2023

Deepfake 2024: Will Citizens United and Artificial Intelligence Together Destroy Representative Democracy?

Richard Painter

Deepfake 2024: Will *Citizens United* and Artificial Intelligence Together Destroy Representative Democracy?

Richard W. Painter*

ABSTRACT

Deepfakes – computer generated counterfeit videos and audios of people saying and doing things they never said or did – are proliferating on social media and increasingly will be used to target candidates in elections. Citizens United v. FEC, and cases decided in its aftermath, have opened the floodgates of dark money funded electioneering communications, and some of this money will be spent on deepfakes made and disseminated by persons unknown. Some deepfakes may originate outside the United States, as they become a new instrument for foreign interference in U.S. elections.

The Federal Election Commission (FEC) has been asked by public interest organizations and members of Congress to do something about deepfakes but has deadlocked on whether to act. Bills are pending in Congress to address the problem, but some of these bills are overbroad and rely on criminal sanctions, exacerbating constitutional problems. No bill addressing deepfakes in elections has passed either house.

Dark money in politics, foreign interference in U.S. elections, and the rise of AI-generated deepfake political ads will become an issue of increasing concern. And, with Congress and the FEC currently deadlocked on pending legislation that could address the problem, the future of U.S. elections remains in jeopardy. However, steps can be taken now to circumvent government inaction. Deepfake political ads could be flagged by publicly or privately funded Deepfake Warnings with a FEC sponsored Deepfake Alert System that could respond quickly to deepfake electioneering communications by identifying them as such in the same social media platforms where they emerge, and other media platforms as well. Enforcement of new regulations prohibiting deepfakes in elections will be hampered by practical and constitutional problems, whereas public and private investment in timely public education about fake video and audio recordings could help reorient voters back toward the real world.

* S. Walter Richey Professor of Corporate Law, University of Minnesota Law School; Former Associate Counsel to the President and chief White House ethics lawyer (2005-2007).

INTRODUCTION

Deepfakes – computer-generated counterfeit videos and audios of people saying and doing things they never said or did – are proliferating on social media, enabling everything from revenge porn to financial fraud. Legal tools against deepfakes are in their infancy, constrained by practical considerations, enforcement problems and the First Amendment.[1]

This article is about deepfakes in elections. *Citizens United v. FEC* and cases decided in its aftermath have opened the floodgates of dark money funded electioneering communications.[2] In the past few years, deep fake communications – almost invariably showing a candidate saying and/or doing something the candidate did not say or do – have emerged.[3] Deepfakes will almost certainly accelerate in the 2024 election cycle.

A Deepfake for purposes of the discussion in this article is a video and/or audio created or altered using digital means with the aid of artificial intelligence (AI) in which identifiable people realistically appear to do or say things that those people did not do or say.[4]Altered video or audio that does not use digital means with the aid of AI is not deepfake. Merely spliced audio or video tape for example is not a deepfake (splicing audio or video of candidates to leave out qualifying or even contradictory statements is an old trick in elections, but that's not the subject here). The "realistic" element also is important – if a reasonable observer or listener would realize that the content does not accurately depict the person saying or doing something, the content is not a deepfake for purposes of this discussion. Also, the focus here is on video and audio that depicts real people, for example, a candidate for public office or another public figure, not video and audio that uses AI to depict nonexistent people – for example a random AI generated image of a generic police officer saying a candidate is soft on crime. Finally, this definition focuses only on video or audio created without the subject's informed consent. A digitally altered video of a candidate giving a speech in which the candidate's staff has removed blemishes from his face, adjusted his hair or edited out "um. . ." and other awkward phrases from his speech may be misleading but it is not deepfake for purposes of this article.

Deepfakes typically includes AI generated content that swaps a person's face and/or voice with the face and/or voice of other people in an existing video. Before AI technology become generally accessible, similar videos could be

1. *See* Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753, 1768-1804 (2019) (providing a general overview of the legal and policy problems with deepfakes and potential legal remedies).

2. *Citizens United v. FEC*, 558 U.S. 310 (2010). *See also* discussion *infra* text accompanying notes 11-13.

3. *See* discussion *infra* text accompanying notes 42-50 (discussing specific examples of recent deepfake ads).

4. *See* Judge Herbert B. Dixon Jr. (Ret.), *Deepfakes: More Frightening Than Photoshop on Steroids*, ABA (Aug. 12, 2019) ("The term deepfakes comes from the name of a Reddit contributor who surprised the technology community in 2017 when, using publicly available AI-driven software, he successfully stitched or imposed the faces of celebrities onto the bodies of people in pornographic videos.").

created with software such as Photoshop (more advanced Photoshop software now uses digital means with the aid of AI). There may or may not be a significant difference between a deepfake video and a pre-AI Photoshopped video, although the producer of an AI generated deepfake needs to put in less time and the video can be created more quickly. AI technology also enables creation of a complete video without inputting an existing video or audio. A single photo of a public figure – such as an official photo of President Biden – and a sample recording of his voice saying anything might be sufficient to generate a deepfake video of him saying anything the creator wants to have him say. Such video generated completely with AI technology, for example Stable Diffusion, can be very realistic. There are of course differences in performance between different AI technology services. Some content is finer than others; some harder to detect than others. Some AI generated technology is superior to non-AI-generated content and some is not.

Deepfakes are also only a subset of content that can be created by AI. For example, an AI generated video could purport to be a news report about President Biden meeting with Chinese leaders to approve an invasion of Taiwan, but the video might only include actual unaltered video of Biden meeting with Chinese leaders and no audio of his voice; the only audio might be the voice of a "news reporter" purporting to report on what was said at the meeting. A CNN or other news organization's logo might appear on the lower part of the screen, but there would be no attempt to impersonate an actual CNN reporter – only the voice of someone pretending to be a generic news reporter. Such an AI generated video would not be a deepfake within the meaning of the definition used in this article.

Some such AI generated content might be just as concerning as a deepfake for reasons similar to those discussed in this article. The FEC or Congress might choose to treat some such content like deepfake for purposes of the early warning system proposed in this article. On the other hand, a lot of AI created video that is not a deepfake – for example a fantasy news clop about what the world will look like after four more years of Biden's presidency – should be beyond the scope of any government involvement, including the early warning system proposed in this article. This article sets aside the question of whether some AI generated content other than deepfakes should be added to the list of content subject to the early warning system proposed here. A regulatory response to such content may or may not be appropriate. The focus of the discussion here will be deepfakes as defined above.

Likewise, there is some non-AI generated content such as pre-AI Photoshop that pose problems like deepfakes and that arguably should be treated like AI generated deepfakes for purposes of this article, particularly if the only official action taken is the early warning proposed in this Article. Congress and the FEC can make that determination. Other non-AI generated content – for example a cartoon of a candidate – should not be subject to government oversight or intervention. An FEC warning Americans about a misleading cartoon of the president would not only be excessive but arguably an illegal and perhaps unconstitutional use of government power. This article sets aside non-AI generated content impersonating or

depicting a public figure to focus on what the FEC should do in the case of AI generated deepfakes.

A regulatory regime that effectively addresses a problem – whether with prohibitions, disclosure requirements or the early warning system proposed in this article – needs to adopt some definition of the problem that will be subject to the regime proposed. The above definition of a deepfake is what this article proposes, and Congress or the FEC might adjust this definition or use a different definition. Definitions, whether the definition of a deepfake or the definition of a security in the federal securities laws, do not always have precise lines and are sometimes subject to contested ex-post interpretation. At the same time, a working definition is needed in a statute, regulation, or policy to proceed to the next step, which is the focus of this article – what to do about deepfakes that fall within this working definition.

The Federal Election Commission (FEC) has been asked by public interest organizations and Members of Congress to do something about deepfakes but has deadlocked on whether to act.[5] Bills are pending in Congress to address the problem, but some of these bills are overbroad and rely on criminal sanctions, exacerbating constitutional problems.[6] In any event, no bill addressing deepfakes in elections has passed either house.

In 2024 voters may be bombarded with images of virtual candidates that do not exist, but that so closely replicate actual candidates that it's impossible to tell the difference. Although it is difficult to know how much money is spent on electioneering communications because only some of it is disclosed, a significant portion of those expenditures could be dedicated to disseminating deepfake ads. Because the United States does not have a system for disclosing dark money in politics, voters will have no idea where these ads are coming from.

Deepfakes are also a national security concern. Our experience with foreign interference in elections, typified by Russian computer hacking and impersonation in the 2016 election,[7] could replicate itself multifold in 2024 with the assistance of AI, so voters don't know who the candidates really are or who's supporting or opposing them. Americans will enter a virtual world resembling a video game where, for over a year, we will watch two teams of players – red and blue – fight primary battles on our screens and then fight each other in a general election. When voters finally get their chance to play, they will make decisions in a single instance – or two instances if they vote in a primary – based on what they think they know about the game at the given moment when they are allowed to play. Winners will be announced, and perhaps take office in a peaceful transition of power, although yet more deepfakes could support efforts to overturn the election result, perhaps even fomenting the type of violence that in 2021 was a feature of post-election-season play.

5. *See* discussion *infra* text accompanying notes 64-65.
6. *See* discussion *infra* text accompanying notes 67-68.
7. *See* discussion *infra* text accompanying notes 27-37.

Part I of this Article outlines the more general problem – unlimited and undisclosed dark money in politics, the almost nonexistent legal remedies for foreign interference in U.S. elections, and AI-generated deepfake video and audio that will inundate voters from 2024 onwards. Part II explores possible legal remedies for deepfakes in elections, none of which are sufficiently robust to address the problem effectively and, with deadlock in both the FEC and Congress, are not likely to become law anyway. Part III explores an alternative: publicly and privately funded Deepfake Warnings that could respond quickly to deepfake electioneering communications by identifying them as such in the same social media platforms where they emerge, and other media platforms as well. This proposal includes an important caveat that publicly endorsed Deepfake Warnings should not respond to the misleading substance of a deepfake ad other than labeling it for what it is – a fake. The FEC also should facilitate rapid issuance of these warnings by establishing an online Deepfake Alert System where deepfake reports can be posted and alleged deepfake content can be quickly analyzed by a panel of computer scientists who announce preliminary findings as soon as possible and a more detailed assessment not long thereafter.

I. CORPORATE MONEY, FOREIGN INTERFERENCE, AND DEEPFAKES IN ELECTIONEERING COMMUNICATIONS

A. Corporate Money in Elections

Corporate money should not influence elections, or so Congress thought over a hundred years ago in 1907 in passing the Tillman Act, which prohibits donations from corporate treasuries to political campaigns.[8] Ever since then, political operatives have exploited loopholes to get around the law. Business interests through trade associations, political action committees, and whatever other means by which they can spend unlimited amounts influencing elections. Wealthy individuals, labor unions, issue advocacy organizations and others join in with "independent expenditures" of their own.[9] Members of Congress, themselves the beneficiaries of this largess, usually do nothing to stop this flood of money in politics and it is difficult to get the House and Senate to agree on a bill that would expose where the money is coming from.[10]

When in rare instances Congress does something to rein in political spending, courts strike down key provisions. Caps on spending by campaigns were struck down in *Buckley v. Valeo*, although caps on individual donations to campaigns

8. *See* Tillman Act of 1907, ch.420, 34 Stat. 864 (1907) (codified as amended at 2 U.S.C. § 441b(a)), *invalidated by Citizens United*, 558 U.S. at 310 (2010).

9. 52 U.S.C. § 30101(17) (2014) ("The term 'independent expenditure' means an expenditure by a person –(A) expressly advocating the election or defeat of a clearly identified candidate; and (B) that is not made in concert or cooperation with or at the request or suggestion of such candidate, the candidate's authorized political committee, or their agents, or a political party committee or its agents.").

10. *See, e.g.*, For the People Act of 2021, H.R. 1, 117th Cong. (1st Sess. 2021), §§ 4206-10 (imposing restrictions and disclosure requirements on electioneering communications, which passed the House but stalled in the Senate).

were upheld.[11] *Citizens United v. Federal Election Commission* struck down a key provision of a 2002 law co-sponsored by Senators John McCain (R AZ) and Russ Feingold (D WI) and signed into law by President George W. Bush.[12] The Court ruled that unlimited expenditures from corporate treasuries on electioneering communications[13] were constitutionally protected speech because corporations are persons under the law just as are natural persons.[14]

Federal courts have not yet struck down the Tillman Act prohibition on direct corporate contributions to political campaigns, but corporate funded independent civic organizations,[15] PACs, and Super PACs have First Amendment protection.[16] The current state of the law is that, while some restrictions on direct contributions to campaigns and political parties are upheld, spending on electioneering communications and on Super PACs is free speech – the sky's the limit.[17] As Justices Sandra Day O'Connor and John Paul Stevens famously wrote in *McConnell v. Federal Election Commission*, one of the few cases upholding campaign finance laws, "money, like water, will always find an outlet."[18] And so it does.

Campaign finance law – what's left of it – is also dominated by strategies to get around the law. Like manipulators who exploit loopholes to get around taxation and environmental regulation, election lawyers combine legal but dubious law avoidance with illegal law evasion – a practice known as law "avoision." This term, introduced by free-market-oriented London School of Economics

11. Buckley v. Valeo, 424 U.S. 1 (1976) (holding that dollar limitations on contributions by individuals to campaigns do not violate the First Amendment but that limitations on spending by political campaigns do violate the First Amendment), *superseded by statute*, Bipartisan Campaign Reform Act of 2002, Pub. L. No. 107-155, 116 Stat. 81, 82, *as recognized in* McConnell v. FEC, 540 U.S. 93, 94 (2003).

12. *Citizens United*, 558 U.S. at 310.

13. An electioneering communication is "any broadcast, cable, or satellite communication" that "refers to a clearly identified candidate for Federal office" made within 60 days before a general election or 30 days of a primary election and that is "targeted to the relevant electorate" if for an office other than President or Vice President. 52 U.S.C. §30104(f)(3)(A).

14. *See Citizens United*, 558 U.S. at 343 ("The Court has thus rejected the argument that political speech of corporations or other associations should be treated differently under the First Amendment simply because such associations are not 'natural persons.'").

15. Many of these organizations, like Citizens United itself, are established under Section 501(c)(4) of the Internal Revenue Code. *See* 26 U.S.C.S. § 501(c)4 (establishing tax exempt status for "[c]ivic leagues or organizations not organized for profit but operated exclusively for the promotion of social welfare . . .. [and] devoted exclusively to charitable, educational, or recreational purposes").

16. *See* Lieu v. FEC, No. 19-5072, 2019 U.S. App. LEXIS 29880, at *2 (D.C. Cir. 2019) (dismissing suit brought against the FEC by Rep. Ted Lieu, Rep. Walter Jones, Sen. Jeff Merkley, State Sen. John Howe, Zephyr Teachout, and Michael Wager asking the Circuit Court to overturn its interpretation of *Citizens United*, 588 U.S. 310, that allows unlimited spending on Super PACs, as found in *SpeechNow.org v. FEC*, 599 F.3d 686 (D.C. Cir. 2010) (en banc), *cert. denied*, 562 U.S. 1003 (2010)).

17. *See* McConnell v. FEC, 540 U.S. 93, 224 (2003) (upholding limits on soft money contributions used to register voters and increase attendance at the polls); McCutcheon v. FEC, 572 U.S. 185, 227 (2014) (striking down aggregate limits on donor contributions to multiple candidates).

18. *McConnell*, 540 U.S. at 224.

professors in a 1979 book Tax Avoision,[19] is now common usage, popularized on the television show *The Simpsons*.[20] Getting around the law is all too often a response to regulation, particularly by people who are skeptical of government regulation to begin with and have resource to find a way around it. Avoision lawyers, for a fee, often lend a helping hand.[21]

Such strategies for getting around the law have spread from the private sector into the public arena including campaign finance and funding of electioneering communications. The recent criminal trial of crypto mogul Sam Bankman-Fried, for example, included not just allegations of fraud on investors but also attempts to conceal millions of dollars in electioneering expenditures and campaign contributions.[22] Whatever legal restrictions are imposed – whether restrictions on coordination between campaigns and independent expenditure organizations, restrictions on foreign funded electioneering communications, or prohibitions on deepfakes – an army of well-paid political operatives and their lawyers can find a way around the law.

B. Foreign Interference in Elections

In *Bluman* v. *FEC*, Judge Brett Kavanaugh, sitting for the federal district court in Washington, D.C. held that the First Amendment protections in *Citizens United* do not apply to electioneering expenditures by foreign entities.[23] Foreign nationals have no constitutional right to spend or contribute money in connection with U.S. elections. The ruling was affirmed without an opinion by the Supreme Court.[24] This means, for now at least, the FEC and Congress can at least try to prevent electioneering expenditures by entities controlled by foreign nationals.

Good luck.

Enforcing a ban on foreign-funded electioneering expenditures is difficult given the close and often concealed ties between American corporations and business entities overseas. Also, Judge Kavanaugh's ruling in *Bluman*, that American corporate money is constitutionally protected in U.S. elections but foreign money is not, draws a distinction that is unusual in First Amendment

19. ALFRED ROMAN ILERSIC, , TAX AVOISION: THE ECONOMIC, LEGAL, AND MORAL INTER-RELATIONSHIPS BETWEEN AVOIDANCE AND EVASION (Arthur Seldon ed., 1979).

20. *The Simpsons: Bart The Fink* (Fox television broadcast Feb. 11, 1996).

21. Avoision in the private sector works the same as avoision in the public sector, and thus an election lawyer (engaging in avoision) is performing a similar function as a corporate lawyer (engaging in avoision). *See* Richard W. Painter The *Moral Interdependence of Corporate Lawyers* and Their Clients. 67 S. Calif. L. Rev. 507-584 (1994) (discussing strategies lawyers use to help business clients get around the law).

22. *See* Indictment at 10-12, United States v. Sam Bankman-Fried, 22 Crim 673 (S.D.N.Y. 2023), https://perma.cc/7Z9F-LPY2 (Count Eight: Conspiracy to Defraud the United States and Violate Campaign Finance Laws).

23. Bluman v. FEC, 800 F. Supp. 2d 281, 288-89 (D.D.C. 2011).

24. Bluman v. FEC, No. 11-275, 2012 U.S. LEXIS 310, at *1 (Jan. 9, 2012) (summarily affirming the three-judge court's decision to grant the Commission's Motion to Dismiss and to deny plaintiffs' Motion for Summary Judgment simply stating "the judgment is affirmed") .

jurisprudence which generally covers foreigners within U.S. borders.[25] Some foreign nationals will perceive this as unjust and seek ways to get around the law. Under *Citizens United* there is a constitutionally protected money party in American politics, but the "Kavanaugh rule" says foreign nationals aren't invited.[26] They will want to come anyway, and will find a way to get in.

The usual way for foreign money to enter U.S. elections is to coordinate with American businesses and individuals that are funding electioneering communications. The money may originate from abroad, but if an American entity makes the spending decision, or the money is at least routed through an American entity, it appears to be constitutionally protected speech. "Straw donor" arrangements can be prosecuted in the case of contributions to campaigns themselves, but are apparently legal and virtually impossible to police in the shadowy world of dark money electioneering expenditures.[27] Whether funneling money through U.S. business joint venturers, wholly owned corporate subsidiaries, consultants, lobbyists, or even lawyers, foreign entities will find a way to join the money party hosted by American political operatives working under the protection of the United States Supreme Court.[28]

Then there's illegal, indeed criminal, foreign interference in U.S. elections in circumstances where the persons responsible almost certainly will never be apprehended.

In 2016 Russia crashed the party. Part I of the Mueller Report,[29] and the indictments of 12 Russian intelligence officers in 2018,[30] show how easy it was for foreign nationals to have a dramatic, perhaps even decisive effect on a U.S. presidential election. Through the clandestine Internet Research Agency, Russian intelligence agents set up accounts on Facebook, Twitter, and other social media platforms, and used false pretenses, impersonation, and other strategies to spread

25. In other contexts, the Supreme Court has ruled that "resident aliens have First Amendment rights." United States v. Verdugo-Urquidez, 494 U.S. 259, 271 (1990); Bridges v. Wixon, 326 U.S. 135, 148 (1945).

26. *See Citizens United*, 558 US at 310.

27. *See, e.g.*, Indictment at 2-3, United States v. D'Souza, No. 14-00034-RMB (S.D.N.Y. Jan. 23, 2014), https://perma.cc/U3LS-2VQM (alleging straw donor scheme to exceed donor limits in U.S. Senate campaign against Hillary Clinton). D'Souza was convicted, sentenced to probation but later pardoned. U.S. DEP'T OF JUST., EXECUTIVE GRANT OF CLEMENCY FOR DINESH D'SOUZA (May 31, 2018), https://perma.cc/PGA6-X2NA (pardoning D'Souza for his conviction in the S.D.N.Y. under 2 U.S.C. §§ 441f and 437g(d)(l)(D) and 12 U.S.C. § 2).

28. *See* RICHARD W. PAINTER, TAXATION ONLY WITH REPRESENTATION: THE CONSERVATIVE CONSCIENCE AND CAMPAIGN FINANCE REFORM, 86-105 (2016).

29. ROBERT MUELLER, U.S. DEP'T OF JUST., OFFICE OF SPECIAL COUNSEL, REPORT ON THE INVESTIGATION INTO RUSSIAN INTERFERENCE IN THE 2016 PRESIDENTIAL ELECTION, vol. 1, at 14-35 (April 18, 2019) (discussing the Russian social media campaign in the 2016 election led by the Russian Internet Research Agency, which controlled Facebook, Twitter and other social media accounts and used false pretenses, impersonation, and similar strategies to spread disinformation).

30. Press Release, U.S. DEP'T OF JUST., Grand Jury Indicts 12 Russian Intelligence Officers for Hacking Offenses Related to the 2016 Election (July 13, 2018), https://perma.cc/RD4B-G95E.

disinformation.[31] Deepfakes were not used, but impersonation was common. Russian Facebook and Twitter trolls took on fake identities reaching millions of Americans.[32] The impersonation was fraudulent – and sufficiently so to result in criminal charges and convictions of the Russian agents (in abstention), even though none of these defendants were apprehended, and Special Counsel Mueller did not find evidence sufficient to charge anyone in the Trump campaign, or any other American for conspiring with them.[33]

The Justice Department press release accompanying the indictments of the Russian agents discussed specific examples of what happened:

> "On the website, defendants claimed to be "American hacktivists" and used Facebook accounts with fictitious names and Twitter accounts to promote the website. After public accusations that the Russian government was behind the hacking of DNC and DCCC computers, defendants created the fictitious persona Guccifer 2.0. On the evening of June 15, 2016 between 4:19PM and 4:56PM, defendants used their Moscow-based server to search for a series of English words and phrases that later appeared in Guccifer 2.0's first blog post falsely claiming to be a lone Romanian hacker responsible for the hacks in the hopes of undermining the allegations of Russian involvement."[34]

Cryptocurrency was used to conceal funding. Again, as the Department of Justice noted:

> "To avoid detection, defendants used false identities while using a network of computers located around the world, including the United States, paid for with cryptocurrency through mining bitcoin and other means intended to obscure the origin of the funds. This funding structure supported their efforts to buy key accounts, servers, and domains. For example, the same bitcoin mining operation that funded the registration payment for DCLeaks.com also funded the servers and domains used in the spearphishing campaign."[35]

Apart from the computer hacking charges for stealing email and other documents, the criminal charges focused on concealment, fraudulent misrepresentation and impersonation: Count One charged conspiracy to defraud the United

---

31. MUELLER, *supra* note 29. Vol. I at 15-28 (describing activities of the Russian Internet Research Agency)

32. *Id*. at 15 ("In November 2017, a Facebook representative testified that Facebook had identified 470 IRA-controlled Facebook accounts that collectively made 80,000 posts between January 2015 and August 2017. Facebook estimated the IRA reached as many as 126 million persons through its Facebook accounts. In January 2018, Twitter announced that it had identified 3,814 IRA-controlled Twitter accounts and notified approximately 1.4 million people Twitter believed may have been in contact with a an IRA-controlled account.") (citations omitted).

33. MUELLER, *supra* note 29, at 2 (finding that the Trump campaign did not coordinate, within the meaning of the federal criminal code, with the Russian government in is election interference activities).

34. U.S. DEP'T OF JUST., *supra* note 30.

35. *Id.*

States[36] by concealing involvement of foreigners in U.S. elections through violations of FEC regulations, the Foreign Agents Registration Act (FARA) and other provisions; Counts Two through Nine charged "aggravated identity theft for using identification belonging to eight victims to further their computer fraud scheme."[37] Count Ten alleged a conspiracy to launder money in which the defendants "laundered the equivalent of more than $95,000 by transferring the money that they used to purchase servers and to fund other costs related to their hacking activities through cryptocurrencies such as bitcoin[.]"[38]

Impersonation was part of this scheme. One of many examples: Paragraph 36 of the Indictment refers to a Twitter Account, "@TEN_GOP," which was an impersonation of the Tennessee Republican Party and attracted more than 100,000 online followers.[39]

These indictments illustrate the legal prohibitions on impersonation and fraud used to gain access to social media platforms for the purpose of influencing elections. The fact that none of the indicted Russian agents showed up in an American courtroom, and none have been apprehended, however, shows how difficult enforcement can be. Part I of the Mueller Report found insufficient evidence to charge anyone in the Trump campaign, or any American, with criminal conspiracy, although several people, including Trump's top national security advisor, Michael Flynn, were convicted of lying about their contacts with the Russians in other contexts.[40] The upshot is that fraudulent electioneering communications originating from outside the United States may involve criminal activity, but are extremely difficult to control and it is likewise extremely difficult to hold anyone inside the U.S. criminally accountable. American political operatives, campaigns and even candidates may know about this fraudulent activity by foreign nationals (Donald Trump even publicly asked Russia to hack Hillary Clinton's email),[41] but support is not the same as criminal conspiracy and proving the elements of a criminal conspiracy is hard.

C. Deepfakes in Elections

The 2024 election cycle will bring a new type of electioneering communication to the fore – "deepfaking." Deepfakes use artificial intelligence and computer

36. *See* 18 U.S.C. § 371 (stating that offense occurs "[i]f two or more persons conspire either to commit any offense against the United States, or to defraud the United States, or any agency thereof in any manner or for any purpose, and one or more of such persons do any act to effect the object of the conspiracy").

37. U.S. DEP'T OF JUST., *supra* note 30.

38. *Id.*

39. Indictment, ¶ 36, United States v. Internet Rsch. Agency L.L.C., No. 18-00032 (D.D.C. Feb. 16, 2018).

40. MUELLER, *supra* note 29.

41. Michael Schmidt, *Trump Invited the Russians to Hack Clinton. Were They Listening?*, N.Y. TIMES (July 18, 2018) (quoting a July 2016 news conference in which Trump said "Russia, if you're listening, I hope you're able to find the 30,000 emails that are missing. .. I think you will probably be rewarded mightily by our press.").

imagery to manipulate an image of a real person. Voice clones dubbed into the video make it seem authentic.

As the watchdog group Public Citizen explains:

> "Extraordinary advances in artificial intelligence now provide political operatives with the means to produce campaign ads and other communications with computer-generated fake images, audio or video of candidates that appear real-life, fraudulently misrepresenting that what candidates say or do. Generative artificial intelligence and deepfake technology – a type of artificial intelligence used to create convincing images, audio and video hoaxes – is evolving very rapidly. Every day, it seems, new and increasingly convincing deepfake audio and video clips are disseminated, including, for example, an audio fake of President Biden, a video fake of the actor Morgan Freeman and an audio fake of the actress Emma Watson reading Mein Kampf."[42]

Days before Chicago's 2023 mayoral election, a deceptive impersonation video of candidate Paul Vallas was posted to Twitter making him appear to be saying that back in his day, "cops would kill 17 or 18 people and nobody would bat an eye" and that Chicago needed to "refund the police."[43] The deepfake apparently was viewed thousands of times before it was taken off Twitter.[44] Ron DeSantis's presidential campaign also used deepfake images to attack Donald Trump, including a video purporting to show Trump hugging and kissing Dr. Anthony Fauci.[45]

The phenomenon is global. On February 7, 2020, a day before legislative assembly elections in Delhi, India deepfake videos of Bharatiya Janata Party (BJP) President Manoj Tiwari criticizing the incumbent government in Delhi went viral on WhatsApp.[46]

Deepfakes are distinguishable from other uses of AI, for example images of real places and unidentified people acting in a fictional world. A Republican Party AI-created ad represents the future if President Joe Biden is re-elected in 2024 – including scenes of China invading Taiwan, a Wall Street crash, immigrants flooding the border at the Rio Grande and San Francisco being overrun by crime and drugs.[47] Such is not the same as a deepfake in which an identifiable person is digitally impersonated doing or saying something they did not do or say. These ads are essentially AI created scenes not that different from commercials, movies, and video games. In an election, they are almost certainly constitutionally

42. Letter from Robert Weissman, President, Pub. Citizen, to Lisa J. Stevenson, Acting Gen. Couns., Fed. Election Comm'n (July 13, 2023), https://perma.cc/4YLZ-X8BR.

43. Megan Hickey, *Vallas Campaign Condemns Deepfake Video Posted to Twitter*, CBS NEWS CHICAGO (Feb. 27, 2023, 6:57 PM), https://perma.cc/644T-5RYE.

44. *Id.*

45. Nicholas Mehamas, *DeSantis Campaign Uses Apparently Fake Images to Attack Trump on Twitter*, N.Y. TIMES (June 8, 2023), https://perma.cc/6QBU-UW3J.

46. Nilesh Christopher, *We've Just Seen the First Use of Deepfakes in an Indian Election Campaign*, VICE (Feb. 18, 2020, 7:27 AM), https://perma.cc/6NPR-Z5C2.

47. GOP, *Beat Biden*, YOUTUBE (April 25, 2023), https://perma.cc/N5ED-HLQ7.

protected speech, although as discussed in the next Part of this article there are proposals that the FEC require them to be labeled as being created with AI.

More problematic, but still very likely protected speech, would be ads that depict past events that didn't happen. This could include for example a scene of the U.S.-Mexico border with hordes of immigrants violently crossing and assaulting border patrol agents with an actual past date of the event appearing in the picture frame, even if the scene didn't happen on that date but instead was created by AI. This ad is not that different from actors reenacting an event differently from what occurred. To varying degrees, almost all documentaries mix actual news footage with reenactments of some sort, although AI makes it much harder to distinguish between the news footage and the reenactment.

Because AI is cheaper than using actors such ads will proliferate (and AI doesn't talk about it afterwards as actors might do).[48] Campaigns already use social media platforms and other data to profile individuals based on issues they care about and AI created ads can be targeted to audiences where they are most likely to be effective.[49] Some of these AI created factual representations may be lies about events that never happened, but like most fantasy, they are probably protected speech. The fact that some people vote on the assumption that these ads depict reality doesn't change the historically broach application of the First Amendment to political speech.

In fact, a lot of AI-generated ads are not that different from traditional campaign ads, with the exception that the actual and the virtual are harder to distinguish. Many scenes in campaign ads are fictional, going back to the famous "Daisy girl" ad that Lyndon Johnson used against Barry Goldwater in 1964 depicting a little girl counting daisy petals in a field right before the world is blown up by atomic bombs, a clear message from LBJ's campaign about what could have happened if Goldwater won the election.[50] The ad was created by the Doyle Dane Bernbach agency, and LBJ's voice ended the ad pronouncing that we must "love each other or die." (There was of course no mention of LBJ's escalation of the Vietnam War). Goldwater, who was against the 1963 Nuclear Test Ban Treaty was never mentioned, and there was no picture of him, but the Daisy

48. Stuart Thompson, *Making Deepfakes Gets Cheaper and Easier Thanks to A.I.*, N.Y. TIMES (Mar. 12, 2023) ("The content they produce, sometimes called cheapfakes by researchers, work by cloning celebrity voices, altering mouth movements to match alternative audio and writing persuasive dialogue.") (Citing Britt Paris & Joan Donovan, *Deep Fakes and Cheap Fakes: The Manipulation of Audio and Visual Evidence*, DATA & SOCIETY (Sep. 18, 2019), https://perma.cc/4PMT-FN38).

49. Minami Funakoshi, Elizabeth Culliford, & Wen Foo, *How Political Campaigns Use Your Data*, REUTERS GRAPHICS (Oct. 12, 2020), https://perma.cc/K9NB-QQ5G (illustrating specific data collection and analysis methods by which "political campaigns use data on more than 200 million voting-age Americans to inform their strategies and tactics").

50. Robert Mann, *LBJ's Mad Men*, POLITICO (Sept. 7, 2014), https://perma.cc/7TVM-EKJV ("[There was] [n]o need to say, 'Barry Goldwater will blow up the world if he's elected president.' The right images would prompt viewers to provide that message themselves.").

Girl message was loud and clear.[51] That was sixty years ago. It will be very hard for AI to top that ad for effective impact on voters (LBJ won in a landslide).

But in one way AI can top the Daisy Girl ad, and it is here that First Amendment protections may give way to legal restrictions on fraudulent impersonation. Imagine an ad depicting film footage of Goldwater saying that he would like to "nuke Russia" and "I don't care if a hundred million Americans die." Unless Goldwater said that and was recorded saying that, without AI it would be extremely difficult to make a realistic ad in which Goldwater says that. A look-alike actor could be used, but it would soon be clear that the actor wasn't Goldwater. That's the difference between 1964 and 2024. Deepfakes bring an additional layer of confusion to elections because absent a disclaimer it is virtually impossible for a viewer to tell the difference between the actual candidate and the virtual candidate. The deepfake ad goes one step further than the Daisy Girl ad, and a potentially dangerous step for the integrity of elections because it's one thing to say that Goldwater would risk nuclear war if elected President; it would be quite another to depict a video of him saying "nuke Russia" if he didn't say that.

Now let's vary this hypothetical AI-generated Daisy Girl ad to explore a greyer area. Barry Goldwater frequently did talk about use of nuclear weapons and apparently did say he wanted to "lob one [a bomb] into the men's room in the Kremlin."[52] But let's say that statement wasn't caught on tape at the time he said it. AI – if LBJ had had AI in 1964 – could have recreated that moment depicting Goldwater saying what he said. The AI-generated ad would be an impersonation of Goldwater's voice – it would not be a real recording of him saying it – but the ad would not be as misleading as it would be if Goldwater had never said it. There are variations on this scenario too – an AI-generated video could alter what Goldwater said slightly (he apparently said "lob *one* into the men's room at the Kremlin" but the ad would be more effective if it said "lob *a nuke* into the men's room at the Kremlin," which would not be entirely misleading if that's what Goldwater presumably meant).

Some AI-generated ads would be more misleading than others, but they share a common characteristic. They replicate a candidate's voice, and often also the candidate's image, as a substitute for a real time audio/visual recording of the candidate. That is a line that should not be crossed if campaign ads are to have any connection with reality. Even in instances where there is evidence that the candidate did say the same thing, or very much the same thing, the deepfake impersonation is still fraudulent. Evidence the candidate said it might be disputed, and subtle alterations can still change meaning. Videos and audios of a person that are not in fact recordings of that person are fraudulent impersonation.

51. Anthony Lewis, *Goldwater Says Test Ban Creates Illusion Of Peace; Tells Senate He Will Risk 'Political Suicide' to Vote Against Ratification Notes Political Factors Hruska Favors Pact Goldwater Says Test Ban Pact Offers Only an Illusion of Peace*, N.Y. TIMES (Sep. 20, 1963).

52. *Id.*

Whether there is a *material* difference between what is depicted in the deepfake ad and what the candidate did or said might be debatable, as it would be in several of the above AI-generated Goldwater hypotheticals. Inquiry into materiality of a misrepresentation involves comparing the substance of the ad with the substance of what the candidate did and spoke. A deepfake of Barry Goldwater saying "nuke Russia" might not be as much of a material misrepresentation as a deepfake of a pacifist candidate saying "nuke Russia."

Such distinctions, however, are hardly a sound basis for government intervention. The government cannot inquire into the material impact on voters of a deepfake misrepresentation without getting embroiled in content-based regulation of political speech. For the FEC or any other regulator, all deepfake impersonations of a candidate probably should be treated similarly, even if the impact on an election could vary widely depending on the candidate and the content of the deepfake. The political content of the ad should be dealt with as it always has been – by the candidate or independent groups producing their own communication rebutting that which they believe to be misleading. But the FEC can identify it as a deepfake ad made with AI technology, and as discussed in Part III of this article, the FEC should do just that.

One more variation on the deepfake theme is worth exploring – a genuine audio of a candidate combined with altered video. For example, the genuine audio of Donald Trump saying to Billy Bush on a bus, "Grab em by the p——,"[53] could be combined with a genuine video of Trump in a cabinet meeting with women cabinet members present, or perhaps standing at the pulpit of a church. A traditional pre-AI campaign ad might have played the recording of Trump's voice saying "grab 'em by the p——" with a video recording of Trump in the cabinet room or church, but it would be obvious that the two were different in time and space. The AI-generated ad, however, would converge the two so viewers would believe they were watching Trump saying "grab 'e, by the p——" in the cabinet room or church.

Deepfakes here too are misleading. As disgusting as it was for Trump to boast about sexual assault to a man on a bus, it would be even worse for him to say the same thing in a cabinet meeting or a church. That would portray Trump not only as having a violent attitude toward women but also as having a psychological problem of being unable to discern how to behave in public. Other genuine audio recordings and videos of Trump in public perhaps could be used to make that second point but superimposing the "grab 'em by the p——" audio onto video of him talking in a cabinet meeting or a church would be materially misleading. That could change decisions of a subset of voters who didn't care about Trump saying what he said on a bus but would not like him saying it elsewhere.

53. *Transcript: Donald Trump's Taped Comments About Women*, N.Y. TIMES (Oct. 8, 2016), https://perma.cc/BW8J-DNTH. (audio recording of Donald Trump and Billy Bush talking and simultaneous video recording from the outside of the bus).

Still, however, this presumed impact on voters' perceptions of Trump, while relevant for assessing the harmfulness of such an AI-generated ad, is not an appropriate inquiry for the FEC or any regulatory regime to use to define deepfakes or to decide whether to act. Otherwise, regulation of deepfakes would be embroiled in content-based inquiry and invite challenges under the First Amendment. An AI-generated video of Trump sitting on the bus next to Billy Bush saying these words would still be a deepfake because the AI-generated video would not be an actual video of Trump inside the bus (there apparently isn't one).[54] The operative inquiry for defining a deepfake thus is whether a reasonable observer of the video would believe they were watching an actual video of Trump talking to Billy Bush on the bus, not whether a reasonable voter would change their mind about what Trump said on the bus or about his fitness for the presidency.[55]

II. CAN THE LAW KEEP DEEPFAKES OUT OF ELECTIONS?

A. Defamation, Tort, and Revenge Porn Law

Political cartoons, many quite vicious, have a long history and are protected speech.[56] Actors have been impersonating politicians for centuries in plays and skits and more recently on Saturday Night Live.[57] Virtually all of this is protected speech, unless a cause of action for defamation can be made, which under *New York Times v. Sullivan*[58] is very difficult for a public figure.

There is, however, a difference between an impersonation of a person by an actor, even a very good look alike, and an impersonation created by AI that is so real that a reasonable person would likely confuse it for the real person. While

54. The definition of deepfake in a bill recently introduced in Congress, the Deep Fakes Accountability Act, however, might not cover this example because Trump did in fact say these words on the bus. *See* Deep Fakes Accountability Act, H.R. 2395, 117th Cong. (2021), https://perma.cc/DLW8-SSP6 (defining "Deepfake"). This would still be a deepfake, however, because there is no known genuine video of Trump inside the bus, and in any event, this wouldn't be it. A law imposing criminal penalties, as would the Deep Fakes Accountability Act if enacted, arguably should not prohibit AI created videos of people saying what they did in fact say. *Id.* A regulatory regime that focuses instead on flagging deepfakes should include this type of AI-created audio-video so all deepfakes are treated similarly. *See* discussion *infra* Part III.

55. Evaluation of a deepfake's materiality premised on voter impact and a reasonable voter standard is confusing and perhaps irrelevant if voters make decisions based on emotion or other factors difficult to incorporate into a workable definition of reasonableness. The definition of a deepfake used in this Article is premised on a narrower inquiry as to whether a reasonable viewer of the video and audio would believe they are watching an actual recording of the real thing.

56. *See, e.g.*, King Andrew the First (illustration), *in Political Cartoons and Public Debate*, LIBRARY OF CONGRESS (1833), https://perma.cc/76XT-7PNE (depicting a caricature of Andrew Jackson as a despotic monarch dressed up in royal robes).

57. Watchmojo.com, *Top 10 Funniest Presidential Impersonations on SNL*, YOUTUBE (Jan. 16, 2021), https://perma.cc/3PB6-X8GQ.

58. New York Times v. Sullivan, 376 U.S. 254, 283-84 (1964) (holding that a suit for libel of a public figure requires a showing of actual malice, which means the defendant either knew the statement was false or showed a reckless disregard for the truth).

there are not yet federal laws explicitly prohibiting such deepfake images, existing law may provide some remedies.

Defamation law in one avenue of relief. Although defamation is easier to plead for plaintiffs who are not public figures, someone who publishes something about a public figure with malice and knowing it to be false, can be sued for defamation.[59] Some deepfake ads attacking political candidates might meet this standard. For example, a deepfake ad showing a candidate accepting a cash bribe from a foreign leader might be made with sufficient malice and disregard for the truth for the ad to be libelous, particularly if there was no credible evidence that the candidate had in fact received a bribe from the foreign leader.

Another possibility is a suit for on-line harassment or a similar tort. The Supreme Court recently held that criminal statutes prohibiting on-line threats must be narrowly construed.[60] Civil suits for harassment might have more leeway, but probably less so in the case of a candidate for public office.

A civil or criminal cause of action for revenge porn is also a possibility if pornographic images are involved. Most laws prohibiting "revenge porn" are premised on the victim expecting privacy and participating in creation of the image,[61] but some states have amended their criminal statutes to include deepfake porn – AI created images of identifiable individuals engaging in sex acts.[62] A bill has been introduced in Congress, the Preventing Deepfakes of Intimate Images Act,[63] which would create a federal cause of action for a "Depicted Individual" who can sue for damages, including liquidated damages of $150,000 for deepfake porn.[64]

---

59. *See id.* at 272.

60. Counterman v. Colorado, 143 s. Ct. 2106, 2113 (June 27, 2023) (holding that for purposes of criminal prosecution for online harassment, to establish that a statement is a "true threat" unprotected by the First Amendment, the State must prove that the defendant had a subjective understanding of the statements' threatening nature, based on a knowledge standard higher than recklessness).

61. *See, e.g.*, Minn. Stat, § 617.261 (2022) ("It is a crime to intentionally disseminate an image of another person who is depicted in a sexual act or whose intimate parts are exposed, in whole or in part, when: (1) the person is identifiable: (i) from the image itself, by the person depicted in the image or by another person; or (ii) from personal information displayed in connection with the image; (2) the actor knows or reasonably should know that the person depicted in the image does not consent to the dissemination; and (3) the image was obtained or created under circumstances in which the actor knew or reasonably should have known the person depicted had a reasonable expectation of privacy.")

62. *See, e.g.*, Va. Code, § 18.2-386.2. (2014) ("Any person who, with the intent to coerce, harass, or intimidate, maliciously disseminates or sells any videographic or still image created by any means whatsoever that depicts another person who is totally nude, or in a state of undress so as to expose the genitals, pubic area, buttocks, or female breast, where such person knows or has reason to know that he is not licensed or authorized to disseminate or sell such videographic or still image is guilty of a Class 1 misdemeanor. For purposes of this subsection, 'another person' includes a person whose image was used in creating, adapting, or modifying a videographic or still image with the intent to depict an actual person and who is recognizable as an actual person by the person's face, likeness, or other distinguishing characteristic.")

63. Preventing Deepfakes of Intimate Images Act, H.R. 3106, 188th Cong. (2023).

64. The bill provides that the term "depicted individual means an individual who, as a result of digitization or by means of digital manipulation, appears in whole or in part in an intimate digital depiction and who is identifiable by virtue of the person's face, likeness, or other distinguishing characteristic, such as a unique birthmark or other recognizable feature, or from information displayed in connection with the digital depiction." *Id.*

B. Campaign Finance Law and AI

The FEC is obligated to safeguard the First Amendment when implementing its statutory directives.[65] FEC regulations require disclaimers on political ads, disclosing whether a candidate or some other organization paid for the ad.[66] So long as disclaimers about who paid for the ad are visible and audible (depending on the mode of communication), the FEC doesn't regulate the content of the ad. An ad that lies about an opposing candidate's record, or the candidate's past, or the candidate's views, is not and probably cannot be prohibited by the FEC .[67] A civil suit for defamation is possible but can only succeed if the high standard for pleading libel against a public figure can be met, and there are no FEC regulations that prohibit libel, which is a private right of action that is difficult to bring for political candidates.[68]

Federal law prohibits fraudulent misrepresentation amounting to impersonation of an opposing candidate, but thus far this rule has not been applied to AI created images or deepfakes. The Federal Election Campaign Act (FECA) 52 U.S.C. §30124. Fraudulent misrepresentation of campaign authority, provides:

> No person who is a candidate for Federal office or an employee or agent of such a candidate shall-
>
> (1) 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; or
>
> (2) willfully and knowingly participate in or conspire to participate in any plan, scheme, or design to violate paragraph (1).

The implementing regulation, 11 CFR § 110.16, is virtually identical.[69]

---

65. *See* Van Hollen v. FEC, 811 F.3d 486, 499 (D.C. Cir. 2016) (noting the FEC's need to safeguard the First Amendment when implementing its statutory directives).

66. 11 CFR § 110.11 (2023).

67. *See* Susan B. Anthony List v. Driehaus, 814 F.3d 466, 476 (6th Cir. 2016) (striking down Ohio law prohibiting lies in political ads); United States v. Alvarez, 567 US 709 (2012) (striking down federal statute prohibiting a person from falsely stating that they had won the Congressional Medal of Honor).

68. *See New York Times*, 376 U.S. at 282.

69. The regulation provides:

> No person who is a candidate for Federal office or an employee or agent of such a candidate shall—
>
> (1) Fraudulently misrepresent the person or any committee or organization under the person's control as speaking or writing or otherwise acting for or on behalf of any other candidate or political party or employee or agent thereof in a matter which is damaging to such other candidate or political party or employee or agent thereof; or
>
> (2) Willfully and knowingly participate in or conspire to participate in any plan, scheme, or design to violate paragraph (a)(1) of this section.

11 CFR § 110.16 (2023).

This provision prohibits candidates and their staff from impersonating each other in contexts such as calling in to radio shows, calls to voters, contacting donors to raise funds, and the like. These rules have never successfully been challenged on constitutional grounds and are a reasonable restriction on political speech in that one candidate pretending he is another candidate is hardly a free expression of the candidates' own views on anything. On the other hand, the FEC has refused to enforce this prohibition when an organization uses a candidate's name, image, and likeness without permission in circumstances where a declaimer or other disclosures make it clear that the message is not authorized by the candidate whose image appears.[70]

52 U.S.C. §30124 and 11 CFR § 110.16 were not drafted with deepfakes in mind but the text appears to cover at least some deepfake ads. A law reform organization, Public Citizen, in May and again in July 2023, asked the FEC to clarify that the regulations do apply "if candidates or their agents fraudulently misrepresent other candidates or political parties through deliberately false AI-generated content in campaign ads or other communications – absent clear and conspicuous disclosure in the communication itself that the content is generated by artificial intelligence and does not represent real events."[71] After the FEC initially deadlocked on the issue, Public Citizen amended its petition, citing another case in which Commissioner Alan Dickerson discussed the FEC's statutory authority to enforce federal law prohibiting a candidate from speaking, writing, or acting on behalf of another candidate for purposes of damaging that other candidate or party.[72] The amended petition asked the FEC to conduct a rulemaking to clarify the meaning of "fraudulent misrepresentation" at 11 C.F.R. §110.16 and 52 U.S.C. §30124. The FEC unanimously determined that this petition met the requirements of 11 C.F.R. § 200.2(b) and published in the *Federal Register* pursuant to 11 C.F.R. § 200.3(a)(1), a Notification of Availability ("NOA") seeking comment up until October 16, 2023, on whether the Commission should initiate full rulemaking on the proposal.[73]

Meanwhile several members of Congress, all Democrats, wrote the FEC requesting action:

---

70. *See, e.g.*, Letter from William A. Powers, Assistant Gen. Couns., FEC, to Mark Braden (Mar. 7, 2014), https://perma.cc/M2YU-6TFV.

71. Weissman, *supra* note 42, at 5.

72. Weissman, *supra* note 42, at 5 (citing FEC, MUR 7140, In the Matter of Americans for Sensible Solutions PAC and David Garrett, Statement of Reasons of Vice Chair Allen Dickerson and Commissioner James E. "Trey" Trainor, III (Apr. 5, 2021), https://perma.cc/R7T8-DBCL (involving an expenditure-only political committee, Americans for Sensible Solutions PAC, that allegedly solicited contributions by fraudulently misrepresenting that it was acting as an agent of a congressional candidate)).

73. Memorandum from Lisa Stevenson, Acting Gen. Couns., FEC, Neven Stipanovic, Assoc. Gen. Couns., FEC, Robert Knop, Assistant Gen. Couns., FEC, and Jennifer Waldman to the Commission (June 15, 2023), https://perma.cc/9DXN-KFCH.

> "As Members of Congress concerned about the ability of generative AI to significantly disrupt the integrity of our elections, we respectfully request that the FEC reconsider its decision and seek comment on whether the Commission should initiate a full rulemaking on a proposal in the Petition for Rulemaking from Public Citizen."[74]

Even if future FEC rulemaking clarifies that 52 U.S.C. §30124 and 11 CFR § 110.16 apply to deepfake ads, there's another problem. This statute and the rule bind candidates for federal office and their agents and employees but does not bind others such as independent expenditure organizations that may support or oppose a candidate. The statute and rule only apply if a candidate, agent of a candidate, or employee of a candidate were to "willingly and knowingly participate in or conspire to participate in" such conduct.[75] Candidates and independent expenditure organizations aren't allowed to coordinate anyway, however, so participation in a "plan, scheme, or design" either does not occur or the participants take precaution not to get caught.[76]

Cutting off funding to organizations making supposedly "independent" electioneering communications, including deepfake communications is extraordinarily difficult. In *Citizens United* the Supreme Court held that corporations have a First Amendment right to bankroll these organizations and their electioneering communications.[77] As discussed in Part I of this article, foreign nationals are getting involved as well. More and more people will combine AI with their "free speech" rights and impersonate candidates in life-like images saying things the candidates didn't say and doing things the candidates didn't do and then disseminate those fake images for the purpose of influencing elections.

C. Proposed Laws

Various bills have been introduced in Congress to regulate deepfakes; so far none have passed. H.R. 2395, the Deep Fakes Accountability Act, provides that any person who produces an "advanced technological false personation record" with intent to distribute such record over the internet or has knowledge it will be so distributed, must make sure the record has an embedded digital watermark and an audible and visible disclaimer.[78] The bill applies to any "advanced

74. Letter from 24 Congressmen to Lisa Stevenson, Acting Gen. Couns., FEC (July 13, 2023), https://perma.cc/WG8J-GYG6.

75. 52 U.S.C. §30124 (a)(2).

76. 52 U.S.C. §30124 (a)(2).

77. *See* discussions *supra* notes 2, 11-13 and accompanying text.

78. Deep Fakes Accountability Act, H.R. 2395, 117th Cong. (2021), https://perma.cc/DLW8-SSP6. Introduced by Rep. Yvette D. Clarke [D-NY-9] on April 8, 2021, the bill provides that "any advanced technological false personation record which contains a moving visual element shall contain an embedded digital watermark clearly identifying such record as containing altered audio or visual elements" and that "any advanced technological false personation records containing both an audio and a visual element shall include— (1) not less than 1 clearly articulated verbal statement that identifies the record as containing altered audio and visual elements, and a concise description of the extent of such alteration; and (2) an unobscured written statement in clearly readable text appearing at the bottom of the

technological false personation record" or deepfakes – not just deepfakes used in elections.[79] The bill was referred to the House Subcommittee on Crime, Terrorism, and Homeland Security, and never advanced further.[80]

The Deep Fakes Accountability Act imposes criminal penalties for violations, which exacerbates existing concerns with overbroad regulation of speech, particularly in the context of political speech where First Amendment scrutiny is the most exacting. Perhaps because of these criminal penalties the bill uses a definition of deepfakes that focuses not just on use of the AI technology but on whether the deepfake depicts an event that did not happen.[81] An AI-generated portrayal of something that did happen – for example Trump having a vulgar conversation with Billy Bush on a bus – presumably would not fit within the statutory definition of a deepfake in the Act and would not be prohibited. The definitions used in the bill do not cover AI depictions of events that took place, even if an AI rendition of such an event would be different from a video recording of the event itself.[82]

Prosecuting a criminal case under this bill thus could be difficult unless it is indisputable that the event portrayed by the AI deepfake did not happen. A defendant who creates a reasonable doubt that the event did happen probably will be acquitted. Because state of mind is an element for most criminal offenses, the defendant's belief that the event did in fact happen also might be sufficient for acquittal. Consider for example the dubious but once widely believed story attributed to the "Steele Dossier" about Donald Trump meeting with prostitutes in a Moscow hotel room.[83] The fact that many people believed the story to be true at

---

image throughout the duration of the visual element that identifies the record as containing altered audio and visual elements, and a concise description of the extent of such alteration." *Id.* at § 1041(b)-(c).

79. *Id.* at §1041 (advanced technological false impersonation record).

80. See Tracker for H.R. 2395 - Deep Fakes Accountability Act, CONGRESS.GOV.

81. H.R. 2395 at § 1041 (n)(3) ("DEEP FAKE.—The term 'deep fake' means any video recording, motion-picture film, sound recording, electronic image, or photograph, or any technological representation of speech or conduct substantially derivative thereof— (A) which appears to authentically depict any speech or conduct of a person who did not in fact engage in such speech or conduct; and (B) the production of which was substantially dependent upon technical means, rather than the ability of another person to physically or verbally impersonate such person.")

82. *See id.* at § 1041(n)(1) ("The term 'advanced technological false personation record' means any deep fake, which .. (A) a reasonable person, having considered the visual or audio qualities of the record and the nature of the distribution channel in which the record appears, would believe accurately exhibits— (i) *any material activity of a living person which such living person did not in fact undertake*; or (ii) any material activity of a deceased person which such deceased person did not in fact undertake, and the exhibition of which is substantially likely to either further a criminal act or result in improper interference in an official proceeding, public policy debate, or election; and (B) was produced without the consent of such living person, or in the case of a deceased person, such person or the heirs thereof." (emphasis added)). *See also id.* at § 1041(n)(3) ("The term 'deep fake' means any video recording, motion-picture film, sound recording, electronic image, or photograph, or any technological representation of speech or conduct substantially derivative thereof— (A) *which appears to authentically depict any speech or conduct of a person who did not in fact engage in such speech or conduct*; and (B) the production of which was substantially dependent upon technical means, rather than the ability of another person to physically or verbally impersonate such person." (emphasis added)).

83. Aaron Blake, *Why We Should All Be Careful About the Lewd Trump-Russian Prostitute Allegation*, WASH. POST (Apr. 14, 2018, 9:00 AM), https://perma.cc/2SGG-SRKL.

the time might be sufficient to defend someone prosecuted under the Deep Fakes Accountability Act for making or disseminating a video of it.[84] The defendant would argue that he reasonably believed that story was true.

A broader approach would be to require a disclaimer for *any* ad that used AI-generated material. Rep. Yvette Clarke (D-NY) has introduced a bill, the *REAL Political Ads Act*, that would extend the Federal Election Campaign Act's disclosure rules for radio and TV ads to online communications and require a disclaimer if AI-generated material is used.[85] The advantage of this approach is that it does not inquire as to whether the events depicted did in fact happen, or whether something close enough to them happened that deepfakes can be excused. Under this approach all video/audio using AI must be identified.

Criminal penalties likely would not ensue for failure to comply with the bill's labeling requirement for AI, although it would be a federal election law violation. Enforcement of the *REAL Political Ads Act* might be too little too late, particularly if a mislabeled AI ad helped win an election. Also, this bill arguably is overbroad in that some AI-generated ads aren't that different from traditional ads; for example, the GOP's AI-generated ad depicting multiple calamities likely to happen if Biden is reelected is probably no more misleading, and probably not as convincing, as LBJ's 1964 Daisy Girl ad. The AI-generated material that is most concerning is AI depicting a public figure, usually a candidate, doing something or saying something that never happened. Such deepfakes should be identified, and as discussed in Part III of this Article, such instances need to be met with rapid and widespread identification in the media.

D. The Enforcement Problem

It is hard to regulate fraudulent misrepresentation in campaign speech.

First any regulation must pass constitutional muster, which is difficult with the Supreme Court's expansive application of the First Amendment in the realm of campaign finance.[86] The conceptual difference between a permissible satire of a candidate, such as impersonation by an actor, and a fraudulent AI-generated impersonation, will invite argument if the focus of inquiry is the perception of voters. Prohibitions on deepfakes could be difficult to apply on a case-by-case basis, and political operatives will explore the boundaries between the permissible and impermissible.

Second, using criminal law to impose a prohibition on speech is even harder as First Amendment scrutiny is likely to be even more exact in criminal cases, and any ambiguity construed in favor of a defendant. This means that Congress and

84. This example might also violate separate criminal statutes prohibiting revenge porn, but this discussion sets aside the elements for prosecuting revenge porn and focuses only on the definition of deepfake in the Deep Fakes Accountability Act which appears to exclude an AI-generated portrayal of an event that happened, or that the defendant reasonably thought did happen. H.R. 2395 at § 1041 (n)(3).

85. REAL Political Ads Act, H.R. 3044, 118th Cong. (2023).

86. *See* discussion *supra* text accompanying notes 11-13.

the FEC would have to rely on civil penalties, which may not be sufficient to deter the illegal conduct.

Third, disseminating the ad is different from creating it. People who create a deepfake ad may not be affiliated with a campaign at all and may even be outside the United States. People who disseminate the ad are not the original publisher or speaker, and social media platforms are protected against defamation suits by Section 230 of the Communications Decency Act of 1996.[87] People who disseminate a deepfake ad also may claim that they didn't know that it was a deepfake, and whether they did know may be impossible to prove.

Fourth, a legal response to deepfakes that is not immediate may come too late. Investigating deepfakes, commencing an FEC proceeding, and even a criminal prosecution, will take far too much time, and the beneficiary of the deepfaking of an opponent could be elected in the meantime. The law should focus on responding quickly and decisively to deepfakes, not so much on the traditional methods of imposing liability and accountability.

In sum, we can try as hard as we might to keep AI out of electioneering communications, but it will likely be a losing battle.

E. Constitutional and Practical Limits on Enlisting Help from Social Media Platforms

The one lever of control the FEC and other federal agencies may be able to use, up to a point, is regulation of social media platforms and other communication venues inside the United States. Regulators have some ability legally and practically to reach major distributors of social media content – Twitter, Facebook, Instagram, YouTube and others – urging them to take down deepfake content as soon as it is identified or to identify it with a disclaimer. Congress has given social media companies broad protection from defamation suits under Section 230 of the Communications Decency Act, and it might be reasonable for Congress and perhaps the Executive Branch to ask for some self-censorship of deepfakes and other fraudulent content in return.[88]

Maybe.

Social media platforms are private and have broader latitude to remove harmful content than the government does because they are not bound by the First Amendment. The difficulty is that the government is bound by the First Amendment and courts have imposed limits on how much the government can pressure social media companies to remove harmful content. For example, on July 4, 2023, U.S. District Judge Terry Doughty in Louisiana entered a preliminary injunction against the Department of Health and Human Services, the FBI,

87. Communications Decency Act, 42 U.S.C. § 230(c)(1) (providing that "[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another content provider" And defining "interactive computer service" to include any system where multiple users can access a single server. This covers just about everyone on the internet, who is either a provider or user of "interactive computer services").

88. *Id.* at § 230.

and dozens of other government agencies and officials from contacting social media companies[89] for the purpose of "encouraging, pressuring, or inducing in any manner the removal, deletion, suppression, or reduction of content containing protected free speech."[90] The Fifth Circuit Court of Appeals in New Orleans modified the preliminary injunction pending appeal. The preliminary injunction issued by the District Court, as modified by the Fifth Circuit on October 3, 2023, was stayed by the Supreme Court on October 20, 2023. Three justices – Alito, Thomas, and Gorsuch – would not have stayed the order, and it is unclear how the full Court will rule on the merits.[91] The outcome of this case could have a profound effect on the ability of any federal agency, including the FEC, to influence the decisions of social media platforms.

In the face of these legal challenges, federal agency action to discourage deepfake video and audio on social media needs to be narrowly tailored, focusing on deepfakes alone, not combined with efforts to discourage other harmful content on social media. Federal agencies, and ultimately courts, will evaluate each

89. Missouri v. Biden, No. 22-01213, 2023 U.S. Dist. LEXIS 114585, at *3 (W.D. La. July 4, 2023) ("Plaintiffs allege that Defendants, through public pressure campaigns, private meetings, and other forms of direct communication, regarding what Defendants described as 'disinformation,' 'misinformation,' and 'malinformation,' have colluded with and/or coerced social-media platforms to suppress disfavored speakers, viewpoints, and content on social-media platforms."). In the Judge's order, "social-media companies" are defined to include "Facebook/Meta, Twitter, YouTube/Google, WhatsApp, Instagram, WeChat, TikTok, Sina Weibo, QQ, Telegram, Snapchat, Kuaishou, Qzone, Pinterest, Reddit, LinkedIn, Quora, Discord, Twitch, Tumblr, Mastodon, and like companies." *Id*. at *213 n.2 (of judgment).

90. *Id*. at *213. The Court enjoined dozens of high ranking officials in the Biden Administration who were named as defendants in this lawsuit, their agents, officers, employees, contractors, and all acting in concert with them from taking the following actions as to social-media companies: "(1) meeting with social-media companies for the purpose of urging, encouraging, pressuring, or inducing in any manner the removal, deletion, suppression, or reduction of content containing protected free speech posted on social-media platforms; (2) specifically flagging content or posts on social-media platforms and/or forwarding such to social-media companies urging, encouraging, pressuring, or inducing in any manner for removal, deletion, suppression, or reduction of content containing protected free speech; (3) urging, encouraging, pressuring, or inducing in any manner social-media companies to change their guidelines for removing, deleting, suppressing, or reducing content containing protected free speech; (4) emailing, calling, sending letters, texting, or engaging in any communication of any kind with social-media companies urging, encouraging, pressuring, or inducing in any manner for removal, deletion, suppression, or reduction of content containing protected free speech; (5) collaborating, coordinating, partnering, switchboarding, and/or jointly working with the Election Integrity Partnership, the Virality Project, the Stanford Internet Observatory, or any like project or group for the purpose of urging, encouraging, pressuring, or inducing in any manner removal, deletion, suppression, or reduction of content posted with social-media companies containing protected free speech; (6) threatening, pressuring, or coercing social-media companies in any manner to remove, delete, suppress, or reduce posted content of postings containing protected free speech; (7) taking any action such as urging, encouraging, pressuring, or inducing in any manner social-media companies to remove, delete, suppress, or reduce posted content protected by the Free Speech Clause of the First Amendment to the United States Constitution; (8) following up with social-media companies to determine whether the social-media companies removed, deleted, suppressed, or reduced previous social-media postings containing protected free speech; (9) requesting content reports from social-media companies detailing actions taken to remove, delete, suppress, or reduce content containing protected free speech; and (10) notifying social-media companies to Be on The Lookout ("BOLO") for postings containing protected free speech." *Id*.

91. Murthy v. Missouri, No. 23-411, 2023 WL 6935337 (U.S. Oct. 20, 2023).

category of allegedly fraudulent content in social media against a presumption in favor of First Amendment protection. Articulating a one-size fits all theory of what fraudulent content regulators can and cannot ask platforms to take down, and how much pressure can be applied, will be exceedingly difficult. Perhaps someday courts will give sufficient guidance that a broader articulation of the law in this area is possible, but we're not there yet, and Judge Doughty's injunction against the Biden Administration is indicative of how haphazard the judicial response can be.

For now, at least, if federal regulators are serious about persuading social media platforms to counter deepfakes, they will have to focus on that problem directly. Deepfake content is unique in that it impersonates another person – it almost always pictures them saying things they didn't say and doing things they didn't do. Most other harmful and/or fraudulent content does not go that far. A false video might say, for example, that President Biden is taking secret bribes from the Chinese Communist Party while showing genuine photos of Biden and Chinese leaders, but the video is not a deepfake unless it is an actual AI-generated fake video or recording of Biden taking a bribe from the Chinese Communist Party. The first video, containing genuine photos of Biden and Chinese leaders, might very well be First Amendment protected speech, as dishonest as it is, but the latter video, which uses AI to show Biden taking a bribe, is very likely not protected speech unless a reasonable viewer and listener would discern that the person taking the bribe in the video is not actually Biden.[92]

Congress and the FEC should work to persuade as many social media platforms as possible to minimize deepfake content, and to take corrective action when it does occur. This can help mitigate its impact, although removing deepfake content from the Web entirely would be impossible. The more cooperation the FEC has from major social media platforms in taking down deepfake content the better, but this focus on social media companies is not a cure all.

III. DEEPFAKE WARNINGS AND A DEEPFAKE ALERT SYSTEM

Deepfakes are part of a broader problem in American elections requiring reexamination of how electioneering communications are paid for and how some electioneering communications are amplified over others.

This essay will not revisit arguments this author and many others have made for fixing a campaign finance system that spends billions of dollars on electioneering communications, many of them misleading. Most electioneering communications aren't deepfakes; many are lies, but they are lies told in this world, not a virtual world. The prospect of future elections revolving around an alternative universe orchestrated by AI, however, is not far off. 2024 is the beginning, and it

92. Theoretically, Biden could sue for defamation for an ad such as this, but his suit would have to meet the very high pleading standard required under *New York Times v. Sullivan*. *New York Times*, 376 U.S. at 254.

will be worse in election cycles afterwards. Who's paying for much of this virtual universe of electioneering communications will probably remain a mystery.

This author[93] and others[94] have proposed broader campaign funding reforms, which fall into three general categories: 1) get special interest money and foreign money out of elections to the extent constitutionally permissible and practical, 2) enhance disclosure requirements for entities responsible for electioneering communications, and 3) promote a counterweight of electioneering communications paid for with public funds, tax breaks for small donors, or other sources more closely aligned with the people our government is supposed to represent.[95]

The response to deepfakes explored here is in the third category – alternative electioneering communications that respond quickly, decisively, and overwhelmingly to deepfakes with warnings ("Deepfake Warnings"). Deepfake Warnings are ideally posted on the same platforms where deepfake communications appear and probably on other media platforms as well.

Deepfake Warnings could be private or public. Private Deepfake Warnings would be paid for by the candidate attacked in a deepfake communication, that candidate's political party, a PAC, Super PAC, 501c4 civic organization, or other private entity. Existing campaign finance laws and tax laws would apply, so Congress could promote private Deepfake Warnings with amendments to existing campaign finance laws. For example, an amendment could allow individual donors who have maxed out on a candidate to make additional contributions escrowed to a Deepfake Response Fund that could only be used to pay for electioneering communications responding to a deepfake attack on the candidate. Deepfakers might think twice if they knew that the target of their attack could respond by exceeding generally allowable campaign fundraising to pay for refutation of a fraudulent impersonation.

Public Deepfake Warnings would be official communications from the FEC, and their content strictly regulated by statute or FEC rule. The FEC's Deepfake Warnings would state something like: "The Federal Election Commission has reviewed this video/audio impersonating Candidate X and has determined that it is a fake made with artificial intelligence. This is not a real [video / audio / video and audio] of [name of person impersonated]." The statement would be accompanied by a still image screenshot from the deepfake video so people can identify it. If the deepfake is audio only, the statement would include an audio recording of the first ten seconds of the audio so listeners could identify it.

The FEC Deepfake Warning could also include boilerplate language explaining the harm from deepfake ads to fair elections, that the funding of their creation and dissemination is often concealed, and that they may originate from outside the United States. This language, however, should be standard in every FEC

93. *See* RICHARD W. PAINTER, TAXATION ONLY WITH REPRESENTATION: THE CONSERVATIVE CONSCIENCE AND CAMPAIGN FINANCE REFORM 154-87 (2016).

94. *See generally* LAWRENCE LESSIG, REPUBLIC, LOST: HOW MONEY CORRUPTS CONGRESS AND A PLAN TO FIX IT (2015).

95. *See* PAINTER, *supra* note 93, at 166-77.

endorsed Deepfake Warning announcement. The FEC should not say anything about the content of the deepfake communication other than that the video, audio or both were determined to be fake and not a real audio and/or video recording of the person depicted.

The FEC Deepfake Warning would not identify the presumed source of the Deepfake communication or say anything else that could influence an election. If persons or organizations responsible for a deepfake communication later can be identified, and the communication is found to violate federal election law, an enforcement proceeding can be commenced with notice and an opportunity to be heard, and after that a finding of a violation. But that is a time-consuming process separate from the Deepfake Warning announcement which should be posted quickly and be limited in content to exposing the deepfake ad without otherwise influencing the election.

A Deepfake Warning should be disseminated broadly on the same media platform as the deepfake ad and perhaps also on other platforms to inform as many voters as possible that the deepfake ad is a fake. FEC regulations, or Congressional legislation, could provide for expediated access to media platforms and broadcasting outlets for private and public Deepfake Warnings with fair compensation for the broadcaster or platform host. The Deepfake Warning should be widely enough disseminated to be a deterrent for candidates and persons producing and disseminating deepfake ads.

The FEC, after approving a public Deepfake Warning, could release rights to the warning, allowing campaigns, political parties, and others to run it as many times as they want on media platforms of their choosing, provided it is not altered, combined with, or run adjacent in time or space to any other campaign ad or electioneering communication for or against any candidate. Such other electioneering communications would not have FEC endorsement and should be kept entirely separate from public Deepfake Warnings. Private Deepfake Warnings, on the other hand, could be combined with other electioneering communications and would be subject to existing regulations for electioneering communications in general, but nothing more.

A parallel approach would be for the FEC, state secretaries of state, or other election officials to run anticipatory public service announcements on broadcast media, print media and on-line social media alerting voters to the high likelihood of deepfake images and audio being used to influence their vote, particularly in the days immediately before an election.

The concept of public Deepfake Warnings, however, runs into legal and practical problems if it is expanded to include publicly financed "warnings" directed at other campaign ads that the FEC deems "misleading" simply because of their content. The Hatch Act prohibits using federal resources or the authority of federal office to influence an election in any way, and this presumably includes official capacity refutation of false campaign ads.[96]

96. *See* 5 U.S.C. § 7323 (providing that a federal employee may not "use his official authority or influence for the purpose of interfering with or affecting the result of an election").

A campaign ad that lies about President Biden and says he took bribes from the Chinese government, but without using deepfake material, thus is best refuted by the Biden Campaign or some other private organization, not by the FEC or another governmental entity. Allowing the FEC or any other part of the government to "rate" campaign ads for their truthfulness and respond is an invitation to government meddling in elections. Deepfake content is different because identifying it is a technological issue and does not require a subjective assessment of the ad's veracity. The FEC's job in issuing a Deepfake Warning should be to make sure voters know what a deepfake is and when they are seeing and hearing it.

Deepfake Warnings, however, won't do much good if they aren't disseminated in time to inform voters before an election. Users of deepfakes know this and will very likely wait until the last minute to disseminate deepfake videos and audios of opponents. Without a rapid response there will be no hope of reversing the harmful impact on a candidate. Some campaigns will have a "deepfake rapid response team" ready to repudiate deepfakes about their candidate, but some will not be organized enough or have sufficient connections with social media platforms and other media outlets to get the story out in time. It is here that the FEC, without taking sides between candidates and without commenting on the content of any political ad or other electioneering communication, could lend a helping hand.

An online FEC "Deepfake Alert System" would allow any campaign registered with the FEC to report a suspected deepfake, and the report would automatically be posted on the FEC alert system website as soon as it were filed.[97] A panel of computer science experts employed by the FEC, either as full-time employees or as part time special government employees, would be notified and immediately analyze the alleged deepfake material, posting their preliminary assessment of it within hours if possible, and a final assessment as soon as possible after that.[98] Color coding next to the report on the website – yellow for "probably contains deepfake" or red for "definitely contains a deepfake" – would alert users of the website of the latest technical assessment. The assessment team would also prepare a written report describing in more detail the specific nature of the AI-generated alterations in the video, focusing only on the technology, not on the truth or falsehood of statements made in the video. The FEC could also issue press releases alerting the media to new postings on its "Deepfake Alert System"

97. State secretaries of state and other election officials should consider a similar Deepfake Alert System for state and local elections where deepfakes have also proliferated, as shown by a recent deepfake video of a Chicago mayoral candidate released days before the election. *See* discussion *supra* text accompanying note 36.

98. Computer scientists working at universities and in private industry have done very important research on deepfake creation and detection. *See, e.g.*, Ghazal Mazaheri & Amit K. Roy-Chowdhury, *Detection and Localization of Facial Expression Manipulations*, 2022 IEEE/CVF WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION, Jan. 2022, at 1, https://perma.cc/W2YW-SWT4 (proposing a new approach to exploit facial expression systems in image/video facial expression manipulation detection).

website with increasing frequency up until the week before the election when FEC press releases should probably be daily. The FEC would also remind media outlets that they should always check the website before reporting on newly released video or audio of a candidate.

Candidates and their campaigns would be allowed, but not required, to report to the FEC "Deepfake Alert System" what they know about an alleged deepfake. The initial report from a candidate would be an example of this, as would be the opponent's campaign saying something like "we have no knowledge of where that alleged deepfake came from." One caveat: lying to voters may be routine, but lying in a communication to the FEC could be a felony under the false statements' statute.[99] Candidates and political operatives who don't want to tell the truth would be well advised to say nothing at all to the FEC. If candidates and political operatives say something to the public, but won't say the same thing to the FEC, then the media and voters should take note of this discrepancy.

In sum, deepfakes cannot effectively be countered with criminal penalties, or other prohibitory regulations. People will find a way around such rules and still use deepfakes to influence elections. Investigations will take far too long, and the response will be too late. Deepfake content needs to be countered rapidly and decisively as soon as it emerges on any media platform. Private and Public Deepfake Warnings and a FEC sponsored Deepfake Alert System would go a long way toward accomplishing that goal. Then it will be up to Americans if they want to live and vote in this world or in some alternative universe.

CONCLUSION

"Confirmation bias" is pervasive in social media; people believe and repeat what they want to hear.[100] Deepfakes are no exception. Once created and released it can spread like wildfire as users predisposed to believe it recirculate it to others who do the same.

We can try to combat deepfakes with regulations, civil causes of action, and even criminal penalties, but such legal remedies may be workable, and constitutional, only in narrow cases such as deepfake porn, fraudulent fundraising using deepfakes, defamation, and other demonstrable injuries to specific persons. Legal prohibitions on electioneering communications are problematic because First Amendment protection is exceptionally robust in the case of political speech. Deepfake images and audio also may be created and disseminated by persons beyond the jurisdiction of U.S. regulators and courts. As pointed out in Part I of the Mueller Report, Russian agents used fraudulent manipulation of social media and impersonation in the 2016 election. Foreign interference in American

99. 18 U.S.C. § 1001 (criminalizing knowing false statements in "any matter within the jurisdiction of the executive, legislative, or judicial branch of the Government of the United States").

100. Marcos R. Fernandes, *Confirmation Bias in Social Networks*, MATHEMATICAL SOCIAL SCIENCES, FORTHCOMING, Feb. 22, 2023, at 2 https://perma.cc/YL76-S56N.

elections is a continued threat, and in future election cycles deepfakes could be part of the plan.

Regulations and remedies are ineffectual if they don't do anything to prevent the potentially irreversible harm from deepfakes in the days leading up to an election. Responding to deepfakes quickly and decisively is the only way to reverse its harmful impact.

The best long-range solution is to revamp the way we fund elections in the United States. Fixing campaign finance will be necessary to defend our independence from corporate interests, foreign governments and organizations using multiple strategies to mislead voters, including AI. The notion that electioneering communications are free speech – no matter what those communications are and who's paying for them – is an invitation to a flood of deepfake ads and other "fake news" influencing elections.

In the short term, we can facilitate, and fund, timely Deepfake Warnings posted on the same social media platforms where voters are exposed to deepfakes. Private Deepfake Warnings are electioneering communications and should be privately funded and regulated as such. Election laws could be adjusted to facilitate political contributions to fund private Deepfake Warnings, although they would not have the endorsement of the FEC. Public Deepfake Warnings from the FEC also are appropriate. They should identify deepfake electioneering communications for what they are – fakes – and should be sufficiently robust to reverse impact of the deepfake on voters.

The FEC also should have a Deepfake Alert System where deepfake reports are posted on a website. The alleged deepfake content should be analyzed on an expedited basis by a panel of computer scientists who announce preliminary findings as soon as possible and a more detailed assessment not long thereafter. Postings on this website and the findings of the FEC technical experts could then be incorporated into public and private Deepfake Warnings used to respond to individual cases.

Until the United States revamps its campaign finance system, deepfakes will be one of our many problems with disinformation funded from sources far from home. Do Americans want to play an election-year video game where the rules are set by persons unknown, and powerful players overwhelm seemingly autonomous decisions of individual voters, or do we want to have real elections with real candidates chosen by real people? That decision is ours to make for now, but we had better make it soon.

***