DEEPFAKES: IDENTITY MISAPPROPRIATION IN THE DIGITAL AGE

BY: BENJAMIN TANDY*

INTRODUCTION

In an April 2018 Public Service Announcement, former President Obama warned the Nation that we were entering an age where our enemies could make it appear as though anyone said anything.[1] President Obama added that our enemies could even make him appear to say anything—even something like, "President Trump is a complete dipshit."[2] The video ultimately revealed that it was not President Obama speaking.[3] Rather, it was comedian Jordan Peele doing a voice impression of the former

* Juris Doctor candidate, Belmont University College of Law, 2025; B.B.A., Baylor University, 2019. I want to thank the Honorable Justice Harold See for his guidance in developing this note, as well as Professors Elizabeth Usman and Amy Moore for their investment in my personal and academic growth and success. Additionally, I am grateful to the entire Belmont Law Review editorial team. Lastly, thank you to my wife, family, and friends for their love and encouragement throughout my law school career.

1. BuzzFeedVideo, *You Won't Believe What Obama Says In This Video!*, YOUTUBE (Apr. 17, 2018), https://www.youtube.com/watch?v=cQ54GDm1eL0 [https://perma.cc/7VSD-B6SD].
2. *Id.*
3. *Id.*

president overlayed with a deepfake video.[4] While this video was merely a harmless deepfake intended to spread awareness of the capabilities of deepfake technology, immense potential exists for this technology to be used by bad actors.[5]

For instance, in January 2023, a popular internet personality and Twitch streamer named Blaire, known online as QTCinderella, publicly became a victim of nonconsensual deepfake pornography by another Twitch streamer, Atrioc.[6] Atrioc purchased the deepfake pornography of Blaire from a website specializing in the creation of such content.[7] One day, when Atrioc was livestreaming, he switched between tabs and the deepfake porn website flashed on screen for a moment.[8] Because this was done on a livestream, viewers could rewind the video to when the tab appeared and take a screenshot.[9] Once viewers screenshotted the image and enhanced it, it was clear that the deepfake pornography was appropriating Blaire's identity.[10]

Blaire's inbox filled with screenshots of what appeared to be a pornographic video of her.[11] She recognized her face but not her body.[12] Seeing her face on the naked body made Blaire pause and question whether the images were real, before realizing that she had never filmed herself in the nude.[13] Blaire spoke out after the incident and called for the implementation of federal laws to combat this type of conduct.[14] To provide a remedy for victims of deepfakes who, like Blaire, had their reputations publicly tarnished, Congress should create laws to punish those that create nonconsensual deepfakes of living persons and provide a private right of action for such victims.

Part I of this note gives a general history of the development and the issues created by deepfakes. Part II discusses the ways in which state legislatures have attempted to address deepfakes. Part III discusses the federal government's attempt to address the use of deepfakes. Part IV proposes that a federal law should be enacted by Congress to prohibit the

---

4. *Id.*
5. *See id.*
6. Andrew Court, *Twitch Star QTCinderella's deepfake porn nightmare: 'F—k the internet,'* N.Y. POST (Feb. 6, 2023, 3:19 PM), https://nypost.com/2023/02/06/twitch-star-tearfully-reveals-shes-victim-of-deepfake-porn-f-k-the-internet/ [https://perma.cc/P56W-MCP8]; Jenna Ryu, *She discovered a naked video of herself online, but it wasn't her: The trauma of deepfake porn*, USA TODAY (Feb. 14, 2023, 7:39 AM), https://www.usatoday.com/story/life/health-wellness/2023/02/14/qtcinderella-deepfake-trauma-nonconsensual-porn/11222588002/ [https://perma.cc/ZGS3-YHA6].
7. Court, *supra* note 6; Ryu, *supra* note 6.
8. Ryu, *supra* note 6.
9. *See id.*
10. *See id.*
11. *Id.*
12. *Id.*
13. *Id.*
14. *Id.*

creation, possession, and dissemination of deepfakes created without the consent of the depicted individual. This proposed law would attach criminal penalties to such conduct and allow for a private right of action for individuals depicted in such media.

I. BACKGROUND

While not the first created, the viral deepfake of President Obama was the general public's first exposure to "synthetic media" designed to intentionally deceive.[15] Synthetic media is generally defined to include all media created through digital or artificial means or media which has been modified or manipulated using technology.[16] As such, synthetic media is a broad category that encompasses several sub-categories of ways to manipulate media.[17]

A. Cheapfakes

Prior to using artificial intelligence to create synthetic media, more modest means of doctoring media existed to create audiovisual (AV) manipulations.[18] To create these AV manipulations, people utilized easily accessible and affordable video editing software to alter and distort (*e.g.*, slow down, speed up, etc.) the content of videos.[19] Additionally, photo editing software allowed users to manually edit documents by faking someone's signature or pictures by drawing mustaches on people.[20] These relatively cheap means of tampering with media are known as "cheapfakes" or "shallowfakes."[21]

Cheapfakes are a quick and affordable alternative to deepfakes that are easily accessible to the average person.[22] Cheapfakes are created using relatively cheap and accessible software like Photoshop.[23] Before editing

15. U.S. DEP'T OF HOMELAND SEC., INCREASING THREAT OF DEEPFAKE IDENTITIES 5–7 (2021), https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf [https://perma.cc/QLM4-63MF].

16. *Id.* at 5.

17. *See id.*

18. Britt Paris, et al., *Deepfakes and Cheap Fakes*, DATA & SOC'Y (Sept. 18, 2019), https://datasociety.net/library/deepfakes-and-cheap-fakes/ [https://perma.cc/K8JF-2DWZ].

19. Hyosun You, *What are Cheapfakes (Shallowfakes)?*, SAMSUNG SDS (May 23, 2022), https://www.samsungsds.com/en/insights/what-are-cheapfakes.html [https://perma.cc/DTQ8-SAUX].

20. *Id.*

21. *Id.*

22. Stuart A. Thompson, *Making Deepfakes Gets Cheaper and Easier Thanks to A.I.*, N.Y. TIMES (Mar. 12, 2023), https://www.nytimes.com/2023/03/12/technology/deepfakes-cheapfakes-videos-ai.html [https://perma.cc/T97K-2VFE].

23. Britt Paris, et al., *Deepfakes and Cheap Fakes*, DATA & SOC'Y (Sept. 18, 2019), https://datasociety.net/library/deepfakes-and-cheap-fakes/ [https://perma.cc/5A3X-DL5R].

software like Photoshop, cheapfakes were created through even more modest means.[24]

A popular example of a cheapfake in action is a 2019 video of, then Speaker of the House, Nancy Pelosi.[25] On its face, the video appeared to show Speaker Pelosi giving a press briefing while intoxicated.[26] In the video, Speaker Pelosi spoke slowly and with slurred speech.[27] In reality, the video was selectively edited and slowed down to give the impression that Speaker Pelosi was intoxicated.[28] This video of Speaker Pelosi is a prime example of a cheapfake, as it was created with easily accessible editing software and required minimal editing skills.

B. Deepfakes

Deepfakes, unlike cheapfakes, require using a certain type of technology—artificial intelligence.[29] Because of the need for artificial intelligence, deepfakes are a new phenomenon, making their first appearance in 2017.[30]

In 2017, a leaked sex tape appeared on the internet depicting actress Gal Gadot, known for her role portraying the comic-book superhero Wonder Woman.[31] In actuality, this video was an existing pornographic video with Gal Gadot's face superimposed over the original female's face.[32] This video is considered to be the first deepfake in the modern era of synthetic media.[33]

A post later appeared on Reddit, an anonymous social media platform.[34] This post, claiming ownership of the video, was made by a user operating under the name "deepfakes."[35] The term "deepfake" is a portmanteau.[36] "Deepfake" blends the words "deep learning" and "fake."[37]

---

24. *See id.*
25. U.S. DEP'T OF HOMELAND SEC., *supra* note 15, at 14.
26. Matthew Brown, *Fact Check: Video of Speaker Nancy Pelosi altered, selectively edited*, USA TODAY (Aug. 11, 2020, 5:37 PM), https://www.usatoday.com/story/news/factcheck/2020/08/11/fact-check-video-pelosi-altered-and-selectively-edited/3332920001/ [https://perma.cc/VK75-MXV6].
27. *Id.*
28. *Id.*
29. *See* U.S. DEP'T OF HOMELAND SEC., *supra* note 15, at 5.
30. *See id.*
31. *Id.*
32. *Id.*
33. *Id.*
34. *Id.*
35. *Id.*
36. HOMELAND SEC. ADVISORY COUNCIL, *Final Report Artificial Intelligence/Machine Learning Emerging Technologies Subcommittee* 1, 12 (Nov. 14, 2019), https://www.dhs.gov/sites/default/files/publications/hsac_emerging_technologies_full_final_report_ai_ml_7.21.20.pdf [https://perma.cc/WBW3-ED8R]. A "portmanteau" is a word that blends the sounds and meaning of two separate words to describe the new word. *Id.*
37. *Id.*

Creating a deepfake requires using deep learning techniques to create manipulated, or fake, media.[38]

Deep learning is a subcategory of machine learning, and both deep learning and machine learning are subcategories of artificial intelligence.[39] Machine learning utilizes training models to become proficient at a specific task.[40] The more holistic the training data is, the better the model becomes.[41] What distinguishes deep learning from normal machine learning is that the deep learning models are able to discover representations of features in the data that allow such models to classify the data.[42] Deep learning models are necessary to create the hyper-realistic facial manipulations associated with deepfakes.[43] These facial manipulations include: "altering expressions, swapping the faces of two real people or generating a nonexistent human face from a dataset that includes thousands of images of real people."[44]

The most common type of deepfake is the "face swap."[45] Face swapping, as the name suggests, occurs when the face or entire head of one person is put onto the body of another person.[46] The capability of creating a face-swapped image first emerged in the 1990s in the form of cheapfakes.[47] Using commercially available image editing software, like Photoshop, an individual could manually create a face-swapped image.[48]

While an individual can still choose to create a face-swapped image manually, advancements in technology have streamlined the process exponentially.[49] Using any number of readily accessible applications on a smartphone (such as Snapchat or TikTok), users have access to face swapping technologies that take advantage of artificial intelligence.[50] These applications use artificial intelligence to offer not only face swapping capabilities in real time, but also allow the user to perform other types of manipulations on a video or image.[51] Because many of these applications are free and roughly eighty-five percent of Americans own a smartphone,

38. *See* U.S. DEP'T OF HOMELAND SEC., *supra* note 15, at 5.
39. *Id.*
40. *Id.*
41. *Id.*
42. *Id.*
43. *Everything You Need to Know About How to Use Deepfake Technology*, DISCOVER DATA SCI., https://www.discoverdatascience.org/articles/everything-you-need-to-know-about-how-to-use-deepfake/ [https://perma.cc/A2CM-UDHS].
44. *Id.*
45. *See* U.S. DEP'T OF HOMELAND SEC., *supra* note 15, at 9.
46. *Id.*
47. *Id.*
48. *Id.*
49. *Id.*
50. *Id.*
51. *Id.*

there is virtually no accessibility barrier to this type of deepfake.[52] Additionally, powerful video libraries, like DeepFakeLab and FaceSwap, are available as public or open-source systems, allowing anyone to utilize the technology.[53]

While early iterations of face swapping technology would likely fool someone at first glance, upon closer inspection, discrepancies become apparent in the form of "artifacts."[54] These artifacts are often left behind in the form of unnatural mouth or eyebrow movements.[55] As technology has evolved, face swapping applications have improved at capturing the subtle details of human movement.[56] Now, modern face swapping technology can track someone's head position and rotating movements, eye movement, eyebrow movements, and even blinking.[57]

A much more extreme type of deepfake is the "puppet" technique.[58] As the name suggests, this technique allows a person to completely control a targeted individual like a puppet, making the targeted individual move in any desired way.[59] These movements can be anything from the smallest of facial twitches to a full body movement.[60] The puppet technique relies on the use of generative adversarial network (GAN) technology.[61]

GANs work by having two machine learning networks develop synthetic media.[62] These networks interact in an adversarial process.[63] One of the networks is known as the "generator."[64] Data is first fed to the generator that is representative of the type of content desired to be created.[65] The generator then tries to create new content which exhibits the same characteristics of the original data the generator was fed.[66] The newly generated content is then presented to the second network, known as the "adversary."[67]

Like the generator, the adversary is also trained.[68] That said, the adversary is trained to learn how to identify the characteristics of the type of data that the generator is attempting to create.[69] The adversary's role is to

---

52. *Mobile Fact Sheet*, PEW RSCH. CTR. (Apr. 7, 2021), https://www.pewresearch.org/internet/fact-sheet/mobile/ [https://perma.cc/4ANZ-9XL6].
53. HOMELAND SEC. ADVISORY COUNCIL, *supra* note 36, at 13.
54. *Id.*
55. *Id.*
56. *See id.*
57. *Id.*
58. U.S. DEP'T OF HOMELAND SEC., *supra* note 15, at 12.
59. *Id.*
60. *Id.*
61. *Id.*
62. *Id.*
63. *Id.*
64. *Id.*
65. *Id.*
66. *Id.*
67. *Id.*
68. *Id.*
69. *Id.*

detect flaws within the examples presented to it.[70] The adversary then rejects those flawed examples that it determines differ from the original data, labeling the examples as "fakes."[71]

Once an example is labeled as a fake, it is returned to the generator, so that the generator can learn to improve its creation process.[72] The generator then creates new content, and the process repeats.[73] This cyclic training continues until the generator creates content that the adversary network mistakenly identifies as real.[74]

Deepfakes encompass not only visual media, such as videos and pictures, but also audio recordings.[75] While audio cheapfakes have been around for some time, they differ greatly from modern audio deepfakes.[76] For instance, as previously mentioned, the video depicting then Speaker of the House Nancy Pelosi as intoxicated was a cheapfake.[77] The original video of Speaker Pelosi was slowed down to create the desired effect on Speaker Pelosi's voice and did not require any special technology.[78] On the other hand, audio deepfakes allow a person to replicate a targeted person's voice, making it sound as though the targeted person was actually speaking.[79]

Audio deepfakes have seen increased use recently in the music industry. Songs created using artificial intelligence, intending to sound like famous singers, are trending on applications like TikTok.[80] On TikTok, users can find popular songs being covered by a plethora of AI artists.[81] For instance, users can hear Frank Sinatra sing Dua Lipa's hit song, "Levitating."[82] Users can also hear Taylor Swift sing Rascal Flatts' "Life is a Highway."[83]

Artists and record labels have begun speaking out against the nonconsensual deepfake covers.[84] For instance, Drake, in a since-deleted

---

70. *Id.*
71. *Id.*
72. *Id.*
73. *Id.*
74. *Id.*
75. *Id.* at 10–11.
76. U.S. DEP'T OF HOMELAND SEC., *supra* note 15, at 14.
77. *Id.*
78. *See id.*
79. Ijaz Ahmed & Muhammad Adnan, et al., *Voice Morphing: An illusion or Reality*, 2018 Int'l Conf. on Advancements in Computational Scis., 1, 1 (2018).
80. Dani Di Placido, *Thanks to AI, Fake Kanye and Drake Songs Are Going Viral on Tiktok*, FORBES (Apr. 24, 2023), https://www.forbes.com/sites/danidiplacido/2023/04/24/ai-generated-songs-that-sound-like-kanye-and-drake-are-going-viral-on-tiktok/?sh=4ca82e273531 [https://perma.cc/966C-YZXA].
81. *Id.*
82. @breakstuffpod, TIKTOK, https://www.tiktok.com/@breakstuffpod/video/7255830375207800069 [https://perma.cc/8MKA-6L8T].
83. @prismxx, TIKTOK, https://www.tiktok.com/@prismxx/video/7272669245467561248 [https://perma.cc/6EY2-F272].
84. Placido, *supra* note 80.

Instagram story, openly disapproved of AI covers after hearing a deepfake cover misappropriating his voice.[85] While already blurring ethical lines, AI covers appropriating the voices of deceased artists raise further ethical concerns.[86]

Audio deepfakes are not confined to only famous musicians.[87] Recently, a high school principal was the target of an audio deepfake.[88] The high school principal opened an investigation into the school's athletic director for the mishandling of school funds.[89] In retaliation for this investigation, the athletic director used artificial intelligence to create an audio deepfake of the principal.[90] In this audio, the voice of the principal can be heard making a variety of disparaging remarks.[91] The voice refers to students and faculty as "dumbasses;" complains about the "ungrateful black kids who can't test their way out of a paper bag;" and notes if he "get[s] one more complaint from one more Jew in this community, [he's] going to join the other side."[92] This audio became public and circulated throughout the community.[93] Finally—after roughly three months and a police investigation—the audio was confirmed to be a deepfake appropriating the principal's voice.[94] During this time, the principal was suspended from his role.[95] The athletic director was ultimately arrested but not charged with anything directly regarding the recording.[96]

As demonstrated above, audio deepfakes, standing alone, are already an extremely powerful technology. But when used in tandem with visual deepfakes, the potential benefits and dangers increase exponentially. One major potential benefit of combining audio and visual deepfakes is the ability to lip-sync dubbed media, such as movies or translated press conferences.[97] Rather than be resigned to read subtitles or listen to dubbed

85. *See id.*

86. *Id.*

87. *See* Maya Yang, *Baltimore Teacher Accused of Using AI to Create Fake, Racist Recording if Principal*, THE GUARDIAN (Apr. 27, 2024, 3:11 PM), https://www.theguardian.com/us-news/2024/apr/27/baltimore-teacher-ai-fake-racist-recording-principal [https://perma.cc/85NL-J9KZ].

88. *Id.*

89. *Id.*

90. Katherine Donlevy, *Baltimore HS Athletic Director Used AI to Make Fake Clip of Principal Spouting Racist Rhetoric: Police*, N.Y. POST (Apr. 25, 2024, 10:21 PM), https://nypost.com/2024/04/25/us-news/baltimore-hs-staffer-used-ai-to-make-fake-clip-of-principal-spouting-racist-rhetoric-police/ [https://perma.cc/WT6S-QFVE].

91. *Id.*

92. Yang, *supra* note 87.

93. *See id.*

94. *See id;* Donlevy, *supra* note 90.

95. Donlevy, *supra* note 90.

96. Jaclyn Diaz, *A Baltimore-Area Teacher Is Accused of Using AI to Make His Boss Appear Racist*, NPR (Apr. 26, 2024, 5:00 AM), https://www.npr.org/2024/04/26/1247237175/baltimore-ai-generated-racist-audio-crime [https://perma.cc/LEW3-9GWP].

97. U.S. DEP'T OF HOMELAND SEC., *supra* note 15, at 10–11.

media, content can be made to sound and look as though the speaker is speaking in the viewer's preferred language.[98]

On the other side of the coin is the potential use by bad actors to deceive people by appropriating the face and voice of someone for their own gains. For instance, in early 2023, a sponsored ad began to circulate on TikTok, showing famous comedian and podcaster Joe Rogan endorsing a "libido-boosting" coffee.[99] This fake advertisement combined real portions of Joe Rogan's podcast with visual and audio deepfakes, giving the impression that Joe Rogan was personally endorsing the sponsoring company's product.[100] This marketing strategy has not been confined to this one video; platforms like TikTok have seen an influx of fake videos, like this one of Joe Rogan, emerging on their platforms.[101]

C. The Issue

"A reputation, like a face, is the symbol of its possessor and creator, and another can use it only as a mask."[102] Thus, when another uses someone's likeness, "he borrows the owner's reputation, whose quality no longer lies within his own control. This is an injury, even though the borrower does not tarnish it[.]"[103] As Judge Learned Hand acknowledged in 1928, there is an inherent injury when someone "borrows" another person's reputation.[104] When a reputation is used as a "mask," the original person loses control of the quality and subsequent associations of his reputation.[105]

While there are many potential benefits to emerging and developing technologies associated with deepfakes, there are also many dangers. As mentioned above, deepfake pornography was the first-known use of modern deepfakes and is still widely prevalent.[106] While not perfect, deepfakes are still convincing enough to trick a portion of the population into thinking they are real or to simply make people indifferent.[107] With the continued development in artificial intelligence, this technology will only become better, cheaper, and more accessible to the public.[108]

Deepfake technology has already become so convincing that when Twitch streamer Blaire saw images of her face on a naked body, it caused

---

98. *See id.*

99. Stewart A. Thompson, *Making Deepfakes Gets Cheaper and Easier Thanks to A.I.*, N.Y. TIMES (Mar. 12, 2023), https://www.nytimes.com/2023/03/12/technology/deepfakes-cheapfakes-videos-ai.html [https://perma.cc/A8MD-EDJJ].

100. *Id.*

101. *Id.*

102. Yale Elec. Corp. v. Robertson, 26 F.2d 972, 974 (2d Cir. 1928).

103. *Id.*

104. *See id.*

105. *See id.*

106. *See* U.S. Dep't of Homeland Sec., *supra* note 15, at 5.

107. *Id.*

108. *See id*. at 10–11.

her to pause and think if the girl she saw was her.[109] Even though Blaire knew that the naked images she saw were not actually images of her, she knew that her face and reputation would be affected by this incident moving forward not only in her public life but also in her private life.[110] Indeed, Blaire relives the trauma of this situation every time she explains the photos to someone, including family members after other people sent the deepfake images to them.[111]

This use of deepfake technology for non-consensual pornography was the catalyst for the boom in deepfake content and is still the vast majority of AI-enabled synthetic media content today.[112] In late 2020, researchers reported finding over 100,000 deepfake pornographic images of women.[113] These images were created without the women's knowledge or consent.[114] In some cases, these images depicted underage individuals as well.[115] Non-consensual pornography makes up roughly ninety to ninety-five percent of deepfake videos created since 2018.[116]

Because the quality of the deepfake depends on the quality and quantity of the training data provided, some people do not think deepfakes pose a great threat to the average individual.[117] This belief is mistaken. In late 2019, Evan Jacoby, a journalist, investigated deepfake pornography and the technology behind it.[118]

Jacoby joined an online face swapping marketplace.[119] He quickly and easily found a creator willing to digitally superimpose a face of Jacoby's choosing onto an existing porn video.[120] Jacoby simply had to pay thirty dollars and provide a link to the desired porn video and a reference of the desired face.[121] The reference could come in the form of a short video with the face he wanted inserted into the pornography.[122] Using his phone's camera, he filmed a thirteen-second video of himself talking and sent it with the other requirements to the content creator.[123] Jacoby did not tell the

109. Ryu, *supra* note 6.
110. *See id.*
111. *See id.*
112. *See* U.S. Dep't of Homeland Sec., *supra* note 15, at 17.
113. *Id.*
114. *Id.*
115. *Id.*
116. *Id.*
117. *See id*. at 5.
118. Evan Jacoby, *I Paid $30 to Create a Deepfake Porn of Myself*, VICE (Dec. 9, 2019), https://www.vice.com/en/article/vb55p8/i-paid-dollar30-to-create-a-deepfake-porn-of-myself [https://perma.cc/S474-QNA5].
119. *Id.*
120. *Id.*
121. *Id.*
122. *Id.*
123. *Id.*

creator that the face he wanted superimposed onto the porn video was his own or that he had consent to use the face.[124]

The next morning, Jacoby received sample screenshots of the deepfake pornography.[125] After receiving the completed video, Jacoby was informed that any future videos he wanted, using the face he sent the creator, would be cheaper.[126] This is because the algorithm was now trained on Jacoby's face.[127] In a world that is now completely intertwined with social media platforms like TikTok, Instagram, and Facebook, one thirteen-second video is all it takes for someone to create a passible deepfake.[128]

As society continues to become more connected through various social media platforms, people are at risk of having their social media used against them.[129] The number of social media users has increased to over 4.9 billion globally.[130] These users do not confine their digital footprint to a single platform.[131] Rather, the average user spreads their attention across six to seven different platforms a month.[132] Additionally, for the younger generation that has grown up with social media, sharing one's life with the world is the norm.[133] For people between the ages of eighteen and twenty-nine, eighty-four percent use at least one social media platform.[134]

Non-consensual deepfake pornography is not the only malicious use of deepfake technology. Deepfakes also pose the threat of spreading misinformation in the political world.[135] In 2019, bad actors began conducting operations using GAN-generated images on social media aimed at influencing viewers, commonly known as "influence campaigns."[136] By using synthetic personas to build credibility with their audience, these bad actors have promoted localized and regional issues with the hope of influencing the political process.[137]

---

124. *Id.*

125. *Id.*

126. *Id.*

127. *Id.*

128. *See id.* Most deepfake software requires hundreds of high-quality images of the person to be depicted in order to have an appropriate dataset to train the algorithm. *Id.* But a single video, such as a 15-second Instagram story, contains approximately 450 individual frames—a sufficient amount to create a dataset to train an algorithm. *Id.*

129. *See* Belle Wong, *Top Social Media Statistics and Trends of 2023*, FORBES (May 18, 2023), https://www.forbes.com/advisor/business/social-media-statistics/ [https://perma.cc/KBF6-FVJZ].

130. *Id.*

131. *Id.*

132. *Id.*

133. *See id.*

134. *Id.*

135. U.S. DEP'T OF HOMELAND SEC., *supra* note 15, at 25.

136. *Id.*; Dr. Brian Kettler, *Influence Campaign Awareness and Sensemaking (INCAS)*, DARPA, https://www.darpa.mil/program/influence-campaign-awareness-and-sensemaking [https://perma.cc/X35C-DHQ7] (last visited Aug. 1, 2024).

137. U.S. DEP'T OF HOMELAND SEC., *supra* note 15, at 16.

Social media platforms and other companies specializing in machine learning and artificial intelligence are often able to detect these synthetic personas and regulate them.[138] But the detection and subsequent action are not always timely—meaning the damage is already done.[139] Additionally, while companies may have confidence in the synthetic personas they identify, it is impossible to know how many other synthetic personas are attempting to deceive the public.[140]

While deepfakes have largely been used by bad actors for use in nonconsensual pornography and influence campaigns, the advancements in this technology pose a danger to the everyday person.[141] As advancements in artificial intelligence and machine learning technologies progress, the more realistic and deceiving deepfakes will become.[142] Additionally, as with all technology, as more advancements are made, deepfakes will become cheaper and more accessible to the public.[143] Once the public has access to this powerful technology, it is likely that the creation of nonconsensual deepfakes targeting normal people—not just public figures—will increase.[144] This will put the reputation of anyone who maintains an online profile at risk of having one's face or voice taken and used without consent.[145]

II. STATES' ATTEMPTS TO LEGISLATE DEEPFAKES

States are beginning to combat AI-assisted identity misappropriation by creating legislation targeting deepfakes. Some have criminalized the use of deepfakes in nonconsensual pornography or to influence elections.[146] Others have created a civil cause of action for victims.[147] However, the overwhelming majority of states still have no laws addressing this new technology.[148] This section walks through the states that have either proposed or enacted legislation targeting the use of deepfake technology, with a particular focus on those states that provide the legislative history and intent of the laws and the public discussions surrounding them.

---

138. *Id.*
139. *See id.*
140. *Id.*
141. *Id.* at 18.
142. *See id.*
143. *See id.*
144. *See id.*
145. *See id.* at 24.
146. *See* VA. CODE ANN. § 18.2-386.2 (2024); *see also* TEX. PENAL CODE ANN. § 21.165 (2023).
147. *See* CAL. ELEC. CODE § 20010 (2024); *see also MINN.* STAT. § 604.32 (2024).
148. *See California Becomes the Second State to Restrict Political "Deepfakes"*, FIRST AMEND. WATCH (Oct. 9, 2019), https://firstamendmentwatch.org/california-becomes-the-second-state-to-restrict-political-deepfakes/ [https://perma.cc/V7U7-J376].

Texas

In September 2019, Texas became the first state to criminalize deepfakes.[149] Specifically, aiming to prevent misinformation during elections, Texas criminalized creating deepfakes that target the political environment.[150] Texas election law now provides that "[a] person commits an offense if the person, with intent to injure a candidate or influence the result of an election: (1) create a deep fake video; and (2) causes the deep fake video to be published or distributed within 30 days of an election."[151] Violating this new election law is a class A misdemeanor with a possible one-year sentence in county jail and a fine of up to $4,000.[152]

This statute defines a deepfake as "a video, created with the intent to deceive, that appears to depict a real person performing an action that did not occur in reality."[153] Consequently, under this broad definition, many things that would not meet the technical definition of a deepfake are encompassed, such as cheapfakes.[154] For instance, the slowed-down video of Speaker Pelosi would be considered a deepfake—subjecting the creator to criminal liability—because it appears to depict Speaker Pelosi as intoxicated, when in reality she was not.[155]

Along with the criminalization of deepfakes in the election setting, Texas has since made it a crime to create non-consensual deepfake pornography.[156] The Texas legislature enacted this law because it was aware of the rapid growth and expansion of artificial intelligence technology and the impact such technology was having on the development of deepfakes.[157] So, the legislature acted to minimize the deceit and harm that could be caused by nonconsensual deepfake pornography.[158]

The law provides that "[a] person commits an offense if, without the effective consent of the person appearing to be depicted, the person knowingly produces or distributes by electronic means a deep fake video that appears to depict the person with the person's intimate parts exposed or engaged in sexual conduct."[159] This criminal law uses the same definition of a deepfake as used in the election law setting.[160] Consequently, deepfake

149. Kenneth Artz, *Texas Outlaws 'Deepfakes'—but the Legal System May Not Be Able to Stop Them*, TEXAS LAWYER (Oct. 11, 2019, 1:20 PM), https://www.law.com/texaslawyer/2019/10/11/texas-outlaws-deepfakes-but-the-legal-system-may-not-be-able-to-stop-them/?slreturn=20231019113318 [https://perma.cc/H77R-CJ5G].

150. *See* TEX. ELEC. CODE ANN. § 255.004 (2023).

151. TEX. ELEC. CODE ANN. § 255.004(d) (2023).

152. TEX. ELEC. CODE ANN. § 255.004(c) (2023).

153. TEX. ELEC. CODE ANN. § 255.004(e) (2023).

154. *See id.*

155. U.S. DEP'T OF HOMELAND SEC., *supra* note 15, at 14.

156. TEX. PENAL CODE ANN. § 21.165 (2023).

157. 2023 Legis. Bill Hist. TX S.B. 1361.

158. *Id.*

159. TEX. PENAL CODE ANN. § 21.165(b) (2023).

160. TEX. PENAL CODE ANN. § 21.165(a)(1) (2023).

pornographic images would not be encompassed under this law because Texas restricts its definition to only include videos.[161]

Hawaii

Hawaii has also worked to criminalize nonconsensual deepfake pornographic images. The Hawaiian legislature established the "twenty-first century privacy task force" and tasked it with providing policy recommendations for ways to protect the privacy interests of Hawaii's residents.[162] This task force was made up of members of the government and private sector who had some expertise and interest in privacy law in the digital era.[163] The task force studied a variety of privacy issues considered by Hawaii and other states.[164] A key factor considered was the advancement and spread of deepfake technology.[165] Along with the rapid advancements in deepfake technology, the legislature considered how easily this type of synthetic media could be shared via social media platforms.[166]

One of the recommendations by the task force was for the State to take measures to protect the privacy of a person's likeness.[167] The State acted on these recommendations and amended the "violation of privacy" crime to include the intentional disclosure or threat of disclosure of certain types of deepfakes. [168] Specifically, a person commits the crime of violation of privacy in the first degree using deepfake technology when:

> The person intentionally creates or discloses or threatens to disclose an image or video of a composite fictitious person depicted in the nude … , or engaged in sexual conduct … , that includes the recognizable physical characteristics of a known person so that the image or video appears to depict the known person and not a composite fictitious person, with intent to substantially harm the depicted person with respect to that person's health, safety, business, calling, career, education, financial condition, reputation, or personal relationships, or as an act of revenge or retribution.[169]

161. *See* TEX. PENAL CODE ANN. §§ 21.165(a)–(b) (2023).
162. 2021 Haw. S.B. 309.
163. *Id.*
164. *Id.*
165. *Id.*
166. *Id.*
167. *Id.*
168. *Id.*
169. HAW. REV. STAT. § 711-1110.9(1)(c) (2024).

This law takes a broader approach to defining what constitutes a deepfake by encompassing both images and videos, making it more inclusive than laws like the one in Texas.[170]

Virginia

Much like Hawaii, Virginia focused on restricting pornographic deepfakes.[171] In 2019, Virginia amended a criminal statute involving the unlawful dissemination or sale of images of another.[172] The legislature redefined the meaning of "another person" under this statute.[173] Under this new definition, "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."[174]

While this statute does not explicitly use the term deepfake, the definition encompasses such technology.[175] This law provides protection to victims of both deepfake videos and images, and it extends further beyond just the realm of deepfakes.[176] Because the definition takes a broad approach, encompassing any modification intended to depict another individual, cheapfakes and other means of manipulating pictures and videos fall under this provision.[177]

Wyoming

In 2021, Wyoming joined the ranks of those states criminalizing nonconsensual deepfake pornography.[178] In the statute criminalizing the unlawful dissemination of intimate images, the legislature took care to define what comprises an "image."[179] An image under this statute includes "a computer generated image that purports to represent an identifiable person."[180]

Wyoming's definition of an image ensures that any use of machine learning or other artificial intelligence to create a deepfake is covered under the statute.[181] But this statute is only applicable when a deepfake image has been disseminated.[182] Thus, if an individual creates a nonconsensual deepfake pornographic video or image, they have not violated the statute

---

170. *Compare* HAW. REV. STAT. § 711-1110.9(1)(c) (2024) ("The person intentionally creates or discloses or threatens to disclose an image or video"), *with* TEX. PENAL CODE ANN. § 21.165(b) (2023) (only encompassing "deep fake videos" in the law).

171. *See* VA. CODE ANN. § 18.2-386.2 (2024).

172. *See* VA. CODE ANN. § 18.2-286.2 (2024).

173. *See* VA. CODE ANN. § 18.2-286.2(A) (2024).

174. *Id.*

175. *Id.*

176. *Id.*

177. *Id.*

178. *See* WYO. STAT. ANN. § 6-4-306 (2024).

179. *See* WYO. STAT. ANN. § 6-4-306(a)(iii) (2024).

180. *Id.*

181. *See id.*

182. *See* WYO. STAT. ANN. § 6-4-306(b)(i) (2024).

unless they sell the synthetic media, post the deepfake to social media, or otherwise make the deepfake available to a third party.[183] But it does not violate the law to share the intimate deepfake image in private to the person depicted in it.[184] Consequently, persons in Wyoming can create as much nonconsensual deepfake pornography as they please, so long as they either keep it to themselves or share it in private with the person whose identity they are appropriating.[185]

California

California has also taken great measures to restrict and regulate deepfakes. California became the second state, after Texas, to prohibit the use of deepfake technology to spread false information during political elections.[186] While California did not use the term "deepfake," the California law entitled "Truth in Political Advertising Act" prohibits the "distribut[ion], with actual malice, [of] materially deceptive audio or visual media, … , of the candidate with the intent to injure the candidate's reputation or to deceive a voter into voting for or against the candidate."[187] Unlike the Texas deepfake election law, California's law casts a wider net, including either "deceptive audio or visual media."[188]

Materially deceptive audio or visual media is defined as "an image or an audio or video recording of a candidate's appearance, speech, or conduct that has been intentionally manipulated" such that it meets a few conditions.[189] First, the deceptive audio or visual media would appear authentic to a reasonable person.[190] And second, the deceptive media "would cause a reasonable person to have a fundamentally different understanding or impression of the expressive content of the image or audio or video recording" than they would if that person saw the original, unaltered version.[191] A candidate targeted by false media in violation of this law may seek both injunctive relief and monetary damages against the offending party.[192]

The law provides several exceptions to this prohibition of deepfakes in the political environment. If the altered media includes a disclosure stating, "[t]his [Image/Video/Audio] has been manipulated[,]" then it does not fall under the general prohibition.[193]

---

183. *See* WYO. STAT. ANN. § 6-4-306(a)(ii) (2024).

184. *Id.*

185. *See* WYO. STAT. ANN. § 6-4-306 (2024).

186. *California Becomes the Second State to Restrict Political "Deepfakes"*, FIRST AMEND. WATCH (Oct. 9, 2019), https://firstamendmentwatch.org/california-becomes-the-second-state-to-restrict-political-deepfakes/ [https://perma.cc/V7U7-J376].

187. CAL. ELEC. CODE § 20010(a) (2023).

188. *Compare* TEX. ELEC. CODE ANN. § 255.004(e) (2023), *with* CAL. ELEC. CODE § 20010(a) (2023).

189. CAL. ELEC. CODE § 20010(e) (2023).

190. CAL. ELEC. CODE § 20010(e)(1) (2023).

191. CAL. ELEC. CODE § 20010(e)(2) (2023).

192. CAL. ELEC. CODE §§ 20010(c)(1)–(2) (2023).

193. CAL. ELEC. CODE §§ 20010(b)(1)–(2) (2023).

Radio and television broadcasting stations that broadcast deceptive media are protected if the media is broadcasted "as part of a bona fide newscast, news interview, news documentary, or on-the-spot coverage of bona fide news events."[194] These types of broadcasts are also required to provide clear disclaimers that can be easily heard or read and indicate "that there are questions about the authenticity of the materially deceptive audio or visual media."[195] Another exception shields radio and television broadcasting stations from the statute when they are "paid to broadcast materially deceptive audio or visual media."[196]

The law also exempts "internet website[s], or a regularly published newspaper, magazine, or other periodical of general circulation" so long as they provide clear disclaimers stating that "the materially deceptive audio or visual media does not accurately represent the speech or conduct of the candidate."[197] Lastly, the statute "does not apply to materially deceptive audio or visual media that constitutes satire or parody."[198]

This statute went into effect January 1, 2020, and was originally only supposed to be in effect until January 1, 2023.[199] But the statute was amended and is now effective until January 1, 2027.[200] While this statute protects individuals engaged in politics from the potentially harmful effects of deepfakes, it also displays the tension between privacy rights and the First Amendment.[201] At the time it was proposed, this law received attention from advocates arguing both sides of the free speech debate.[202]

Constitutional Law scholar and Dean of the School of Law at the University of California, Berkeley, Erwin Chemerinsky, wrote a letter in support of an early version of the bill, arguing that it fits squarely within First Amendment jurisprudence.[203] On the other hand, the law was criticized by the California News Publishers Association (CNPA) and the American Civil Liberties Union (ACLU).[204] When the bill was first presented to the California Legislature, a staff attorney for the CNPA noted that this law is a "content-based regulation of speech" that is not narrowly tailored and it may be unnecessary as defamation laws already address fake political advertisements.[205]

---

194. CAL. ELEC. CODE § 20010(d)(2) (2023).

195. *Id.*

196. CAL. ELEC. CODE § 20010(d)(3) (2023).

197. CAL. ELEC. CODE § 20010(d)(4) (2023).

198. CAL. ELEC. CODE § 20010(d)(5) (2023). While an exception is created for satire and parody, these words are not defined for the purposes of this section. *See id.*

199. CAL. ELEC. CODE § 20010 (2023).

200. *Id.*

201. *See* Ethan Jones, *Assembly Bill Policy Committee Analysis* 5, https://aelc.assembly.ca.gov/sites/aelc.assembly.ca.gov/files/AB%20730%20Berman%20CSA%2009132019.pdf [https://perma.cc/7VMG-UK4A].

202. *Id.* at 5–7.

203. *Id.* at 5.

204. *Id.* at 6–7.

205. *Id.* at 8.

In tandem with the CNPA, the ACLU of California argued against the bill for three reasons.[206] First, the ACLU argued that the bill "[c]reates a false expectation that voters can trust images and videos unless they are labeled as manipulated[.]"[207] This false sense of security would be created because the bill only applies to a fraction of the media that could affect an election.[208] Next, the ACLU argued the law will have a minimal effect because it does not apply where voters are most likely to encounter deepfake media.[209] Because federal law preempts this state law, it does not apply to manipulated media shared on social media and many of the political ads on television.[210] Lastly, the ACLU feared that the law would be misused and abused.[211] The ACLU noted that courts will be unable to resolve claims before election day because of the short timeframe to which the statute applies.[212] Therefore, "the bill's procedures can be weaponized to add legitimacy to claims that real images and recordings are 'fake news.'"[213] Despite these concerns, the California Legislature enacted the bill.[214]

Along with California prohibiting the use of deepfakes to spread misinformation around elections, California has also targeted nonconsensual deepfake pornography.[215] In 2020, individuals gained a private right of action to recover against those who either: (1) created and disclosed sexually explicit material when they knew or reasonably should have known the depicted person did not consent to its creation or disclosure, or (2) intentionally disclosed explicit material they did not create and they knew the depicted individual did not consent to its creation.[216] A "depicted individual" is defined as "an individual who appears, as a result of digitization, to be giving a performance they did not actually perform or to be performing in an altered depiction."[217]

Like the statute prohibiting misinformation in elections, this law does not explicitly use the term "deepfake," but it is clear by the definition of "digitization" that such technology is covered.[218] Like other states, this statute provides exceptions to the general rule.[219] A person is not liable if

206. Jones, *supra* note 201, at 7.
207. *Id.*
208. *Id.*
209. *Id.*
210. *Id.*
211. *Id.*
212. *Id.*
213. *Id.*
214. *See generally* CAL. ELEC. CODE § 20010 (2023).
215. CAL. CIV. CODE § 1708.86 (2024).
216. CAL. CIV. CODE. § 1708.86(b)(1)–(2) (2024). While private individuals have a right of action, there is currently no California criminal law specifically aimed at prohibiting or regulating deepfakes. *Id.*
217. CAL. CIV. CODE § 1708.86(a)(4) (2024).
218. *See id.*
219. CAL. CIV. CODE § 1708.86(c) (2024).

they disclose this sexually explicit material when: reporting unlawful activity; exercising their law enforcement duties; or at hearings, trials, or other legal proceedings.[220] A person is also not liable if the material is: a matter of legitimate public concern; a work of political or newsworthy value or similar work; or commentary, criticism, or disclosure that is otherwise protected by the California Constitution or the United States Constitution.[221]

While the law takes care to define consent under the section and how it can be given, it is silent on what happens when consent is rescinded.[222] For instance, it is unclear what would happen if a person were to consent to the creation of the deepfake but rescind that consent before any distribution of the altered media.[223]

Minnesota

Minnesota has enacted three statutes in its efforts to address deepfakes.[224] First, grouped under "Crimes Against Reputation," Minnesota has criminalized the use of deepfakes to influence an election.[225] An individual commits this offense when they disseminate or enter into a "contract or agreement to disseminate" a deepfake that they know, or reasonably should know, is a deepfake.[226] Additionally, the dissemination, or intent to disseminate, must take place within ninety days before an election; the deepfake must be created without the target's consent; and the deepfake must be created to either injure the candidate or influence the election.[227]

A person who violates this law is exposed to potential criminal penalties, such as a fine and potential jail time.[228] The law also contemplates the potential for repeat offenders. Consequently, if a perpetrator has been convicted of violating this same statue within five years, that person is subject to more severe penalties.[229]

Along with criminal sanctions, the law provides a cause of action for various people to seek an injunction against a person who has violated or is about to violate the law.[230] Those empowered to seek an injunction include: the attorney general, a country or city attorney, the depicted

220. CAL. CIV. CODE § 1708.86(c)(1)(A)(i)–(iii) (2024).
221. CAL. CIV. CODE § 1708.86(c)(1)(B)(i)–(iii) (2024).
222. CAL. CIV. CODE § 1708.86(a)(3) (2024).
223. *See id.*
224. MINN. STAT. § 604.32 (2024); MINN. STAT. § 609.771 (2024); MINN. STAT. § 617.262 (2024).
225. MINN. STAT. § 609.771 (2024).
226. MINN. STAT. § 609.771(Subdiv. 2) (2024).
227. MINN. STAT. § 609.771(Subdiv. 2)(1)–(3) (2024).
228. MINN. STAT. § 609.771(Subdiv. 3) (2024).
229. MINN. STAT. § 609.771(Subdiv. 3)(1) (2024).
230. MINN. STAT. § 609.771(Subdiv. 4) (2024).

individual, or a candidate in the election who is injured or likely to be injured by the dissemination of the false media.[231]

In addition to prohibiting the use of deepfakes to influence elections, Minnesota has criminalized the dissemination of deepfake pornography.[232] In this statute, Minnesota defines a deepfake to include "any technological representation of speech or conduct … that is so realistic that a reasonable person would believe it depicts speech or conduct of an individual."[233] The statute later specifies that the production of the deepfake must be "substantially dependent" upon technical means "rather than the ability of another individual to physically or verbally impersonate such individual."[234] While this carve out shields those that may be able to impersonate the targeted individual, it does not define or give an example of what "substantially dependent" looks like for the production of the deepfake.[235] Thus, it is unclear whether an impression aided by technological means would fall under this statute.[236]

Like other states criminalizing nonconsensual deepfake pornography, Minnesota provides a law whereby a crime is not committed until the deepfake is intentionally disseminated.[237] But this statute does not consider the media to be disseminated if the only person it is sent to is the person depicted in the deepfake.[238] Thus, the statute does not criminalize the creation of the deepfake—rather, it is a crime to: share a deepfake when the depicted individual is identifiable; realistically depict intimate parts of the depicted individual or depict that individual engaging in a sexual act; and disseminate the deepfake when the person knows or reasonably should know that the individual depicted did not consent to the dissemination.[239]

While consent is ordinarily a defense to this type of statute, Minnesota law does not allow for a consent defense when the individual portrayed consented only to the creation of the deepfake or private transmission of it, not the public dissemination of it.[240] This law also contains several exceptions to avoid culpability.[241] For instance, a person does not violate the statute if "the deep fake was obtained in a commercial setting for the purpose of the legal sale of goods or services, including the creation of artistic products for sale or display, and the depicted individual knew, or should have known, that a deep fake would be created and

231. MINN. STAT. § 609.771(Subdiv. 4)(1)–(4) (2024).
232. MINN. STAT. § 617.262 (2024).
233. MINN. STAT. § 617.262(Subdiv. 1)(b)(1) (2024).
234. MINN. STAT. § 617.262(Subdiv. 1)(b)(2) (2024).
235. *See id.*
236. *See id.*
237. MINN. STAT. § 617.262(Subdiv. 2) (2024).
238. MINN. STAT. § 617.262(Subdiv. 1)(d) (2024).
239. MINN. STAT. § 617.262(Subdiv. 2)(1)–(3) (2024).
240. MINN. STAT. § 617.262(Subdiv. 3a) (2024).
241. MINN. STAT. § 617.262(Subdiv. 6) (2024).

disseminated[.]"[242] Other exceptions exist where the deepfake "relates to a matter of public interest and dissemination serves a lawful public purpose" or when "the dissemination is for legitimate scientific research or educational purposes[.]"[243]

Along with the criminal penalties associated with the dissemination of nonconsensual deepfake pornography, Minnesota has created a private right of action for victims of this offense.[244] This private right of action mirrors the criminal statute but allows the victim to recover damages and seek injunctive relief.[245]

Washington

The state of Washington enacted legislation seeking to prevent the use of synthetic media in electioneering communications.[246] If a candidate's "appearance, action, or speech [was] altered through the use of a synthetic media," the candidate can seek injunctive relief, monetary damages, or any other equitable relief to prohibit the publication of the forged media.[247] The law defines synthetic media to include "an image, an audio recording, or a video recording of an individual's appearance, speech, or conduct that has been intentionally manipulated with the use of generative adversarial network [(GAN)] techniques or other digital technology in a manner to create a realistic but false image, audio, or video."[248]

The media created by the GAN technique must make a reasonable observer think the depiction is of a real person doing something that they did not do in reality.[249] Additionally, it must produce a "fundamentally different understanding or impression" of the content of the media that a reasonable person would have from seeing the original media.[250]

Washington law does provide someone who violates this law an affirmative defense if the electioneering communication, that uses synthetic media, includes a clear disclosure that the media was manipulated.[251] While ordinarily only the sponsor, the person paying for the synthetic media,[252] is liable, the medium can also be liable if it removes the disclaimer that the media has been manipulated or if the medium further manipulates the content such that the finished product qualifies as synthetic media, without providing a new disclosure.[253]

---

242. MINN. STAT. § 617.262(Subdiv. 6)(4) (2024).
243. MINN. STAT. § 617.262(Subdiv. 6)(5)–(6) (2024).
244. MINN. STAT. § 604.32 (2024).
245. *Compare* MINN. STAT. § 604.32 (2024), *with* MINN. STAT. §617.262 (2024).
246. WASH. REV. CODE § 42.62.010 et seq (2024).
247. WASH. REV. CODE § 42.62.020(2)–(3) (2024).
248. WASH. REV. CODE § 42.62.020(1) (2024).
249. WASH. REV. CODE § 42.62.020(1)(a) (2024).
250. WASH. REV. CODE § 42.62.020(1)(b) (2024).
251. WASH. REV. CODE § 42.62.020(4) (2024).
252. WASH. REV. CODE § 42.17A.005(47)(a) (2024).
253. WASH. REV. CODE § 42.62.030(2)(a)–(b) (2024).

Georgia

Georgia, in its attempt to protect against deepfake technology, amended an existing law that prohibited the transmission of nude or sexually explicit content.[254] In its amended version, a person violates the law when they electronically transmit or post, or cause the electronic transmission or posting, of a "falsely created videographic or still image" without the consent of the depicted person.[255]

Violating this law makes a person vulnerable to criminal liability.[256] Depending on how the fake media is disseminated, a person can either be charged with a misdemeanor or a felony.[257] On top of potential imprisonment, fines are also a potential penalty for violating the law.[258]

Before the amendment to the statute, the Georgia Court of Appeals clarified that there is no private right of action under this criminal statute.[259] The amendment did not create a new private right of action, nor has there been a new statute to create such a private right of action.

New York

New York has been attempting to address deepfakes with legislation for several years with no success.[260] Since 2021, members of the New York legislature have been trying to pass a bill that would amend the penal code to criminalize aggravated harassment by digital means.[261] The phrase "digital means" within the proposed legislation includes producing, distributing, publishing, or broadcasting material that contains "a picture, photograph or image of the person or persons or a deep fake into which the image of another person or persons is superimposed as a deep fake."[262]

In addition to this proposed legislation criminalizing the use of deepfakes for harassment, the proposed legislation also seeks to amend the civil rights laws in New York to create a private right of action for victims of digital harassment by deepfakes.[263] This private right of action allows victims of deepfakes, where the deepfake depicts intimate images, to seek injunctive and monetary (both compensatory and punitive) damages.[264]

The other avenue by which lawmakers in New York are attempting to combat deepfakes is by amending the unlawful dissemination or publication of intimate image statute.[265] This amendment would allow the

254. GA. CODE. ANN. § 16-11-90 (2024).
255. GA. CODE. ANN. § 16-11-90(b)(1) (2024).
256. GA. CODE. ANN. § 16-11-90(c) (2024).
257. GA. CODE. ANN. § 16-11-90(c)(1)–(2) (2024).
258. Id.
259. Somerville v. White, 787 S.E.2d 350, 352–53 (Ga. Ct. App. 2016).
260. *See* S. 6829, 2021 Leg., Reg. Sess. (N.Y. 2021).
261. *See* 2021 Bill Text N.Y. S.B. 6829.
262. *See id.*
263. *Id.*
264. *Id.*
265. *Id.*

term "image" to encompass "[i]mage[s] created or altered by digitization, where such person may reasonably be identified."[266]

The purpose of each of these pieces of proposed legislation is to protect citizens from the "malicious dissemination" of materials that have been "digitally altered or produced" to resemble them.[267]

Louisiana

As of August 1, 2023, Louisiana criminally penalized nonconsensual deepfake pornography.[268] This new law specifically targets deepfakes that depict minors engaging in sexual conduct.[269] A person violates the law by knowingly creating or possessing media that depicts a minor engaging in sexual conduct and that person has knowledge that the media is a deepfake.[270] Additionally, the law criminalizes the dissemination of a deepfake of another person, or a minor, engaging in sexual activities.[271]

The law defines deepfakes as any media that "is created, altered, or digitally manipulated in a manner that would falsely appear to a reasonable observer to be an authentic record …"[272] The law does not consider altered media to be a deepfake if it "constitutes a work of political, public interest, or newsworthy value, including commentary, criticism, satire, or parody[.]"[273] Another exception exists where a clear disclosure is visible throughout the recording such that a reasonable person would understand the media is not depicting a real event.[274]

Illinois

In early 2020, Illinois introduced a bill to amend the state election code to prohibit the use of fake media to interfere with state or local elections.[275] A person would violate this proposed law by "knowingly using cheap fake or deep fake media in a state or local election."[276] The phrase "cheap fake" would encompass any "photo shopped imagery that implies a situation occurred that did not happen."[277] The phrase "deep fake" would cover "the use of artificial intelligence to create inauthentic photographs or videos of a person."[278] But this law was never enacted.[279]

Several years later, in July 2023, Illinois passed a bill, which took effect January 1, 2024, which amends the current laws regulating the

266. *Id.*
267. *Id.*
268. LA. STAT. ANN. § 14:73.13 (2023).
269. LA. STAT. ANN. § 14:73.13(A) (2023).
270. *Id*.
271. LA. STAT. ANN. § 14:73.13(B)(1)–(2) (2023).
272. LA. STAT. ANN. § 14:73.13(C)(1) (2023).
273. *Id.*
274. *Id.*
275. H.R. 5321, 101st Gen. Assemb., Reg. Sess. (Ill. 2019).
276. *Id.*
277. *Id.*
278. *Id.*
279. *See id.*

nonconsensual dissemination of private sexual images.[280] The law first amends the definition of a "sexual image" to include media that "falsely appears to show" some sexual activity.[281]

Next, the law adds a civil cause of action for victims who had private "sexual images" disseminated.[282] This law provides a private right of action against an individual who disseminates, or threatens to disseminate, a sexual image if the person knew or recklessly disregarded the possibility that the image was intentionally digitally altered.[283] If a case centers on digitally altered sexual images, "disclosing that the images were digitally altered shall not be a defense to liability."[284]

Illinois law provides some exceptions to liability under this section.[285] Examples of exceptions include if the dissemination, or threat of dissemination, was made in good faith by law enforcement in a legal proceeding, or for medical education or treatment.[286] Another exception exists where the dissemination, or threat of dissemination, relates to a matter of public concern.[287] But the law does provide that dissemination of such media "is not a matter of public concern solely because the depicted individual is a public figure or the image is accompanied by a political message."[288]

New Jersey

New Jersey introduced a bill in March 2023 that targets the use of deepfakes in nonconsensual pornography.[289] The drafters of the bill noted that "[d]eepfakes have been intentionally used to embarrass or harass" people or to cast them in a false light.[290] The bill would amend the current criminal law to protect individuals' privacy from "revenge porn" created using deepfakes.[291]

The bill protects against deepfakes used in revenge porn by making it a violation of the statute to disclose "any deceptive audio or visual media" to which the depicted individual has not consented.[292] The proposed bill defines "deceptive audio or visual media" broadly, encompassing almost any media that "appears to authentically depict any speech or conduct," that the person did not actually do, where the creation "was substantially dependent upon technical means, rather than the ability of

280. H.R. 2123, 103rd Gen. Assemb., Reg. Sess. (Ill. 2023).
281. 740 ILL. COMP. STAT. 190/5(14) (2024).
282. 740 ILL. COMP. STAT. 190/10(a) (2024).
283. *Id.*
284. 740 ILL. COMP. STAT. 190/10(c) (2024).
285. 740 ILL. COMP. STAT. 190/15 (2024).
286. 740 ILL. COMP. STAT. 190/15(a)(1)(A)–(C) (2024).
287. 740 ILL. COMP. STAT. 190/15(a)(3) (2024).
288. 740 ILL. COMP. STAT. 190/15(d) (2024).
289. Assemb. 5333, 2022 Gen. Assemb., Reg. Sess. (N.J. 2022).
290. *Id.*
291. *Id.*
292. *Id.*

another person to physically or verbally impersonate the person."[293] The bill would also add similar language to the child endangerment statute, making it an offense to knowingly distribute, possess with the intent to distribute, or store and maintain items that depict the sexual exploitation or abuse of a child—even when the items fall under the definition of "deceptive audio or visual media" as above defined.[294]

Along with these criminal penalties, the proposed bill provides a person depicted by a deepfake used in revenge porn a private right of action.[295] This private right of action would allow the depicted individual access to both equitable relief and monetary damages.[296] Nothing in this proposed bill would allow a private right of action to children depicted in deepfakes in violation of the proposed changes to the child endangerment statute.[297]

Massachusetts

In 2019, Massachusetts introduced a bill that would protect against the use of deepfakes that were created to facilitate criminal or tortious conduct.[298] It would violate the proposed law to create, with the intent to distribute, a deepfake or audiovisual record, when the creator has "actual knowledge that the audiovisual record is a deep fake," when the intent to distribute such media would "facilitate criminal or tortious conduct."[299] Massachusetts sought to criminalize such conduct as identity fraud, but the proposed legislation was never enacted.[300]

At the beginning of 2023, Massachusetts introduced legislation that would empower the Governor to appoint a "Massachusetts State Deepfake and Digital Provence Task Force."[301] The purpose of this task force would be to evaluate the "proliferation of deepfakes impacting state government, Massachusetts-based businesses, and residents."[302] Additionally, the task force would assess the privacy risks, the effects of digital content forgery technologies, and the best practices for preventing and mitigating the use of such technology.[303]

Tennessee

On March 21, 2024, the Ensuring Likeness Voice and Image Security (ELVIS) Act was signed into law in Tennessee.[304] The ELVIS Act

293. *Id.*
294. *Id.*
295. *Id.*
296. *Id.*
297. *See id.*
298. H.R. 3366, 191st Gen. Ct., Reg. Sess. (Mass. 2019).
299. *Id.*
300. *Id.*
301. H.R. 72, 193rd Gen. Ct., Reg. Sess. (Mass. 2023).
302. *Id.*
303. *Id.*
304. H.R. 2091, 113th Gen. Assemb., Reg. Sess. (Tenn. 2024). Although enacted on March 21, 2024, the law took effect on July 1, 2024. *Id.*

primarily seeks to protect musicians by amending Tennessee's right of publicity statutes to include protections for people's voices.[305] First, the Act adds a definition for the term "voice."[306] "Voice" is defined as a sound "that is readily identifiable and attributable to a particular individual, regardless of whether the sound contains the actual voice or a simulation of the voice of the individual."[307] Existing definitions use the same "readily identifiable" language with regards to an individual's likeness in a photograph.[308]

The Act then goes through existing right-of-publicity laws and amends them to include the newly defined term "voice."[309] As such, under Tennessee law, "[e]very individual has a property right in the use of that individual's name, photograph, voice, or likeness in any medium in any manner."[310] With this property interest, musicians will be able to bring a civil action against those that infringe on their voice.[311] Because of the definition given to the term "voice" by the Tennessee legislature, artists are now able to bring an action against those that use deepfakes to appropriate their voice.[312]

In sum, a minority of States thus far have attempted to address the ever-growing issue of deepfakes.[313] Of those that have addressed it, there is a complete lack of uniformity as to what constitutes a deepfake, what deepfakes are regulated, and what remedies are available to victims, if any.[314] As more States seek to regulate deepfakes, this varying legal landscape will only continue to grow in the absence of a national standard.

III. FEDERAL ATTEMPTS TO ADDRESS DEEPFAKES

The federal government has made some attempts at addressing the growing threat of deepfake technology. At the end of 2018, a bill was introduced to the Senate that sought to criminalize deepfakes in connection with fraud.[315] This bill, while ultimately unsuccessful, would have made it a crime to "create, with the intent to distribute" a deepfake where the intent to distribute would "facilitate criminal or tortious conduct under Federal, State, local, or Tribal law."[316] This proposed legislation contained a broad

305. *See id.*
306. *Id.*
307. *Id.*
308. *See* TENN. CODE ANN. § 47-25-1102 (2024).
309. H.R. 2091, 113th Gen. Assemb., Reg. Sess. (Tenn. 2024).
310. *Id.* The previous version of this statute was nearly identical, but did not include the term "voice." *See* TENN. CODE ANN. § 47-25-1103 (2024).
311. *See* H.R. 2091, 113th Gen. Assemb., Reg. Sess. (Tenn. 2024).
312. See *id.*
313. *Infra* Part II.
314. *Infra* Part II.
315. Malicious Deep Fake Prohibition Act of 2018, S. 3805, 115th Cong. (2018).
316. *Id.*

exception for protected First Amendment speech and included no civil right of action for an individual.[317]

Just a few years later, in 2020, a bill was introduced in the House of Representatives that would have required a study to be performed.[318] The study would have focused on how artificial intelligence could be used to combat harms occurring online.[319] One such focus of this proposed law was how artificial intelligence could be used to address "[m]anipulated content intended to mislead individuals, including deepfake videos[.]"[320] While the proposed law sought to study the use of artificial intelligence to combat deepfakes, it made no attempt to define or clarify what constitutes a deepfake.[321]

One year later, in 2021, another bill was introduced in the House.[322] This proposed bill, dubbed the "DEEP FAKES Accountability Act," sought to cast a much wider protective net against the harms of deepfakes.[323] The DEEP FAKES Accountability Act defined a deepfake as:

> 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.[324]

Using this definition of deepfake, the proposed Act defined another phrase—"advanced technological false personation record."[325] This phrase was defined as any deepfake that a reasonable person would believe accurately exhibits (1) any "material activity of a living person" that the living person did not actually do; or (2) any "material activity of a deceased person" which the deceased person did not actually do, 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

317. *Id.* The proposed legislation provides the following exception: "No person shall be held liable under this section for any activity protected by the First Amendment to the Constitution of the United States." *Id.*

318. Countering Online Harms Act, H.R. 6937, 116th Cong. (2020).

319. *Id.*

320. *Id.*

321. *See id.*

322. Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2021, H.R. 2395, 117th Cong. (2021).

323. *Id.*

324. *Id.*

325. *Id.*

election."[326] Additionally, an "advanced technological false personation record" must have been produced without the consent of the depicted individual.[327] If the depicted individual is deceased, the record must have been produced without the consent of the depicted individual or the consent of the individual's heirs.[328]

Operating under these definitions, the proposed Act would have required a watermark on all advanced technological false personation records.[329] For records that were visual, a clearly identifiable digital watermark would have been required.[330] But if a record was purely made up of audio, a clear and unambiguous verbal disclaimer would have been required.[331] Records that were a mix of audio and visual deepfakes required both a watermark and a verbal disclaimer.[332] Failing to meet the disclosure portion of the proposed Act, by either failing to disclose or altering a disclosure, would have opened an offending party to both criminal and civil liability, including civil suits by private individuals .[333]

Together with the watermark requirement, the proposed Act would have revised the criminal offense of fraud in connection with certain identification to include deepfakes.[334] The proposed Act would have also directed the Secretary of Homeland Security to establish the "Deep Fakes Task Force."[335] This task force would have been charged with "advanc[ing] [the] efforts of the United States Government to combat the national security implications of deep fakes[.]"[336] Unfortunately, the DEEP FAKES Accountability Act failed and was never enacted.[337]

In September 2023, the "Preventing Deep Fake Scams Act" was introduced in the House.[338] This proposed Act could create a task force on artificial intelligence focused on the financial services sector.[339] The purpose of the task force is to determine the steps financial institutions are taking to protect themselves from fraud enabled by artificial intelligence and to comprehend the possible risks associated with bad actors exploiting artificial intelligence to commit fraud.[340]

In early October 2023, members of Congress proposed the "Nurture Originals, Foster Art, and Keep Entertainment Safe Act of 2023," referred

326. *Id.*
327. *Id.*
328. *Id.*
329. *Id.*
330. *Id.*
331. *Id.*
332. *Id.*
333. *Id.*
334. *Id.*
335. *Id.*
336. *Id.*
337. *Id.*
338. Preventing Deep Fake Scams Act, H.R. 5808, 118th Cong. (2023).
339. *Id.*
340. *Id.*

to as the "NO FAKES Act of 2023."[341] This legislation was proposed because members of Congress recognized that artificial intelligence is able to use "name, image, likeness, [and] voice cloning" to infringe on the livelihood of entertainers and creators.[342] This proposed Act could essentially create a federal right of publicity for deepfake content.[343]

A perpetrator is liable under the NO FAKES Act if they either (1) produce a "digital replica" with the consent of the individual or rights holder, or (2) publish, distribute, transmit, or otherwise make available a "digital replica" if the person know it was not consented to by the individual or rights holder.[344] The proposed legislation defines a "digital replica" to be "a newly-created, computer-generated, electronic representation of the image, voice, or visual likeness of an individual that" (A) is nearly indistinguishable from the depicted person's real name, image, likeness, or voice; and (B) is a part of an audio recording or audiovisual work that the depicted person did not actually perform or appear in.[345]

A victim whose name, image, likeness, or voice is appropriated under this proposed Act could have the ability to bring a private action against the offender and recover either $5,000 per violation or any damage suffered as a result of the injury—whichever is greater.[346] Additionally, if the injured party can prove the offender acted with "malice, fraud, or oppression," punitive damages may be awarded.[347]

The right protected by this Act is a property right that is descendible and licensable in whole or in part.[348] Most importantly, the proposed Act goes on to state that this right should be considered a "law pertaining to intellectual property for the purposes of section 230(e)(2) of the [Communications Decency Act]."[349] This means that if a third party posted or advertised a digital replica in violation of this proposed act on a third-party platform, like Facebook or Instagram, the injured party could bring an action against the third-party platform (Facebook or Instagram) for the violation.[350]

---

341. Nurture Originals, Foster Art, and Keep Entertainment Safe Act of 2023, "NO FAKES Act of 2023," https://www.coons.senate.gov/imo/media/doc/no_fakes_act_draft_text.pdf [https://perma.cc/FB7U-46JW][hereinafter NO FAKES Act].

342. 169 Cong. Rec. 5226 (2023).

343. NO FAKES Act, *supra* note 341.

344. *Id.*

345. *Id.*

346. *Id.*

347. *Id.*

348. *Id.*

349. *Id.*

350. *See* 47 U.S.C. § 230(e)(2) (2024). Currently no federal right of publicity law exists and courts are split as to whether state right of publicity laws can be brought under 47 U.S.C. § 230(e)(2) (2024). *Compare* Perfect 10, Inc. v. CCBill LLC, 488 F.3d 1102 (9th Cir. 2008) (holding § 230(e)(2) precludes state right of publicity claims against platforms), *with* Hepp v. Facebook, 14 F.4th 204 (3d Cir. 2021) (holding § 230(e)(2) allows state right of publicity claims against platforms). The NO FAKES Act, in addition to protecting against deepfakes,

A short time later, on October 30, 2023, President Biden signed an executive order focused on artificial intelligence.[351] The purpose of this executive order is to foster the safe development of artificial intelligence while trying to mitigate the potential harms.[352] As part of this executive order, the Secretary of Commerce shall submit a report identifying "existing standards, tools, methods, and practices, as well as the potential development of further science-backed standards and techniques, for" authenticating, labeling, and detecting synthetic content.[353] Additionally, the report shall include tools, methods, and practices for "preventing generative AI from producing child sexual abuse material or producing non-consensual intimate imagery of real individuals (to include intimate digital depictions of the body or body parts of an identifiable individual)."[354]

Once the report has been submitted, the Secretary of Commerce shall develop and update periodically "guidance regarding the existing tools and practices for digital content authentication and synthetic content detection measures."[355]

In January 2024, a week after deepfake pornographic images of Taylor Swift emerged on the internet, the Disrupt Explicit Forged Images and Non-Consensual Edits (DEFIANCE) Act was introduced.[356] The DEFIANCE Act could expand the means of recovery for victims of non-consensual intimate "digital forgeries" by amending various existing federal laws. The Act defines a "digital forgery" as:

> any intimate visual depiction of an identifiable individual created through the use of software, machine learning, artificial intelligence, or any other computer-generated or technological means, including by adapting modifying, manipulating, or altering an authentic visual depiction, to appear to a reasonable person to be indistinguishable from an authentic visual depiction of the individual, regardless of whether the visual depiction indicates, through a label or

---

would allow private individuals to bypass the § 230(c) immunity given to platforms and recover for right of publicity violations. *See* NO FAKES Act, *supra* note 341; *see generally* Samantha P. McCaleb, *Paws Off My Profile: Protecting the Persona in a Modern Digital Age*, 27 MARQ. INTELL. PROP. & INNOVATION L. REV. 107 (2023) (discussing the current tension between state right of publicity laws and § 230 and the need for a federal right of publicity law.).

351. Exec. Order No. 14110, 3 C.F.R. 75191 (2023).

352. Exec. Order No. 14110, 3 C.F.R. 75191 § 1 (2023).

353. Exec. Order No. 14110, 3 C.F.R. 75202 § 4.5(a)(i)–(iii) (2023).

354. Exec. Order No. 14110, 3 C.F.R. 75203 § 4.5(a)(iv) (2023).

355. Exec. Order No. 14110, 3 C.F.R. 75203 § 4.5(b)–(c) (2023).

356. Solcyré Burga, *How a New Bill Could Protect Against Deepfakes*, TIME (Jan. 31, 2024, 4:34 PM), https://time.com/6590711/deepfake-protection-federal-bill/ [https://perma.cc/H54Q-4CSU].

> some other form of information published with the visual depiction, that the visual depiction is not authentic.[357]

This definition ensures that any synthetic media is encompassed so long as it depicts an identifiable individual, regardless of if there is a disclosure attached to the media.[358]

The Act could allow "an identifiable individual who is the subject of a digital forgery" to bring an action against anyone who "knowingly produced or possessed the digital forgery" who intended to disclose, did disclose, or solicited the digital forgery if three elements are met.[359] First, the identifiable individual must not have consented to the production, disclosure, solicitation, or possession of the digital forgery.[360] Second, the alleged perpetrator must know or recklessly disregard that the individual depicted did not consent.[361] Lastly, "the production, disclosure, solicitation, or possession is in or affects" interstate commerce.[362]

The proposed Act allows plaintiffs to preserve their privacy when bringing civil actions.[363] Plaintiffs can proceed under a pseudonym and have personal identifying information redacted (or have the documents filed under seal), and a protective order may be issued, for purposes of discovery, to allow "any intimate visual depiction or digital forgery" to remain in the care of the court.[364] A successful plaintiff may recover monetary damages, reasonable attorney's fees and other litigation costs, and obtain equitable relief (in the form of temporary restraining orders or injunctions).[365] Under the proposed Act, a plaintiff has ten years to file an action from either the date the plaintiff reasonably discovers the violation or the date the plaintiff turns eighteen, whichever is later.[366]

IV. SOLUTION

With the continued growth and development of artificial intelligence technologies, deepfakes pose an ever-growing threat, particularly to the everyday person. So, Congress should adopt laws to address the use of nonconsensual deepfakes intended to misappropriate the identity of a living person. Congress should focus on creating legislation

357. Disrupt Explicit Forged Images And Non-Consensual Edits Act of 2024, S. 3696, 188th Cong. (2024).
358. *See id.*
359. *Id.*
360. *Id.*
361. *Id.*
362. *Id.*
363. *Id.*
364. *Id.*
365. *Id.*
366. *Id.*

that protects citizens from the reputational harm that is inherent from the creation of a deepfake.

While reputational harm is apparent in cases involving nonconsensual pornography or misinformation connected with elections, it is not confined to only these instances. As Judge Learned Hand stated, "a reputation, like a face, is the symbol of its possessor and creator, and another can use it only as a mask."[367] As a result, when another uses someone's likeness, "he borrows the owner's reputation, whose quality no longer lies within his own control. This is an injury, even though the borrower does not tarnish it[.]"[368] Therefore, any federal laws should prohibit the use of deepfakes generally and not be confined to a certain category, like nonconsensual pornography, because even if no apparent harm is done to one's reputation, there is an inherent injury in not being able to control one's reputation and what it is associated with. However, before a specific solution can be discussed, First Amendment implications should be briefly acknowledged.

A. Potential First Amendment Implications

The First Amendment states that "Congress shall make no law … abridging the freedom of speech[.]"[369] As highlighted by the mixed reactions to the California law, laws that seek to restrict deepfake technology may implicate the First Amendment's protection of free speech because deepfakes potentially involve some expressive activity.[370]

Prominent Constitutional Law scholar, Erwin Chemerinsky, noted that government restrictions on deepfakes would likely comport with both the Constitution and Supreme Court precedent.[371] Because one of the main purposes of the First Amendment, as revealed through Supreme Court precedent, is to further the "marketplace of ideas."[372]

Indeed, the Supreme Court has held that "the right to free speech is not absolute at all times and under all circumstances."[373] The Court has found "certain well-defined and narrowly limited classes of speech" that are thought not to raise any Constitutional concerns.[374] Some of the relevant classes of speech that fall beyond the First Amendment's protections include speech that is lewd and obscene or speech that is libelous.[375] These types of speech fall outside the First Amendment's protections because "such utterances are no essential part of any exposition of ideas, and are of

367. Yale Elec. Corp. v. Robertson, 26 F.2d 972, 974 (2d Cir. 1928).
368. *Id.*
369. U.S. CONST. amend. I.
370. Jones, *supra* note 201, at 5–7.
371. *Id.* at 5–6.
372. *Id.*
373. Chaplinsky v. New Hampshire, 315 U.S. 568, 571–72 (1942).
374. *Id*.
375. *Id.* at 572.

such slight social value as a step of truth that any benefit that may be derived from them is clearly outweighed by the social interest in order and morality."[376]

Relying on this reasoning from the Supreme Court, some believe that deepfakes "add nothing to the marketplace of ideas and indeed detract from it."[377] This is because nothing is likely gained from realistic manipulations of media making people say, or do, things that never happened.[378] Truly, as Dean Chemerinsky remarked, the use of nonconsensual deepfakes "can cause great harm to [one's] reputation."[379]

Additionally, some aspects of tort law, specifically defamation, support the argument that deepfakes are not protected by the First Amendment.[380] For instance, defamatory speech has no First Amendment protection if the speaker knows the statements are false or acts with reckless disregard of the truth.[381] "The [Supreme Court] has explained that the importance of preventing wrongful harm to reputation and of protecting the marketplace of ideas justifies the liability for the false speech."[382] As such, the remainder of this section will assume that deepfakes generally fall outside the protections of the First Amendment and can therefore be regulated.

B. Proposed Statute

To protect citizens from the reputational dangers of deepfakes and to provide citizens a uniform cause of action across the country, Congress should create a law that regulates deepfakes as a category rather than regulating individual subcategories of deepfakes, like pornography or election laws.

First, this new law will need to define what constitutes a "deepfake." As seen from the various state laws, every state that regulates deepfakes defines the term in a different way.[383] Some States cabin the term "deepfake" to only include altered videos, while others include a wide array of media.[384] Likewise, some States classify media as a "deepfake" if it is

376. *Id.*
377. Jones, *supra* note 201, at 5.
378. *Id.* at 5–6.
379. *Id.* at 6.
380. *Id.*
381. *Id.*
382. *Id.*
383. *Infra* Part II.
384. *Compare* TEX. ELEC. CODE ANN. § 255.004(e) (2023) (defining a deepfake as "a video, created with the intent to deceive, that appears to depict a real person performing an act that did not occur in reality."), *with* WYO. STAT. ANN. § 6-4-306(a)(iii) (2024) (defining "image" as "a photograph, film, videotape, recording, digital file or any other recording, including a computer generated image that purports to represent an identifiable person.").

created by digital means, while others do not have any technological prerequisite, classifying any altered media as a deepfake.[385]

When defining the term "deepfake," for purposes of the law, Congress should adopt the definition used in the proposed DEEP FAKES Accountability Act. This proposed Act defined a "deepfake" as:

> 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.[386]

Rather than limit deepfakes to include only pictures or videos, this definition contemplates the potential use of audio deepfakes as well, fully encompassing the various types of media that could be used to create deepfakes.

To remedy some of the ambiguity created by this definition of "deepfake," Congress should alter part (B) of this definition. The altered definition of part (B) should instead read: "the production of which involved the use of artificial intelligence." This revision to the DEEP FAKES Accountability Act definition ensures that only synthetic media that relies on artificial intelligence will be encompassed. The original definition, as written, is ambiguous because it fails to define what "substantially dependent" means. Additionally, the definition regulates those deepfakes that relied on "technical means," potentially encompassing altered media that is merely a cheapfake, which uses no artificial intelligence at all. Consequentially, the original definition is likely overly broad, reaching media that is likely protected speech. As such, making the revision to the definition ensures it is narrowly tailored to target only synthetic media that is created using modern artificial intelligence.

Along with the change to part (B) of the definition, Congress should define "person" for purposes of this law to only include a living person. Defining "person" in this way would best serve this law's purpose. As mentioned above, the purpose of this law is to prevent and remedy the reputational damage deepfakes inflict on the individual. Much like privacy

---

385. *Compare* WYO. STAT. ANN. § 6-4-306(a)(iii) (2024) (including "computer generated image[s]" as a deepfake), *with* TEX. ELEC. CODE ANN. § 255.004(e) (2023) (classifying a video as a deepfake so long as it "appears to depict a real person performing an act that did not occur in reality.").

386. Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2021, H.R. 2395, 117th Cong. (2021).

torts, this right is personal; and thus, should not be descendible.[387] As such, only a living person affected by a deepfake should have a cause of action under this law.

As a practical matter, restricting "person" to include only living persons makes sense as well. Deepfakes are typically created to spread misinformation, to defraud someone, or simply to harm someone's reputation.[388] As a result, bad actors who create deepfakes will typically only target living people. In instances where a deepfake is created of a deceased person, it is likely because that person is famous and can provide commercial benefits to a person or company. In these instances, right of publicity laws are a more appropriate vehicle for redress.[389]

With the important terms defined, the next step is to specify what conduct the law prohibits. Many deepfake laws focus on regulating the dissemination of deepfakes to third parties not depicted in the deepfake. Because of this misplaced focus, in those States a bad actor can create nonconsensual deepfake pornography and show it to the person depicted in the video and not violate the state's deepfake law. To combat this phenomenon, the prohibited conduct in a federal law should read as follows: "A person violates this law if he or she creates, disseminates, or stores any deepfake that the person knows or should have known was created without the consent of the person depicted."

Writing the law to encompass creating, disseminating, and storing deepfakes will deter people from creating and storing deepfakes for personal use—something that few state laws address. As noted earlier, because this law would only apply to living persons depicted in deepfakes, individuals could still create deepfakes depicting deceased persons without violating the law. This issue is certainly an aspect of the suggested law where reasonable minds may differ, but for an initial law to address deepfakes, this is likely the simplest line to draw.

If consent is required in order to depict a deceased person in a deepfake, it becomes unclear whose consent is required. It is possible that the deceased person's property right in their likeness belongs to a single person, but it is also possible that it belongs to a number of people. This leaves it unclear as to whether making a deepfake requires the approval of every individual with an interest in the property right, merely the majority of them, or perhaps just one of them.

---

387. Joshua L. Simmons & Miranda D. Means, *Split Personality: Constructing a Coherent Right of Publicity Statute*, ABA, https://www.americanbar.org/groups/intellectual_property_law/publications/landslide/2017-18/may-june/split-personality/ [https://perma.cc/UEM6-NTJD] (last visited Oct. 17, 2024) (stating "the right to privacy is a right to one's own "inviolate personality," and is therefore personal to each individual).

388. *See* U.S. DEP'T OF HOMELAND SEC., *supra* note 15, at 16–18.

389. Simmons & Means, *supra* note 387 (noting that the descendability of rights of publicity "properly rewards hard work and investment in the development of a famous personality.").

An example of this potential problem would be if a person's mother died and left each of her three children a one-third share in her property right to her likeness. Who would have to consent if one of the kids wanted to create a deepfake of his mother singing his favorite childhood song? If the children have a strained relationship, is the consent of one child sufficient to comply with the law? Or does the child need to convince one or both of his siblings to consent in order to be in compliance? For reasons such as this, a bright-line rule: protecting against deepfakes only of living persons is the simplest solution, and Congress can debate moving forward the best way to handle instances of deepfakes involving deceased individuals.

The next component in crafting this deepfake law is outlining the penalties for violating it. A person who violates this law should be subject to criminal penalties, which will provide a sufficient deterrent for bad actors. These criminal penalties should provide a judge or jury with appropriate discretion to both punish the offending behavior and deter future behavior. Thus, violating the law should subject an individual to a fine. This fine should be scalable to account for repeat violations. Along with a monetary penalty, imprisonment should be an option for the judge or jury to impose, based on the nature of the offense. As with the fine, the more times the offender has violated the law, the greater weight imprisonment should be given to deter the conduct.

For each of these penalties, there should be special provisions escalating the penalty for violations depicting minors or sexual acts. Violating the law in these ways should go to the weight of the offense and afford the judge or jury the ability to impose stricter penalties—even in situations where it is a first-time offender, depending on the gravity of the offense.

Aside from exposing an offender to criminal liability, a person who violates this law should be exposed to civil liability in order to compensate the injured person for any reputational damage. The law should provide a section stating that "a victim of a nonconsensual deepfake, as defined by this law, has a private right of action to seek injunctive relief, or any other equitable relief, monetary damages, and reasonable attorney's fees." These damages should include the ability to be awarded any money the offender earned from the creation, storage, or dissemination of the deepfake. There should be a rebuttable presumption that the mere creation of a deepfake without consent caused some reputational damage, but the depicted individual should be required to prove the extent of the reputational injury and be awarded appropriate damages. Along with compensatory damages, a judge or jury should be allowed to award punitive damages, based on the severity and nature of the offense. Allowing attorney's fees to be recoverable is also important because it removes a potential barrier to the courts for those individuals who are victims but may not be able to afford an attorney.

Lastly, the law should address any exceptions and how it may interact with other laws. There should, however, be exceptions so the law should not be construed to interfere with any lawful exercise of someone's First Amendment rights. As mentioned earlier, most deepfakes are likely not protected under the First Amendment, but there may be some instances when a deepfake is protected by an individual's free speech rights.

This suggested law should have a section stating the following: "For purposes of section 230 of the Communications Decency Act (47 U.S.C. § 230), this law shall be considered to be a 'law pertaining to intellectual property' under subsection (e)(2) of section 230." This provision would allow those with a right of action under this law to bring an action against internet platforms that host the deepfakes. Allowing platforms to be held liable for deepfake content on their applications will encourage them to take steps to ensure no deepfake content is present.

It should be noted that platforms would not be strictly liable for deepfake content on their platform; rather, platforms must meet the scienter requirement, meaning that they must have known or should have known that the depicted person did not consent. Practically speaking, this standard means that once a platform is put on notice that some content is a nonconsensual deepfake, the platform must remove it, or they will be subject to suit. Platforms could eventually gain constructive notice if deepfakes continue to appear on their application and no steps are taken to identify them prior to posting.

CONCLUSION

What once seemed like something out of a science-fiction movie—using technology to appropriate someone's identity—is more of a reality each and every day. Although nonconsensual deepfakes are still mostly confined to the realms of pornography and politics, it is only a matter of time until the everyday person finds that they are a victim and had their reputation harmed. While some states have tried to protect citizens against nonconsensual deepfakes, the vast majority have not. Therefore, Congress should take action and enact legislation, like the one proposed above, to criminalize the creation and dissemination of deepfakes and provide individuals with a private right of action when they become victims of identity misappropriation by deepfakes. This legislation would sufficiently deter bad actors from creating and disseminating deepfakes while simultaneously providing victims with adequate avenues for recovery.