2018, <https://techcrunch.com/2018/04/11/reddit-bans-944-accounts-linked-to-russian-internet-research-agency/>.

160 Office of the Director of National Intelligence, *Background to "Assessing Russian Activities and Intentions in Recent US Elections": The Analytic Process and Cyber Incident Attribution*, January 6, 2017, https://www.dni.gov/files/documents/ICA_2017_01.pdf.

161 Young Mie Kim, "New Evidence Shows How Russia's Election Interference Has Gotten More Brazen," Brennan Center for Justice, last modified March 5, 2020, <https://www.brennancenter.org/our-work/analysis-opinion/new-evidence-shows-how-russias-election-interference-has-gotten-more>.

162 U/con_commenter, "Changes to Reddit's Political Ads Policy," *r/announcements*, last modified April 13, 2020, https://www.reddit.com/r/announcements/comments/g0s6tn/changes_to_reddits_political_ads_policy/.

163 U/con_commenter, "Changes to Reddit's," *r/announcements*.

164 Spandana Singh, "Reddit's Intriguing Approach to Political Advertising Transparency," *Slate*, May 1, 2020, <https://slate.com/technology/2020/05/reddit-political-advertising-transparency.html>.

165 Singh, "Reddit's Intriguing".

166 U/worstnerd, "Misinformation and COVID-19: What Reddit is Doing," *r/ModSupport*, last modified April 15, 2020, https://www.reddit.com/r/ModSupport/comments/g21ub7/misinformation_and_covid19_what_reddit_is_doing/.

- 167 U/worstnerd, "Misinformation and COVID-19," r/ModSupport.
- 168 U/worstnerd, "Misinformation and COVID-19," r/ModSupport.
- 169 U/worstnerd, "Misinformation and COVID-19," r/ModSupport.
- 170 Spandana Singh, *Rising Through the Ranks How Algorithms Rank and Curate Content in Search Results and on News Feeds*, October 21, 2019, <https://www.newamerica.org/oti/reports/rising-through-ranks/>.
- 171 Candice Wang, "Revisiting Reddit's Attempt to Stop 'Secondary Infection' Misinformation Campaign from Russia," The Free Internet Project, last modified July 20, 2020, <https://thefreeinternetproject.org/blog/revisiting-reddits-attempt-stop-secondary-infection-misinformation-campaign-russia>.
- 172 Wang, "Revisiting Reddit's," The Free Internet Project.
- 173 U/worstnerd, "Additional Insight into Secondary Infection on Reddit," r/redditsecurity, last modified April 8, 2020, https://www.reddit.com/r/redditsecurity/comments/xfi5s/additional_insight_into_secondary_infection_on/.
- 174 U/worstnerd, "Reddit Security Report - June 18, 2020," r/redditsecurity, last modified June 18, 2020, https://www.reddit.com/r/redditsecurity/comments/hbiuas/reddit_security_report_june_18_2020/.
- 175 U/worstnerd, "Reddit Security," r/redditsecurity.
- 176 Josh Constine, "Snapchat beats in Q3, adding 7M users & revenue up 50%," TechCrunch, October 22, 2019, <https://techcrunch.com/2019/10/22/snapchat-earnings-q3-2019/>
- 177 "Register to vote on Snapchat!," Snap.Inc, September 25, 2018, <https://www.snap.com/en-US/news/post/register-to-vote>
- 178 Sara Fischer, "Snapchat preps young users to vote in November," Axios, May 14, 2020, <https://www.axios.com/snapchat-young-users-vote-november-2020-1f0159ed-5894-4d15-949d-f7d1c63f9e3a.html>
- 179 Kim Lyons, "Snapchat is planning in-app voter awareness tools to help get its users ready for November," The Verge, August 6, 2020, <https://www.theverge.com/2020/8/6/21357023/snapchat-voter-tools-election-app-registration>
- 180 Casey Newton, "Snap announces Minis to bring other apps into Snapchat," The Verge, June 11, 2020, <https://www.theverge.com/2020/6/11/21286537/snap-minis-snapchat-partner-summit-2020-announcements>
- 181 Lyons, "Snapchat is planning in-app," The Verge
- 182 Brad Bennett, "Get-out-the-vote effort goes digital for COVID-19 pandemic," Southern Poverty Law Center, April 11, 2020 <https://www.splcenter.org/news/2020/04/11/get-out-vote-effort-goes-digital-covid-19-pandemic>
- 183 Andrew Hutchinson, "Snapchat Says Political Ads on its Platform Are Fact-Checked, Adding to Debate," Social Media Today, November 20, 2019, <https://www.socialmediatoday.com/news/snapchat-says-political-ads-on-its-platform-are-fact-checked-adding-to-deb/567641/>
- 184 Hutchinson, "Snapchat says Political," Social Media Today
- 185 "Snap Advertising Policies," Snap Inc., <https://www.snap.com/en-US/ad-policies#prohibited-content>

- 186 “Snap Political & Advocacy Advertising Policies,” Snap Inc., Effective May 15, 2019, <https://www.snap.com/en-US/ad-policies/political>
- 187 “Snap Political & Advocacy,” Snap.Inc.
- 188 “Snap Political & Advocacy,” Snap.Inc.
- 189 “Snap Political & Advocacy,” Snap.Inc.
- 190 “Snap Political & Advocacy,” Snap.Inc.
- 191 “Snap Political & Advocacy,” Snap.Inc.
- 192 “Snap Political Ads Library,” Snap Inc., <https://www.snap.com/en-US/political-ads/>
- 193 “Community Guidelines,” Snap Inc., September 2020, <https://www.snap.com/en-US/community-guidelines>
- 194 “Community Guidelines,” Snap Inc
- 195 “Community Guidelines,” Snap Inc
- 196 “Community Guidelines,” Snap Inc
- 197 Michael Nuñez, “Snapchat and TikTok Embrace ‘Deepfake’ Video Technology Even As Facebook Shuns It,” *Forbes*, January 8, 2020, <https://www.forbes.com/sites/mnunez/2020/01/08/snapchat-and-tiktok-embrace-deepfake-video-technology-even-as-facebook-shuns-it/#7da5557042c0>
- 198 Catherine Thorbecke, “Snapchat users can see who’s behind political ads with new feature,” *ABC News*, September 16, 2019, <https://abcnews.go.com/Business/ahead-2020-election-snapchat-unveils-political-ads-library/story?id=65639999>
- 199 Over the past several months, TikTok has drawn scrutiny from U.S. lawmakers for its ties to China. On August 6, President Trump issued an Executive Order that effectively bans TikTok in the United States by blocking any transactions with its parent company ByteDance. President Trump also signed an Executive Order on August 14 requiring ByteDance to sell or spin off its TikTok U.S. business within 90 days. As this report is being published, news accounts indicate that Oracle has agreed to become a “trusted technology provider” for TikTok. The deal is still waiting for approval from the U.S. government. This issue is not a focus of this report, although it is later mentioned at a high-level. <https://www.whitehouse.gov/presidential-actions/executive-order-addressing-threat-posed-tiktok/> <https://home.treasury.gov/system/files/136/EO-on-TikTok-8-14-20.pdf> <https://www.cnbc.com/2020/09/15/cramer-says-trump-administration-will-approve-oracle-tiktok-deal-tuesday.html>
- 200 Maryam Mohsin, “10 TikTok Statistics That You Need to Know in 2020 [Infographic],” *Oberlo*, last modified July 3, 2020, <https://www.oberlo.com/blog/tiktok-statistics>.
- 201 Andrew Marino, “How TikTok Could Be A Player In Election Disinformation,” *The Verge*, May 12, 2020, <https://www.theverge.com/2020/5/12/21255823/alex-stamos-interview-election-2020-security-tiktok-zoom-vergecast>.
- 202 Tony Romm and Drew Harwell, “TikTok Revamps Content Rules, Aiming to Clear Up Which Videos It Allows or Blocks,” *The Washington Post*, January 8, 2020, <https://www.washingtonpost.com/technology/2020/01/08/tiktok-revamps-content-rules-aiming-clear-up-which-videos-it-allows-or-blocks/>.
- 203 Lavanya Mahendran and Nasser Alsherif, “Adding Clarity To Our Community Guidelines,” *TikTok Newsroom*, last modified January 8, 2020, <https://newsroom.tiktok.com/en-us/adding-clarity-to-our-community-guidelines>.
- 204 TikTok, “Community Guidelines,” *TikTok*, last modified August 2020, <https://www.tiktok.com/community-guidelines?lang=en>.
- 205 TikTok, “Community Guidelines,” *TikTok*.

- 206 Vanessa Pappas, "Combating Misinformation and Election Interference on TikTok," TikTok Newsroom, last modified August 5, 2020, <https://newsroom.tiktok.com/en-us/combating-misinformation-and-election-interference-on-tiktok>.
- 207 Pappas, "Combating Misinformation," TikTok Newsroom.
- 208 Pappas, "Combating Misinformation," TikTok Newsroom.
- 209 Sarah Perez, "TikTok Explains Its Ban on Political Advertising," *TechCrunch*, October 3, 2019, <https://techcrunch.com/2019/10/03/tiktok-explains-its-ban-on-political-advertising/>.
- 210 Mahendran and Alsherif, "Adding Clarity," TikTok Newsroom.
- 211 Angela Chen, "A Leaked Excerpt of TikTok Moderation Rules Shows How Political Content Gets Buried," *MIT Technology Review*, November 25, 2019, <https://www.technologyreview.com/2019/11/25/102440/tiktok-content-moderation-politics-protest-netzpolitik/>.
- 212 Chen, "A Leaked".
- 213 Mariel Soto Reyes, "TikTok Could Be the Next Battleground for Political Content," *Business Insider*, January 7, 2020, <https://www.businessinsider.com/tiktok-could-be-battleground-for-political-content-2020-1>.
- 214 Georgia Wells and Emily Glazer, "TikTok Wants to Stay Politics-Free. That Could Be Tough in 2020.," *The Wall Street Journal*, January 5, 2020, <https://www.wsj.com/articles/tiktok-wants-to-stay-politics-free-that-could-be-tough-in-2020-11578225601>.
- 215 Wells and Glazer, "TikTok Wants".
- 216 Common Sense Media, *2020 Social Media Voter Scorecard*, 2020, <https://www.commonsensemedia.org/social-media-voter-scorecard2020>.
- 217 Reyes, "TikTok Could".
- 218 Reyes, "TikTok Could".
- 219 Stephanie Hind and Tara Wadhwa, "TikTok's 'Be Informed' Series Stars TikTok Creators to Educate Users About Media Literacy," TikTok Newsroom, last modified July 16, 2020, <https://newsroom.tiktok.com/en-us/tiktoks-be-informed-series-stars-tiktok-creators-to-educate-users-about-media-literacy>.
- 220 Alfred Ng, "US Officials in Contact With TikTok Over Political Disinformation," *CNET*, March 3, 2020, <https://www.cnet.com/news/us-officials-in-contact-with-tiktok-over-political-disinformation/>.
- 221 Ng, "US Officials".
- 222 Ben Kesling and Georgia Wells, "U.S. Military Bans TikTok Over Ties to China," *The Wall Street Journal*, January 3, 2020, <https://www.wsj.com/articles/u-s-military-bans-tiktok-over-ties-to-china-11578090613#:~:text=Representatives%20for%20the%20Coast%20Guard,declare%20the%20app%20off%2Dlimits>.
- 223 Riya Bhattacharjee, Amanda Macias, and Jordan Novet, "Trump Says He Will Ban TikTok Through An Executive Action," *CNBC*, August 1, 2020, , <https://www.cNBC.com/2020/07/31/trump-says-he-will-ban-tiktok-through-executive-action-as-soon-as-saturday.html>.
- 224 TikTok, *TikTok Transparency Report*, July 9, 2020, <https://www.tiktok.com/safety/resources/transparency-report?lang=en>.
- 225 Marino, "How TikTok".

- 226 Omnicore Agency, "Twitter by the Numbers: Stats, Demographics & Fun Facts," Omnicore Agency, last modified February 10, 2020, <https://www.omnicoreagency.com/twitter-statistics/>.
- 227 Twitter, "Civic Integrity Policy," Twitter, last modified September 2020, <https://help.twitter.com/en/rules-and-policies/election-integrity-policy>.
- 228 Twitter, "Civic Integrity," Twitter.
- 229 Cristiano Lima, "Twitter Forces Democratic Candidate To Delete Post Flouting Voter Suppression Rules," *Politico*, September 1, 2020, <https://www.politico.com/news/2020/09/01/texas-democrat-deleted-twitter-post-407031>.
- 230 Twitter, "Civic Integrity," Twitter.
- 231 Twitter, "Civic Integrity," Twitter.
- 232 Twitter, "Civic Integrity," Twitter. Twitter, "Violent Threats Policy," Twitter, last modified March 2019, <https://help.twitter.com/en/rules-and-policies/violent-threats-glorification>.
- 233 Twitter, "Civic Integrity," Twitter.
- 234 Twitter Safety, "Expanding Our Policies to Further Protect the Civic Conversation," *Twitter Blog*, entry posted September 10, 2020, https://blog.twitter.com/en_us/topics/company/2020/civic-integrity-policy-update.html.
- 235 Twitter, "Civic Integrity," Twitter.
- 236 Twitter, "About Public-Interest Exceptions on Twitter," Twitter, <https://help.twitter.com/en/rules-and-policies/public-interest>.
- 237 Twitter, "Civic Integrity," Twitter.
- 238 Twitter, "Civic Integrity," Twitter.
- 239 Twitter, "Civic Integrity," Twitter.
- 240 Kanishka Singh, "Twitter to Expand Rules Against Misinformation on Mail-In Ballots, Early Voting," *Reuters*, August 13, 2020, <https://www.reuters.com/article/usa-election-twitter/twitter-to-expand-rules-against-misinformation-on-mail-in-ballots-early-voting-idUSL4N2FF19K>.
- 241 Bridget Coyne and Sam Toizer, "Helping You Find Accurate US Election News and Information," *Twitter Blog*, entry posted September 15, 2020, https://blog.twitter.com/en_us/topics/company/2020/2020-election-news.html Helping you find accurate US Election News and Information.
- 242 Newman, "Tech Platforms".
- 243 Steven Overly, "Twitter Users Can Now Report Voter Suppression, Misinformation," *Politico*, January 29, 2020, <https://www.politico.com/news/2020/01/29/twitter-misinformation-voter-suppression-109282>.
- 244 Pranav Dixit, "Twitter Will Let Users Report Tweets That Mislead Voters," *Buzzfeed News*, April 24, 2019, <https://www.buzzfeednews.com/article/pranavdixit/twitter-tool-report-tweets-mislead-voters-elections>.
- 245 Amnesty International, *Toxic Twitter*, March 2018, <https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-4/>.
- 246 Twitter, *Rules Enforcement Report*, <https://transparency.twitter.com/en/reports/rules-enforcement.html#2019-jul-dec>.
- 247 Twitter, *Rules Enforcement*.
- 248 Twitter, *Rules Enforcement*.
- 249 Amnesty International, *Toxic Twitter*.
- 250 Twitter, "Updating Our Advertising Policies on State Media," *Twitter Blog*, entry posted August 19, 2019, https://blog.twitter.com/en_us/topics/

company/2019/
advertising_policies_on_state_media.html.

251 Twitter, "Updating Our Advertising," *Twitter Blog*.

252 Jack Dorsey (@jack), "We've made the decision to stop all political advertising on Twitter globally. We believe political message reach should be earned, not bought. Why? A few reasons...Thread," Twitter, October 30, 2019, 4:05 PM, <https://twitter.com/jack/status/1189634360472829952>.

253 Kim Lyons, "You Can Now Report Voter Suppression on Twitter Ahead of the 2020 Election," *The Verge*, January 30, 2020, <https://www.theverge.com/2020/1/30/21115487/twitter-elections-facebook-youtube-warren-iowa-caucuses>.

254 Twitter Safety, "Defining Public Interest on Twitter," *Twitter Blog*, entry posted June 27, 2019, https://blog.twitter.com/en_us/topics/company/2019/publicinterest.html.

255 Dan Mangan and Kevin Breuninger, "Twitter Fact-Checks Trump, Slaps Warning Labels On His Tweets About Mail-In Ballots," *CNBC*, May 26, 2020, <https://www.cnbc.com/2020/05/26/twitter-fact-checks-trump-slaps-warning-labels-on-his-tweets-about-mail-in-ballots.html>.

256 Holmes, "Trump Made".

257 Emily Saltz et al., "It Matters How Platforms Label Manipulated Media. Here are 12 Principles Designers Should Follow.," *The Startup* (blog), entry posted June 9, 2020, <https://medium.com/swlh/it-matters-how-platforms-label-manipulated-media-here-are-12-principles-designers-should-follow-438b76546078>.

258 Twitter Support, "New Labels for Government and State-Affiliated Media Accounts," *Twitter Blog*, entry posted August 6, 2020, https://blog.twitter.com/en_us/topics/product/2020/new-labels-for-government-and-state-affiliated-media-accounts.html.

labels-for-government-and-state-affiliated-media-accounts.html.

259 Twitter Support, "New Labels," *Twitter Blog*.

260 Twitter, "About Government and State-Affiliated Media Account Labels on Twitter," Twitter, <https://help.twitter.com/en/rules-and-policies/state-affiliated>.

261 Twitter, *Twitter Transparency Report*, <https://transparency.twitter.com/en/reports/countries/us.html#2019-jul-dec>.

262 Radsch, "Tech Platforms," Committee to Protect Journalists.

263 Nix and Wagner, "Social Media".

264 Spandana Singh, *Charting a Path Forward: Promoting Fairness, Accountability, and Transparency in Algorithmic Content Shaping*, September 9, 2020, <https://www.newamerica.org/oti/reports/charting-path-forward/>.

265 J. Clement, "Most Popular Global Mobile Messenger Apps As Of July 2020, Based on Number of Monthly Active Users," Statista, last modified July 24, 2020, <https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/>.

266 Manish Singh, "WhatsApp's New Limit Cuts Virality of 'Highly Forwarded' Messages by 70%," *TechCrunch*, April 27, 2020, <https://techcrunch.com/2020/04/27/whatsapps-new-limit-cuts-virality-of-highly-forwarded-messages-by-70/>.

267 Facebook, "New Steps," Facebook Newsroom.

268 Pranav Dixit, "WhatsApp Is Now Letting Users Know When A Message Has Been Forwarded Too Many Times," *BuzzFeed News*, August 2, 2019, <https://www.buzzfeednews.com/article/pranavdixit/whatsapp-double-arrows-forwarded-messages-misinformation>.

- 269 Manish Singh, "WhatsApp Now Lets You Control Who Can Add You To A Group," *VentureBeat*, April 3, 2019, <https://venturebeat.com/2019/04/03/whatsapp-now-lets-you-control-who-can-add-you-to-a-group/>.
- 270 WhatsApp, "Search the Web," *WhatsApp Blog*, entry posted August 3, 2020, <https://blog.whatsapp.com/search-the-web>.
- 271 Julio C.S. Reis et al., "Can WhatsApp Benefit From Debunked Fact-Checked Stories To Reduce Misinformation?," *Harvard Kennedy School Misinformation Review* 1, no. 5 (August 20, 2020): <https://misinfreview.hks.harvard.edu/article/can-whatsapp-benefit-from-debunked-fact-checked-stories-to-reduce-misinformation/>.
- 272 In addition, civil society experts have raised numerous concerns around the use of hash databases, as internet platforms often provide little transparency around how they vet content before including it as hashes in these databases, and what impact these databases have on user expression. Svea Windwehr and Jillian C. York, "One Database to Rule Them All: The Invisible Content Cartel that Undermines the Freedom of Expression Online," Electronic Frontier Foundation, last modified August 27, 2020, <https://www.eff.org/deeplinks/2020/08/one-database-rule-them-all-invisible-content-cartel-undermines-freedom-1>.
- 273 WhatsApp, *Stopping Abuse: How WhatsApp Fights Bulk Messaging and Automated Behavior*, February 6, 2019, https://scontent.whatsapp.net/v/t61.22868-34/69510151_652112781951150_6923638360331596993_n.pdf/WA_StoppingAbuse_Whitepaper_020418_Update.pdf?_nc_sid=41cc27&_nc_ohc=l-n5XK5SywkAX_XhWXS&_nc_ht=scontent.whatsapp.net&oh=192fd642b27b6e717a4ab8a91bf19fd3&oe=5F627A7F.
- 274 Private meeting with Facebook representative, August 5, 2020
- 275 Private meeting with Facebook representative, August 5, 2020
- 276 WhatsApp, *Stopping Abuse*.
- 277 WhatsApp, *Stopping Abuse*.
- 278 WhatsApp, *Stopping Abuse*.
- 279 Maryam Mohsin, "10 Youtube Stats Every Marketer Should Know in 2020 [Infographic]," Oberlo, May 11, 2020, <https://www.oberlo.com/blog/youtube-statistics>
- 280 Lata Nott, "Political Advertising on Social Media Platforms," American Bar Association, June 26, 2020, https://www.americanbar.org/groups/crsj/publications/human_rights_magazine_home/voting-in-2020/political-advertising-on-social-media-platforms/
- 281 Leslie Miller, "How YouTube supports elections," YouTube: Official Blog, February 3, 2020, <https://youtube.googleblog.com/2020/02/how-youtube-supports-elections.html>
- 282 "Spam, deceptive practices & scams policies," YouTube Help, <https://support.google.com/youtube/answer/2801973?hl=en>
- 283 "Spam, deceptive practices," YouTube Help
- 284 "YouTube Trusted Flagger program," YouTube Help, https://support.google.com/youtube/answer/7554338?ref_topic=7124235
- 285 "YouTube Trusted Flagger," YouTube Help
- 286 Accountable Tech, *Election Integrity*.
- 287 "The Four Rs of Responsibility, Part 2: Raising authoritative content and reducing borderline content and harmful misinformation," YouTube Official Blog, December 3, 2019 <https://blog.youtube/inside-youtube/the-four-rs-of-responsibility-raise-and-reduce>

288 “Expanding fact checks on YouTube to the United States,” YouTube:Official Blog, April 28, 2020, <https://blog.youtube/news-and-events/expanding-fact-checks-on-youtube-to-united-states>

289 “Information panel providing publisher context,” YouTube Help, <https://support.google.com/youtube/answer/7630512?hl=en>

290 Paul Lewis, “Fiction is outperforming reality: how YouTube’s algorithm distorts truth,” The Guardian, February 2, 2018, <https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth> Spandana Singh, *Why Am I Seeing This? How Video and E-Commerce Platforms Use Recommendation Systems to Shape User Experiences*, March 25, 2020, <https://www.newamerica.org/oti/reports/why-am-i-seeing-this/>.

291 Andrew Hutchinson, “YouTube Updates Recommendations Algorithm to Lessen the Spread of ‘Borderline Content’,” Social Media Today, January 26, 2019, <https://www.socialmediatoday.com/news/youtube-updates-recommendations-algorithm-to-lessen-the-spread-of-borderline/546906/>

292 Paige Cooper, “How Does the YouTube Algorithm Work? A Guide to Getting More Views,” Hootsuite, August 18, 2020, <https://blog.hootsuite.com/how-the-youtube-algorithm-works/>
#:~:text=According%20to%20YouTube%2C%20the%20algorithm,get%20viewers%20to%20keep%20watching

293 “Continuing our work to improve recommendations on YouTube,” YouTube: Official Blog, January 25, 2019, <https://blog.youtube/news-and-events/continuing-our-work-to-improve>

294 Common Sense Media, *2020 Social*.

295 “Protecting our extended workforce and the community,” YouTube Official Blog, March 16, 2020,

<https://youtube-creators.googleblog.com/2020/03/protecting-our-extended-workforce-and.html>

296 “YouTube Community Guidelines enforcement,” Google Transparency Report, <https://transparencyreport.google.com/youtube-policy/removals?hl=en>

297 “Community Guidelines strike basics,” Google, <https://support.google.com/youtube/answer/2802032?hl=en>

298 “Political advertising in the United States,” Google Transparency Report, <https://transparencyreport.google.com/political-ads/region/US>

299 “YouTube Community Guidelines,” Google Transparency Report,

300 Gaurav Laroia and David Brody, “Privacy Rights Are Civil Rights. We Need to Protect Them.,” FreePress, last modified March 14, 2019, <https://www.freepress.net/our-response/expert-analysis/insights-opinions/privacy-rights-are-civil-rights-we-need-protect-them>.



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.

If you have any questions about citing or reusing New America content, please visit www.newamerica.org.

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