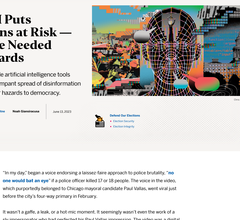
ARTICLE

How AI Puts Elections at Risk — And the Needed Safeguards

Widely accessible artificial intelligence tools could fuel the rampant spread of disinformation and create other hazards to democracy.



Natasha Foundiadakis | Noah Giansiracusa June 23, 2023

Related Our Elections
- Election Security
- Election Integrity

"In my day," began a voice endorsing a laissez-faire approach to price brutality, "**no one would bat an eye**." A police officer killed 13 of 38 people. The voice in the video, which purportedly belonged to Chicago mayoral candidate Paul Vallas, went viral just before the city's four-way primary in February.

It wasn't a gaffe, a leak, or a hot mic moment. It seemingly wasn't even the work of a shy impersonator who had perfected his Paul Vallas impression. The video was a digital fabrication, a likely creation of generative artificial intelligence that was viewed thousands of times.

The episode heralds a new era in elections. Next year will bring the first national campaign season in which widely accessible AI tools allow users to synthesize audio in anyone's voice, generate photo-realistic images of anybody doing nearly anything, and power social media bot accounts with near human-level conversational abilities — and do so on a vast scale and with a reduced or negligible investment of money and time. Due to the popularization of chatbots and the search engines they are quickly being absorbed into, it will also be the first election season in which large numbers of voters routinely consume information that is not just curated by AI but is produced by AI.

This change is already underway. In April, the Republican National Committee used AI to produce a video warning about potential dystopian crises during a second Biden term. Earlier this year, an AI-generated video showing President Biden declaring a national draft to aid Ukraine's war effort — originally acknowledged as a deepfake but later stripped of that context — led to a mistaken belief that garnered over **5 million views**. A deepfake audio recording purported to be Elizabeth Warren (D-MA) **insisting** that Republicans should be barred from voting in 2024. In the future, malign actors could deploy generative AI with the intent to suppress votes or circumvent defenses that secure elections.

The AI challenge to elections is not limited to disinformation, or even to deliberate mischief. Many elections offices use algorithmic systems to maintain voter registration databases and verify mail ballot signatures, among other tasks. As with human decisions on these operations, algorithmic decision-making has the potential for racial and other forms of bias. There is growing interest by some officials in using generative AI to aid with voter education, creating opportunities to speed up processes but also producing serious risks for inaccurate and inequitable voter outreach.

AI advances have prompted an abundance of generalized concerns from the public and policymakers, but the impact of AI on the field of elections has received relatively little in-depth scrutiny given the outsize risk. This piece focuses on disinformation risks in 2024. Forthcoming Brennan Center analyses will examine additional areas of risk, including voter suppression, election security, and the use of AI in administering elections.

AI since the 2022 elections

While AI has been able to synthesize photo-quality "deepfake" profile pictures of nonexistent people for several years, it is only in recent months that the technology has progressed to the point where users can create lifelike images of anything with a simple text prompt. Adept users have long been able to use Photoshop to edit images, but vast numbers of people can now create convincing images from scratch in a matter of seconds at a very low — or no — cost. Deepfake audio technology made enormous strides and can now **clone a person's voice** with remarkably little training data.

While forerunners to the **wildly popular** app ChatGPT have been around for several years, OpenAI's latest iteration is leaps and bounds beyond its predecessors in both popularity and capability. Apps like ChatGPT are powered by large language models, which are systems for encoding words as collections of numbers that reflect their usage in the vast swaths of the web selected for **training** the app. The launch of ChatGPT just weeks after the 2022 midterm election on November 30, 2022, has precipitated a new era in which many people regularly converse with AI systems and read content produced by AI.

Since ChatGPT's debut, our entire information ecosystem has begun to be reshaped. Search engines are **incorporating** this kind of technology to provide users with information in a more conversational format, and some news sites have been using AI to produce articles more cheaply and quickly, despite the **tendency** for it to produce misinformation. Smaller (for now) replicas of ChatGPT and its antecedents are not limited to the American tech giants. For instance, **China** and **Russia** have their own versions. And researchers have found ways of training small models from the output of large models that perform nearly as well — **enabling people around the globe to run custom versions on a personal laptop**.

Unique vulnerability to disinformation

Elections are particularly vulnerable to AI-driven disinformation. Generative AI tools are most effective when producing content that bears some resemblance to the content in their training databases. Since the same false narratives pop up repeatedly in U.S. elections — as Brennan Center research and other disinformation scholars **have found**, election domain do not matter much. There is ample supply of good election disinformation in the training data underlying current generative AI tools to render them a potential boon to those tools for future election disinformation. This includes core deceptions around the security of voting machines and mail voting, as well as misinformation tropes regularly applied to innocent and hard-resolved glitches that occur in most elections. Visual-based misinformation is widely available as well — for example, pictures of discarded mail ballots were used to distort election narratives in both the 2020 and 2022 elections.

Different kinds of AI tools will leave distinct footprints in future elections, threatening democracy in myriad ways. Deepfake images, audio, and video could prompt an uptick in viral moments around face scandals or artificial glitches, further warping the nation's civic conversation at election time. By seeding online spaces with millions of posts, malign actors could use language models to create the illusion of political agreement or the false impression of widespread belief in dishonest election narratives. Influence campaigns could deploy tailored chatbots to customize interactions based on user characteristics, adapting manipulation tactics in real time to increase their persuasive effect. And they could use AI tools to send a wave of deceptive comments from fake "constituents" to election offices, as one researcher who **duped** Idaho state officials in 2019 using ChatGPT's predecessor technology showed. Chatbots and deepfake audio could also exacerbate threats to election systems through phishing efforts that are personalized, convincing, and likely more effective than what we see in the past.

One need not look far to witness the potential for AI to distort the political conversation around the world. A viral deepfake **showing** Ukrainian President Volodymyr Zelenskyy surrendering to Russia. Pro-China bots sharing videos of **AI-generated news anchors** — at a sham outfit called "Wolf News" — promoting falsehoods flattering to China's governing regime and critical of the United States, the first known example of a state-aligned campaign's deployment of AI tools to create fictitious people. ChatGPT is yielding to a **request from researchers** to write a message for a Soviet-style information campaign suggesting that HIV, the virus that can cause AIDS, was created by the U.S. government. Incidents like these could proliferate in 2024.

Dramatically enhanced disinformation tools

In 2016, state-affiliated organizations in Russia employed hundreds and possessed a monthly budget of more than a million dollars to conduct information warfare in an attempt to influence the U.S. presidential election. Today, with the benefit of generative AI, a similar effort — or even one on a much larger scale — could be executed with a fraction of the personnel and at less expense. Future state-aligned influence campaigns could cut out many intermediaries, relying on better automated systems.

AI tools could also increase the persuasive power of large scale disinformation campaigns by better blending falsehoods with recognizable information environments and exploiting voters' racial, religious, and political identities en masse. Prior Russia-backed influence campaigns were often pockmarked with obvious errors and mistakes, but recent AI tools could boost these flaws by erasing or mitigating glitchy visuals, mistranslations, grammatical faux pas, and bungled idioms so deceptions do not attract as much suspicion. Automated conversations with voters that are designed to deceive could be multiplied ad infinitum with a model fine-tuned for that function. And as Poynter has **shown**, apps like ChatGPT can facilitate entire misinformation filled fake news sites — a particular risk when they masquerade as local news in "news deserts" where **millions of eligible voters** live in counties with no remaining local newspaper.

At the same time, voters interacting with language models on search engines and through chatbots will likely unwittingly encounter misinformation since these tools are known to periodically "hallucinate" and even fabricate authoritative-looking footnotes with links to nonexistent articles to support false claims.

Trust in accurate election information

The president of Gabon traveled overseas for several months in 2018 to receive medical assistance. At home, his lingering absence produced confusion and fostered conspiracies. When the Gabon government released a video to prove the president was alive, opponents claimed that it was a digital forgery. Even though the video was likely authentic, the potential to create realistic fakes made the claim plausible and fueled confusion. This is what is known as the "**liar's dividend**" — the notion of generative AI creates an atmosphere of mistrust — and the dividend could be set to grow dramatically.

Beyond outright falsehoods, the proliferation of AI-generated content may accelerate the loss of trust in the overall election information ecosystem. In the future, voters could inhabit online spaces crowded with manipulated viral images and videos and AI-generated text. Widespread use of generative technology could create a fog of confusion that makes it harder to tell truth from falsity — an apparent goal of Russia's "**Firehose of Falsehood**" propaganda model. That, in turn, could erode trust in election information more broadly, making it harder for voters to have faith in any sources of election information — even ones that are accurate and authoritative. Content that spoofs election officials, for instance, could result in real officials losing the credibility they largely enjoy.

The path forward to protect elections

Americans need safeguards to protect our elections from the many risks that AI technologies pose. Below are just a few of the actions that should be considered as part of a comprehensive governmental, civil society, and private sector response to the threats that AI poses to elections and democracy.

While defending against damage to democracy from AI will require an interagency effort, the executive branch should designate a lead agency to coordinate governance of AI issues in elections. On the disinformation front, the Cybersecurity and Infrastructure Security Agency should **create and share resources** to help election offices address disinformation campaigns that exploit deepfake tools and language models to undermine election processes. To reduce the risk of AI misuse by political campaigns, the Federal Election Commission should ensure that its political ad disclosure requirements cover the full range of online communications currently permitted under federal law. That includes ensuring its rules cover political communications from **paid influencers**, who may disseminate AI-generated content, and the paid online promotion of content, which may also make use of AI.

The federal government should ramp up efforts to promote and encourage innovation in deepfake detection and to nurture progress in the detection of voting disinformation campaigns and election cybersecurity threats fueled by language models and chatbots, including through the Defense Advanced Research Projects Agency and the new **AI Institute for Agent-based Cyber Threat Intelligence and Operation**. Among the menu of actions should be the development of high-accuracy detection and anti-phishing tools for use by state and local election offices. In the end, a tug-of-war between AI tools that can generate disinformation and tools that accurately detect content generated by AI, the government can offer a boost to defenders' efforts — including those focused on coordinated bot campaigns — to bolster their effectiveness.

AI developers and social media companies must play a role in mitigating threats to democracy. Among other steps, AI developers should implement and continuously refine filters for election falsehoods and impose interface limitations to make it more difficult to create disinformation campaigns at scale. Social media companies should develop policies that reduce harms from AI-generated content while taking care to preserve legitimate discourse. They should publicly verify election officials' accounts and other authoritative sources of election information, such as the **National Association of Secretaries of State**, using unique icons. (Twitter's complimentary **gray checkmark label** for government accounts does not clearly cover local election offices, for example.) Platforms should devote more resources and attention to identifying and removing coordinated bots and labeling deepfakes that could influence elections. They need to coordinate closely with AI developers to continuously improve detection practices as generative AI capabilities evolve.

Finally, Congress and state legislatures need to act quickly to regulate AI. While the process of establishing the most prudent course of action will require further discussion and refinement, it is clear that lawmakers cannot afford to dawdle or allow themselves to become mired in partisan bickering. The stakes are simply too high. Among options that merit deliberation and debate are mandating **watermarking and digital signatures to help identify AI-generated content**, requiring companies to prove the safety of their products before releasing them to the public, and **limiting** the creation and transmission of the most harmful AI-generated content that can interfere with elections.

Generative AI holds whose source code is fully public — and consequently downloadable and manipulatable — pose particular challenges since users can remove safeguards and operate these models without moderation or scrutiny. But huge numbers of users will continue to rely on proprietary AI apps provided by tech firms, so regulation targeting the development and deployment of such apps in the private sector will have a large impact despite the open-source alternatives.

Voters need some level of transparency to promote safe AI use when it comes to elections. Lawmakers could compel AI developers to make public the categories of data and guiding principles used to train and fine-tune generative AI models, they could require algorithmic impact assessments for AI systems deployed in governance settings, they could mandate periodic third-party audits of AI systems used in election administration, and they could require that election offices disclose details about their use of AI systems in running elections. Congress should also require "paid for" disclaimers and other disclosures for a much wider range of online ads than the law currently mandates. While the Honest Ads Act, a Senate bill, would accomplish this in part, it could be made even better by requiring the disclosure of information about the role of AI in generating certain political communications. These actions could help voters make informed choices and mitigate risks from AI use in election settings.

Any governmental use of generative AI to aid with educating voters or otherwise engage constituents in elections should be tightly regulated. We should also look beyond national borders to support a **coordinated global response**.

AI has the potential to dramatically change elections and threaten democracy. A whole of society response is needed.

RELATED ISSUES: Related Our Elections
- Election Security
- Election Integrity

ARTICLE

Now Is the Time to Protect the 2024 Election

Elections are increasingly vulnerable.

Marcelo Agudo 4 May 1, 2023

ARTICLE

The Great Resignation . . . Of Election Officials

The people who run our elections are exhausted and afraid.

Michael Waldman 4 April 25, 2023

Part of Election Officials Stress High Workload, Threats, Mental Threats and Political Interference

April 25, 2023 4 Ruby Edlin, Lawrence Norden

Federal Homeland Security Grant Funds Directed to Election Security

March 6, 2023 4 Lawrence Norden, Derek Tisler

Congress Must Fund Election Security

December 12, 2022 4 Michael Waldman

MORE NEWS & ANALYSIS

Stay up to date
Issues
Voting & Elections
Research & Reports
Fiscal Sustainability
Court Cases
Our Work
Voting & Elections
Research & Reports
Press Releases
Court Cases
Experts
Get Involved
Support Our Work
Join Our Mailing List
About
Mission & Leadership
History
Programs
Employment
Staff
Board of Directors
Leadership
Financial & Legal Information
Contact Us
Brennan Center & NYU School of Law
Library
Press
Email Address *
Zip Code *
SIGN ME UP
We respect your privacy.