POLICY TIP SHEET

Election-Related Deepfakes

Understanding the Threat and Exploring Potential Solutions

By Brady Smith & Jack McPherrin

August 2024

THE PROBLEM

- Deepfakes use AI to create false audio, photo, and video representations, making individuals appear to say or do things they never did, undermining public debate and trust.
- These manipulations pose a significant threat to election integrity, spreading misinformation, influencing voter perceptions, and undermining trust in election outcomes.
- As deepfake technology improves, detecting these forgeries becomes increasingly difficult, making it easier for malicious actors to deceive the public.
- Countries like India, Brazil, Gabon, and Slovakia have experienced deepfake-related issues in their elections, highlighting the potential for similar problems in the United States.
- A significant percentage of Americans believe AI has contributed to eroding trust and spreading misinformation, exacerbating the challenge of maintaining credible elections.

THE SOLUTIONS

- Policymakers could enact state laws requiring election advertisements to include clear notices when deepfakes are used, promoting transparency.
- Legislators could encourage dissemination of information from organizations like truemedia.org to educate voters about deepfakes and how to identify them.
- Legislators could encourage platforms to establish formal procedures for receiving and addressing deepfake complaints, with clear explanations of their policies.
- State lawmakers could consider mandating that platforms label content that is known or suspected to be machine-generated, helping users identify potential deepfakes.
- Laws dealing with legal penalties for creating and disseminating deepfakes intended to influence elections could be created and/or strengthened, deterring malicious actors.
- Legislators could encourage integrating policy and ethical literacy into the education of aspiring engineers, preparing them to develop responsible AI solutions and mitigate deepfake risks.

EST. 2020 AMERICAN JOURNEY EXPERIENCE

THE HEARTLAND INSTITUTE

The Henry Dearborn LIBERTY NETWORK

PFLN PRO-FAMILY LEGISLATIVE NETWORK BY WALLBUILDERS

Election-Related Deepfakes

Understanding the Threat and Exploring Potential Solutions

INTRODUCTION

Election-related deepfakes involving audio, photo, and video manipulations have become a significant concern in the digital age. Deepfakes use artificial intelligence (AI) to create convincing but false representations of individuals, making them appear to say or do things they never said or did. These manipulations pose significant threats to the integrity of public debate and electoral processes, as they can be used to spread false information, influence voter perceptions, and undermine trust in election outcomes.

Deepfakes are typically created using generative adversarial networks (GANs), which pit two AI models against each other to improve the authenticity of the final product.[1] As AI technology has advanced exponentially in recent years, deepfakes have become more sophisticated, making it substantially more difficult to detect them. Advanced detection methods rely on analyzing inconsistencies in digital content, such as unnatural facial movements or audio-visual mismatches.[2]

Deepfakes can undermine electoral processes by spreading false information about candidates, creating confusion among voters, and damaging the reputations of political figures. The ability to produce convincing forgeries can lead to a loss of trust in legitimate information sources, making it harder for voters to make informed decisions.

Even worse, creating a deepfake using established AI models is a relatively simple process; some sources claim that the average individual can create a deepfake in less than 30 seconds.[3] As a result, Americans have become increasingly concerned about how to decipher what is real and what is not. For instance, one survey conducted in 2021 found that a majority of Americans believe the proliferation of AI has led to a loss of trust in elections (57 percent), a loss of trust in institutions (56 percent), and the spread of misinformation (58 percent).[4]

These concerns are not unfounded or based on speculation. Deepfakes have already caused significant controversies in elections around the world.

> **"Deepfakes can undermine electoral processes by spreading false information about candidates, creating confusion among voters, and damaging the reputations of political figures."**

NOTABLE DEEPFAKE INCIDENTS

The use of AI-generated deepfakes has already created substantial problems in the electoral processes of many countries, including India, Brazil, Gabon, Slovakia, and the United States.

During several recent Indian elections, deepfakes were used to manipulate speeches of political leaders, spreading false messages.[5]

In Brazil, deepfakes were deployed in recent elections to create fake news stories about candidates, leading to widespread misinformation.[6]

In Gabon, a false video of President Ali Bongo was released to dispel rumors about his health. The video showed Bongo addressing the nation, but it was widely speculated to be a deepfake due to his unnatural appearance and movements. The release of the video led to political unrest and a military coup attempt, as opposition parties and the public questioned its authenticity and the true state of the president's health.[7]

Two days before the 2023 presidential election in Slovakia, an audio recording purportedly showing one of the leading candidates discussing rigging elections went viral. Though the recording was later revealed

to be a deepfake, the candidate, Michael Šimečka, lost the election. Experts believe the false recording impacted the results.[8]

The United States has not been immune to deepfakes, either. Though deepfakes have likely not yet affected the result of a major election, the potential to do so is evidenced by numerous recent events involving deepfakes.

For instance, many AI-generated audio and video deepfakes appeared on social media platforms in advance of the 2020 election, leading to increasing concern about misinformation.[9] During Chicago's 2023 mayoral election, deepfake technology was used to clone the voice of candidate Paul Vallas and make it appear as if Vallas condoned police violence.[10]

There have been several deepfake instances related to the 2024 presidential election as well. For instance, in 2023, Florida Gov. Ron DeSantis—who was still in the running to be the Republican presidential candidate at the time—produced an attack ad against former President Donald Trump that contained AI-generated deepfake images of Trump hugging former White House chief medical advisor Anthony Fauci.[11] During the January 2024 New Hampshire primary, a Democratic Party political consultant created a deepfake of President Joe Biden's voice that was used in a statewide robocall, in which "Biden" encouraged primary voters to stay home and abstain from voting.[12] In June 2024, a video was created and widely circulated on X that featured Trump saying: "If we want to make America great, we've got to make antisemitism great again!" The video was a deepfake.[13]

Ultimately, deepfakes have the potential to undermine the individual liberty of everyone by creating realistic but fake content that can cause severe reputational damage. Public figures and political candidates are particularly at risk due to their large public profiles. Victims of deepfakes may find their personal and professional lives disrupted by false representations that they have little ability to refute convincingly; in the case of politicians, this can have massive effects upon the rest of society. And yet, there are other ways that deepfakes can be used to impact elections, such as by burnishing a candidate's reputation.

'POSITIVE' DEEPFAKES

Though much of the concern about deepfakes is related to how they can be used to deleteriously impact a political candidate's reputation and widely spread "negative" information, they can also be used by political campaigns to create false information that enhances politicians' reputations. In some cases, candidates or their supporters have already begun leveraging this technology to augment the appeal of their campaigns.

For example, during the 2021 South Korean presidential election, candidates used deepfake technology to create an AI-generated version of themselves that appeared in video ads.[14] These digital doppelgangers of the candidates mimicked their voice, gestures, and appearance, allowing them to seem more relatable, approachable, and in touch with younger voters. The campaign used the "AI candidate" to communicate in a more informal, humorous manner—while removing some of the tics or habits that the candidate may have been criticized for previously—and engage with voters in ways the real candidate might not have been able to do.

To illustrate the potential misuse of this technology on the "positive" side of a campaign, imagine if deepfake technology could have been similarly used in the United States during the 2024 election cycle to benefit candidates like President Joe Biden prior to his dropping out of the race. Concerns about Biden's age and mental sharpness were prominent talking points among his opponents. Theoretically, his campaign could have utilized deepfakes to present a more vigorous, energetic image of Biden in campaign

ads, smoothing out any gaffes or stumbles to depict a candidate who was consistently articulate and poised. This could have involved altering his voice or expressions to make him appear more confident and youthful, thereby countering negative perceptions and appealing to a broader electorate.

The ethical implications of such uses are significant. While these technologies can humanize candidates and make them seem more personable, they also blur the line between reality and manipulation, potentially misleading voters and undermining the authenticity of political discourse.

Deepfakes pose a direct threat to the integrity of election institutions. By disseminating false information, deepfakes can manipulate voter perceptions, spread misinformation, and erode trust in the electoral process. Deepfakes can create confusion among voters, making it difficult to distinguish between genuine statements and fabricated ones, thereby undermining our election process.

As noted in a paper published by the Brennan Center for Justice, Josh Goldstein and Andrew Lohn explain that "If politicians or their proxies can successfully use false claims to deceive the public, then they can undermine the public's ability to affect those informed preferences, opinions, and decisions. In the starkest of terms, disinformation becomes a threat to deliberative democracy itself."[15] Echoing Goldstein and Loeb, Renée DiResta, technical research manager at the Stanford Internet Observatory, stated in an article for *Wired*: "As synthetic media of all types—text, video, photo, and audio—increases in prevalence, and as detection becomes more of a challenge, we will find it increasingly difficult to trust the content that we see."[16]

The threat posed by election-related deepfakes is real and growing. Policymakers should consider proactive steps to address this issue by implementing robust state and federal solutions. Transparency, education, and technological advancements are key to safeguarding the integrity of elections and ensuring that voters can trust the information they receive.

By implementing these policy recommendations, we can better protect our public discourse and legislative process from the dangers of deepfakes as well as ensure that elections remain free, fair, and trustworthy.

POLICY RECOMMENDATIONS

To mitigate the risks posed by election-related deepfakes, several policy solutions and educational efforts could be considered by legislators at both state and federal levels.

Mandatory Disclosure Laws: Legislators could pass state laws requiring that election advertisements include clear notices when deepfakes are used. This transparency will help voters distinguish between genuine and manipulated content. In 2019, California passed a law requiring disclosure of manipulated content.[17]

Educational Campaigns: Legislators could support the dissemination of information from organizations such as truemedia.org to educate voters about the existence and risks of deepfakes.[18] Providing resources on how to identify deepfakes can empower voters to critically assess the content they encounter.

Complaint Procedures and Transparency: Legislators could encourage platforms that host user-generated content to establish a formal procedure for receiving and addressing complaints about deepfake content. This procedure could also be accompanied by a clear and concise overview of the principles and standards behind the platform's policies on deepfakes.

Content Labeling: Legislators could encourage platforms to label content that is known or suspected to be machine generated. This labeling will help users identify potential deepfake content and make informed decisions about the veracity of the information they consume.

Legal Penalties for Deepfake Use in Elections: Legislators could establish stringent penalties for the creation and dissemination of deepfakes intended to influence election outcomes. For instance, Texas passed a law in 2019 making it illegal to "create a deepfake video and publish it within 30 days of an election with the intent to injure a candidate or influence the result of an election,"[19] punishable by up to a year in jail and a $4,000 fine.[20]

Endnotes

1 Mika Westerlund, "The Emergence of Deepfake Technology: A Review," Technology Innovation Management Review, Volume 9, Issue 11, November 2019, https://www.researchgate.net/publication/337644519_The_Emergence_of_Deepfake_Technology_A_Review

2 Michela Gravina et al., "FEAD-D: Facial Expression Analysis in Deepfake Detection," *Image Analysis and Processing—ICIAP 2023*, September 5, 2023, https://link.springer.com/chapter/10.1007/978-3-031-43153-1_24

3 Lutz Finger, "Overview of How to Create Deepfakes—It's Scarily Simple," Forbes, September 8, 2022, https://www.forbes.com/sites/lutzfinger/2022/09/08/overview-of-how-to-create-deepfakesits-scarily-simple/

4 Stevens Institute of Technology, *2021 Techpulse Report*, accessed July 30, 2024, https://www.stevens.edu/stevens-techpulse-report/2021-techpulse-report

5 Meryl Sebastian, "AI and deepfakes blur reality in India elections," BBC, May 15, 2024, https://www.bbc.com/news/world-asia-india-68918330

6 France24.com, "Brazil seeks to curb AI deepfakes as key elections loom," August 3, 2024, https://www.france24.com/en/live-news/20240308-brazil-seeks-to-curb-ai-deepfakes-as-key-elections-loom

7 Ali Breland, "The Bizarre and Terrifying Case of the 'Deepfake' Video that Helped Bring an African Nation to the Brink," Mother Jones, March 15, 2019, https://www.motherjones.com/politics/2019/03/deepfake-gabon-ali-bongo/

8 Curt Devine *et al.*, "A fake recording of a candidate saying he'd rigged the election went viral. Experts say it's only the beginning," CNN, February 1, 2024, https://edition.cnn.com/2024/02/01/politics/election-deepfake-threats-invs/index.html

9 John Villasenor, "Deepfakes, social media, and the 2020 election," The Brookings Institution, June 3, 2019, https://www.brookings.edu/articles/deepfakes-social-media-and-the-2020-election/

10 Megan Hickey, "Vallas campaign condemns deepfake video posted to Twitter," CBS News, February 27, 2023, https://www.cbsnews.com/chicago/news/vallas-campaign-deepfake-video/

11 N. David Bleish, "Deepfakes and American Elections," American Bar Association, May 6, 2024, https://www.americanbar.org/groups/public_interest/election_law/american-democracy//resources/deepfakes-american-elections/

12 Nick Robertson, "Political consultant indicted in fake Biden robocall in New Hampshire," The Hill, May 23, 2024, https://thehill.com/policy/technology/4681403-joe-biden-fake-robocall-new-hampshire-political-consultant-indicted/

13 Aleksandra Wrona, "Video of Trump Saying 'Make Antisemitism Great Again' Is a Deepfake," Snopes, June 7, 2024, https://www.snopes.com/fact-check/trump-make-antisemitism-great-again/

14 Jo He-rim, "Election 2022: AI spokesman, avatars enter election campaigns," *The Korea Herald*, December 8, 2021, https://www.koreaherald.com/view.php?ud=20211208000709

15 Josh A. Goldstein and Andrew Lohn, "Deepfakes, Elections, and Shrinking the Liar's Dividend," Brennan Center for Justice, January 23, 2024, https://www.brennancenter.org/our-work/research-reports/deepfakes-elections-and-shrinking-liars-dividend

16 Renee Diresta, "AI-Generated Text Is the Scariest Deepfake of All," Wired, July 31, 2020, https://www.wired.com/story/ai-generated-text-is-the-scariest-deepfake-of-all/

17 California AB 730, 2019-2020 Regular Session, Legiscan.com, https://legiscan.com/CA/text/AB730/id/2041326

18 Truemedia.org, "Identifying Political Deepfakes in Social Media using AI," accessed August 19, 2024, https://www.truemedia.org/

19 N. David Bleish, "Deepfakes and American Elections."

20 Texas SB 751, 86th Legislature, 2019-2020, Legiscan.com, https://legiscan.com/TX/bill/SB751/2019