DEEP CONCERN: SAFEGUARDING ELECTIONS IN THE AGE OF DEEPFAKES

Jack Wampler*

"You can't trust anything you see or hear." This is the context in which future elections exist. Indeed, deepfakes—exceptionally realistic AI-generated pictures, videos, and audio recordings—are poised to upend established election norms by facilitating character assassination of political candidates, enabling the proliferation of disinformation regarding voting procedures, eroding public trust in election institutions, and creating what scholars call "the liar's dividend." Given the inadequacy of current legislative, judicial, administrative, and corporate safeguards, this Note aims to provide a fresh suggestion to address the seemingly insurmountable dangers deepfakes pose to our electoral process.

TABLE OF CONTENTS

---

* J.D. Candidate, University of Arizona James E. Rogers College of Law, 2025. I am deeply grateful to the incomparable Professor Diana Simon for supervising this Note and for welcoming me into the law school community when she recognized me from a picture in which I was dressed as a pumpkin. I would also like to thank the brilliant Professor Cas Laskowski for her insights on deepfakes and AI, and my wonderful peers at *Arizona Law Review* for their help at every stage of the writing process. Finally, I would like to thank my friends and family for their unwavering love and support—past, present, and future.

INTRODUCTION

"You can't trust anything that you see or hear," warned former Google CEO Eric Schmidt regarding the 2024 U.S. election cycle.[1] The 2023 explosion in generative artificial intelligence ("AI") technology such as ChatGPT precipitated Schmidt's grim prediction.[2] Generative AI refers to artificial intelligence that uses massive datasets to create new content, including text, images, and video.[3] One specific application of generative AI, deepfakes, has garnered significant attention in recent years, as experts increasingly fear that bad actors may use deepfakes to upend the democratic processes of countries worldwide.[4]

A deepfake is a synthetic but exceptionally realistic video, image, or audio recording created using generative AI techniques.[5] Thus, a deepfake allows its creator to portray a person doing or saying something that the person never actually did or said.[6] For example, in 2018, filmmaker Jordan Peele and Buzzfeed CEO Jonah Peretti famously produced a video purporting to show former President Barack Obama call Donald Trump a "total and complete dip****."[7] While Peele went on to admit that the video was manipulated, concurrently warning the public of the rapidly evolving threat of misinformation in the digital age,[8] this example

---

1. John Frank, *The 2024 Presidential Race is the AI Election*, AXIOS (June 27, 2023) (cleaned up), https://www.axios.com/2023/06/27/artificial-intelligence-ai-2024-election-biden [https://perma.cc/2347-YSP7].

2. *Id.*

3. Kim Martineau, *What is Generative AI?*, IBM (Apr. 20, 2023), https://research.ibm.com/blog/what-is-generative-AI [https://perma.cc/6XLV-BKYQ].

4. *See* Frank, *supra* note 1 ("Top technologists are portraying a dystopian landscape in 2024 in which misinformation and disinformation proliferate with . . . speed and ease."); *see also* Hasan Chowdhury, *AI Deepfakes Threaten to Wreak Havoc with the World's Year of Elections*, BUS. INSIDER (Jan. 25, 2024, 3:01 AM), https://www.businessinsider.com/ai-deepfakes-threaten-havoc-with-the-worlds-year-of-elections-2024-1 [https://perma.cc/QJD2-G5BY]; Julia Mueller & Jared Gans, *Fears Grow Over AI's Impact on the 2024 Election*, THE HILL (Dec. 25, 2023, 12:14 PM), https://thehill.com/homenews/campaign/4371959-ai-artificial-intelligence-2024-election-deepfake-trump/ [https://perma.cc/CL5G-YVWS].

5. Dave Johnson & Alexander Johnson, *What are Deepfakes? How Fake AI-Powered Audio and Video Warps Our Perception of Reality*, BUS. INSIDER, https://www.businessinsider.com/guides/tech/what-is-deepfake [https://perma.cc/ZP3F-QMMX] (June 15, 2023, 7:58 AM).

6. *Id.*

7. David Mack, *This PSA About Fake News from Barack Obama is Not What It Appears*, BUZZFEED NEWS (Apr. 17, 2018, 8:26 AM), https://www.buzzfeednews.com/article/davidmack/obama-fake-news-jordan-peele-psa-video-buzzfeed [https://perma.cc/7CZ2-L6YA].

8. *Id.*

underscores how individuals can leverage deepfake technology to generate deceptive content with remarkable realism.

Concerns about deepfakes extend to various other contexts.[9] Notably, as of 2019, 96% of online deepfakes were nonconsensual pornographic deepfakes.[10] These deepfakes harass and dehumanize their targets, which overwhelmingly tend to be women.[11] Female celebrities are targeted most frequently,[12] but pornographic deepfakes targeting noncelebrities are becoming increasingly common.[13] Furthermore, the realistic and deceptive capabilities of deepfakes allow for their seamless integration into the repertoire of scammers, paving the way for more convincing and harmful schemes.[14] For example, a scammer could trick your family member into sending ransom money by using a deepfake of your voice to convince the family member that you were kidnapped.[15] Similarly, scammers have created deepfake celebrity endorsements to trick people into disclosing personal information and signing up for expensive subscriptions.[16] Finally, deepfakes may present novel challenges to judges and lawyers, as litigants may attempt to introduce "deepfaked" content into evidence, potentially compromising courts' ability to ascertain the truth.[17]

Nevertheless, many technologists are quick to point out the potential beneficial applications of deepfake technology.[18] For instance, educators can use deepfakes in the classroom to liven their lectures by creating interactive portrayals of historical figures.[19] Some museums have already upgraded their exhibits in a similar manner.[20] For example, a life-size deepfake of deceased artist Salvador Dalí

---

9. *See* Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753, 1771–86 (2019).

10. Arwa Mahdawi, *Nonconsensual Deepfake Porn is an Emergency that is Ruining Lives*, THE GUARDIAN (Apr. 1, 2023, 9:00 AM), https://www.theguardian.com/commentisfree/2023/apr/01/ai-deepfake-porn-fake-images [https://perma.cc/E6KN-8HFP].

11. *Id.* ("[O]f those, 99% featured women.").

12. *See, e.g.*, Samantha Murphy Kelly, *Explicit, AI-Generated Taylor Swift Images Spread Quickly on Social Media*, CNN (Jan. 25, 2024, 3:19 PM), https://www.cnn.com/2024/01/25/tech/taylor-swift-ai-generated-images/index.html [https://perma.cc/EJ4K-PK5P]; Rashi Agarwal, *Horrifying Porn Deepfakes of Scarlett Johansson and Emma Watson Dominate 'Predatory' Website*, UNILAD (Mar. 31, 2023, 7:25 PM), https://www.unilad.com/news/scarlett-johansson-emma-watson-deep-fake-ai-380897-20230331 [https://perma.cc/GM3A-257D].

13. Mahdawi, *supra* note 10.

14. *See* Chesney & Citron, *supra* note 9, at 1772.

15. *Id.*

16. Thomas Orsolya, *Don't Get Scammed by The Fake Jennifer Aniston MacBook Giveaway*, MALWARETIPS (Dec. 20, 2023), https://malwaretips.com/blogs/jennifer-aniston-macbook-giveaway-scam [https://perma.cc/S545-L98S].

17. *See* William Sasse, *Deepfakes and the Courtroom*, 2 MD. BAR J. 88, 88 (2020).

18. *See* Chesney & Citron, *supra* note 9, at 1769.

19. *Id.*

20. Simon Chandler, *Why Deepfakes Are a Net Positive for Humanity*, FORBES (Mar. 9, 2020, 12:33 PM), https://www.forbes.com/sites/simonchandler/2020/03/09/why-deepfakes-are-a-net-positive-for-humanity/?sh=4da40e172f84 [https://perma.cc/N2QR-C7BE].

greets and talks to visitors at the Dalí Museum in Florida.[21] Moreover, deepfakes may be a powerful tool for enhancing accessibility. For instance, individuals suffering from permanent disabilities can use deepfakes to digitally engage in activities that their disabilities previously precluded.[22] Likewise, the entertainment industry can use deepfake dubbing to seamlessly translate films and television, adapting the content for new audiences.[23]

Reflecting the rapid evolution of AI technology over the last decade, "barriers to deepfake creation have lowered dramatically."[24] Creating a convincing deepfake was once a complex, time-consuming, and expensive process.[25] First, the creator needed a graphics processing unit ("GPU") to run the software.[26] While the software is "free, open source, and easily downloadable,"[27] a GPU can cost as little as a couple hundred dollars and as much as a couple thousand.[28] DeepFaceLab and FaceSwap are the most commonly used software and can be found on GitHub.[29] In

---

21. *Id.* Notably, postmortem deepfakes of this kind have generated significant controversy. *See, e.g.*, Catherine Shoard, *Peter Cushing is Dead.* Rogue One*'s Resurrection is a Digital Indignity*, THE GUARDIAN (Dec. 21, 2016, 12:55 PM), https://www.theguardian.com/commentisfree/2016/dec/21/peter-cushing-rogue-one-resurrection-cgi [https://perma.cc/K5UF-WSR2]; Helen Rosner, *The Ethics of a Deepfake Anthony Bourdain Voice*, NEW YORKER (July 17, 2021), https://www.newyorker.com/culture/annals-of-gastronomy/the-ethics-of-a-deepfake-anthony-bourdain-voice [https://perma.cc/NZW9-GB2T]. For a comprehensive discussion of the ethical concerns surrounding postmortem deepfakes, see Olivia Wall, *A Privacy Torts Solution to Postmortem Deepfakes*, 100 WASH. U. L. REV. 885 (2023).

22. Chesney & Citron, *supra* note 9, at 1771 ("[D]eep-fake audio technology holds promise to restore the ability of persons suffering from certain forms of paralysis, such as ALS, to speak with their own voice.").

23. James Vincent, *Deepfake Dubs Could Help Translate Film and TV Without Losing an Actor's Original Performance*, THE VERGE (May 18, 2021, 7:13 AM), https://www.theverge.com/2021/5/18/22430340/deepfake-dubs-dubbing-film-tv-flawless-startup [https://perma.cc/P63G-94CA].

24. Dr. A. Shaji George & A.S. Hovan George, *Deepfakes: The Evolution of Hyper Realistic Media Manipulation*, 1 PARTNERS UNIVERSAL INNOVATIVE RSCH. PUBL'N 58, 62 (2023).

25. Catherine Bernaciak & Dominic A. Ross, *How Easy Is It to Make and Detect a Deepfake?*, CARNEGIE MELLON UNIV.: SOFTWARE ENG'G INST. (Mar. 14, 2022), https://insights.sei.cmu.edu/blog/how-easy-is-it-to-make-and-detect-a-deepfake [https://perma.cc/ZQ9W-449Z].

26. *Id.*

27. *Id.*

28. *See* Jacob Roach & Monica J. White, *GPU Prices and Availability (Q1 2024): How Much Are GPUs Today?*, DIGITAL TRENDS (Feb. 8, 2024), https://www.digitaltrends.com/computing/gpu-price-tracking/ [https://perma.cc/8HXR-DH2Q]; Stewart Bendle, *GPU Price Index 2024: Lowest Price on Every Graphics Card From Nvidia, AMD, and Intel Today*, TOM'S HARDWARE, https://www.tomshardware.com/news/lowest-gpu-prices [https://perma.cc/L4G7-9KNN] (July 8, 2024).

29. Bernaciak & Ross, *supra* note 25. GitHub is a web-based platform that allows software developers to create, store, and manage their code. GitHub is fundamental to modern software development, as it fosters seamless collaboration between developers, enabling them to contribute to projects while simultaneously keeping track of the changes made. *See* Ben

addition, deepfake creation required large volumes of high-quality footage of the source and the destination, a "minimum of several minutes" of video.[30] Finally, the creator needed ample time, several weeks to a few months, and a considerable amount of skill.[31] However, the proliferation of user-friendly deepfake apps has disrupted these established norms.[32] Now, creating a deepfake is as simple as downloading an app on your phone and uploading a single source image,[33] "putting sophisticated manipulation tools into the hands of the masses."[34] Alternatively, anyone can now use text-to-image generators such as OpenAI's DALL-E to create realistic deepfake images.[35] Disturbingly, the "trajectory of deepfake advancement shows no signs of slowing down."[36] While app-made deepfakes are not nearly as realistic as those made by professionals, researchers continually "achieve new milestones in AI-synthesized media, [and] consumer apps integrate these" advancements.[37]

Given the manipulative capabilities of deepfakes, it is easy to see how they readily lend themselves to influencing voters and disrupting elections.[38] This Note explores the dangers of deepfakes in this context. Part I evaluates the harms deepfakes pose to elections and the level of risk they pose to the United States. Part II discusses the legal and practical considerations limiting the government's ability to effectively regulate malicious election deepfakes. Part III summarizes the safeguards imposed to protect U.S. elections from deepfakes and analyzes the vulnerabilities of these safeguards. Finally, Part IV puts forth a two-pronged proposal to mitigate these harms.

Lutkevich & Meredith Courtemanche, *GitHub*, TECHTARGET, https://www.techtarget.com/searchitoperations/definition/GitHub [https://perma.cc/NS8C-G3U4] (July 2024).

30. Bernaciak & Ross, *supra* note 25.

31. *Id.*

32. George & George, *supra* note 24, at 63.

33. Geoffrey A. Fowler, *Anyone With an iPhone Can Now Make Deepfakes. We Aren't Ready for What Happens Next.*, WASH. POST (Mar. 25, 2021, 8:00 AM), https://www.washingtonpost.com/technology/2021/03/25/deepfake-video-apps/ [https://perma.cc/Z5QY-BZGE].

34. George & George, *supra* note 24, at 62.

35. Nitasha Tiku, *AI Can Now Create Any Image in Seconds, Bringing Wonder and Danger*, WASH. POST (Sept. 28, 2022, 4:20 PM), https://www.washingtonpost.com/technology/interactive/2022/artificial-intelligence-images-dall-e/ [https://perma.cc/M9A9-N28B]. Further, OpenAI's text-to-video generator, Sora, is in its beta testing phase prior to a wider release to the public. Maxwell Zeff, *OpenAI's New Video Generator Sora is Both Incredible and Concerning*, QUARTZ (Feb. 16, 2024), https://qz.com/openai-sora-video-generator-1851263858 [https://perma.cc/F7YB-EEWT]. Technically, text-generated images and videos do not fall under the original definition of deepfakes, which only covered "sophisticated manipulations of existing audio-visual content," whereas text-to-image and text-to-video generators create entirely new content from textual prompts. Rebekah Robinson, *AI Image Generators Enable the Creation of Fake Pictures to Support Fake News*, CODA (Oct. 31, 2022), https://www.codastory.com/disinformation/ai-image-generators-fake-news/ [https://perma.cc/B6X4-GR8K]. Nonetheless, the term "deepfake" is often used as an umbrella term for all forms of AI-manipulated audio-visual media. *Id.*

36. George & George, *supra* note 24, at 63.

37. *Id.*

38. *See* Frank, *supra* note 1.

I. HARMS POSED BY DEEPFAKES TO U.S. ELECTIONS: A DEEP DIVE

Deepfakes can undermine elections in four distinct ways: (1) by facilitating character assassination of political candidates;[39] (2) by enabling the proliferation of disinformation regarding voting procedures;[40] (3) by eroding public trust in election institutions;[41] and (4) by creating what scholars call "the liar's dividend," a phenomenon in which consumers begin to distrust authentic content.[42]

First, a deepfake depicting a political candidate in a false but compromising situation can injure that candidate's reputation, thus influencing voters and potentially affecting the outcome of an election.[43] For example, a notorious deepfake video depicting Nancy Pelosi slurring her words, likely created with the intention of portraying her as drunk and incompetent, garnered 2.5 million views on Facebook.[44] Similarly, in June 2023, the Ron DeSantis presidential campaign distributed three deepfake images depicting former President Donald Trump embracing Dr. Anthony Fauci, a move likely aimed at conflating Trump with Dr. Fauci, whom was disliked within conservative circles.[45] Furthermore, well-timed distribution could amplify the effects of these deepfakes.[46] To illustrate this point, imagine a deepfake that goes viral on election eve; such a deepfake would be difficult to debunk before voters head to the polls.[47]

Second, deepfakes containing disinformation regarding election procedures can influence the outcome of an election by misleading voters.[48] This situation unfolded during the New Hampshire primaries in January 2024 when voters received robocall messages, ostensibly from President Joe Biden, urging them not to vote in the primary election because doing so would preclude them from voting in the general election in November.[49] It is easy to imagine similar deepfakes featuring false voting locations or incorrect dates aimed at preventing people from

---

39. *See* Chesney & Citron, *supra* note 9, at 1778–79.

40. *See, e.g.*, Tiffany Hsu, *New Hampshire Officials to Investigate A.I. Robocalls Mimicking Biden*, N.Y. TIMES (Jan. 22, 2024), https://www.nytimes.com/2024/01/22/business/media/biden-robocall-ai-new-hampshire.html [https://perma.cc/4AJP-TFGQ].

41. Chesney & Citron, *supra* note 9, at 1779.

42. *Id.* at 1785.

43. *See id.* at 1778–79.

44. *Doctored Nancy Pelosi Video Highlights Threat of "Deepfake" Tech*, CBS NEWS, https://www.cbsnews.com/news/doctored-nancy-pelosi-video-highlights-threat-of-deepfake-tech-2019-05-25/ [https://perma.cc/JJH3-XQJ2] (May 26, 2019, 9:26 AM).

45. Nicholas Nehamas, *DeSantis Campaign Uses Apparently Fake Images to Attack Trump on Twitter*, N.Y. TIMES (June 8, 2023), https://www.nytimes.com/2023/06/08/us/politics/desantis-deepfakes-trump-fauci.html [https://perma.cc/EN84-KE7R].

46. Chesney & Citron, *supra* note 9, at 1778.

47. *See* Rebecca Green, *Counterfeit Campaign Speech*, 70 HASTINGS L.J. 1445, 1488–89 (2019) (discussing the so-called "election-eve problem").

48. *See* Hsu, *supra* note 40. Similarly confusing is the public's tendency to use the terms "misinformation" and "disinformation" interchangeably. While misinformation refers to inaccurate information—"getting the facts wrong"—disinformation refers to "false information which is deliberately intended to mislead." *Misinformation and Disinformation*, AM. PSYCH. ASS'N, https://www.apa.org/topics/journalism-facts/misinformation-disinformation [https://perma.cc/M26G-SG93] (last visited Feb. 1, 2024).

49. Hsu, *supra* note 40.

voting at the polls on election day or mailing their ballots in on time. Such deceptive tactics amount to voter suppression, effectively disenfranchising large segments of eligible voters and compromising the election results.[50]

Third, a deepfake depicting government corruption can erode public trust in electoral institutions.[51] For example, a deepfake portraying officials discussing the commission of voter fraud may cause voters to question the legitimacy of an election. This scenario materialized in October 2023 when a deepfake audio recording, featuring the leader of the Progressive Slovakia party talking about buying votes, circulated on social media just days before the 2023 Slovakian election.[52] Furthermore, these kinds of deepfakes have a greater impact on voters "where strong narratives of distrust already exist."[53] Such narratives undoubtedly exist in the United States: a poll conducted in 2022 found that 58% of Americans have little or no confidence that elections in America reflect the will of the people.[54] The events of January 6, 2021, illustrate the potentially violent effects of widespread distrust in election institutions.[55]

Fourth, as deepfakes increase public skepticism of online content, the more likely the public is "to doubt the authenticity of real audio and video evidence."[56] Professors Bobby Chesney and Danielle Citron coined the term "liar's dividend" to describe this phenomenon.[57] Leveraging the liar's dividend, dishonest candidates can avoid accountability by claiming that any authentic audio-visual content portraying them in a negative light is merely a deepfake.[58] Unsurprisingly, world-class accountability deflector Donald Trump has already caught on to this tactic.[59] In December 2023, he claimed that an ad featuring him struggling to pronounce the

---

50. *See* Alice Nunwick, *Audio Deepfakes Could Significantly Damage Democratic Process, Warns Analyst*, VERDICT (Jan. 23, 2024), https://www.verdict.co.uk/audio-deepfakes-could-significantly-damage-democratic-process-warns-analyst/?cf-view [https://perma.cc/8WUN-NQTK].

51. Chesney & Citron, *supra* note 9, at 1779.

52. Morgan Meaker, *Slovakia's Election Deepfakes Show AI Is a Danger to Democracy*, WIRED (Oct. 3, 2023, 7:00 AM), https://www.wired.co.uk/article/slovakia-election-deepfakes [https://perma.cc/J5H5-AQ9M].

53. Chesney & Citron, *supra* note 9, at 1779.

54. Jennifer Agiesta & Ariel Edwards-Levy, *CNN Poll: Percentage of Republicans Who Think Biden's 2020 Win Was Illegitimate Ticks Back up Near 70%*, CNN, https://www.cnn.com/2023/08/03/politics/cnn-poll-republicans-think-2020-election-illegitimate/ [https://perma.cc/MNN9-7ZU9] (Aug. 3, 2023, 10:18 AM).

55. On January 6, 2021, a riotous mob of Trump supporters stormed the U.S. Capitol to disrupt Congress's certification of the 2020 election results, believing that widespread fraud robbed Trump of reelection. For a comprehensive overview of the assault and the events leading up to it, see Alan Feuer et al., *Jan. 6: The Story So Far*, N.Y. TIMES https://www.nytimes.com/interactive/2022/us/politics/jan-6-timeline.html [https://perma.cc/5BDC-YK9K] (last visited Sept. 13, 2024).

56. Chesney & Citron, *supra* note 9, at 1785.

57. *Id.*

58. *Id.* at 1785–86.

59. Pranshu Verma & Gerrit De Vynck, *AI is Destabilizing the 'Concept of Truth Itself' in 2024 Election*, WASH. POST., https://www.washingtonpost.com/technology/2024/01/22/ai-deepfake-elections-politicians/ [https://perma.cc/56U6-CL22] (Jan. 22, 2024, 1:50 PM).

word "anonymous" at a rally in Montana was AI generated, despite thorough documentation of the event.[60] Conversely, opponents of a candidate may create scandals where none exist by claiming that favorable but genuine depictions of that candidate are, in fact, deepfakes. Arguably, this is the most concerning election-related deepfake harm, as the liar's dividend "destabilizes the concept of truth itself."[61] Chesney and Citron argue that this destabilization of truth creates "greater space for authoritarianism."[62]

Having explored the election-related harms of deepfakes, it is important to identify the individuals or entities likely to perpetuate these harms and those who are susceptible to the deception. There are three distinct groups that are likely to use deepfakes to undermine U.S. elections: (1) foreign governments;[63] (2) political candidates[64] and associated entities;[65] and (3) political candidates' supporters.[66] Each of these groups, driven by a desire to give their preferred candidate (or in the case of political candidates, themselves) a competitive edge, have a strong incentive to use deepfakes for character assassination and to proliferate voting procedure disinformation to influence the outcome of an election in that candidate's favor. Additionally, the National Intelligence Council's findings that foreign governments aimed to "undermine public confidence in the electoral process" during the 2020 U.S. election[67] highlight the likelihood that foreign governments will employ deepfakes in a similar manner. Finally, the liar's dividend is a unique harm that is effectuated by society at large as deepfakes become more prevalent.

In contrast, one's susceptibility to deepfakes is influenced by various demographic factors, albeit to a limited extent. Although empirical research on the subject remains limited,[68] existing studies suggest that older individuals are more susceptible to deepfake trickery.[69] This finding aligns with broader research on

---

60. *Id.*

61. *Id.*

62. Chesney & Citron, *supra* note 9, at 1786.

63. Ali Swenson & Kelvin Chan, *Election Disinformation Takes a Big Leap with AI Being Used to Deceive Worldwide*, AP NEWS, https://apnews.com/article/artificial-intelligence-elections-disinformation-chatgpt-bc283e7426402f0b4baa7df280a4c3fd [https://perma.cc/Q34Q-K5W4] (Mar. 14, 2024, 12:46 AM) ("FBI Director Christopher Wray recently warned about the growing threat, saying generative AI makes it easy for 'foreign adversaries to engage in malign influence.'").

64. *See, e.g.*, Nehamas, *supra* note 45.

65. Campaigns, PACs, donors, interest groups, etc.

66. *See, e.g.*, Marianna Spring, *Trump Supporters Target Black Voters with Faked AI Images*, BBC (Mar. 3, 2024), https://www.bbc.com/news/world-us-canada-68440150 [https://perma.cc/K96M-SFHC].

67. NAT'L. INTEL. COUNCIL, INTEL. CMTY. ASSESSMENT 2020-00078D, FOREIGN THREATS TO THE 2020 US FEDERAL ELECTIONS 2 (2021), https://www.dni.gov/files/ODNI/documents/assessments/ICA-declass-16MAR21.pdf [https://perma.cc/DV5C-RRMC].

68. Markus Appel & Fabian Prietzel, *The Detection of Political Deepfakes*,J. COMPUT.-MEDIATED COMMC'N, July 27, 2022, at 3 (noting that the number of studies on deepfakes has increased since 2020 but that empirical evidence still remains limited).

69. *See* Juniper Lovato et al., *Lifelikeness is in the Eye of the Beholder: Demographics of Deepfake Detection and Their Impacts on Online Social Networks*, CEUR

misinformation indicating that older individuals may be more susceptible to believing false information found online.[70] Higher rates of social media use among younger generations is one potential explanation for this disparity.[71] Despite these trends, susceptibility to deepfakes is not exclusive to any particular age group, as extensive research shows that people are more prone to accept information that aligns with their preexisting views.[72] Therefore, transcending other demographic factors, anyone with strong ideological beliefs is a susceptible target for deepfake manipulation.[73]

Despite growing concern over the impact of deepfakes on elections,[74] some believe that the panic surrounding political deepfakes is a false alarm.[75] For example, some commentators note that even though deepfakes have existed for nearly a decade, they have not yet "taken off as a propaganda technique."[76] However, this argument ignores the evolution of deepfake technology, which has undergone significant advancement in accessibility and sophistication in recent years. First, coinciding with the dramatic increase in accessibility to deepfake technology,[77] deepfakes have become exponentially more prevalent.[78] Specifically, a 2020 report by Sensity AI, a firm specializing in AI threat intelligence,[79] found that the number of deepfakes online was nearly doubling every six months.[80] Second, advances in the sophistication of deepfake technology have resulted in marked improvements in

---

2 (2022), https://ceur-ws.org/Vol-3461/2022-invited-abstract.pdf [https://perma.cc/DE43-CZ3E].

70. Dora-Olivia Vicol, *Who is Most Likely to Believe and to Share Misinformation*, FULL FACT 5 (Feb. 2020), https://fullfact.org/media/uploads/who-believes-shares-misinformation.pdf [https://perma.cc/VFF9-QBRH].

71. Lovato et al., *supra* note 69, at 2. *But see* Juniper Lovato et al., *Diverse Misinformation: Impacts of Human Biases on Detection of Deepfakes on Networks*, ARXIV, 7, https://arxiv.org/abs/2210.10026 [https://perma.cc/G52G-DSMZ] (Jan. 13, 2024) (finding only weak evidence that frequent social media usage impacts one's ability to discern deepfakes).

72. Holly Kathleen Hall, *Deepfake Videos: When Seeing Isn't Believing*, 27 CATH. U. J.L. & TECH. 51, 55 (2018) ("A Pew study from 2016 analyzing '376 million Facebook users' interactions with over 900 news outlets found that people tend to seek information that aligns with their views.' This mindset makes consumers susceptible to misinformation.") (footnotes omitted); Vicol, *supra* note 69, at 9 (explaining what psychologists call "motivated reasoning," a form of cognitive bias).

73. Vicol, *supra* note 70, at 9 (using climate change as an example to show that one's political affiliation has a far greater impact than education, age, levels of training, income, or sex on that individual's beliefs).

74. *See supra* note 4 and accompanying text.

75. *See, e.g.*, Russell Brandom, *Deepfake Propaganda is Not a Real Problem*, THE VERGE (Mar. 5, 2019, 10:25 AM), https://www.theverge.com/2019/3/5/18251736/deepfake-propaganda-misinformation-troll-video-hoax [https://perma.cc/A8YR-LFKZ].

76. *Id.*

77. *See supra* Introduction.

78. Kavyasri Nagumotu, *Deepfakes Are Taking over Social Media: Can the Law Keep Up?*, 62 IDEA: L. REV. FRANKLIN PIERCE CENTER FOR INTELL. PROP. 102, 107 (2022).

79. *Why Sensity*, SENSITY (Mar. 27, 2024), https://sensity.ai/ why-sensity/ [https://perma.cc/7L7N-24PU].

80. Nagumotu, *supra* note 78, at 107.

realism.[81] Newer professionally made deepfakes, lacking the obvious markers that revealed the inauthenticity of their predecessors, "can deceive even expert human reviewers without forensic analysis."[82] Thus, while bad actors in the past seemingly decided that manufacturing deepfakes as a part of their misinformation campaigns "wasn't worth the trouble,"[83] the rapid evolution of deepfake technology is likely to produce a different decision in the future.

Another argument suggests that deepfakes are simply next in a long line of technology with the potential to deceive.[84] However, this argument rests on a false equivalence, as changing times and technology reveal radical differences between deepfakes and previous means of deception.[85] Unlike previous eras in which people obtained their news from traditional news sources, 50% of Americans today obtain their news from social media at least sometimes.[86] While social media companies moderate content to an extent,[87] news found on social media is generally not scrupulously verified for validity like news found through traditional sources.[88] Moreover, "fake news spreads faster online because of how social media has prioritized virality."[89] Thus, the modern information space facilitates and expedites the spread of deepfakes, whereas traditional forms of deceptive media were restrained by established journalistic safeguards. In addition, unlike the "fairly straightforward" detection of images altered in Photoshop, the detection of deepfakes is extremely difficult and is, in some ways, getting harder.[90]

In sum, the fear surrounding political deepfakes is not overstated. In fact, election-related deepfakes have been popping up with increasing regularity, impacting voters in several major 2023 elections, such as those of Slovakia,[91] Poland,[92] and Argentina.[93] This surge has prompted fear that other major 2024

---

81. George & George, *supra* note 24, at 63.

82. *Id.*

83. Brandom, *supra* note 75.

84. *See* N.Y.U. School of Law, *Legislative Solutions, Individual Rights, and the Question of Government Intervention*, YOUTUBE (June 30, 2020), https://youtu.be/81Ppe7Vmo8o?si=el_IusLvRRdyFAgA [https://perma.cc/K2KZ-A9HE].

85. *See infra* note 88 and accompanying text.

86. Jacob Liedke & Luxuan Wang, *Social Media and News Fact Sheet*, PEW RSCH. CTR. (Nov. 15, 2023), https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/ [https://perma.cc/42VF-GTD3].

87. *See infra* Part III.C.

88. Bonai Fan et al., *Why Do You Trust News? The Event-Related Potential Evidence of Media Channel and News Type*, FRONTIERS PSYCH., April 2021, at 6 ("[T]raditional media strictly follows an inherent standard and offers more serious, exhaustive, and in-depth information, which is more credible and trustworthy.").

89. Nagumotu, *supra* note 78, at 118.

90. Chesney & Citron, *supra* note 9, at 1759; *see also infra* Part II.C.

91. *See* Meaker, *supra* note 52 and accompanying text.

92. Antoaneta Roussi, *European Election at Risk From AI, Says EU's Cyber Agency*, POLITICO (Oct. 19, 2023, 12:51 PM), https://www.politico.eu/article/european-union-election-risk-artificial-intelligence-interference-cybersecurity-agency-enisa/ [https://perma.cc/5PJ8-368M].

93. Valerie Wirtschafter, *The Impact of Generative AI in a Global Election Year*, BROOKINGS (Jan. 30, 2024), https://www.brookings.edu/articles/the-impact-of-generative-ai-in-a-global-election-year/ [https://perma.cc/F6B4-YSVK].

elections[94] are the next targets.[95] Thus, the question is not whether deepfakes will be present in the next U.S. election—they are already here. Indeed, the question is whether the United States is prepared for them.

II. THE POLICY DE(EP)BATE: LIMITATIONS TO REGULATING DEEPFAKES

At first glance, banning election-related deepfakes or imposing liability on platforms to stop the spread of election-related deepfakes seem like obvious solutions. However, the First Amendment[96] and Section 230 of the Communications Decency Act[97] ("Section 230") severely limit the U.S. government's ability to take such measures. In addition, several practical considerations similarly limit the potential solutions to the political deepfake problem.

A. The First Amendment

The first limit to regulating deepfakes is the First Amendment to the U.S. Constitution, which protects speech from government interference.[98] Notably, the scope of the First Amendment extends far beyond speech in a literal sense, encompassing "pictures, films, paintings, drawings, engravings, oral utterance and the printed word."[99] Moreover, even false speech is protected.[100] Given these precedents, "there is a growing sense among legal scholars that deepfakes are a form of First Amendment expression."[101] Thus, any law regulating the creation of deepfakes would be subject to a First Amendment analysis.[102]

Under a First Amendment analysis, courts first consider whether the regulation is content based or content neutral.[103] The importance of this distinction

---

94. "2024 has been dubbed a super election year[,]" as more than sixty countries worldwide will hold critical elections. Katharina Buchholz, *2024: The Super Election Year*, STATISTA (Jan. 19, 2024), https://www.statista.com/chart/31604/countries-where-a-national-election-is-was-held-in-2024/ [https://perma.cc/8TEW-H7QW]. Moreover, just eight of those countries—Taiwan, Indonesia, Russia, India, South Africa, the United Kingdom, the European Union and the United States—make up over 50% of the global GDP. *Eight Key Elections to Watch in 2024*, BRUNSWICK GROUP 2 (Sept. 2023), https://www.brunswickgroup.com/media/11259/geopolitical_eightelections2024_091523_final.pdf [https://perma.cc/R87N-M5LJ].

95. Roussi, *supra* note 92.

96. U.S. CONST. amend. I.

97. 47 U.S.C. § 230 (shielding online platforms from liability for user-generated content and for good-faith content moderation).

98. U.S. CONST. amend. I ("Congress shall make no law . . . abridging the freedom of speech . . . .").

99. Kaplan v. California, 413 U.S. 115, 119 (1973) (cleaned up).

100. United States v. Alvarez, 567 U.S. 709, 723 (2012) (holding that the Stolen Valor Act, which made it a crime to falsely claim to have received military decorations, is unconstitutional because criminalizing false speech would cast a chilling effect on the First Amendment).

101. Yinuo Geng, *Comparing "Deepfake" Regulatory Regimes in the United States, the European Union, and China*, 7 GEO. L. TECH. REV. 157, 164 (2023).

102. Nina I. Brown, *Deepfakes and the Weaponization of Disinformation*, 23 VA. J.L. & TECH. 1, 35 (2020).

103. Reed v. Town of Gilbert, 576 U.S. 155, 163–64 (2015).

cannot be understated, as it determines the standard of review a court will use. Generally, "[g]overnment regulation of speech is content based if a law applies to particular speech because of the topic discussed or the idea or message expressed."[104] Content-based laws are subject to strict scrutiny and thus are "presumptively unconstitutional and may be justified only if the government proves that they are narrowly tailored to serve compelling state interests."[105] In contrast, government regulation of speech is content neutral when the regulation applies to all expression without regard to the content of the expression.[106] Content-neutral laws are subject to intermediate scrutiny, and thus "will be sustained if [they] further[] an important or substantial governmental interest; if the governmental interest is unrelated to the suppression of free expression; and if the incidental restriction on alleged First Amendment freedoms is no greater than is essential to the furtherance of that interest."[107] Courts rarely uphold content-based laws because the strict scrutiny standard of review places an enormous burden on the government.[108] In sum, legislators must keep these principles in mind when crafting any law that regulates the creation of election-related deepfakes.

B. Section 230 of the Communications Decency Act

Section 230 of the Communications Decency Act also significantly hinders the regulation of election deepfakes. Because online platforms are "the main medium through which deepfakes circulate,"[109] imposing liability on the platforms is an obvious way to mitigate the election-related harms posed by deepfakes. However, Section 230 generally precludes such a measure.

In the internet's early days, the New York Supreme Court's holding in *Stratton Oakmont, Inc. v. Prodigy Services Co.*[110] ignited fears that online platforms' liability exposure would inhibit the internet's growth.[111] In response, Congress enacted Section 230, providing broad immunity to online platforms.[112] Section

---

104. *Id.* at 163.

105. *Id.*

106. Ward v. Rock Against Racism, 491 U.S. 781, 782 (1989).

107. Turner Broad. Sys. v. FCC, 512 U.S. 622, 662 (1994) (citation omitted).

108. *See* Bolger v. Youngs Drug Prods. Corp., 463 U.S. 60, 65 (1983) ("With respect to noncommercial speech, this Court has sustained content-based restrictions only in the most extraordinary circumstances.").

109. Geng, *supra* note 101, at 165.

110. Stratton Oakmont, Inc. v. Prodigy Servs. Co., No. 31063/94, 1995 WL 323710, at *4 (N.Y. Sup. Ct. May 24, 1995) (holding that Prodigy, the operator of online bulletin boards, is a publisher rather than a distributor, and is thus subject to liability for libelous content posted by its users because Prodigy exercised control over the content posted on its bulletin boards). Interestingly, the film *The Wolf of Wall Street* depicted the notorious fraudulent activities of Stratton Oakmont. *See* THE WOLF OF WALL STREET (Paramount Pictures 2013). In fact, Stratton Oakmont sued Prodigy for libel only after an anonymous user accused the firm of fraud on one of Prodigy's message boards. *Stratton Oakmont*, 1995 WL 323710, at *1.

111. Matt Reynolds, *The Strange Story of Section 230, the Obscure Law That Created Our Flawed, Broken Internet*, WIRED (Mar. 24, 2019, 2:00 AM), https://www.wired.co.uk/article/section-230-communications-decency-act [https://perma.cc/RHT5-4TV5].

112. *Id.*

230(c)(1) provides 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 information content provider."[113] This provision shields online platforms from liability for most[114] user-generated content.[115] Section 230(c)(2)(A) protects online platforms from liability for "any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers . . . objectionable, whether or not such material is constitutionally protected."[116] This provision shields platforms from liability for good-faith content moderation.[117]

These two provisions represent the twin aims of Section 230. First, the Act incentivized investment into internet startups by ensuring that platforms could host content from ordinary users without incurring liability for what those users posted.[118] Second, the Act promoted civil discourse by allowing platforms to engage in good-faith content moderation without incurring liability for doing so.[119] For better or worse, the Act enabled the internet to become what it is today.[120]

In recent years, politicians on both sides of the aisle have taken aim at Section 230.[121] On the one hand, Republicans argue that Section 230(c)(2)(A) provides online platforms too much leeway when it comes to content moderation, resulting in "anti-conservative bias" on social media.[122] On the other hand, Democrats argue that Section 230(c)(1) is too permissive, allowing social media companies to host illegal content on their sites with impunity.[123] No matter which argument you find more persuasive, the outcome is the same: a potential restriction on or repeal of the monumental law.[124] In response, social media companies could do one of two things.[125] First, to avoid liability for the content they host, social media companies could over-moderate their content, resulting in the deterioration of

---

113. 47 U.S.C. § 230(c)(1).

114. The statute carves out exceptions for content that violates any federal criminal statute, any law pertaining to intellectual property, any state law consistent with Section 230, the Electronic Communications Privacy Act, and any sex trafficking law. § 230(e).

115. Chesney & Citron, *supra* note 9, at 1796.

116. § 230(c)(2)(A).

117. Chesney & Citron, *supra* note 9, at 1796.

118. Reynolds, *supra* note 111.

119. *Id.*

120. *See generally* JEFF KOSSEFF, THE TWENTY-SIX WORDS THAT CREATED THE INTERNET (2019).

121. Adi Robertson, *Lots of Politicians Hate Section 230 — But They Can't Agree on Why*, THE VERGE (June 24, 2020, 7:28 AM), https://www.theverge.com/21294198/section-230-tech-congress-justice-department-white-house-trump-biden [https://perma.cc/Y5RC-VE2D].

122. *Id.*

123. *Id.*

124. *Id.*

125. Barbara Ortutay, *What You Should Know About Section 230, the Rule That Shaped Today's Internet*, PBS (Feb. 21, 2023, 10:55 AM), https://www.pbs.org/newshour/politics/what-you-should-know-about-section-230-the-rule-that-shaped-todays-internet [https://perma.cc/4CYS-4KSZ].

functions we know and love.[126] Alternatively, the companies could abandon moderation altogether, leading to the proliferation of problematic or, in some cases, illegal content.[127]

Significantly, courts interpret Section 230 immunity broadly.[128] While there are no legal precedents involving election-related deepfakes, U.S. courts' favorable treatment of Section 230 suggests that the Act would shield social media companies from liability for the harms caused by election-related deepfakes that circulate on the companies' platforms. Thus, absent a revision or repeal of Section 230, measures to impose liability on platforms for facilitating the spread of deepfakes harmful to elections remain off limits. Considering the potential consequences of amending or repealing Section 230, exploring other strategies to address election-related deepfakes is necessary.

C. Practical Considerations

Finally, several practical considerations constrict policymakers as they erect safeguards to protect against political deepfakes. For example, it is imperative that the regulations do not inadvertently stifle the beneficial applications of deepfakes.[129] More broadly, lawmakers should be cognizant of the implications of any regulation on the continued development and advancement of AI technology. A nuanced approach is essential to preserve beneficial uses and foster innovation while simultaneously protecting our elections.[130]

Notably, issues related to the detection of deepfakes severely limit the measures available to lawmakers. First, the efficacy of current deepfake detection tools is somewhat limited.[131] While these tools have undergone significant accuracy improvements in recent years,[132] they are still plagued by many issues, particularly when they encounter low-quality video.[133] Moreover, the mass production of

---

126. *Id.*

127. *Id.*

128. *See* Brown, *supra* note 102, at 42 ("Courts have applied this immunity to claims for defamation, negligence, intentional infliction of emotional distress, privacy, terrorism support, and more.").

129. *See supra* Introduction.

130. *Approaches to Regulating Artificial Intelligence: A Primer*, NAT'L CONF. OF STATE LEGISLATORS, https://www.ncsl.org/technology-and-communication/approaches-to-regulating-artificial-intelligence-a-primer [https://perma.cc/77RV-MA97] (Aug. 10, 2023) ([M]ost AI proponents agree some balance between the innovations of AI with basic human values must be achieved.").

131. Saddam Hossain Mukta et al., *An Investigation of the Effectiveness of Deepfake Models and Tools*, J. SENSOR & ACTUATOR NETWORKS, Aug. 4, 2023, at 35 https://www.mdpi.com/2224-2708/12/4/61 [https://perma.cc/S5H2-6VQK].

132. For example, Sensity AI's popular detection software has demonstrated accuracy scores of up to 95%. *Id.* at 27. In contrast, the winner of Facebook's Deepfake Detection Challenge in 2019 could detect only 82% of the deepfakes it encountered. Nagumotu, *supra* note 78, at 108.

133. Mukta et al., *supra* note 131, at 35.

deepfakes renders even the most precise detection tools incapable of detecting a substantial number of deepfakes.[134]

Second, it's becoming progressively more difficult to detect deepfakes. Each improvement in detection software is met by "a counter effort from creators to bypass" that software.[135] Likewise, "[e]ach new technique for creating realistic fakes spawns new forensic techniques for exposing them,"[136] resulting in an ongoing cycle—in other words, a deepfake arms race. Experts predict the arms race is likely to endure absent a major breakthrough.[137]

Third, issues with scaling existing detection methods pose a formidable challenge. For example, the algorithms underlying these methods rely on massive high-quality data sets, which are inherently limited in availability.[138] Moreover, the algorithms typically have extensive time requirements.[139] Both of these factors prevent efficient replication of current methods on a mass scale, contributing to their limited widespread adoption.[140]

III. IN TOO DEEP? CURRENT SAFEGUARDS DO NOT ADEQUATELY PROTECT U.S. ELECTIONS FROM DEEPFAKES

Current measures to protect U.S. elections from deepfakes are either largely nonexistent or ill-suited to serve their purpose. Five categories of safeguards are addressed below.

A. State and Federal Government Efforts

As of 2024, there is no federal law regulating deepfakes, but legislators have proposed several. The Malicious Deep Fake Prohibition Act of 2018, which would have criminalized the knowing distribution or intent to distribute a deepfake, never made it out of committee.[141] The DEEP FAKES Accountability Act, proposed in 2019, would have imposed penalties for the failure to disclose that a deepfake was artificially manipulated and would have created a task force to enforce the Act, but it similarly never left committee.[142] The bill's author revived the proposal in September 2023, but the bill has failed to gain traction.[143] Finally, the Deepfake Report Act of 2019 would have required the U.S. Department of Homeland Security

134. Alex Engler, *Fighting Deepfakes When Detection Fails*, BROOKINGS (Nov. 14, 2019), https://www.brookings.edu/articles/fighting-deepfakes-when-detection-fails/ [https://perma.cc/XT2D-99PY] ("Further, even highly successful deepfake detection methods are difficult to scale. Identifying 90% of deepfakes may sound excellent, but keep in mind that these materials can be mass-produced and mass-distributed[.]").

135. Nagumotu, *supra* note 78, at 142.

136. George & George, *supra* note 24, at 71.

137. *Id.*

138. Mukta et al., *supra* note 131, at 35.

139. *Id.*

140. *Id.*

141. Malicious Deep Fake Prohibition Act of 2018, S. 3805, 115th Cong.

142. DEEP FAKES Accountability Act, H.R. 3230, 116th Cong. (2019).

143. Emmanuelle Saliba, *Bill Would Criminalize 'Extremely Harmful' Online 'Deepfakes'*, ABC NEWS (Sept. 25, 2023, 11:20 AM), https://abcnews.go.com/Politics/bill-criminalize-extremely-harmful-online-deepfakes/story?id=103286802 [https://perma.cc/R6JH-VJN8].

to report on "digital content forgery," but the proposed bill failed to move forward after passing in the Senate.[144]

While no federal law directly regulates deepfakes, Congress has touched on deepfakes in a few enacted statutes. The National Defense Authorization Acts of each fiscal year since 2020 have included a provision instructing the Director of National Intelligence to submit a report to Congress regarding the national security risks of deepfakes and the use of deepfakes by foreign governments to engage in malicious activities.[145] Additionally, the Identifying Outputs of Generative Adversarial Networks Act instructed the National Science Foundation to research generative adversarial networks[146] and other AI techniques used to create deepfakes.[147] In addition to federal legislative research instructions, the Defense Advanced Research Projects Agency has two programs dedicated to researching deepfake detection.[148]

At the executive level, the Federal Election Commission ("FEC") appears willing to regulate deepfake political advertisements.[149] In August 2023, the FEC unanimously voted to advance a petition submitted by advocacy group Public Citizen asking the agency to regulate advertisements that use AI to "misrepresent political opponents as saying or doing something they didn't."[150] Following a 60-day public comment period, the FEC has not yet taken action.[151] However, even if the FEC implemented rules limiting deepfakes in political advertisements, most harmful deepfakes would be unaffected because the FEC's regulatory authority extends only to campaigns and their parties.[152]

Given the federal government's failure to protect our elections from deepfakes, states have taken matters into their own hands, albeit insufficiently. The legislatures of California and Texas led the way, enacting laws in 2019 that banned the creation and distribution of deepfakes with the intent to hurt a political candidate or influence an election within a designated time period before the election.[153] The legislatures of Minnesota, Washington, and Michigan enacted similar laws in 2023,

---

144. Deepfake Report Act of 2019, S. 2065, 116th Cong.

145. *See* 50 U.S.C. § 3369a.

146. General adversarial networks ("GANs"), sophisticated machine learning models using two neural networks, are partially responsible for the proliferation of increasingly convincing deepfakes. Chesney & Citron, *supra* note 9, at 1760.

147. 15 U.S.C. § 9201.

148. KELLEY M. SAYLER & LAURIE A. HARRIS, CONG. RSCH. SERV., IF11333, DEEPFAKES AND NATIONAL SECURITY 2 (2023).

149. Ali Swenson, *FEC Moves Toward Potentially Regulating AI Deepfakes in Campaign Ads*, AP NEWS (Aug. 10, 2023, 6:48 PM), https://apnews.com/article/fec-artificial-intelligence-deepfakes-election-2024-95399e640bd1e41182f6c631717cc826 [https://perma.cc/K5D8-K9XT].

150. *Id.*

151. *FEC Remains Woefully Behind as States Take Action on AI Deepfakes*, PUBLIC CITIZEN (Feb. 16, 2024), https://www.citizen.org/news/fec-remains-woefully-behind-as-states-take-action-on-ai-deepfakes/ [https://perma.cc/9KVP-L9RX].

152. Swenson, *supra* note 149.

153. CAL. ELECT. CODE § 20010 (2023); TEX. ELEC. CODE § 255.04 (2023), *held unconstitutional in part by Ex parte* Stafford, 667 S.W.3d 517 (Tex. Ct. App. 2023) (holding that a different subsection violated the First Amendment).

while those in 44 additional states and the District of Columbia have put forth similar bills in the current legislative session, several of them having been enacted.[154] However, these laws suffer from a host of problems. First, ascertaining the identity of a deepfake's creator can be very challenging, and in many cases impossible.[155] Thus, laws imposing civil or criminal liability on deepfake creators lack any deterrent effect because anonymous creators will escape liability. Second, the government may face challenges proving the intent element of each statute[156] given the range of reasons individuals create deepfakes.[157] Third, many scholars opine that these laws would fail to survive a First Amendment challenge.[158]

B. Voluntary Commitments by AI Firms

In July 2023, the Biden Administration secured voluntary commitments from seven AI firms to manage the dangers posed by AI.[159] The firms included industry leaders such as OpenAI, Google, Meta, and Microsoft.[160] Regarding deepfakes, the companies pledged to develop mechanisms to ensure that users know when content is AI generated, such as the development of a watermarking system.[161] In September 2023, eight more companies including Adobe and IBM joined the pledge.[162] While cooperation between the government and tech firms is undoubtedly necessary to protect our elections from deepfakes, the watermarking proposal is insufficient because the open source nature of deepfake creation software permits

---

154. *Tracker: State Legislation on Deepfakes in Elections*, PUBLIC CITIZEN, https://www.citizen.org/article/tracker-legislation-on-deepfakes-in-elections/ [https://perma.cc/2HQW-JRTG] (July 29, 2024).

155. Nagumotu, *supra* note 78, at 125.

156. *See, e.g.*, CAL. ELEC. CODE § 20010 (2023) (requiring that the deepfake was made with "the intent to injure the candidate's reputation or to deceive a voter[.]"); TEX. ELEC. CODE § 255.04 (2023) (requiring that the deepfake was made "with intent to injure a candidate or influence the result of an election[.]); MINN. STAT. § 609.771 (2023) (requiring that the deepfake was "made with the intent to injure a candidate or influence the result of an election.").

157. *See, e.g.*, *infra* note 165.

158. *See, e.g.*, Alexandra Tashman, *"Malicious Deepfakes"—How California's A.B. 730 Tries (and Fails) to Address the Internet's Burgeoning Political Crisis*, 54 LOY. L.A. L. REV. 1391, 1408–16 (2021) (arguing that California's law, subject to a strict scrutiny review because it governs political speech specifically, is neither narrowly tailored nor in furtherance of a compelling government interest).

159. Press Release, The White House, FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI (July 21, 2023), https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/ [https://perma.cc/K9WM-KGNR].

160. *Id.*

161. *Id.*

162. Cecilia Kang, *8 More Companies Pledge to Make A.I. Safe, White House Says*, N.Y. TIMES (Sep. 12, 2023), https://www.nytimes.com/2023/09/12/technology/white-house-ai-tech-pledge.html [https://perma.cc/RM8X-KMBA].

modification of the source code, thereby allowing sophisticated actors to bypass any watermarking requirement imposed by the firms.[163]

C. Self-Regulation by Online Platforms

The three social media sites from which Americans are most likely to obtain news—Facebook, YouTube, and X, formerly known as Twitter[164]—have each promulgated guidelines limiting deepfakes on their platforms. Under Facebook's policy, the platform removes AI-manipulated content that is likely to mislead; however, the policy creates an exception for parody and satire.[165] YouTube bans artificially manipulated content that may mislead viewers and poses a serious risk of egregious harm.[166] Finally, under X's policy, the platform may label manipulated content that is likely to deceive or confuse users, while content likely to cause widespread confusion or serious harm is subject to removal.[167] Although willingness on the part of these platforms to manage deepfakes is promising, gauging the effectiveness of these policies remains challenging due to the platforms' reluctance to disclose internal data.

In addition, in November 2023, Google platforms, including YouTube, began requiring disclaimers on political advertisements that feature "synthetic

---

163. R. Michael Alvarez et al., *Generative AI and the Future of Elections*, CTR. FOR SCI., SOC'Y, AND PUB. POL'Y (July 21, 2023), https://lindeinstitute.caltech.edu/documents/25475/CSSPP_white_paper.pdf [https://perma.cc/FPX6-AL8Q].

164. Elisa Shearer & Elizabeth Grieco, *Americans Are Weary of the Role Social Media Sites Play in Delivering the News*, PEW RSCH. CTR. (Oct. 2, 2019), https://www.pewresearch.org/journalism/2019/10/02/americans-are-wary-of-the-role-social-media-sites-play-in-delivering-the-news/ [https://perma.cc/3MPT-4ADG].

165. Monika Bickert, *Enforcing Against Manipulated Media*, FACEBOOK (Jan. 6, 2020), https://about.fb.com/news/2020/01/enforcing-against-manipulated-media/ [https://perma.cc/J2W5-MK9G]. Regarding parody and satire, many good-intentioned creators have used deepfake technology to make truly hilarious satirical and parodical content. *See, e.g.*, @ZackDAbrams, X (Oct. 20, 2023, 2:13 PM), https://twitter.com/ZackDAbrams/status/1715476256932692238 [https://perma.cc/Z6A9-TH4S] (depicting former President Barack Obama on the body of rapper Ice Spice addressing his haters); r/midjourney, REDDIT (Mar. 24, 2023), https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/ [https://perma.cc/4A7W-PB8N] (depicting Pope Francis wearing a trendy white Balenciaga puffer jacket); @WhatEverWillie, X (July 9, 2023, 6:55 PM), https://twitter.com/WhatEverWillie/status/1677858958936006657 [https://perma.cc/XTC3-XUD6] (depicting rapper Nicki Minaj and actor Tom Holland as a newlywed couple describing a recent home invasion by Meta CEO Mark Zuckerberg).

166. *Misinformation Policies*, YOUTUBE, https://support.google.com/youtube/answer/10834785 [https://perma.cc/W5JV-QVZ9] (last visited Apr. 19, 2024).

167. *Synthetic and Manipulated Media Policy*, X (Apr. 2023), https://help.twitter.com/en/rules-and-policies/manipulated-media [https://perma.cc/HL3J-AS5T]. Although this remains X's official policy regarding deepfakes, Elon Musk, upon taking ownership of the platform, took steps to dismantle its content moderation system, opting instead for a reliance on crowdsourced "community notes" written by users. Jim Rutenberg & Kate Conger, *Elon Musk Is Spreading Election Misinformation, but X's Fact Checkers Are Long Gone*, N.Y. TIMES, https://www.nytimes.com/2024/01/25/us/politics/elon-musk-election-misinformation-x-twitter.html [https://perma.cc/9AV3-FUCU] (Jan. 29, 2024). Recent deepfake scandals on the platform have highlighted the failures of the policy. *See supra* note 12.

content that inauthentically depicts real or realistic-looking people or events."[168] Meta, the owner of Facebook, quickly followed suit on all of its platforms, requiring political advertisements to disclose whether they were created using AI.[169] Unfortunately, these rules are unlikely to have a big impact because they apply exclusively to political advertisements and not user-generated content, leaving a majority of the harmful deepfakes unaffected.

D. Tort Law

Some scholars suggest that existing torts—specifically defamation, false light invasion of privacy, intentional infliction of emotional distress, and right of publicity—adequately deter malicious election deepfakes and hold their creators accountable.[170] However, upon closer inspection, these causes of action are insufficient to address the political deepfake problem.

First and foremost, the difficulty in ascertaining the identity of a deepfake creator[171] renders any cause of action unfeasible. Identifying the account that first posted a deepfake presents its own challenge due to the rapid spread of content on online platforms.[172] The identification problem is exacerbated by the ability of deepfake creators to anonymize their identities using easily accessible technologies.[173] The absence of an identifiable creator leaves potential plaintiffs with no one to sue. Moreover, many deepfake creators live outside of the United States and thus are beyond the jurisdiction of U.S. courts.[174]

In the rare circumstances that a creator is identified and within U.S. jurisdiction, success in any of these causes of action is questionable at best. The defense of parody provides a formidable roadblock to plaintiffs in defamation, false light invasion of privacy, and intentional infliction of emotional distress suits. Success in a defamation or false light invasion of privacy suit requires proof that the

---

168. Michael Bayes et al., *Google to Require Disclaimer on Political Ads with AI-Generated Content: Move Follows Proposed Legislation and Efforts to Prompt FEC Action*, JD SUPRA (Sept. 13, 2023), https://www.jdsupra.com/legalnews/google-to-require-disclaimer-on-3980445/ [https://perma.cc/CHE5-69XT].

169. David Klepper, *To Help 2024 Voters, Meta Says It Will Begin Labeling Political Ads That Use AI-Generated Imagery*, AP NEWS (Nov. 8, 2023, 6:02 PM), https://apnews.com/article/meta-facebook-instagram-political-ads-deepfakes-2024-c4aec653d5043a09b1c78b4fb5dcd79b [https://perma.cc/SFZ9-V7VU].

170. Zachary Shapiro, *Deep Fakes Accountability Act: Overbroad and Ineffective*, B.C. INTELL. PROP. & TECH. F. 2020, at 16 ("[C]ourts should use existing . . . state tort liability to prevent the harmful use of deep fakes.").

171. *See supra* note 155 and accompanying text.

172. Nagumotu, *supra* note 78, at 125.

173. Chesney & Citron, *supra* note 9, at 1794. There are three common methods by which internet users may remain anonymous online—proxy servers, Tor, and VPNs. While each method differs significantly, they all essentially work by preventing the receiving web server from learning the web user's IP address and location. Jacob Roach, *VPN vs Proxy vs Tor: Remaining Anonymous Online in 2023*, CLOUDWARDS, https://www.cloudwards.net/vpn-vs-proxy-vs-tor/ [https://perma.cc/EGZ4-G8QB] (July 17, 2024).

174. Chesney & Citron, *supra* note 9, at 1792.

defendant made a false statement of fact.[175] Parody, however, cannot constitute a false statement of fact.[176] Thus, parody may be an absolute defense to a defamation or false light invasion of privacy claim if "the creator can show that the deepfake cannot be perceived as real."[177] To succeed in an intentional infliction of emotional distress suit, a plaintiff must prove that the defendant intentionally or recklessly caused the emotional distress giving rise to the claim.[178] While not an absolute defense, plaintiffs may encounter difficulty proving the intent element where one may reasonably construe the deepfake as parody or satire.

In addition, individuals harmed by malicious election deepfakes would be unlikely to prevail in a right of publicity suit due to the commercial gain element. To prevail in a right of publicity misappropriation suit, the plaintiff must prove that the defendant appropriated the plaintiff's likeness for the defendant's commercial advantage.[179] But in the electoral context, creators generally use deepfakes to effectuate their political agenda, not for commercial gain. Without this essential element, plaintiffs are left without recourse.

Moreover, several other considerations make tort law an impractical solution to the deepfake problem. Significantly, tort remedies only provide redress to plaintiffs.[180] Thus, while applicable when a deepfake is used for character assassination, tort law does little to address the other election-related harms of deepfakes, which affect society at large.[181] Finally, the legal system is slow,[182] whereas the harms posed by deepfakes often manifest quickly.[183] Therefore, by the time an affected individual is afforded a remedy, the damage will have already been done.

---

175. *See* RESTATEMENT (SECOND) OF TORTS § 558(a) (AM. L. INST. 1965); *Id.* § 652E.

176. Erik Gerstner, *Face/Off: "Deepfake" Face Swaps and Privacy Laws*, 87 DEF. COUNS. J. 1, 5 (2020).

177. Eric Kocsis, *Deepfakes, Shallowfakes, and the Need for A Private Right of Action*, 126 DICK. L. REV. 621, 640 (2022).

178. RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM § 46 (AM. L. INST. 2010).

179. Fifty-Six Hope Rd. Music, Ltd. v. A.V.E.L.A., Inc., 778 F.3d 1059, 1072 (9th Cir. 2015) ("State publicity right claims protect a plaintiff when the defendant uses the plaintiff's identity for commercial advantage, without permission.")

180. *See, e.g.*, Kenney v. Liston, 760 S.E.2d 434, 445 (W. Va. 2014) ("The primary unifying principle of tort law is one of corrective justice, that is, the law establishes a legal duty for a tortfeasor to repair any damage or losses carelessly inflicted upon a victim.").

181. *See supra* Part I.

182. Jeffrey Kluger, *Why Is the Court System So Slow?*, TIME (June 30, 2016, 7:58 AM), https://time.com/4389196/why-is-the-court-system-so-slow/ [https://perma.cc/SG9Q-K4NY] ("Civil cases in federal courts take an average of nearly two years to reach a resolution from the time of the initial filing[.]").

183. *See* Green *supra* note 47 and accompanying text.

E. Copyright Law

Next, some scholars argue that copyright law can counter malicious election deepfakes.[184] However, individuals harmed by malicious election deepfakes pursuing copyright claims face the same identification, jurisdiction, and timing issues as those pursuing tort claims.[185] Notwithstanding these issues, copyright law would afford such an individual relief only in exceptional circumstances.[186] The images, videos, and audio recordings used to create a deepfake are all subject to copyright protection.[187] Thus, the copyright owner may initiate an action for copyright infringement.[188] However, the copyright owner is generally the creator of the copyrighted content and not its subject.[189] Accordingly, the subject of the deepfake would generally not have a claim for copyright infringement. Moreover, the creator of the copyrighted content has little incentive to bring a suit because the subject, and not the creator, suffers the reputational harm associated with malicious election deepfakes.

Even if the copyright owner initiates a suit for copyright infringement, "[t]he prospects for success . . . are uncertain."[190] To determine whether the deepfake creator infringed the copyright, courts may apply the fair use doctrine.[191] Under this doctrine, courts consider the following factors:

> (1) the purpose and character of the use . . . ; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work.[192]

Under this doctrine, outcomes will differ depending on the circumstances, but it is easy to see how deepfakes in this context could constitute fair use. The first factor considers whether the work is primarily commercial and whether the work is transformative.[193] As noted previously, deepfakes in the electoral context are generally used to effectuate one's political goals and not for commercial purposes.

---

184. Shapiro, *supra* note 170, at 16 ("[C]ourts should use existing copyright . . . liability to prevent the harmful use of deep fakes.").

185. *See supra* notes 171–74, 182–83 and accompanying text.

186. *See* Chesney & Citron, *supra* note 9, at 1793.

187. *See* 17 U.S.C. § 102(a)(5)–(7).

188. *See* § 501.

189. *See* § 201.

190. Chesney & Citron, *supra* note 9, at 1793. Notably, the U.S. Copyright Office plans to put out three reports in 2024 addressing the question of whether machine-generated works are eligible for copyright protections. Cecilia Kang, *The Sleepy Copyright Office in the Middle of a High-Stakes Clash Over A.I.*, N.Y. TIMES, https://www.nytimes.com/2024/01/25/technology/ai-copyright-office-law.html [https://perma.cc/QX2Q-MRUS] (Jan. 26, 2024). Moreover, the *New York Times* sued OpenAI, alleging that the AI company violates copyright laws by training its models on copyrighted content. *Id.* The findings of the reports and the outcome of the case will play a crucial role in determining whether an individual has a viable copyright claim against a deepfake creator.

191. *See* U.S.C. § 107.

192. *Id.*

193. Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 578–79 (1994).

Additionally, as the purpose of a deepfake is often different than that of the corresponding original work, many election-related deepfakes may constitute transformative work.[194] Thus, in many cases, the first factor would favor a finding of fair use.[195] Second, the copyrighted content in this context, usually media depicting elected officials, has a more informative rather than creative nature.[196] Thus, the second factor also favors a finding of fair use. The third factor is fact-specific, so outcomes will differ based on the specific facts of the deepfake at issue. Finally, there is no reason to believe that a deepfake using copyrighted election-related content would have any impact on that content's profitability. Weighing these four factors, a court, in many circumstances, is likely to find that a harmful deepfake in the electoral context constitutes a fair use.

IV. THE PATH FORWARD: A DE(E)PENDABLE STRATEGY TO MITIGATE HARM

Several scholars have noted that there is no "silver-bullet" solution to mitigate the harms associated with election-related deepfakes.[197] Thus, an effective strategy must be multifaceted such that it can adapt to the complex and ever-evolving landscape of deepfakes. This Note advocates for a two-pronged strategy based on: (1) robust media authentication, and (2) widespread digital literacy education. Because this proposal does not restrict deepfake creation nor regulate the platforms on which deepfakes spread, it is not limited by the First Amendment or Section 230.[198] Moreover, because this proposal is not judicial in nature, it avoids any identification, jurisdiction, and timing issues faced by proposed judicial remedies.[199]

A. Media Authentication

The first prong of this proposal requires the establishment of a robust media authentication framework. Avoiding the problems of deepfake detection, media authentication works by verifying the authenticity of genuine content rather than identifying the fake. Blockchain, the innovative technology at the heart of most cryptocurrencies, underlies this system.[200]

---

194. *See* Ty, Inc. v. Publ'ns Int'l, Ltd., 333 F. Supp. 2d 705, 712 (N.D. Ill. 2004) ("In other words, a work is not considered transformative if it serves the same purpose as plaintiff's original or derivative works.").

195. *See* Hustler Mag., Inc. v. Moral Majority Inc., 796 F.2d 1148, 1152 (9th Cir. 1986) ("If the work is used for a commercial or profit-making purpose, the use is presumptively unfair.").

196. *See id.* at 1153–54 ("The scope of fair use is greater when 'informational' as opposed to more 'creative' works are involved.").

197. *See* Chesney & Citron, supra note 9, at 1758; Nicholas O'Donnell, *Have We No Decency? Section 230 and the Liability of Social Media Companies for Deepfake Videos*, 2021 U. ILL. L. REV. 701, 715 (2021).

198. *See supra* Parts II.A, II.B.

199. *See supra* notes 171–74, 182–83, 185 and accompanying text.

200. Evin Cheikosman et al., *Blockchain Can Help Combat the Threat of Deepfakes. Here's How*, WORLD ECON. F. (Oct. 12, 2021), https://www.weforum.org/agenda/2021/10/how-blockchain-can-help-combat-threat-of-deepfakes/ [https://perma.cc/2W57-C8RX].

Blockchain is a digital, decentralized, and immutable ledger that securely records and verifies data across a network of computers.[201] Under this system, media could be uploaded to a public blockchain and stored there as a unique hash value.[202] Utilizing the immutability of the blockchain, the resulting hash value would become "a tamper-proof reference of the digital content at a specific point in time."[203] Any alteration to the original media, no matter how minor, would result in a completely different hash value when someone uploads the altered media to the blockchain.[204] Thus, "[a]nyone could then compare a file and its metadata to the blockchain version to prove or disprove authenticity."[205] The election-related harms of deepfakes rely on significant amounts of people being deceived by them. If individuals have the means to verify the authenticity of media, they cannot fall victim to the deception, thereby preventing the potential harms from materializing.

Notably, OpenAI CEO Sam Altman called for the creation of a new federal agency dedicated to regulating AI.[206] At first glance, such an agency seems like an ideal candidate to establish a media authentication system. However, the rapidly evolving nature of deepfakes necessitates a flexible approach to AI regulation.[207] Moreover, deepfakes already implicate multiple established federal agencies—the Federal Trade Commission, the Federal Communications Commission, and the Federal Election Commission, among others. Accordingly, a coordinated effort among multiple agencies is more practical. Specifically, creating and implementing the system should be a joint venture of the chief AI officers[208] of the relevant agencies.

B. Digital Literacy

Seeking to educate citizens to mitigate the dissemination of harmful election deepfakes, the second prong of the proposal requires a widespread digital

---

201. David Rodneck & Michael Adams, *Understanding Blockchain Technology*, FORBES, https://www.forbes.com/advisor/investing/cryptocurrency/what-is-blockchain/ [https://perma.cc/X337-SD7Q] (May 23, 2023).

202. Cheikosman et al., *supra* note 200.

203. *Id.*

204. *Id.*

205. *Science & Tech Spotlight: Combatting Deepfakes*, U.S. GOV'T ACCOUNTABILITY OFF. (Mar. 11, 2024), https://www.gao.gov/products/gao-24-107292 [https://perma.cc/G7D4-5BRK].

206. Brian Fung, *US Senator Introduces Bill to Create a Federal Agency to Regulate AI*, CNN (May 18, 2023, 5:00 AM), https://www.cnn.com/2023/05/18/tech/bennet-digital-regulator-bill-ai-provisions/index.html [https://perma.cc/N25K-594F].

207. Alexandra Kelley, *AI Regulation Will Come from Existing Frameworks – Not a New Agency – lawmakers say*, NEXT GOV (Feb. 8, 2024), https://www.nextgov.com/artificial-intelligence/2024/02/ai-regulation-will-come-existing-frameworks-not-new-agency-lawmakers-say/394030/ [https://perma.cc/N958-UKJG] ("'I'm not sure that they need a brand new agency,' [House Representative Jay Obernolte] said, explaining that future legislative efforts will need to be flexible.").

208. Kathryn Watson, *White House Orders Federal Agencies to Name Chief AI Officers*, CBS NEWS, https://www.cbsnews.com/news/white-house-chief-ai-officers-federal-agencies-artificial-intelligence/ [https://perma.cc/3RX3-W58R] (Mar. 28, 2024, 10:49 AM).

literacy campaign promulgated by the executive branch of the federal government. Empirical research backs up the efficacy of such a campaign.[209]

Drawing from the successful[210] model of the COVID-19 public education campaign, the Executive should tailor the campaign "to meet the needs of diverse communities," coordinate "across national, state, and local levels," and "engage with the private and public sector."[211] However, the COVID-19 public education campaign still encountered its fair share of public backlash.[212] Addressing this issue necessitates the engagement of local leaders, tapping into their established credibility[213] to effectively disseminate key messages.

First and foremost, the campaign should educate the public about the existence of deepfakes and their potential for harm. Additionally, it should equip the public with the necessary skills to interact with the media authentication system.[214] Furthermore, the campaign should provide the public tips to distinguish real from manipulated content.[215] Research indicates that individuals who successfully identify deepfakes often rely on technological glitches and content anomalies.[216] For instance, the subjects of deepfakes rarely blink.[217] The campaign should strongly encourage citizens to pay close attention to these markers.[218] Beyond content identification, the campaign should also focus on broader aspects of digital literacy,

---

209. Yoori Hwang et al., *Effects of Disinformation Using Deepfake: The Protective Effect of Media Literacy Education*, 24 CYBERPHYSCHOLOGY, BEHAV., & SOC. NETWORKING 188, 192 (Mar. 17, 2021), https://www.liebertpub.com/doi/full/10.1089/cyber.2020.0174 [https://perma.cc/S4P4-BQCE] ("This study suggests that media literacy education focusing on disinformation literacy in general could have protective effects for addressing the negative effects of deepfake videos.").

210. *See* Benjamin Denison et al., *Evaluation of the "We Can Do This" Campaign Paid Media and COVID-19 Vaccination Uptake, United States, December 2020–January 2022*, 28 J. HEALTH COMMC'N 573 (Aug. 1, 2023) (finding a statistically significant relationship between the "We Can Do This" media campaign and increased vaccination rates).

211. THE WHITE HOUSE, NATIONAL STRATEGY FOR THE COVID-19 RESPONSE AND PANDEMIC PREPAREDNESS 8, 10 (Jan. 2021), https://www.whitehouse.gov/wp-content/uploads/2021/01/National-Strategy-for-the-COVID-19-Response-and-Pandemic-Preparedness.pdf [https://perma.cc/8FQY-NA6X].

212. *See, e.g.*, Natasha Korecki & Sarah Owermohle, *Attacks on Fauci Grow More Intense, Personal and Conspiratorial*, POLITICO (June 4, 2021, 11:00 AM), https://www.politico.com/news/2021/06/04/fauci-attacks-personal-conspiratorial-491896 [https://perma.cc/GEH3-TLG9].

213. Jeffrey M. Jones, *Americans Trust Local Government Most, Congress Least*, GALLUP (Oct. 13, 2023), https://news.gallup.com/poll/512651/americans-trust-local-government-congress-least.aspx [https://perma.cc/4WNQ-6RSG].

214. *See supra* Section IV.A.

215. Note that, while these tips would provide a means for individuals to detect rudimentary or app-generated deepfakes, professional deepfakes can rarely be distinguished without forensic analysis. *See supra* notes 81–82 and accompanying text. Moreover, as deepfake technology advances, the utility of these tips will diminish. *See id.*

216. Appel & Prietzel, *supra* note 68, at 9.

217. Wirtschafter, *supra* note 93 (listing questions to conduct a "glitch analysis" of content using the technological glitches that reveal a deepfake's falsity).

218. Appel & Prietzel, *supra* note 68, at 10.

such as fact-checking and source identification, empowering citizens to intelligently navigate the online environment in the age of deepfakes.

C. Additional Considerations

Considering the global nature of the deepfake problem,[219] a true solution requires international cooperation. Some experts have proposed the creation of an international AI agency, drawing inspiration from the International Atomic Energy Agency created to curb the proliferation of nuclear weapons.[220] Furthermore, nations should consider entering into international agreements to refrain from the use of deepfakes to effectuate political agendas, mirroring regional agreements like the European Union's AI Act.[221] Similarly, sanctions should be imposed on nations that engage in disinformation campaigns using deepfakes.[222] While international agreements of this kind lack strong enforcement mechanisms, these measures would represent a pivotal step toward establishing international norms to counter the harmful effects of election deepfakes.

CONCLUSION

Like it or not, deepfakes are here, and they are here to stay. While the full extent of their impact remains uncertain, it is undeniable that they will exert influence and perhaps even lead our electoral system to the brink of destruction. Unfortunately, the U.S. government's response has been lackluster, marked by a failure to enact any comprehensive safeguards against malicious election deepfakes. Furthermore, the efficacy of current judicial remedies is compromised by a variety of issues with respect to identification, jurisdiction, and timing. In light of these inadequacies, it is imperative to adopt a more proactive stance. The two-pronged approach proposed here is meant to circumvent these systemic deficiencies. Moreover, it is intended to lay the groundwork for a safe and sustainable digital environment for all elections to come.

---

219. *See supra* note 174 and accompanying text.

220. Cecilia Kang & Adam Satariano, *Five Ways A.I. Could Be Regulated*, N.Y. TIMES (Dec. 6, 2023), https://www.nytimes.com/2023/12/06/technology/artificial-intelligence-regulation.html [https://perma.cc/CXB4-RVAS]. Interestingly, commentators have frequently drawn comparisons between artificial intelligence and nuclear weapons. *See, e.g.*, Ian Prasad Philbrick & Tom Wright-Piersanti, *A.I. or Nuclear Weapons: Can You Tell These Quotes Apart?*, N.Y. TIMES (June 10, 2023), https://www.nytimes.com/2023/06/10/upshot/artificial-intelligence-nuclear-weapons-quiz.html [https://perma.cc/6ES8-SSCV].

221. *See* EUR. PARLIAMENTARY RSCH. SERV., PE 690.039, TACKLING DEEPFAKES IN EUROPEAN POLICY 61 (Jul. 2021).

222. *See id.*

***