https://epic.org/safeguarding-democratic-elections-in-the-age-of-ai/

Safeguarding Democratic Elections in the Age of AI

April 4, 2024 | Sophie Nyombi Nantanda, EPIC Spring Intern

As national elections approach in the United States, concerns about AI-generated robot calls and manipulated media have grown more urgent. The 2024 elections mark a pivotal moment in American history, as it stands as the first presidential election year poised at the intersection of the burgeoning advancements in AI-generated content. Both legal and technological interventions will be necessary to mitigate these threats in advance of November.

There is reason to believe the U.S. is not adequately equipped to address this problem. State and local election officials need additional funding and resources to address threats ranging from AI to personal violence. Further underscoring the urgency of addressing these issues, trust in the electoral system is declining, leaving Americans more susceptible to misinformation and disinformation campaigns.

In the digital age, combating AI-generated misinformation, robocalls, deepfakes demands attention. Technologies like AI-powered content verification tools and deepfake detection algorithms digital watermarking may help to address this threat, albeit with imperfections. While they provide a starting point, the role of public engagement in reporting disinformation is also essential. Dedicated platforms have emerged for this purpose. These platforms serve as vital tools for empowering individuals to flag misleading content, fostering a collaborative effort in combating misinformation and safeguarding the integrity of our digital landscape. However, to truly fortify our information ecosystem, comprehensive legislation for AI governance may emerge as the most effective solution. Exploring these multifaceted approaches paves the way towards safeguarding democracy in elections.

Calls and Manipulated Media

A few months ago, AI voice cloning technology was employed to mimic the voice of President Joe Biden with the intention to dissuade voters from participating in the state's Democratic presidential primary. Such misinformation poses a threat to democracy, and in some respects the legal system responded swiftly. The suspects were served with notice; the Federal Communications Commission clarified that calls using an AI-powered voice could violate the Telephone Consumer Protection Act (TCPA) if made without the called party's consent; the Federal Trade Commission proposed liability for impersonating individuals using AI; and private litigants sued the parties behind the calls. But

all of these measures came long after the offending message had already been broadcast. More needs to be done to prevent these types of AI-generated calls before they happen in the first place.

The rise of AI-generated calls represents a significant challenge to the integrity of electoral processes. Although automated calling systems have traditionally been used by political campaigns to reach out to voters and disseminate campaign messages, the malicious use of robocalls in elections is not itself new. Robocalls have previously been used to disseminate false information, manipulate public opinion, and suppress voter turnout, undermining the democratic process. For example, the Michigan Supreme Court recently heard a case in which two political operatives are alleged to have used 85,000 robocalls to disseminate false information and dissuade people from participating in the 2020 presidential election. But AI threatens to radically intensify the problem of misleading election robocalls. Advances in AI technology enable the automated generation of hyper-realistic voice simulations, making it easier to create persuasive and deceptive phone calls and media at scale. For a sense of a scale: YouTube recently deleted 1,000 videos of celebrity AI scams.

AI-generated images, videos, and audio clips also exacerbate concerns about misinformation and disinformation in political advertising. Deepfake technology in particular enables the creation of highly realistic and difficult-to-detect synthetic media, such as videos of political candidates saying or doing things they never actually said or did. AI models like Midjourney are being used to generate misleading images for use as political disinformation. Midjourney has announced some proactive steps to prevent its users from fabricating counterfeit images of both President Joe Biden and former President Donald Trump, but these alone will not come close to solving the problem of misinformation and disinformation.

In one notable case, Donald Trump supporters created and shared AI-generated fake images of Black voters to encourage them to vote Republican. This type of misinformation poses a significant threat to the integrity of electoral campaigns, as AI-generated fake media can sway public opinion, damage candidates' reputations, and erode trust in democratic institutions. Deepfakes can also sow confusion among the public, blurring the lines between what is authentic and what is fabricated. This may cause individuals to unintentionally label genuine information as false—or enable others to do so intentionally. Professors Danielle Citron and Bobby Chesney have labeled this the liar's dividend: i.e., the phenomenon by which the existence of believable deepfakes enables and encourages bad actors and authoritarian leaders to label truthful content as fake.

The circulation of AI-generated misinformation and disinformation through social media platforms further compounds the challenge, as seen in recent presidential elections. Organizations like The Knight Institute and Stanford Cyber Policy Center have shown how false narratives can spread rapidly and unchecked via algorithmic recommendation systems, amplifying polarization and undermining democratic discourse. Fact checking mechanisms can be helpful for combating the spread of misinformation. For example, Meta has implemented a range of tools to deal with misinformation of its platforms. However, these tools and mechanisms can be overwhelmed by the sheer volume and sophistication of AI-generated fake content and may carry their own downsides (discussed below). Moreover, Meta is preparing to discontinue CrowdTangle, a data analysis tool utilized for identifying misinformation across Facebook and Instagram. The tool will be deactivated just three months before the U.S. presidential elections, curtailing the ability of researchers, journalists, and others to identify dangerous misinformation trends at a key moment.

Legislative and Technological Measures

In response to these challenges, policymakers, tech companies, and civil society organizations must work to develop comprehensive strategies for addressing AI-related threats to electoral integrity. State legislatures are weighing and enacting laws that would regulate deepfakes in electoral processes, often garnering bipartisan support. Implementing such laws would help combat the spread of AI-generated misinformation and disinformation, enhance transparency and accountability in political advertising, and promote media literacy to empower voters to discern fact from fiction. The FCC also has the authority to crack down on robocalls disseminating false information, as it has done in the past (though the violation must be related to the TCPA and consent rather than to the content of the calls). Stricter penalties may help the Commission better deter such deceptive practices.

Technological solutions like AI-powered content verification tools and deepfake detection algorithms may help detect and mitigate the impact of AI-generated fake media, but they cannot solve the problem on their own and may introduce additional problems. For instance, although deepfake detectors can scan for distinct biometric indicators within a video, such as an individual's heartbeat or a voice produced by natural human vocal organs rather than synthesized ones, their effectiveness is not guaranteed. Currently available algorithms struggle to consistently detect high-quality deepfakes produced through advanced AI technologies. Moreover, these detectors represent potential privacy and equity risks as well. Regulatory guardrails such as AI watermarking legislation could serve as an effective—albeit imperfect—measure to curb the spread of

misinformation and disinformation. Skepticism stems from concerns surrounding the standardization and widespread adoption of digital watermarking practices. Without a universally agreed-upon framework for implementing and recognizing watermarks, its effectiveness could be limited.

Nevertheless, despite these challenges, the adoption of AI watermarking legislation would signify a step forward from the current regulatory landscape. It would represent a great improvement in the fight against misinformation and disinformation, offering authorities an additional tool to safeguard the integrity of digital content and protect public discourse.

There are also sources encouraging the general public to report disinformation to prevent its dissemination. Empowering individuals to flag misleading content helps authorities and platforms address such instances swiftly. Alongside reporting mechanisms, promoting media literacy equips people with tools to discern credible information. Leveraging technology for efficient reporting enhances participation and defense against disinformation, fostering a collaborative approach for a more resilient information ecosystem.

According to the Cybersecurity and Infrastructure Security Agency (CISA), the most effective strategies for addressing generative AI-enhanced threats in elections aligns with longstanding cybersecurity best practices, which may already have been implemented. These measures include hardening social media accounts, implementing robust email security protocols to combat phishing attacks, enhancing network perimeter defenses to detect and prevent unauthorized access, conducting regular security audits and vulnerability assessments, and fostering greater collaboration and information sharing among government agencies, election officials, and cybersecurity experts. More cybersecurity measures can be found here. However, while these strategies may have already been implemented to varying degrees, it's important to recognize that cybersecurity is an evolving field, and new threats continually emerge. Therefore, while progress may have been made in certain areas, there is still ample room for improvement and the implementation of more comprehensive and proactive cybersecurity measures to safeguard electoral integrity effectively.

The convergence of AI and elections presents challenges for democratic societies. As we confront the threats posed by AI-generated robocalls and manipulated media, it is imperative for government to take the lead on working towards upholding the principles of transparency, integrity, and accountability in electoral processes. The World Economic Forum in the Global Risks Report 2024, forecasted that misinformation and disinformation could significantly disrupt electoral processes in multiple economies including Bangladesh, India, Indonesia, Mexico, Pakistan, the United Kingdom within the next two years.

Similar concerns helped motivate the European Union to adopt the Digital Services Act, which requires social media platforms to combat politically motivated campaigns and the spread of false information. Meanwhile, the European Union AI Act (passed by both the European Parliament and the Council of the European Union and awaiting a formal adoption) represents a groundbreaking effort to regulate the entire lifecycle of AI development, deployment, and usage. One of the most critical provisions in the AI Act pertains to classifying AI systems that aim to influence electoral processes as high-risk AI with stringent transparency, accountability, and responsible usage obligations. States are swiftly enacting legislation aimed at addressing the production of AI-generated deepfakes in anticipation of the 2024 presidential elections. Over 100 bills have been introduced or passed in 40 state legislatures this year alone. However, by embracing a comprehensive regulatory model similar to that of European Union, the United States can fortify its electoral integrity, ensuring fairness and fidelity to voters' choices.