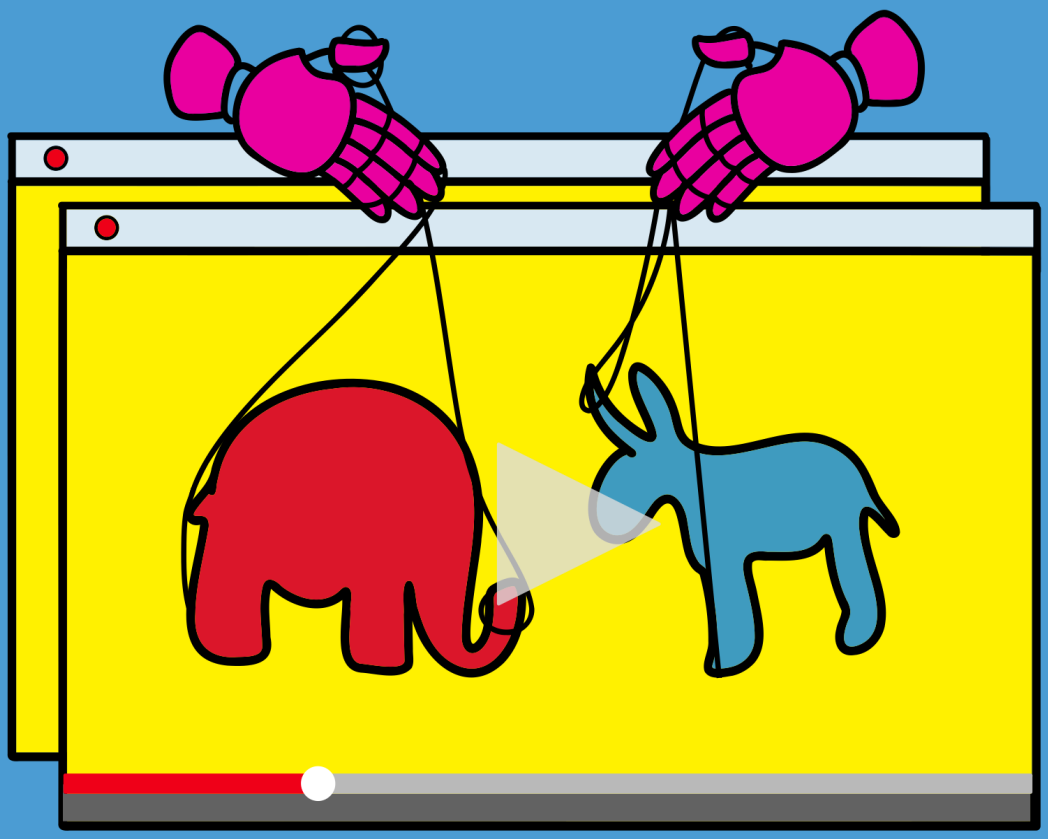


The new political ad machine:

Policy frameworks for political ads in an age of AI

Scott Babwah Brennen & Matt Perault

Center on Technology Policy at the University of North Carolina at Chapel Hill

About the Center on Technology Policy

UNC's Center on Technology Policy seeks to craft public policy for a better internet. Utilizing an interdisciplinary academic framework, CTP works to identify knowledge gaps and develop actionable policy frameworks that will enable us to realize the potential benefits of technology while minimizing its harms. By working closely with students and expanding the University's offerings in technology policy analysis, we seek to cultivate and train the field's future practitioners.

About the Authors

J. Scott Babwah Brennen is head of online expression policy at the Center on Technology Policy at UNC-Chapel Hill.

Matt Perault is the director of the Center on Technology Policy at UNC-Chapel Hill, a professor of the practice at UNC's School of Information and Library Science, and a consultant on technology policy issues at Open Water Strategies.

This report was designed by Le Ha.

Executive Summary

As we approach the 2024 presidential election, policymakers, practitioners, and scholars are assessing the promise and pitfalls of generative artificial intelligence (GAI) in elections. While some practitioners have observed that GAI may help optimize and improve political ad production and targeting, there has been far more concern that GAI will lead to wide-scale disruption of political life.

This brief examines the use of GAI in political ads to date, assesses the potential risks and benefits of its use, reviews what existing empirical research can teach us about those risks, and then uses those insights as the basis for a set of recommended policy interventions.

Although the use of GAI in political ads use has been limited thus far, we anticipate increased usage in the 2024 election cycle and beyond. Despite the limited use of GAI in political ads to date, a great deal of public commentary has speculated on the potential harm that GAI might bring to political advertising. Those concerns fall into four main categories:

1. **Scale:** GAI may facilitate an increase in the volume of deceptive content in political ads by lowering the cost and difficulty of producing manipulated content.
2. **Authenticity:** GAI may produce falsehoods that look more realistic or that appear to come from authentic sources.
3. **Personalization:** GAI may allow advertisers to better personalize targeted content to smaller audience segments, increasing the effectiveness of deceptive ads.
4. **Bias:** GAI may exacerbate bias and discrimination in political ads.

Policymakers have moved quickly to introduce proposals to address these concerns. Most proposals have focused on three interventions: watermarks on all GAI content, disclaimers on political ads containing GAI content, and bans on deceptive GAI content in political ads.

While there is limited empirical research on GAI in political ads, our reading of the literature considering online misinformation, political ads, and bias in AI models offers five important insights into the potential harm of GAI in political ads:

- First, research suggests that the persuasive power of both political ads and online misinformation is often overstated. Political ads likely have more of an effect on behavior – such as voter turnout and fundraising – than on persuasion.
- Second, political ads likely have the greatest impact in smaller, down-ballot races where there is less advertising, oversight, or familiarity with candidates.
- Third, GAI content has the potential to replicate bias, including racial, gender, and national biases.
- Fourth, research on political disclaimers suggests that watermarks and disclaimers are unlikely to significantly curb risks.
- Fifth, significant holes in the research remain.

These insights from the literature help to formulate recommendations for policymakers that can mitigate the potential harm of GAI without unduly constraining its potential benefits. Research suggests that policy should focus more on preventing abuse in smaller, down-ballot races and in mitigating bias than on banning deceptive GAI content or requiring disclaimers or watermarks. Although the research points in this direction, holes in the literature remain. The result is that we should approach its insights from a position of curiosity, rather than certainty, and conduct additional research into the impact of GAI on the electoral process.

Building on our assessment of the academic literature, we offer ten recommendations for policymakers seeking to limit the potential risks of GAI in political ads. These recommendations fall into two categories: First, public policy should target electoral harms rather than technologies. Second, public policy should promote learning about GAI so that we can govern it more effectively over time.

Recommendations

Public policy should target electoral harms, not technologies.

1. Congress and states should outlaw voter suppression.
2. Governments should allocate additional funding for law enforcement to enforce existing civil rights law that protects the electoral process.
3. Local and state governments should "flood the zone" with factual content.
4. Governments should fund digital literacy programs focused on detecting and contextualizing false online content.
5. The FEC should publish guidance for political advertisers on identifying and mitigating bias in political ads, with special emphasis on bias introduced by GAI models.

Public policy should promote learning about GAI.

6. Governments should fund empirical studies on the impact of GAI in political ads and the effectiveness of GAI-related interventions.
7. Governments should establish policy experiments to test interventions aimed at mitigating the negative impact of GAI in political ads.
8. Congress should learn about the impact of different GAI content liability regimes and incorporate these learnings into the law.
9. Federal and state governments should establish political ad archives and require advertisers to submit copies of ads along with regular disclosure reports.
10. The Federal Election Commission (FEC) should require that campaigns report their vendors' advertising spending.

Background

Following the release of OpenAI's ChatGPT in November 2022, public attention on GAI skyrocketed. As the 2023 election approached, policymakers began to examine the potential impact of GAI in elections, and their attention has already turned to the likely increase in GAI during the 2024 primary and general elections. In a recent Senate hearing on AI and elections, Senator Chuck Schumer (D-NY) observed:

> *If we don't act, we could soon live in a world where political campaigns regularly deploy totally fabricated but also totally believable images and footage of Democratic or Republican candidates distorting their statements and greatly harming their election chances.*

Many people see either the savior or demon of the political process in GAI models: the tool that will widen participation or that which will pervert it, amplify the voices of smaller candidates, or those of liars and bad actors. This section assesses the potential harms that GAI in political ads pose to the electoral process. To do so, it first discusses how GAI has been used in political ads to date, how platforms govern its use, and the interventions policymakers have proposed to regulate GAI in political ads.

How are campaigns using GAI in political ads?

While we are likely to see additional examples as the 2024 primary and general elections proceed, to date there have been few confirmed examples of GAI in political ads. Yet, the examples provide some indications of how campaigns plan to use GAI to advance their objectives. Based on our observations, we group them into three main categories: publicity, alteration, and fabrication.

Figure 1. Ad released by the Republican National Committee imagining life under a second Biden administration.

Figure 2. Ad released by DeSantis-aligned PAC. The jets were likely added to the video.

Publicity

In late April 2023, the Republican National Committee (RNC) released an ad it described as an "AI-generated look into the country's possible future if Joe Biden is re-elected in 2024." Unsurprisingly, the ad shows a particularly dark future (Figure 1). The RNC explicitly stated that the ad was created using a GAI tool. In fact, it is likely that the RNC referenced its use of GAI to try to increase publicity of the ad.

Alteration

During the summer of 2023, a DeSantis-aligned political action committee (PAC) released an ad that likely used a GAI tool to synthesize audio of Trump speaking the words of one of his tweets regarding Iowa Governor Kim Reynolds. While the tweet was real, the audio was synthesized.

In May 2023, the same PAC released an ad that included a shot of DeSantis waving to a crowd as military jets flew overhead (Figure 2). According to news reports, the jets had been artificially added to the image. Whether the PAC used GAI or other video editing tools is unclear, but the ad raises the possibility that campaigns could use GAI to make subtle changes to content.

Fabrication

In June 2023, the DeSantis campaign released an ad that included GAI images of Donald Trump hugging Anthony Fauci. The still images were included alongside real images of Trump and Fauci together (Figure 3).

Figure 3. Ad released by the DeSantis campaign that included false GAI images of Donald Trump and Anthony Fauci hugging.

How are platforms addressing GAI in political communications?

Following the introduction of deepfakes around 2017, tech platforms began to revise their policies to address GAI media. Many platforms have since updated their policies to address GAI more broadly. Several of the GAI platforms that have burst on the scene in the past year have released their own policies. There remains wide variation in how platforms treat GAI content.

Digital Advertising Platforms

For tech platforms that run political ads, we observe at least three key differences in their policies on GAI media. Note that in this analysis, we do not consider platforms and ad networks that prohibit all political ads, including Amazon, LinkedIn, Pinterest, TikTok, Twitch, and Xandr. While these platforms do have synthetic media policies, we do not consider them here because they do not permit advertisers to run political ads.

First, Google is the only platform to require disclaimers on ads that include GAI content. In September 2023, Google announced that it would require advertisers to "prominently disclose when their ads contain synthetic content that inauthentically depicts real or realistic-looking people or events. This disclosure must be clear and conspicuous..." In the announcement, Google indicated that the policy would take effect in "mid-November 2023," meaning that it will not be in place for the 2023 election.

Notably, however, the forthcoming policy states that "[a]ds that contain synthetic content altered or generated in such a way that is inconsequential to the claims made in the ad will be exempt from these disclosure requirements."

It is unclear how exactly Google will enforce this exception. For example, would the policy apply to the recent DeSantis campaign ad that may have used GAI to add military jets (see Figure 2), or would Google exempt the ad because it would deem the presence of military jets to be inconsequential? At the same time, as noted below, Google's advertising policy already prohibits "[m]anipulating media to deceive, defraud, or mislead others." Taken together, Google seems likely to require disclosures only for consequential claims that are not intended to deceive or mislead.

Second, there is a difference in how current policies define GAI media. On one hand, Snap's policy covers all manipulated media—whether created by an AI or by a human. On the other hand, Meta's policy only applies if a "video is the product of artificial intelligence or machine learning."

Third, and most importantly, there is wide variation in the scope of the prohibitions. Several platforms prohibit misleading content, but do not address GAI-produced media specifically. For example, Hulu does not address GAI media, but notes that "Advertisements must not be misleading."

Other platforms do not address this issue in their ad policies, but instead tackle it in content policies that apply to both paid and organic content. For example, X's content policies prohibit GAI-produced or manipulated media, but only if they are "likely to cause harm." Reddit's content policy prohibits "deepfakes or other manipulated content presented to mislead, or falsely attributed to an individual or entity." Reddit is silent on the issue in its ad policies, however, like other platforms, Reddit explicitly states that all ads must follow general content policies.

Snap offers one of the broadest prohibitions of manipulated and GAI-produced media. Snap prohibits "manipulating content for false or misleading purposes," specifying that it disallows all "content that is misleading, deceptive, impersonates any person or entity, or otherwise misrepresents your affiliation with a person or entity." Depending on how it is interpreted, prohibiting all "misleading" content regardless of severity of harm could amount to a near total ban on GAI-produced media.

Google's ad policy prohibits "[m]anipulating media to deceive, defraud, or mislead others," specifically including "[d]eceptively doctoring media related to politics, social issues, or matters of public concern." However, YouTube's content policies forbid GAI media only if it "may pose a serious risk of egregious harm." This means that a political advertiser could still post an ad with manipulated content, even if Google or YouTube will not serve it as an ad.

Meta offers perhaps the most specific GAI media policy. Meta prohibits:

- *Videos that have been edited or synthesized, beyond adjustments for clarity or quality, in ways that are not apparent to an average person, and would likely mislead an average person to believe:*
  - *A subject of the video said words that they did not say, AND*
  - *The video is the product of artificial intelligence or machine learning, including deep learning techniques (e.g., a technical deepfake), that merges, combines, replaces, and/or superimposes content onto a video, creating a video that appears authentic.*

The policy is limited to videos that show people speaking. It likely permits a wide range of other GAI-produced media, such as a fake image of President Trump hugging Anthony Fauci included in the ad described above (Figure 3).

Finally, of the Demand-Side Platforms (DSPs) we reviewed, none offer explicit public policies on GAI content in ads.

Generative AI Platforms

OpenAI's ChatGPT and Anthropic's Claude-2, two text-based platforms, both expressly prohibit the use of GAI for "political campaigning or lobbying," including "generating campaign materials personalized to or targeted at specific demographics." However, a recent Washington Post investigation found that, at least as of several months ago, ChatGPT produced text targeted to specific demographic groups.

Meta's Llama 2 prohibits use of the model to "[i]ntentionally deceive or mislead others." Google's Bard prohibits generating and distributing "content intended to misinform, misrepresent or mislead."

For image generators, OpenAI's DALL-E 2 and Midjourney both prohibit generating content for political ads. Stability AI's Stable Diffusion prevents "[g]enerating or facilitating large-scale political advertisements, propaganda, or influence campaigns."

Adobe's Firefly 2 does not bar use for political ads, but prohibits "misleading, fraudulent, or deceptive content that could lead to real-world harm."

Notably, DALL-E 2 refuses to generate images of famous political figures, including President Donald Trump, President Joe Biden, Senator Mitch McConnell (R-KY), or Representative Alexandria Ocasio-Cortez (D-NY). However, it did return images for lesser known federal and state politicians. For example, it returned an image of U.S. Senators Bob Menendez and Alex Padilla hugging in September (Figure 4a), but refused in mid-October. It did return an image of Representatives Greg Stanton (AZ-D) and David Scheikert (AZ-R) hugging (Figure 4b).

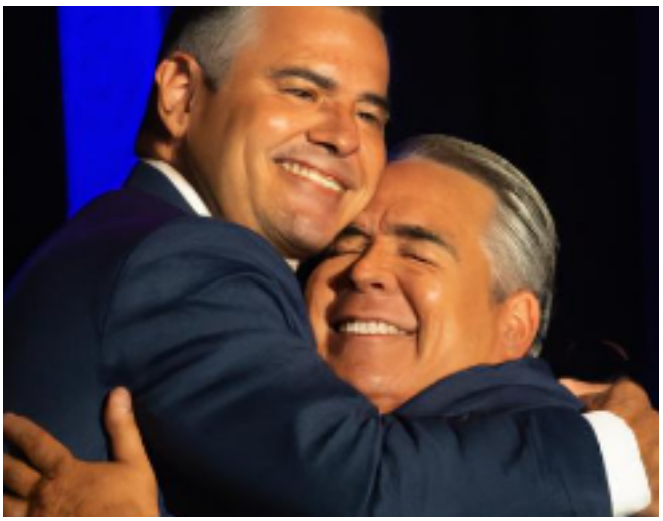


Figure 4a. DALL-E 2 image of US Senators Bob Menendez (D-NJ) and Alex Padilla (D-CA) hugging.

Figure 4b. DALL-E 2 image of US Representatives Greg Stanton (AZ-D) and David Scheikert (AZ-R) hugging.

Figure 5a. Stable Diffusion-created image of Donald Trump and Joe Biden hugging.

Figure 5b. Stable Diffusion-created image of NC Governor Roy Cooper eating a hamburger.

Stable Diffusion returned images for all queries for all levels of politician, creating images of North Carolina Governor Roy Cooper eating a hamburger and of Biden and Trump hugging — although with some errors (Figure 5a).

What are the concerns about GAI in political ads?

Following the appearance of the examples of GAI content in political ads discussed above and in anticipation of increased use of GAI in ads in the 2024 election, digital strategists, companies, and commentators have raised concerns about the risks of this technology. Their commentary focuses on four primary harms: authenticity, scale, personalization, and bias.[1]

Scale

Commentators suggest that new GAI-based tools will make it far easier and cheaper for advertisers to create false content. Photo manipulation has existed as long as photographs and this problem has increased in prominence with the introduction of deep fake technology.

GAI could enable existing advertisers to create a higher volume of false content and new advertisers to enter the market and run false ads.

Authenticity

Second, commentators have expressed concern that GAI will allow bad actors to create more realistic deceptive ads. While off-the-shelf GAI models continue to struggle to produce photo-realistic images, those images are rapidly improving. More realistic content could potentially increase the effectiveness of broad influence campaigns by making deceptive ads more persuasive. Alternatively, regardless of how prevalent deceptive GAI ads are, GAI could decrease trust in the electoral process if voters believe that political ads are rife with deceptive GAI content.

Personalization

Third, other commentators are concerned that GAI will allow advertisers to more easily create personalized content. Personalization has been a major component of political advertising campaigns for two decades, but it is possible that advertisers could use GAI to significantly increase their personalization of ads by quickly generating versions of an advertisement that could then be targeted to different audience segments.

1 Notably, while others have discussed how AI in general may impact elections, we focus narrowly on the question of generative AI in political ads. Therefore, while there is much to say about how new AI-based models may lead to improvements in campaign management, that is outside of our scope here.

Campaigns could use these different versions to test out which ads are most effective for a specific segment. Coupled with the authenticity and scale concerns identified above, this enhanced personalization could make it easier to develop targeted, false ad content that is distributed more widely and that is more impactful.

As discussed above, while some GAI models explicitly prohibit this use, there are holes in their enforcement of those restrictions.

Bias

Finally, GAI may perpetuate or exacerbate bias. AI models and systems have exhibited biases in areas like race, gender identity, and sexuality. GAI models are likely no different; accordingly, it is possible that GAI content in political ads may also exhibit bias. Biased ads have the potential to deepen social division, exacerbate stereotypes, fuel racism and other prejudices, and even provoke real-world violence.

What are the potential benefits of GAI in political advertising?

Alongside these concerns, commentators have also pointed to potential benefits of using GAI tools in political advertising. The potential downsides of scale, authenticity, and personalization all have potential benefits as well. Lowering the barriers to advertising – by reducing the cost and difficulty of producing them – and making ads more impactful may benefit advertisers looking to use online ads as a means of communicating their message, garnering support, and changing behavior. A wide array of diverse political causes could benefit from these improvements, from climate change and the pro-choice movement on the left to border security and the pro-life movement on the right. Reducing barriers to paid online expression may mean that more people will be able to use this communication tool to express themselves.

GAI improvements are likely to be particularly impactful for smaller campaigns and challengers. Those types of campaigns often operate with more limited funding than larger campaigns and incumbents, and so reducing advertising barriers may make it easier for them to compete with organizations that have deeper pockets.

What interventions have policymakers proposed to address the concerns and promote the benefits of GAI?

Policymakers have responded to these concerns with three main proposals: requiring watermarks on GAI images, requiring disclaimers for GAI ads, and banning deceptive GAI content from political ads.

In this section, we review these proposals at both the state and federal level.

Federal Proposals

Federal legislators have not passed any legislation restricting GAI media in political ads. However, at least two proposals were introduced in Congress in 2023: one would ban GAI audio and video content intended to deceive, and the other would require disclaimers.

In September, Senators Amy Klobuchar (D-MN), Chris Coons (D-DE), Josh Hawley (R-MO), and Susan Collins (R-ME) introduced the Protect Elections from Deceptive AI Act. The bill would prohibit a person from "knowingly distribut[ing] materially deceptive AI-generated audio or visual media" with the intent to "(1) influence an election; or (2) solicit funds." The bill is not limited to ads and would apply to both paid and organic content meant to influence an election or solicit funds.

The REAL Political Ads Act was introduced by Representative Yvette Clarke (D-NY) in the House, and Senators Klobuchar (D-MN), Cory Booker (D-NJ), and Michael Bennet (D-CO) in the Senate. In addition to expanding the definition of "electioneering communication" to include internet and digital advertising, the bill states that if an ad "contains an image or video footage which was generated in whole or in part with the use of artificial intelligence (generative AI), the communication shall include, in a clear and conspicuous manner, a statement that the communication contains such an image or footage."

The executive branch has also been active on this issue. On October 30, 2023, Biden signed an executive order on AI. While the EO does not directly address GAI in political ads, it supports new standards and assessments of AI models, addresses the prosecution of civil rights violations, and directs the Department of Commerce to develop standards for content authentication and watermarking.

The Biden administration published a Blueprint for an AI Bill of Rights in October 2022. In July 2023, it announced that it had brokered an agreement with seven companies who voluntarily committed to "developing robust technical mechanisms to ensure that users know when content is AI generated, such as a watermarking system."

In August, the FEC opened a public comment process on a petition for rulemaking submitted by Public Citizen. The proposed rule would outlaw "deliberately deceptive AI-produced content in campaign communication." Existing law includes the following prohibition:

> *[No candidate shall] fraudulently misrepresent himself or any committee or organization under his control as speaking or writing or otherwise acting for or on behalf of any other candidate or political party or employee or agent thereof on a matter which is damaging to such other candidate or political party or employee or agent thereof.*

The proposed rule would clarify that this prohibition includes deliberately deceptive GAI media. The Republican members of the FEC have expressed skepticism that this provision should be expanded to ban GAI media without legislative change to the FEC's powers.

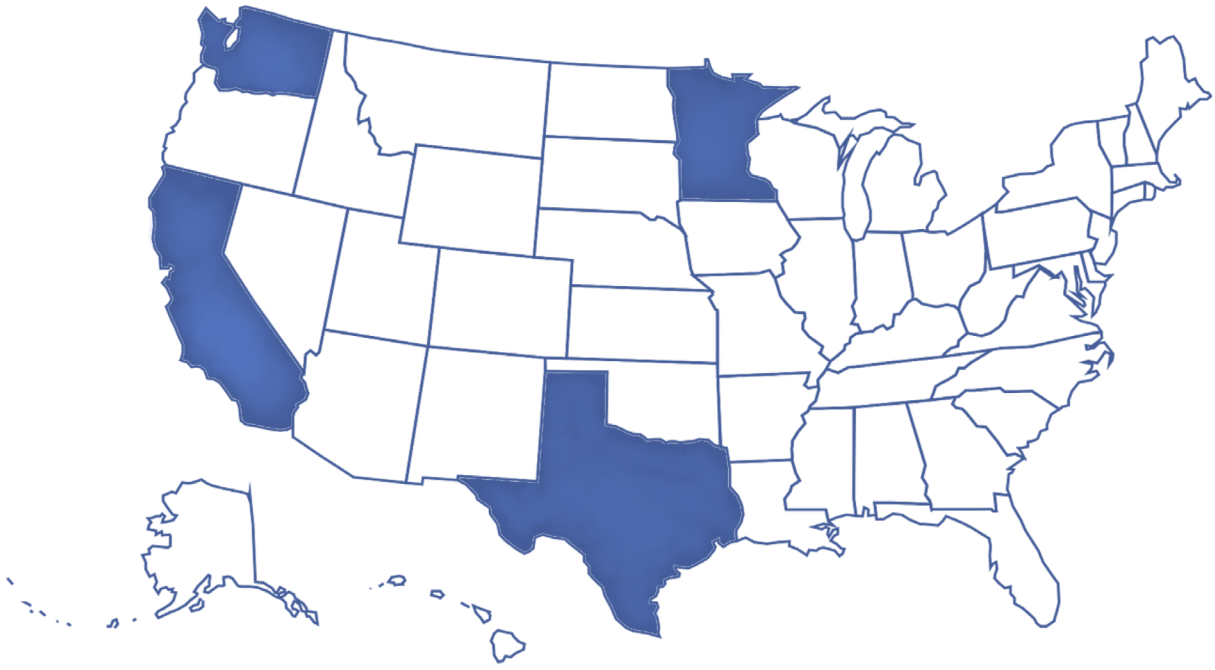
State Laws

Four states have passed laws that restrict GAI-produced media in political ads.

In 2019, Texas passed a law that banned the creation and distribution of GAI media "with intent to injure a candidate or influence the result of an election." The law, however, narrowly applies to "video[s] created with artificial intelligence that, with the intent to deceive, appears to depict a real person performing an action that did not occur in reality."

Minnesota also enacted a similar law banning deceptive GAI media in political communication meant to influence an election. That law, however, adopts a more expansive definition than the Texas law, prohibiting any technologically created media that results in a "a realistic but false image, audio, or video" meant to deceive voters. The Minnesota law, which went into effect on August 1, 2023, also allows for injunctive relief or prior restraint "against any person who is reasonably believed to be about to violate or who is in the course of violating this section."

Additionally, California, Washington, and Michigan have all enacted laws that require political ads containing deceptive GAI content to include disclaimers stating that the ads have been manipulated. The California law, which was passed in 2019, expired as of January 1, 2023, and has not been renewed.



Research

What can research tell us about the potential impact of GAI on political ads?

The rise of GAI is so new that we have very limited public research on the impact of GAI in political ads. However, there is more research on the effectiveness of misinformation and political ads that offers relevant insight into the potential harm of GAI in political ads. In this section, we review existing empirical literature to help assess the four concerns about GAI political ads identified above.

In sum, the research provides five key insights about the potential impact of GAI on political ads:

- First, the persuasive power of both political ads and online misinformation is often overstated. Political ads likely have more of an effect on behavior – such as voter turnout and fundraising – than on persuasion.
- Second, political ads likely have the most impact in smaller, down-ballot races where there is less advertising, oversight, or familiarity with candidates.
- Third, GAI content has the potential to replicate bias, including racial, gender, and national biases.
- Fourth, research on political disclaimers suggests that watermarks and disclaimers are unlikely to significantly curb risks.
- Fifth, significant holes in the research remain. As a result, we caution against relying too heavily on the insights above, and suggest additional research is necessary to review and expand upon these initial assessments.

Harms

SCALE

As discussed above, many commentators are concerned that by lowering the costs and the technical requirements to produce false or deceptive imagery, GAI may permit more bad actors to create a higher volume of problematic content for political ads.

The core concern is that false content will be more impactful if more people see more of it. What does existing research tell us about this claim?

The persuasive effect of both online misinformation and political ads is likely minor. However, political ads can motivate behavior, such as voter turn-out, donation, and volunteering. While increasing the supply of deceptive ads is unlikely to impact national elections, it may have more impact in smaller, more local elections.

Decades of communication and media effects research have shown that media messages are not like "hypodermic needles" or viruses with strong, clear effects on people's beliefs, attitudes, or behaviors. "Information is not passed from brain to brain like a virus is passed from body to body. When humans communicate, they constantly reinterpret the messages they receive, and modify the ones they send."[2] Rather than simply believing what they encounter offline or online, many users approach what they read skeptically and critically, often making sense of that content in unexpected ways.

This is not to say that political ads or online misinformation cannot persuade or motivate voters in certain circumstances. There is evidence of misinformation's persuasive power both in experimental settings and in the wild. Yet, misinformation's power to persuade is hard to assess empirically, and is likely limited and conditional on a wide range of other factors, such as political affiliation and alignment between the falsehood and a person's political beliefs.

2 Altay, S., Berriche, M., & Acerbi, A. (2023). Misinformation on misinformation: Conceptual and methodological challenges. *Social Media+ Society, 9*(1), p. 8.

Falsehoods may also be most impactful when part of cross-media campaigns or influence operations. However, the small amount of misinformation in the wild is often "eclipsed by content from domestic news media and politicians."[3]

Similarly, research suggests that the persuasive effect of political ads is minimal. One review of 49 field experiments in political ads concluded that "the best estimate of the effects of campaign contact and advertising on Americans' candidate choices in general elections is zero."[4]

But the research on the persuasive impact of ads is not definitive. One researcher concluded that "ad effects are conditional, and generally small," and that ads may be more impactful in "ballot measure and down ballot elections, earlier in the campaign, and under conditions where persuadable voters are more numerous and information is low."[5]

Social scientists and practitioners also have more confidence that political ads "have been much more productive at mobilising action, like getting out the vote and improving donation rates, than at persuasive goals of getting someone to support a candidate."[6]

Misinformation may also be more impactful when it is repeated. Repetition has been a key element in misinformation and political campaigns, and it may be impactful. Steve Bannon famously told author Michael Lewis that his approach to dealing with the media is to "flood the zone with shit." The Russian Internet Research Agency used repetition from many different accounts to seed false and inflammatory content.

Experimental research suggests this strategy of repetition works in some circumstances.

It is well established that "[a]dults are more likely to believe information is true if they have repeated exposure to the information, whether the information is true or false" an effect called the "truth by repetition effect," "repetition-induced truth effect," or "illusory truth effect."[7] This study suggested that repetition has an effect not only for neutral content but also for false content, however, some have questioned the validity of these experimental results.

While repetition is a key strategy in information operations, its impact may still be limited. Misinformation remains a minor part of most users' media diets — as little as 0.15% of the typical American's media diet.

A very small number of users are responsible for sharing and viewing the vast majority of false content. In this context, one recent study found that "reshares" on social media "do not detectably affect beliefs or opinions." Another recent article concludes that even if GAI does increase the supply of misinformation, it is unlikely to impact demand and therefore "will likely have limited influence on the diffusion of misinformation."[8]

There is also limited data on the impact of scale in political ads specifically. Campaigns certainly use the tactic, "aiming to have more ads on the air than an opponent, with the hope that such imbalances will leave voters with a more favorable impression of the candidate."[9] This can be difficult amid the massive amount of political ads released during election season. However, this tactic of repetition may be most impactful at the state or local level, where races typically receive less public attention and attract a narrower advertising base. The result may be that one side can dramatically outspend a rival.

---

3 Eady, G., Paskhalis, T., Zilinsky, J., Bonneau, R., Nagler, J., & Tucker, J. A. (2023). Exposure to the Russian Internet Research Agency foreign influence campaign on Twitter in the 2016 US election and its relationship to attitudes and voting behavior. *Nature communications*, *14*(1), p. 62.

4 Kalla, J. L., & Broockman, D. E. (2018). The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. *American Political Science Review, 112*(1), p. 148.

5 Fowler, E. F., Franz, M. M., & Ridout, T. N. (2021). *Political advertising in the United States*. Routledge, p. 117.

6 Baldwin-Philippi, J. (2019). Data campaigning: between empirics and assumptions. *Internet Policy Review*, *8*(4). p. 2.

7 Morgan, J. C., & Cappella, J. N. (2023). The Effect of Repetition on the Perceived Truth of Tobacco-Related Health Misinformation Among US Adults. *Journal of Health Communication*, *28*(3), p. 183.

8 Simon, F, Altay, S., and Mercier, H (2023). Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown. *Harvard Kennedy School (HKS) Misinformation Review*.

9 Fowler, E. F., Franz, M. M., & Ridout, T. N. (2021). *Political advertising in the United States*. Routledge. p. 120.

Authenticity

Beyond increasing the number of ads, many commentators are concerned that GAI will enable bad actors to create more authentic deceptive content in political ads, and that the more authentic deceptive content will have a greater impact on elections. What does existing research tell us about this claim?

GAI may be impactful in enabling advertisers to create more realistic false visual content, as visuals are often seen as more credible and memorable than text.

But research does not yet paint a clear picture of how this extends to deceptive advertisements, since it has focused on less sophisticated content and some studies have shown little difference between the persuasive effect of video and text.

We focus our discussion of authenticity on visual content, since inauthentic text content is primarily addressed in the section on misinformation above. Decades worth of scholarship suggests that audiences process images differently than text. Images are often seen as more credible, more memorable, may provoke more of an emotional response, and have greater "framing effects on opinions and behavioral intentions than text."[10] Rather than focusing on narrow media effects, cultural scholars and media theorists have suggested the wholesale refashioning of American culture around visual media, giving us "more opportunities to live at the surface, continually posing, to see and measure ourselves by the images we make and the images others make of us."[11]

While these studies suggest that images may be more impactful than text in general, there is no current consensus on the persuasiveness of visual misinformation, and the research on the impact of deceptive visual content remains limited. For example, while one study observed that visual disinformation is more credible than text, others have seen small or no meaningful difference between the persuasive effect of video compared to text.

Limitations in the research may be due in part to a lack of sophisticated content to date. Most deceptive visual content to date has been shallow or "cheap fakes," content doctored using simple techniques and widely available tools, or content with misleading labels that suggest a visual is something it is not.

Beyond direct effects, there is some evidence that visual disinformation can have broader impacts on political life. There is some empirical support for Chesney and Citron's argument that misinformation may present a "liar's dividend," where people can claim that real video evidence is false.

This tactic speaks to the more general and persistent decline in trust across most institutions. Relatedly, several studies have shown that even if people do not believe deepfakes or visual disinformation, their existence induces more general uncertainty, reducing trust in media or the political process.

That is, widespread public discussion of deepfakes "increases disbelief in both real and manipulated videos without improving participants' ability to successfully identify deepfakes."[12]

Personalization

Commentators have expressed concerns that GAI will allow advertisers to create personalized content that will impact elections. What does existing research tell us about this claim?

The research does not provide clear guidance about the effect of additional personalization on political ads. Given these gaps, it is difficult to assess the impact of GAI personalization on political ads.

Digital advertising is built on targeting. The ability to target ads to specific audience segments is the foundation of the more than $500 billion digital ad industry. Many researchers, major companies, and startups are working to understand and improve tools and models for better audience segmentation.

10 Powell, T. E., Boomgaarden, H. G., De Swert, K., & De Vreese, C. H. (2015). A clearer picture: The contribution of visuals and text to framing effects. *Journal of communication, 65*(6), p. 997.

11 Adatto, K. (2008). *Picture Perfect: Life in the Age of the Photo Op*. Princeton University Press. p. 7.

12 Ternovski, J., Kalla, J., & Aronow, P. (2022). The negative consequences of informing voters about deepfakes: Evidence from two survey experiments. *Journal of Online Trust and Safety, 1*(2), p. 3.

Yet some studies have introduced skepticism about the impact of personalization, claiming that there are profound errors in targeting data sets and significant quantities of fraud. In recent years, a handful of large advertisers, including eBay and Procter & Gamble, have reduced digital advertising spending, and reported no negative impact on sales. One recent book by a former Google employee alleges that the systems underwriting the effectiveness of digital advertising are flawed. He claims that the "data used in targeting ads," the "algorithm being used to deliver advertising" and "the 'attention' that ads do receive" are all "garbage."[13]

Personalization and targeting have become essential in digital political advertising. However, there is some indication that more complex targeting models add little to more basic models built on the voter file. More generally, "the evidence on the effectiveness of political advertising personalized to target, for instance, people with different personalities is mixed, with at best small and context-dependent effects."[14]

Beyond political advertising, researchers have been examining the impact of personalization in social media news feeds.

Over the past decade many researchers have claimed that personalization in social media search and news feeds worsens ideological and affective polarization. However, recent experimental evidence suggests that algorithmic curation does not result in greater affective or ideological polarization, compared to the less personalized option of a reverse chronological feed.

Taken together, this research suggests that claims about the impact of personalization of ads may be overstated. However, the research also emphasizes that these conclusions merit further study. Unfortunately, researchers have encountered challenges in studying this area, since digital advertising is opaque, and researchers often struggle to get the data they need to make these assessments.

Bias

Finally, there is a risk that GAI may reproduce and amplify existing biases in political ads. What does the existing research tell us about this claim?

Research suggests that AI models display bias. This finding extends to GAI models, which is a good reason to be concerned about GAI models reproducing bias in political ads.

Decades worth of research has shown that some algorithms and AI models display different forms of bias. Thanks to work like Safiya Noble's *Algorithms of Oppression*, we understand better how algorithms often reinforce racism and sexism.

The Department of Commerce's National Institute of Standards and Technology (NIST) identifies a range of technical means through which bias is introduced into algorithms and AI models, including selection and sampling of data, data processing and validation, and model use and interpretation. But the bias in AI models also derives from systemic social and human bias as well.

While researchers have been examining LLMs and NLP models for several years, ChatGPT's release last November led to a huge explosion in both research and reporting on bias in LLMs. Many have found that LLMs, including both text-based and image models, reinforce racial and gender biases.

For example, one analysis by Bloomberg found that when asked to generate images of professionals, Stable Diffusion disproportionately used images of subjects presenting as white men to represent traditionally higher paying and higher status jobs, such as "politician," "lawyer," or "CEO." Similarly, a recent analysis by Rest of World found that leading text-to-image systems often reinforced national stereotypes. For example "an Indian person" tended to return "an old man with a beard" while querying "a Mexican person" usually returned an image of a man wearing a sombrero.

13 Hwang, T. (2020). *Subprime attention crisis: Advertising and the time bomb at the heart of the Internet*. Farrar, Struass and Giroux. p. 7, p. 53.

14 Simon, F, Altay, S., and Mercier, H (2023). Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown. *Harvard Kennedy School (HKS) Misinformation Review*.

Many researchers have attempted to assess political bias in LLMs. One study of ChatGPT concluded that the model "presents a significant and systematic political bias toward the Democrats in the US, Lula in Brazil, and the Labour Party in the UK." Another study, which assessed more than a dozen models, also concluded OpenAI's models were more left and libertarian leaning (on a four quadrant, left-right and authoritarian-liberal graph). It saw other models exhibiting a broad range of political orientations and biases.

Interventions

Policymakers have introduced proposals that rely heavily on watermarks and disclaimers to mitigate concerns about the use of GAI in political ads. What does existing research tell us about the likely efficacy of these interventions?

There is little public scholarship that speaks to the effectiveness of the interventions being currently discussed by federal and state policymakers, such as watermarks and disclaimers. Research to date suggests that neither intervention is likely to be a silver bullet.

For example, even though policymakers have suggested watermarks as a tool to address potential harm associated with GAI, we do not know if audiences would observe watermarks, understand what they are, or alter how they view GAI content as a result. There has been some recent anecdotal evidence that watermarks can themselves be manipulated.

However, there is a small amount of existing empirical work on the effectiveness of disclaimers in political advertising. There is also research suggesting that disclaimers have the opposite effect.

There is some indication that even if users recognize a disclaimer, it does not necessarily mean they understand they are viewing an ad, and "only in certain conditions did it help them recall the source of the ad correctly."[15] In the context of political ads, one study found that including a disclaimer at the start of an ad made audiences think better of the source of an ad, rather than worse.

There is also mixed evidence about the ability of political disclosures to increase "persuasion knowledge": audiences' understanding that they are viewing an advertisement that is trying to persuade them. More generally, persuasion knowledge can help minimize persuasive effects on audiences, and result in less favorable evaluations of an ad and source. However, that effect may be contingent on source and design.

Ultimately, research suggests that these interventions may have minimal impact in mitigating the potential harms of GAI in political ads.

15 Binford, M. T., Wojdynski, B. W., Lee, Y.-I., Sun, S., & Briscoe, A. (2021). Invisible transparency: Visual attention to disclosures and source recognition in Facebook political advertising. *Journal of Information Technology & Politics, 18*(1), p. 80.

Recommendations

Based on the concerns raised about the use of GAI in political ads and existing empirical research on these concerns, we offer ten recommendations. Our recommendations are designed to incorporate insights from relevant literature and are centered on two main ideas.

First, we suggest that public policy should target electoral harms, not technologies. Several of the recommendations proposed by federal and state lawmakers single out GAI tools for regulation, even though the alleged harm is not specific to GAI. Should we care about deceptive political advertising only when it is created using GAI technologies? What is the rationale for labeling false speech that is produced by GAI, but not false speech produced by a human? In our view, public policy should focus on identifying and mitigating harms, independent of what tool or technology produced them.

In our reading of the research, we identified two types of harm that public policy should prioritize addressing. One is the potential for GAI misinformation to impact smaller, down-ballot races. The second is the potential for GAI to reinforce biases. Several of our recommendations are targeted at these types of harm.

Second, we suggest that public policy should promote learning about GAI. Learning will help us to develop a better understanding of the potential benefits and harms of GAI technologies in the context of political ads, and they will enable us to use this understanding to inform future public policy. Since our review of the literature revealed several gaps, we emphasize the value of additional research to fill those holes. Several of our recommendations focus on transparency; building a strong governance regime is impossible if it is constructed on top of a weak data foundation.

This approach is consistent with our prior work on experimentation and "regulatory curiosity," the idea that regulators should institute policy mechanisms to gather data and learn, which will in turn help them to develop smarter governance regimes in the future. There is much we don't know, and we should use public policy to inform us.

For any regulations related to speech, policymakers should be cognizant of the constraints imposed by the First Amendment. These constraints are particularly robust when a regulation is content-based, such as singling out deceptive GAI content. Policymakers restricting election speech must ensure a strong nexus between the interest they are promoting and the specific elements of the restriction. One benefit of our second category of recommendations — policy measures focused on learning — is that they may produce information that could help policymakers to establish this nexus.

As with any policy recommendations, we believe that our suggestions here should be rigorously tested, such as by conducting cost-benefit analysis to evaluate their potential impact. Where the costs of a proposed reform exceed its benefits, we should be skeptical of the value of that change. Where the benefits exceed the costs, our bias should be in favor of altering the status quo. Consistent with this experimental approach to public policy, several of our recommendations are aimed at producing information, so as to enable more accurate assessments of impacts over time.

Public policy should target electoral harms, not technologies.

1. Congress and states should outlaw voter suppression.

As we have argued previously, we believe the federal government should pass new legislation outlawing voter suppression. This law should prohibit all forms of falsehoods about voting locations, places, or dates, and should not exclude GAI content from the prohibition. Several proposals for this type of law have already been introduced in Congress.

States should also outlaw voter suppression. Some states currently prohibit voter suppression, but many do not. These states should pass laws prohibiting falsehoods about voting locations, places, and dates. GAI content should not be excluded from these prohibitions.

2. Governments should allocate additional funding for law enforcement to enforce existing civil rights laws that protect the electoral process.

Existing federal civil rights law prohibits the use of misinformation to deprive people of their right to vote. It also protects against some conduct that exacerbates historic bias and discrimination.

Law enforcement officials should have the resources necessary to enforce these laws. To ensure that this existing body of law is adequately enforced to protect against bias in the electoral process – including the use of GAI to perpetuate or exacerbate that bias – Congress and the states should earmark additional funding for the Department of Justice and state Attorneys General to enforce existing civil rights protections related to the electoral process. This funding should include resources to help law enforcement identify problematic content, such as educational training on how advertisers could use GAI to deprive voters of their civil rights.

3. Local and state governments should "flood the zone" with factual content.

Research suggests that inauthentic content may be more impactful at the state and local level, in part because those races often generate less online speech and receive less press attention. To combat the risk that GAI's effect on scale, authenticity, personalization, and bias will harm state and local elections, state and local governments should generate a large amount of factual content to combat potential misinformation. Specifically, states should earmark increased funding for election boards to produce and distribute factual non-partisan information about voting and about candidates. They should run ads and also distribute content organically.

4. Governments should fund digital literacy programs focused on detecting and contextualizing false online content.

Generations of potential voters will be exposed to false online content, so citizens must have some baseline understanding of how to assess veracity, including the veracity of GAI content.

Although paeans to digital literacy are sometimes disregarded as either woefully insufficient or as thinly veiled efforts to block meaningful regulatory reform, those critiques are not a reason to shy away from ensuring that students receive training in smart, ethical, safe online conduct.

Congress and state legislatures should make more funding available for developing and instituting curricula on digital literacy that help students identify and contextualize false online content. These curricula should include tools for assessing GAI content. Policymakers should also support academic research into the efficacy of different digital literacy curricula, and integrate the insights from this research into the programs that schools implement.

In addition, as noted above, governments should also fund training for civil rights law enforcers so that they can understand how to identify GAI-related harms.

5. The FEC should publish guidance for political advertisers on identifying and mitigating bias in political ads, with special emphasis on bias introduced by GAI models.

Limiting bias from GAI in political ads will require concerted effort from platforms, advertising firms, and advertisers. The FEC should work with digital literacy and digital equity organizations to develop resources for advertisers on identifying and mitigating bias in political ads. This guidance should build off NIST's "AI Risk Management Framework" and any standards or best practices developed by NIST, the Department of Commerce, and other agencies in the wake of the recent executive order on AI. Over time, this guidance could help to inform platform policies as well.

Public policy should promote learning about GAI.

6. Governments should fund empirical studies of impacts of GAI in political ads and the effectiveness of GAI-related interventions.

Despite growing concern, there has been very little research that empirically assesses the impact that GAI content in political ads has on persuasion, behavior change, extremism, trust, and expression. Similarly, there has been little research on the effectiveness of interventions like disclaimers or watermarks for GAI. While the Biden administration's recent executive order on AI provides some new resources for AI research, it would not provide research funding on the harms of GAI or the effectiveness of interventions. To better translate between the policy world and the academy, additional research should assess the following questions:

- Scale: What is the impact of repeating false statements in advertisements on persuasion, mobilization, trust, and extremism? Are those impacts greatest in smaller, down-ballot races?
- Authenticity: Does deceptive visual imagery affect voters and if so, how?
- Personalization: What is the effectiveness of personalization of political advertising on candidate choice, turnout, and donations?
- Bias: How does GAI text or visuals in political ads reinforce racial, gender, political, or other forms of bias? How does the presence of those biases impact voters?
- Interventions: Do watermarking or disclaimers alter a user's understanding of the accuracy of content? Do they impact user behavior? Do they pose feasibility challenges for platforms that hinder competition by erecting barriers that disproportionately burden smaller companies?

7. Governments should establish policy experiments to test interventions aimed at mitigating the negative impacts of GAI in political ads.

We recommend the government use time-limited experiments to test the effectiveness of watermarking, disclosures, and other interventions. Audit committees should review the performance of those experiments, including their impact on bias and on smaller, down ballot races. Experiments could enable governments to trial new regulations and platforms to test new products and features.

8. Congress should learn about the impact of different GAI content liability regimes and incorporate these learnings into the law.

The use of GAI in online communication tools has raised questions about who bears liability for GAI-created speech. While Section 230 has helped to protect platforms from liability for speech by their users since it was passed in 1996, the law may not protect platforms that use GAI. There is also significant uncertainty about the legality of training models that ingest copyrighted work without permission.

Congress should clarify online speech liability in an era of GAI. Before it institutes a permanent change in the law, it should learn more about the impact of potential liability regimes. For instance, it could use policy experiments to test different liability regimes and assess their costs and benefits.

The objective should be to avoid instituting rules that constrain speech simply because it employs GAI, but rather to enable platforms to offer features that utilize GAI to facilitate speech, while also constraining harmful speech and conduct.

9. Federal and state governments should establish political ad archives and require advertisers to submit copies of ads along with regular disclosure reports.

As we argued in previous work, we believe that a public government archive of political ads would further transparency, allowing journalists and researchers the ability to track and analyze ad content, and would provide a more comprehensive picture of the political ad ecosystem than relying on a patchwork of archives from individual companies.

10. The FEC should require that campaigns report their vendors' advertising spending.

Currently, when committees report spending to the FEC, they are required to report only how much they pay to ad consultancies, not how those consultancies disburse money to ad platforms. In a previous report, we estimated that as much as 94% of ad spending goes through advertising consultancies.

This gap in public reporting will make it harder to track the impact of GAI on political advertising, since it will be difficult to assess how ad spending shifts as platforms introduce GAI integrations. To evaluate the impact of GAI in political ads, the FEC should close this loophole.

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL