Preparing for Generative AI in the 2024 Election:
Recommendations and Best Practices Based on Academic Research

This white paper is the result of a partnership between the University of Chicago Harris School of Public Policy and the Stanford Graduate School of Business intended to generate non-partisan, expert recommendations on various critical AI governance issues. In late August 2023, the two institutions hosted a convening of experts from academia, technology, industry, and civil society who explored the challenges of generative artificial intelligence for the 2024 U.S. Election and opportunities for establishing best practices and governance guidelines.

An initiative of:

THE UNIVERSITY OF CHICAGO
HARRIS SCHOOL
OF PUBLIC POLICY

STANFORD
BUSINESS
GRADUATE SCHOOL OF

Preparing for Generative AI in the 2024 Election:
Recommendations and Best Practices Based on Academic Research

Ethan Bueno de Mesquita*, Brandice Canes-Wrone[†], Andrew B. Hall[†], Kristian Lum*,
Gregory J. Martin[†], and Yamil Ricardo Velez[‡]

Executive Summary

The rapid development of generative AI technology is transforming the political landscape, presenting both challenges and opportunities for the 2024 US election. This document provides a research-based overview of the potential impact of generative AI and offers best practices to safeguard the electoral process. The purpose of our review is not to recommend many specific legal or policy actions, but rather to promote clear understanding among voters, journalists, civil society, tech leaders, and other stakeholders about both the risk and promise of AI for electoral democracy in the hope of fostering a more productive public discussion of these issues.

Challenges and Opportunities

1. Degrading Information Environment: As has been well documented, generative AI can create highly convincing deepfakes, posing risks of deceptive content, especially when released close to election day. There is also concern that voters will seek factual election information from chatbots that are not reliable sources of such information.
2. Manipulation and Microtargeting: There is concern that AI-powered micro-targeting or emotionally manipulative chatbots will persuade voters to behave contrary to their interest or polarize the electorate. However, social scientific evidence suggests mass persuasion or manipulation of voters is unlikely and that the greater concern is the perception of this manipulation, rather than an actual direct effect on electoral outcomes.
3. Positive Uses: Generative AI technologies offer opportunities for positive applications in politics, such as generating accessible summaries of policies, helping voters assess candidates, aiding citizen-to-lawmaker communication, and leveling the playing field for under-resourced campaigns.
4. Information Centralization: To the extent that generative AI tools concentrate in a small number of tech firms, there will be concerns about the power they wield over political information. This will create thorny issues of content moderation, bias, and distrust.

* University of Chicago
[†] Stanford University
[‡] Columbia University

Best Practices
- For tech companies:
  - Continue building on recent efforts to watermark and label AI-generated content while clearly communicating that these efforts are not a panacea.
  - Plan table-top exercises to anticipate best responses to late-breaking "October surprises" that rely on AI-generated content.
  - Study how labeling AI-generated material affects users' understanding of the information environment, building on recent academic research in this area.
  - Ensure that AI chatbots point users toward authoritative information from official state sources regarding other narrowly defined issues of clear fact, such as how and where to vote. Monitor uses of chatbots for microtargeting and misleading content intended to deter participation—but primary focus should remain on deepfakes and the information environment.
- For journalists:
  - Disincentivize misinformation and manipulation by avoiding covering stories whose only case for newsworthiness is the use of AI-generated content.
  - Aid with AI-literacy through non-sensationalist coverage of AI's role in elections.
  - Partner with civil society to evaluate the provenance of widely distributed political content, and participate in tabletop exercises to consider how to cover and rapidly evaluate October surprises.
- For political actors:
  - Political parties, politicians, and campaigns should publicly pledge not to use deceptive AI-generated content.
  - Political actors should consider partnering with tech in building AI-based tools that empower voters to become better informed about candidates and policy.
- For voters:
  - Voters should remain skeptical of sensational or scandalous political information, especially when released close to election day.
- For the future:
  - Both government and big tech companies should limit centralization of power over content generation and moderation in a small number of platforms.
  - Encourage a diversity of generative AI tools and involve users and external stakeholders in setting content guardrails and governance structures.
  - The political risk associated with power centralization should be among the considerations when designing AI safety regulations that might increase market concentration, such as licensing requirements.

Introduction

The rapid development of generative AI technology over the past year has created new challenges as well as some new opportunities for democracy. This document provides a research-based overview of the potential impact of generative AI and offers best practices to safeguard the electoral process. The purpose of our review is not to recommend many specific legal or policy actions, but rather to promote clear understanding among voters, journalists, civil society, tech leaders, and other stakeholders about both the risk and promise of AI for electoral democracy in the hope of fostering a more productive public discussion of these issues.

Democracy relies on electoral accountability. Voters are asked to elect candidates they support and to fire politicians they think are not doing a good job. Democracy therefore requires a healthy information environment in which voters can monitor what politicians are doing, learn what candidates promise to do if elected, and assess what policies might be needed in response to societal challenges.

By facilitating the automated creation of highly detailed text, image, and videos that are difficult to distinguish from human-generated content, generative AI could negatively impact the information environment. It could be used to create false or misleading information, feed a general sense of radical skepticism or nihilism about the reliability of information, and even allow political actors to dismiss real, damaging information as AI-generated fake news.

By assisting in the creation of micro-targeted political content (for example, for use in online political ads), generative AI could polarize voters or contribute to a sense of a fractured polity in which people live in different informational environments from one another or are manipulated into voting for the wrong candidate or into not voting at all.

Generative AI also powers chatbots capable of engaging in human-like conversation. Startups, campaigns, and other actors are likely to use these chatbots to engage in political conversations. This may not be inherently problematic—indeed, one could imagine increased engagement leading to better-informed voters. But some may object to the inauthentic nature of these conversations, the potential for emotional manipulation, or the intentional or unintentional spread of false or misleading information. Moreover, extremist groups might use these chatbots in an effort to recruit people to their causes.

Finally, and more generally, the growth of generative AI could lead to further centralization of the online information ecosystem, with a small number of large tech companies deciding which ideas and values can be expressed and which are out of bounds. This could threaten free expression and further erode Americans' trust in the political environment.

As an extraordinarily promising new technology, though, generative AI also holds promise for democratic governance. Researchers are already studying ways to create chatbots that can discuss politics with people. Because generative AI is good at synthesizing large amounts of information, and if

conversing through chatbots proves to be a more natural and engaging way for people to consume political information, then training chatbots to understand political issues and the positions of parties and candidates may be an effective way to help Americans become better informed about politics. These tools could also help under-resourced campaigns to communicate more effectively with voters, and could help voters to communicate with their representatives more easily. While it's essential that we safeguard the 2024 election and the democratic process more generally in the face of disruptive new technologies, we should also seek to maximize these opportunities for generative AI to improve our democracy.

Voters' Information Environment

Generative AI has the capacity to alter the voters' information environment in at least two ways. First, generative AI allows for the creation of highly convincing deepfakes-—images and videos that are difficult for non-specialists to distinguish from genuine content. Second, AI chatbots are a new, direct source of information for certain voters, especially younger ones.

Deepfakes

Unscrupulous actors are likely to disseminate deepfakes, perhaps experimentally tested and tuned for maximal impact, unconstrained by norms of truth telling or public accountability. Early instances of these concerns have already arisen in real world elections. Deepfakes played a role in the recent Turkish elections. At a large political rally, President Recep Tayyip Erdoğan [showed a fake video](#) linking his chief opponent, Kemal Kılıçdaroğlu, to the leader of the PKK, a Kurdish group classified by the State Department's as a foreign terrorist organization. Separately, an online Kılıçdaroğlua supporter used AI to generate a video that appeared to show the [candidate delivering a campaign speech in perfect English](#). And here in America, the Republican National Committee [released a campaign video](#) that used AI-generated images of President Biden and Vice President Harris to show a dystopian future following a Biden-Harris victory in 2024. This video was not deceptive—it seems unlikely any viewers thought it showed current reality—but it did use AI to try to manipulate voters' perceptions of political opponents and to garner press attention and notoriety.

There are several reasons to be concerned about the use of deepfakes in campaigns. First, of course, voters may well be deceived by such content. It is particularly concerning that AI-manufactured content could be released very close to election day in order to generate fake scandals within a time frame that makes fact checking difficult. These "October surprises" may be especially difficult to respond to if they are generated or shared by major political candidates. Tech companies, civil society, and journalists may want to consider plans to rapidly analyze these October surprise deepfakes as they arise, focusing only on the most salient ones that gain the most reach online. Second, wide-spread circulation of manufactured content may undermine voters' trust in the broader information environment. If voters come to believe that they cannot trust any digital evidence, it becomes difficult to seriously evaluate those who seek to represent them. Third, politicians may use this undermining of the credibility of the information environment to dismiss genuine information. Late in the Turkish election, a tape came to

light showing compromising images of a candidate, Muharrem İnce. While İnce eventually withdrew, he also [claimed the video was a deepfake](#). If voters genuinely can't tell the difference between what is fake and what is real, it is not hard to imagine that such denials will become a commonplace.

Much of the public discussion around how to address such concerns is focused on establishing content provenance. The most prominent approach involves watermarking and labeling, something that major AI companies including Meta, Google, and OpenAI [have publicly pledged to implement](#). TikTok has [announced](#) a similar approach. The idea is that the large AI companies should build-in code that "watermarks" content–making content produced or modified by generative AI detectable–and that media and social media outlets should then label such content as "AI generated". [Other approaches](#) involve building technologies that make it possible for creators to cryptographically sign content in ways that are not preserved if the content is altered without permission, thereby allowing users to attribute responsibility for content to a particular creator. The goal of all of these steps is to allow voters to make better inferences about the reliability of information.

White watermarking, labeling, and signing are certainly worthwhile, they are not a silver bullet due to several practical constraints. Only a small percentage of online content can be expected to be cryptographically signed. And, while AI companies are exploring a variety of watermarking technologies, all current approaches can be worked around by sophisticated actors. [For generative text models, shorter texts are harder to detect. Even advanced cryptographic techniques for watermarking require "a few hundred" tokens for reasonable detection, which is well over the length of a typical Tweet.](#) Moreover, because of the many approaches being pursued, the industry has not agreed on a set of standards; this lack of harmonization makes labeling more difficult for media and social media companies.

Finally, labeling content as AI generated requires making judgment about substantial gray areas. In addition to being used to create deceptive or false content, generative AI tools are used for a host of innocuous tasks, such as red-eye reduction, editing for fluency, and the like. Watermarking and labeling any content that has been touched by an AI risks being so over-inclusive that labels become meaningless. But the alternative requires laborious and contentious judgment calls. Moreover, the effects of labeling on human behavior and perception are not yet well understood. It is possible that users will respond to seeing labels alerting them that content is "AI generated" by assuming it is untrue. Such a response would become problematic if there is useful, informative AI-generated content that gets systematically dismissed, or if users fallaciously conclude that any content *not* AI-generated must therefore be true. [Recent research](#) offers helpful evidence on what labels might be most effective, and further research by academics and platforms could help to refine best practices for labeling.

[Without effective methods to establish the veracity of digital content, an alternative path forward is improved verification for content providers](#). Even if we can't know whether digital content has been altered, account verification and verifiable content signing can increase confidence that the content that is released by an entity is truly coming from them, not an imposter. If content is determined to be misleading or false, being able to confidently trace the content to its origin is a useful step towards

creating accountability. Recent events have shown us the impacts of misleading account verification. In the midst of recent changes to its verification program, X (formerly Twitter) users were able to create accounts that appeared to be verified and official but were, in fact, imposters. One user created a [fake but official-looking Eli Lilly account](#) that tweeted "We are excited to announce insulin is now free." In the wake of this tweet, Eli Lilly's stock fell over 4%.

Voter Information

Another concern in the information environment is that voters might seek technical facts about elections—for instance, voting dates, polling locations, registration procedures, voting eligibility—from AI chatbots. [Recent polling](#) suggests that seeking such factual information is the election-related activity voters are most likely to use generative AI for. But [research by the bipartisan policy center finds that chatbots are neither a reliable nor an appropriate tool for this use case](#). In recent years, [internet search has been the most common source of such information for voters](#). To the extent that AI chatbots are viewed as substitutes for search, this could be a substantial concern.

Chatbots have not, at this point, displaced search as a source of information. [Only about one quarter of adult Americans had accessed ChatGPT by July of this year](#). However, that number is substantially higher among those of young voting age—over 40% of Americans age 18-29 have done so. The two most common uses were for entertainment and to access new information. Thus, while it is unlikely that chatbots will be a source of widespread misinformation on technical election information, it is important that the creators of such chatbots take steps to reduce the risk. The most straightforward approach is to train generative AIs to inform users that the chatbot is not a reliable source of technical election information and to direct users seeking such information to authoritative sources, such as the websites of state or local election administrators. Several of the large AI companies are already pursuing such an approach.

Recommendations and Best Practices
- **Campaigns and political parties** should publicly pledge not to use deceptive AI generated content in campaign materials.
- **Tech companies** should continue to build on their recent efforts to collaborate around watermarking and labeling AI-generated content, and support further research on the effects of labeling this content on users' overall understanding of the information environment.
- **Tech companies** should ensure that chatbots inform users that they are not a reliable source of technical election information and point users toward authoritative sources, as they have done in past election cycles.
- **Tech companies, journalists, and civil society** should carry out tabletop exercises exploring how to respond to the release of late-breaking deepfake content, especially under scenarios in which the content is generated or promoted by a major political candidate.
- **State election boards** should emphasize that existing voter intimidation and deception laws apply to AI generated content an outside group or campaign may use; the fact that the content was generated by AI is not a defense for voter intimidation or deception.

- **Journalists** should disincentivize misinformation and manipulation by avoiding covering stories whose only case for newsworthiness is the use of AI-generated content.
- **Journalists and Civil Society** should collaborate to provide trustworthy information on content authenticity.

Microtargeting and Manipulation

In an article published in The Conversation in June, political scientist Archon Fung and law professor Lawrence Lessig describe a dystopian scenario in which political campaigns use generative AI to develop highly persuasive messages tailored to individual voters. AI tools use reinforcement learning to get better and better at changing minds, far surpassing the capability of human campaign strategists. Fong and Lessig warn that well-resourced candidates with access to such technology could win elections despite holding unpopular policy views, undermining electoral accountability.

Although the specific AI-fueled context in this scenario is new, the general fear of candidates using persuasive technology to convince voters to vote against their preferences is not. Similar arguments were made about targeted digital advertising when it became widespread in campaigns in the late 2000s through mid 2010s, and about broadcast advertising on television and radio when those technologies were young.

Political scientists have evaluated the plausibility of similar claims to Fung and Lessig's in the context of existing persuasive technologies – direct mail, television, and non-AI-generated digital ads – and have generally found persuasive effects close to zero. While there is some evidence that television advertising can have small but nonzero effects on vote shares in general elections, this effect is believed to operate mainly through changing the composition of the electorate (by motivating partisans to turn out or not), not by persuading voters to change their minds on policy issues. Two further checks on the potential power of any communication technology, including AI, come from the competitive nature of elections–which implies that any persuasive arguments developed by one campaign are likely to be countered by its opponent–and the fact that any messaging voters receive from campaigns is swamped in volume by information from the mass media, social media, and friends and family.

To the extent that generative AI-produced ads are different from human-produced ads, the difference lies in the potential for mass customization on a scale not feasible with human-produced content. Existing technologies do allow for some customization, so this difference is one of degree rather than of fundamental character. Evidence about candidates' use of existing targeting and message-tailoring technologies is, therefore, informative about how candidates would be likely to use generative AI. Interestingly, candidates do not appear to take much advantage of the large difference in targeting precision between broadcast TV ads and digital ads on social media platforms like Facebook; candidates run very similar campaign messages in both settings and do not present significantly different versions of their policy positions to different groups of platform users. It does not appear that candidates perceive large gains from targeting, perhaps because of the potential for news outlets or their

opponents to expose and broadcast any narrowly-targeted messages outside of their intended audiences.

We are, accordingly, skeptical of the potential for generative AI to undermine political accountability by duping voters into supporting candidates whose policy agendas they oppose. While it is possible that AI-enabled campaign advertising is substantially more effective than other existing campaign advertising technology, our accumulated experience and knowledge of previous waves of technological innovation in campaigns suggests that this is unlikely. Nonetheless, there are some potential consequences about which some concern is warranted.

First, one reason that campaigns find it difficult to persuade voters is that their opponents and the news media provide a counterweight to the information and arguments they offer. Were either the competitiveness of elections or the quality of political information provided by news media to decline, the scope for campaign persuasion by any means, and in particular through AI-generated targeted messages, would increase.

Second, generative AI is costly to use. If it came to be perceived by candidates as a necessary tool to mount a serious campaign, it could raise the financial cost required to enter a race. The arms-race effect of such competitive pressure could impact the pool of candidates willing to run, even in the absence of any net effect on voter behavior. There is now good evidence from [multiple](#) [settings](#) that increases in campaign spending levels tend to shrink the candidate pool, advantage wealthier candidates, and increase incumbency advantages.

Third, AI-enabled campaigning might be perceived by voters as unfair or illegitimate, particularly if its adoption is not symmetric across the parties. News media might contribute to such perceptions by uncritically reporting claims about AI campaigns' effectiveness from self-interested technology vendors. Losing candidates could use such claims, were they widely believed, to cast doubt on the legitimacy of their opponent's victory.

Recommendations and Best Practices

- **Journalists** should inform the public that the likelihood of generative AI ads or chatbot conversations massively affecting electoral outcomes is low.
- **Journalists, civil society, and voters** should exercise extreme skepticism when startups or other companies claim to be able to use new technology to persuade large numbers of voters to switch their votes or to not turn out.
- **Tech companies** should continue to monitor concerning uses of chatbots for microtargeting, persuasion, and particularly misleading content intended to deter participation. Where appropriate, they should consider forbidding such behavior in their terms of services as some already do. However, in the short amount of time before the 2024 election, issues with deepfakes and the information environment are likely to be more pressing and should be the primary focus.

- **Campaigns and political parties** should avoid allocating undue resources to microtargeting and chatbots, which may foster mistrust even if they are not effective at political persuasion.

Opportunities

Though some of the most salient examples of generative AI being used in politics involve nefarious applications such as deepfakes and inauthentic news stories, these technologies also have the potential to be harnessed for civic engagement and education. Positive use cases include summarizing policies and salient political events for lay audiences, developing more accessible voting advice applications, and facilitating communication between citizens and lawmakers. Beyond the potential to empower voters, generative AI could also enable under-resourced campaigns to improve their outreach and communication strategies.

Distilling policies and presenting them in an accessible format is a capability of large language models (LLMs) that could enhance voter learning. For example, one can ask Anthropic's Claude, an LLM with a 100,000 token context window, to generate summaries of House and Senate resolutions. Doing so for H.R. 5745, a 57-page bill introduced in the 118th Congress aimed at regulating digital assets, produces a summary correctly highlighting its aim to "create new regulations for cryptocurrencies, stable coins, and other digital money." If one continues the chat by claiming to be a "cryptocurrency expert," Claude can return details about new reporting requirements introduced in the bill. The ability for these tools to tailor information based on user knowledge provides an opportunity for citizens to engage with the legislative process in a more transparent and efficient way.

LLMs could also enhance voter learning in electoral settings. Voting advice applications (VAAs) have become ubiquitous in the US and Europe. VAAs solicit voters' stances across several issues and match voters to politically congruent candidates. Though previous research has found positive effects on voter turnout and vote choice, VAAs assume that voters possess stable issue preferences, and might appeal to those who are already politically engaged. Integrating generative AI into the development of VAAs could allow lower propensity voters to pose simple questions about political parties (e.g., "where does the Republican Party stand on abortion?") and receive information about party positioning within seconds. Instead of responding to an extensive list of policy items, as is the case with existing VAAs, AI-guided conversations could help make interacting with VAAs less intimidating and more accessible to a broader spectrum of voters.

In implementing tools such as VAAs or even summarizing policies, a natural question is whether LLMs will be "biased" toward certain viewpoints. Methods like retrieval-augmented generation (RAG) can be used to mitigate these concerns by grounding the generation of content in a variety of sources that span the political spectrum. When a user interacts with an LLM, RAG ensures that relevant sources are fetched from a preloaded database, inserted into the prompt, and used to populate the answer. It is possible to load an external database with sources such as party platforms and instruct the LLM to only summarize the information contained in those sources. This produces outputs that are generally free of editorializing and accurately represent the source material.

Outside of electoral settings, extensions of RAG could be used to help voters learn about topics discussed in city council meetings, scour publicly available but underutilized civic data, and understand local policies. Using RAG is akin to using a search engine that retrieves relevant information and summarizes it in an accessible fashion. This makes it an ideal method for simplifying vast amounts of data, which can be useful for voters making choices in political contexts where knowledge is limited due to a lack of media coverage. For example, one could use RAG to summarize recent approvals of housing developments or a list of policies considered by a city council in the previous week.

Beyond political learning, generative AI could also be used to facilitate communication between citizens and elected officials. Writing to lawmakers can be an intimidating process for many citizens, as they might know how to articulate their concerns or they may be unsure about the relevant norms for such communication. LLMs can improve this process by enabling citizens to craft more persuasive messages. They can also provide real-time assistance with respect to grammar, style, and tone, allowing citizens to communicate more effectively. In the future, these tools may also be used by legislators to proactively identify common concerns among citizens, further improving responsiveness.

Generative AI could also help level the playing field with respect to political campaigns. Under-resourced campaigns may face challenges in creating content that appeals to voters. Generative image and text tools can enable these campaigns to draft more compelling speeches, press releases, social media posts, and other materials. These methods can also be used to create tailored materials for different audiences. To the extent that Generative AI allows financially constrained campaigns to maintain a veneer of professionalism, it could reduce imbalances between lesser-known candidates and more established politicians.

In sum, although negative use cases involving generative AI have received significant media attention, there are positive use cases that could enhance civic engagement and education. From powering political information-retrieving chat bots to improving citizen-to-legislator communication, the summarization and style transfer capabilities of LLMs could be used to improve democratic processes. Encouraging the public to use these tools to productively cut through the morass of everyday politics can serve as a powerful counterweight to misinformation, whether human or AI-generated.

Recommendations and Best Practices
- Encourage campaigns, journalistic outlets, and technology companies to create pathways that help voters to use Generative AI to summarize complex policies and political platforms.
- Political parties, technology companies, and civic organizations should help under-resourced campaigns learn to use AI tools to generate higher quality and more informative content.
- Lawmakers and other actors should create pathways for citizens to use AI tools to communicate directly with their offices.

Information Centralization

There are [good reasons](#) to believe that the AI market is likely to end up concentrated in the hands of a small number of dominant firms who build and deploy foundation models—the large, flexible models that can be used for a wide variety of applications. To the extent that generative AI tools end up clustered in a small number of large tech platforms, it may further exacerbate concerns that tech companies have too much power over the information environment.

Already, OpenAI, Microsoft, Meta, and other tech companies offering AI-related services have faced concerns over what content their chatbots and other generative AI tools will and won't countenance, with some critics worrying that the tools will say offensive things, while others worry that the generative AIs exhibit a left-wing political bias. More generally, there are hard questions about (a) what content is allowed to enter a mode's training algorithms, (b) what prompts users are permitted to ask and which are rejected for being offensive, dangerous, or otherwise off limits, and (c) what output the tools will and won't provide.

As generative AI use expands, the power these companies have over these rules could become problematic in much the same way that the power of social media companies over online speech has over the past 10 years.

There are two main risks. First, companies could err, and create content guardrails that are unfair, biased, or confusing in ways that perturb the information environment.

Second, people might further lose trust in the information environment, concluding that big tech companies who don't share their values are unduly influencing what ideas are generated or gain widespread distribution. In the US, these concerns could come from both sides of the ideological spectrum. Just as we've seen with regards to social media, the right will worry that generative AI tools are biased against conservative ideas and values, while the left will worry that these tools are biased towards conservative ideas and values.

These concerns rest on the widespread use of generative AI and so may not come to fruition before the 2024 election. Indeed, currently, [polling shows](#) that there is bipartisan support among Americans for a variety of interventions by government and industry to reduce the risk that generative AI poses to elections. However, it is entirely possible that public attitudes will polarize once Americans experience their first important "AI election" in 2024. As such, these concerns bear careful watching, and companies and society would be wise to think now about how to resist the centralization of information provision before the problem becomes more developed.

There are at least two potential ways to mitigate this issue.

First, society could take steps to ensure that there are many providers of generative AI tooling, so that no one company has undue power over content generation guardrails. This could include exploring

policies to prevent market concentration in the generative AI industry, and perhaps is a reason to be cautious about endorsing policies that envision restricting the set of companies licensed to provide generative AI tools (though these policies may have other important benefits).

Second, generative AI platforms could avoid taking on this power over content moderation by giving decision-making power concerning guardrails over to users, industry-level self-regulatory bodies, civil-society based regulatory bodies, or other external stakeholders.

Already a number of companies have announced efforts along these lines. Building on their experiment with the Oversight Board to make binding content decisions for the company, Meta has recently announced a "community forum" concerning generative AI guardrails, in which users are randomly recruited and paid to learn about, debate, and make recommendations concerning AI policy issues. OpenAI has announced funding for a similar exploration. Meanwhile, Anthropic has implemented an effort to develop a "constitution" through consultation with the community.

While these efforts are interesting, they are still nascent. The recruited users can only consider a small number of issues and, at most, make recommendations rather than binding decisions. For these self governance efforts to reassure skeptical Republicans and Democrats that tech companies are not shaping the expression of values in society, users would have to be given substantially greater power over more issues, to make binding decisions on them, and to communicate these decisions to the public so that the public could understand that tech companies have truly given over this power. Expanding the power of self-governance institutions like these is difficult, though, due to the expense and the practical challenges associated with recruiting and paying users to deliberate and decide.

Conclusion

We have reviewed a number of essential ways that generative AI will affect the 2024 election, focusing both on the ways that it threatens to disrupt the information environment and affect the outcome of the election and voters' faith in the process itself, as well as the ways it can be used to inform voters and improve democratic functioning.

If there is a single theme to our review, it is that voters, journalists, and everyone who cares about elections should regard claims about new technology with great skepticism. In some instances, the threat to the election is caused by generative AI itself—such as latebreaking "October surprises" driven by highly persuasive fake information. People should be skeptical of these surprises given these threats, and they should find ways to seek out information verification from sources they trust.

In other instances, the threat is not from the technology itself but from over-credulousness about its power. When startups claim to be able to use generative AI to change the minds of large numbers of voters, you should approach the claim with a great deal of skepticism.

Finally, you should also be skeptical about the promise of quick fixes to new technological problems. While the idea of "watermarking" generative AI content is a valuable one, it is clear it will not be a panacea for the 2024 election. Watermarking is simultaneously too easy to evade, and also aimed at only one part of the problem: nothing about watermarking will tell you whether or not you should believe the claims and information in a piece of content. There is ample misleading content that is not AI-generated, and there will be plenty of perfectly accurate AI-generated content. Ultimately there will be no substitute for your skepticism, common sense, and trusted sources.

Author Bios and Disclosures

Ethan Bueno de Mesquita is the interim dean and Sydney Stein Professor in the Harris School of Public Policy at the University of Chicago. Bueno de Mesquita discloses that he receives consulting income from Meta Platforms, Inc for work related to governance issues.

Brandice Canes-Wrone is Professor of Political Science and Maurice R. Greenberg Senior Fellow of the Hoover Institution, Stanford University.

Andrew B. Hall is the Davies Family Professor of Political Economy in the Graduate School of Business at Stanford University, and a Professor of Political Science. He is a Senior Fellow at the Stanford Institute for Economic Policy Research and a Senior Fellow (courtesy) at the Hoover Institution. Hall discloses that he receives consulting income from Meta Platforms, Inc for work related to Augmented Reality strategy, and from Andreessen-Horowitz for work related to web3 governance.

Kristian Lum is Research Associate Professor in the Data Science Institute at the University of Chicago. She was previously a Senior Staff Machine Learning Researcher on the ML Ethics, Transparency and Accountability Team at Twitter.

Gregory J. Martin is Associate Professor of Political Economy in the Graduate School of Business at Stanford University.

Yamil Ricardo Velez is an Assistant Professor of Political Science at Columbia University.