

RECOMMENDATIONS FOR REGULATION OF  
DEEPPAKES IN THE U.S.:  
DEEPPAKE LAWS SHOULD PROTECT EVERYONE NOT  
ONLY PUBLIC FIGURES

By Lourdes Vazquez

## TABLE OF CONTENTS

CONTENTS	1
INTRODUCTION	2
I. History of Deepfakes	2
A. What is a Deepfake?	2
B. Who Created Deepfakes?	2
C. How are Deepfakes Created?	3
II. The Importance of Regulating Deepfakes	3
A. The Issues Deepfakes Present Currently and in the Future	3
III. Relevant Considerations in Deepfake Regulation	6
A. First Amendment as it Applies to Deepfake Laws	6
B. Defamation Law	7
a. Definitions of the Four Classes of Individuals	7
b. Standards of Proving Fault for the Four Classes of Individuals	9
c. Is Defamation Law Enough?	11
IV. Deepfake Laws Landscape	13
A. Federal Legislation	14
a. National Defense Authorization Act for Fiscal Year 2020	14
B. State Legislation	14
a. California	14
i. AB 602	14
ii. AB 730	15
iii. AB 1280	16
b. Texas	17
c. Virginia	18
C. Recommendations	18
a. Federal Deepfake Definition	18
b. Five Takeaways for State Deepfake Laws	19
CONCLUSION	19
BIBLIOGRAPHY	21

## INTRODUCTION

What is fake? This has become an increasingly difficult question to answer due to the emergence of deepfake technology. It is not just a question of what is true and what is fake because deepfakes pose harms at a large-scale level and individual level, particularly defamation harms. Protections and reparations from deepfake harms at the individual level are not guaranteed to everyone in our society. Thus, it has become increasingly important to find a way to handle this new source of defamation harm.

This paper gives an overview of deepfakes, including what they are and their creation. Then this paper provides reasons why regulation of deepfakes is important, including the concern that defamation law may not be enough to address deepfake harms. Third, this paper addresses relevant considerations in deepfake regulation, such as the First Amendment and defamation law, specifically how defamation law disparately protects individuals based on their classification as private figures or public figures. Fourth, this paper provides an overview of the deepfake legislative landscape at the federal level and state level, and the strengths and weaknesses of such deepfake laws. Lastly, it addresses the issue that current deepfake laws do not protect private figures to the same extent that deepfake laws protect public figures and provides recommendations for deepfake laws to protect private figures equally to public figures.

### I. History of Deepfakes

#### A. What is a Deepfake?

Deepfakes can take the form of audio manipulation or visual manipulation. Deepfakes that manipulate audio involve changing what was said or removing what was said. Deepfakes that manipulate videos involve superimposing images of a person or thing onto another person or thing.<sup>1</sup> Deepfakes are false yet highly realistic artificial intelligence-created media, such as a video showing people saying things they never said and doing things they never did.<sup>2</sup> Deepfake technology can “seamlessly stitch anyone in the world into a video or photo they never actually participated in.”<sup>3</sup>

#### B. Who Created Deepfakes?

Deepfakes were created in 2017 by a Reddit user dubbed ‘deepfakes.’<sup>4</sup> The term to characterize “such superimposed realistic images were obtained when” said Reddit user “superimposed actresses’ faces on porn performers’ faces.”<sup>5</sup> The term deepfake has since then stuck.<sup>6</sup>

---

<sup>1</sup> Sally Adee, *What Are Deepfakes and How Are They Created?*, TECH TALK BLOG (April 29, 2020 16:00 GMT), <https://spectrum.ieee.org/tech-talk/computing/software/what-are-deepfakes-how-are-they-created>.

<sup>2</sup> Wilmer Hale, *First Federal Legislation on Deepfakes Signed Into Law*, JD SUPRA (Dec. 24, 2019), <https://www.jdsupra.com/legalnews/first-federal-legislation-on-deepfakes-42346/>.

<sup>3</sup> Sally Adee, *What Are Deepfakes and How Are They Created?*, TECH TALK BLOG (April 29, 2020 16:00 GMT), <https://spectrum.ieee.org/tech-talk/computing/software/what-are-deepfakes-how-are-they-created>.

<sup>4</sup> *Id.*

<sup>5</sup> *Id.*

<sup>6</sup> *Id.*

## C. How are Deepfakes Created?

Deepfakes are created by algorithms, in which one algorithm produces the deepfake and the other algorithm attempts to detect the deepfake, thereby continually improving the fake.<sup>7</sup> The majority of today's deepfakes are made using a combination of artificial intelligence (AI) and non-AI algorithms.<sup>8</sup>

Generative adversarial networks (GANs) are “a class of statistical models that can generate new data.”<sup>9</sup> Although deepfakes may be created using GANs “such an attempt would be impractical because GANs require a huge amount of training data, and a longer time for models to generate the preferred images.”<sup>10</sup> For example, a GAN can generate new photos of animals that look like real animals if they are provided with enough images of the animals.<sup>11</sup> While GANs are good at synthesizing images so that it is “near impossible to tell apart a deepfake face from a real face, more specifically GAN models are not good for making videos because GANs have a difficult time keeping the same image aligned from one frame to the next.”<sup>12</sup>

Rather, the main ingredient to produce deepfakes is machine learning.<sup>13</sup> Machine learning has produced deepfakes at a much faster and lower cost.<sup>14</sup> The machine learning process for deepfake creation can be summarized in three steps: understanding, combination, and superimposition. To make a deepfake video of someone “a creator would first train a neural network on many hours of real video footage of the person to give it a realistic “understanding” of what he or she looks like from many angles and under different lighting.”<sup>15</sup> This leads to the creation of a trained network from which the machine works off of. The creator then “combines the trained network with computer graphic techniques to superimpose a copied image of the person” onto a different actor, body, or being.<sup>16</sup>

## II. The Importance of Regulating Deepfakes

### A. The Issues Deepfakes Present Currently and in the Future

Some deepfakes are concerning and some deepfakes are not concerning. Deepfakes that are concerning and present an issue today are deepfakes that have been used for a myriad of

---

<sup>7</sup> *Id.*

<sup>8</sup> *Id.*

<sup>9</sup> Google, *Generative Adversarial Networks Introduction*, (Oct. 8, 2019), <https://developers.google.com/machine-learning/gan>.

<sup>10</sup> Sally Adee, *What Are Deepfakes and How Are They Created?*, TECH TALK BLOG (April 29, 2020 16:00 GMT), <https://spectrum.ieee.org/tech-talk/computing/software/what-are-deepfakes-how-are-they-created>.

<sup>11</sup> Google, *Generative Adversarial Networks Introduction*, (Oct. 8, 2019), <https://developers.google.com/machine-learning/gan>.

<sup>12</sup> Sally Adee, *What Are Deepfakes and How Are They Created?*, TECH TALK BLOG (April 29, 2020 16:00 GMT), <https://spectrum.ieee.org/tech-talk/computing/software/what-are-deepfakes-how-are-they-created>.

<sup>13</sup> *Id.*

<sup>14</sup> *Id.*

<sup>15</sup> *Id.*

<sup>16</sup> *Id.*

purposes such as bullying, revenge porn, video manipulation, audio manipulation, and extortion.<sup>17</sup> Such use of deepfakes harms individuals' reputations. In addition to the above examples, deepfakes pose a real danger of harming U.S. democracy because deepfakes could negatively affect elections, increase instances of identity fraud, spread fake news, and lead individuals to try to escape accountability by claiming that evidence of what the individual said or did is a deepfake when it is not. Hence, there is a need for action to be taken to address the issues created by deepfakes at the individual level.<sup>18</sup>

The first instance of deepfakes was nonconsensual pornography in which the faces of famous actresses were superimposed on the bodies of other women.<sup>19</sup> According to Deeptech, a deepfake detection firm in Amsterdam, "the clearest threat that deepfakes pose right now is to women."<sup>20</sup> The data provided by Deeptech shows that nonconsensual pornography accounts for 96% of deepfakes.<sup>21</sup> While most attention has focused on nonconsensual pornography affecting celebrities, there is an increasing number of deepfake reports regarding the use of deepfakes for creating fake revenge porn of private individual figures.<sup>22</sup> While women have been the major targets of deepfake bullying and revenge porn, women will not be the sole targets of deepfake bullying and revenge porn.<sup>23</sup>

Deepfakes pose a danger to democracy because deepfakes can lead to false information as to what a political candidate did or did not say. On May 20, 2020, Speaker of the House Pelosi addressed President Trump's tweets about MSNBC host Joe Scarborough.<sup>24</sup> This video was later edited to reflect Pelosi slurring her words in an attempt to appear as if she was drunk.<sup>25</sup> While some have considered this to be a deepfake, it is not a deepfake. At best it is a 'cheapfake.' There was no superimposition of a face onto another person's body. To create the slurring speech effect, the video was slowed down.<sup>26</sup> Although this is not a deepfake, it demonstrates the intent of malicious use of videos for misleading and misinforming people. If a 'cheapfake' can harm Pelosi's reputation, imagine how much more deepfakes can harm an individual's reputation, regardless of whether the individual is a private figure or a public figure.

In addition to video manipulation, deepfakes also include audio manipulation where data of a person's voice is manipulated to say things the person has not said.<sup>27</sup> Examples of audio deepfakes include humorous recordings of George W. Bush reading "In Da Club" by 50 Cent,<sup>28</sup> Queen Elizabeth II reading "God Save the Queen,"<sup>29</sup> and an altered Public Service

---

<sup>17</sup> *Id.*

<sup>18</sup> *Id.*

<sup>19</sup> *Id.*

<sup>20</sup> *Id.*

<sup>21</sup> *Id.*

<sup>22</sup> *Id.*

<sup>23</sup> *Id.*

<sup>24</sup> *Another Fake Video of Pelosi Goes Viral*, WASHINGTON POST (Aug. 3, 2020), <https://www.washingtonpost.com/technology/2020/08/03/nancy-pelosi-fake-video-facebook/>.

<sup>25</sup> *Id.*

<sup>26</sup> *Id.*

<sup>27</sup> Luke Dormehl, *These Amazing Audio Deepfakes Showcase Progress of A.I. Speech Synthesis*, DIGITAL TRENDS (June 6, 2020), <https://www.digitaltrends.com/news/best-audio-deepfakes-web/>.

<sup>28</sup> Dave Johnson, *Audio Deepfakes: Can Anyone Tell If They're Fake?*, HOW TO GEEK (Aug. 3, 2020, 6:40 AM EDT), <https://www.howtogeek.com/682865/audio-deepfakes-can-anyone-tell-if-they-are-fake/>.

<sup>29</sup> Luke Dormehl, *These Amazing Audio Deepfakes Showcase Progress of A.I. Speech Synthesis*, DIGITAL TRENDS (June 6, 2020), <https://www.digitaltrends.com/news/best-audio-deepfakes-web/>.

Announcement of former President Obama.<sup>30</sup> As for deepfakes leading to extortion, there have been unconfirmed reports of deepfake audio of CEOs being used to swindle employees into sending money to fraudsters.<sup>31</sup>

Deepfakes also present future issues. Governments are worried that deepfakes pose a danger to democracy.<sup>32</sup> Governments are concerned that a person can make a politician running for reelection appear in a porn video, or that deepfakes may be used to spark an attempted coup of the current government.<sup>33</sup> Meanwhile, corporations are concerned about deepfakes used for identity fraud purposes. Corporations are worried that deepfakes would be used to make fraudulent online payments and hack into personal banking services.<sup>34</sup> It is unclear now how deepfakes could obtain access to such personal information.<sup>35</sup>

There is also concern that deepfakes may be used to circulate fake news.<sup>36</sup> Most recently, the creators of South Park have created a series called “Sassy Justice with Fred Sassy” in which a character named Fred Sassy reports fake news out of Cheyenne, Wyoming.<sup>37</sup> While these web series are a comedic effort to make fun of deepfakes and make it less scary, it shows the possibility of using deepfakes for fake news reporting.<sup>38</sup> At the moment, fake news reporting using deepfakes is costly, which might serve as a deterrence to many people and companies.<sup>39</sup>

Professors Bobby Chesney and Danielle Citron have stated that one danger that deepfakes pose is what is called the “Liar’s Dividend.”<sup>40</sup> This is because “the decreasing possibility that any image of a politician or celebrity is genuine ‘makes it easier for liars to avoid accountability for things that are in fact true.’”<sup>41</sup> This means that the liar’s dividend can be used as a way to get out of trouble because people can deny the evidence of wrongdoing, and shield themselves by claiming the evidence of wrongdoing is a deepfake.<sup>42</sup>

Most importantly, there should be a concern regarding the lack of an adequate and equitable mechanism to address the defamation harms caused to private individuals by deepfakes. While individuals defamed by deepfakes can seek legal redress under existing defamation law, currently, defamation law contains disparities in the protections it affords to public figures and private figures. Hence, relying on defamation law to address deepfake

---

<sup>30</sup> Joseph Foley, *10 Deepfake Examples that Terrified and Amused the Internet*, CREATIVE BLOQ (Mar. 23, 2020), <https://www.creativebloq.com/features/deepfake-examples>.

<sup>31</sup> Sally Adee, *What Are Deepfakes and How Are They Created?*, TECH TALK BLOG (April 29, 2020 16:00 GMT), <https://spectrum.ieee.org/tech-talk/computing/software/what-are-deepfakes-how-are-they-created>.

<sup>32</sup> *Id.*

<sup>33</sup> *Id.*

<sup>34</sup> *Id.*

<sup>35</sup> *Id.*

<sup>36</sup> Randall Colburn, *The Creators of South Park Launched an Entire Deepfake Studio to Make Sassy Justice*, AV CLUB (Nov. 30, 2020 10:10 AM), <https://news.avclub.com/the-creators-of-south-park-launched-an-entire-deepfake-1845529116>.

<sup>37</sup> *Id.*

<sup>38</sup> *Id.*

<sup>39</sup> *Id.*

<sup>40</sup> Richard L. Hasen, *Deep Fakes, Bots, and Siloed Justices: American Election Law in a Post-Truth World* (July 10, 2019), St. Louis University L. J., 2019, Forthcoming, U.C. Irvine School of L. Research Paper No. 2019-36, Available at SSRN: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3418427](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3418427).

<sup>41</sup> *Id.*

<sup>42</sup> Sally Adee, *What Are Deepfakes and How Are They Created?*, TECH TALK BLOG (April 29, 2020 16:00 GMT), <https://spectrum.ieee.org/tech-talk/computing/software/what-are-deepfakes-how-are-they-created>.

defamation can result in disparate treatment and outcomes for private figures compared to public figures.

### III. Relevant Considerations in Deepfake Regulation

There are two relevant considerations for deepfake regulations: the First Amendment as it applies to deepfake laws, and defamation law as a possible tool to address the harms caused by deepfakes at an individual level.

#### A. First Amendment as it Applies to Deepfake Laws

The First Amendment of the U.S. Constitution states “Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.”<sup>43</sup> The First Amendment applies to the U.S. government and people living in the U.S.

The connection between deepfake laws and the First Amendment is that regulating deepfakes may pose a First Amendment free speech issue.<sup>44</sup> The Constitution’s First Amendment states that “Congress shall make no law ... abridging the freedom of speech.”<sup>45</sup> This means that restricting speech based on its content is generally not allowed with exceptions.<sup>46</sup> The Supreme Court has stated that the First Amendment’s free speech guarantee does not apply to the following: “the lewd and obscene, the profane, the libelous, and insulting words or ‘fighting’ words.”<sup>47</sup> Hence, restricting deepfakes because their content is misinformation likely will not be allowed unless deepfakes fall into the Supreme Court’s stated exceptions.<sup>48</sup> Not all deepfakes would meet these exceptions so banning deepfakes completely would be a form of government censorship prohibited under the First Amendment.<sup>49</sup>

In instances where deepfakes are used for defamatory speech, deepfakes are not protected under the First Amendment.<sup>50</sup> Two possible defenses to deepfake creators are “parody” and “satire” because, in instances where deepfakes are used for parody or satire, deepfakes are protected under the First Amendment.<sup>51</sup>

Given that this paper is about the defamatory harm deepfakes cause to individuals, both private and public figures, this paper will work within the framework that deepfakes are defamatory to both private and public figures, and so no First Amendment issue is present in creating deepfake legislation that will protect individuals such as public figures and private

---

<sup>43</sup> U.S. Const. amend. I.

<sup>44</sup> Lourdes Vazquez, *Deepfakes And Fake News: Is Regulation of Deepfakes Possible?*, IP BRIEF AT AMERICAN U. WASHINGTON C. OF L. (Jan. 30, 2020), <http://www.ipbrief.net/2020/01/30/deepfakes-and-fake-news-is-regulation-of-deepfakes-possible/>.

<sup>45</sup> *Id.*

<sup>46</sup> *Id.*

<sup>47</sup> *Id.*

<sup>48</sup> *Id.*

<sup>49</sup> Hasen, *Deep Fakes, Bots, and Siloed Justices: American Election Law in a Post-Truth World*.

<sup>50</sup> Jessica Ice, *Defamatory Political Deepfakes and the First Amendment*, 70 Case W. Res. L. Review 417, 419 (2019), <https://scholarlycommons.law.case.edu/cgi/viewcontent.cgi?article=4854&context=caselrev>.

<sup>51</sup> *Id.*

figures from defamation. That said if legislation as a tool to regulate deepfakes were found to be questionably unconstitutional under the First Amendment, the courts would or could balance constitutionally protected speech with an important governmental interest.<sup>52</sup>

## B. Defamation Law

So, how can current defamation laws handle deepfakes, and how does current defamation law treat private figures and public figures? First, it is necessary to provide background information about defamation lawsuits. Defamation includes libel and slander.<sup>53</sup> Libel is a written defamatory statement.<sup>54</sup> Slander is a verbal defamatory statement.<sup>55</sup> One of the issues posed by deepfakes is whether it would fall under libel or slander. “Libel is defamatory statements and/or pictures published in print or writing; or broadcast in the media, such as over the radio, on TV or in film ... [I]t must be represented as a fact, not an opinion.”<sup>56</sup> The most common deepfakes are composed of images so these deepfakes would most neatly fit under libel. Hence, the below analysis focuses on libel lawsuits brought by public figures, limited-purpose public figures, involuntary public figures, and private figures.

Defamation law seems to protect individuals from defamation on a four-tiered basis. Before delving into the legal standards provided to those four tiers, it is important to define the classes of individuals those tiers represent.

### a. Definitions of the Four Classes of Individuals

The four tiers represent classes of individuals and are as follows: public figures are the most protected, limited-purpose public figures are less protected, involuntary public figures even less so, and private individuals are the least protected.<sup>57</sup> According to *Doe v. Pittsylvania City*, the first three tiers represent the types of public figures in defamation law.

*Doe* defines public figures to be those “who achieve such pervasive fame or notoriety that they become public figures for all purposes and in all contexts.”<sup>58</sup> In other words, public figures are those who “occupy positions of ... persuasive power and influence.”<sup>59</sup> Public figures can include appointed or elected public officials.<sup>60</sup> Other examples of public figures include actresses, actors, celebrities, and well-known athletes.

*Doe* defines limited-purpose public figures as those “who voluntarily inject themselves into a particular public controversy and thereby become public figures for a limited range of

---

<sup>52</sup> *Id.*

<sup>53</sup> *Libel vs. Slander: Different Types of Defamation*, NOLO, <https://www.nolo.com/legal-encyclopedia/libel-vs-slander-different-types-defamation.html> (last visited Nov. 30, 2020).

<sup>54</sup> *Id.*

<sup>55</sup> *Id.*

<sup>56</sup> *Defamation Law - Guide to Libel and Slander Law*, HG LEGAL RESOURCES, <https://www.hg.org/defamation.html#:~:text=Libel%20is%20defamatory%20statements%20and,a%20fact%2C%20not%20an%20opinion> (last visited Nov. 30, 2020).

<sup>57</sup> *Doe v. Pittsylvania Cty.*, 844 F. Supp. 2d 724 (W.D. Va. 2012).

<sup>58</sup> *Id.*

<sup>59</sup> *Defamation*, CORNELL LEGAL INFORMATION INSTITUTE, <https://www.law.cornell.edu/constitution-conan/amendment-1/defamation> (last visited Nov. 27, 2020).

<sup>60</sup> *Id.*

issues.”<sup>61</sup> Limited-purpose public figures are “individuals who, while not having any generalized fame, influence, or power, nonetheless ha[ve] one or more of those attributes in some limited area or category to be rightfully regarded as a public figure in that area or category.”<sup>62</sup> Another way individuals can be categorized as limited-purpose public figures are by “voluntarily becom[ing] a key figure in a particular controversy, or ... gain[ing] prominence in a particular, limited field,” but by not reaching such fame that their fame is not at “an all-encompassing level.”<sup>63</sup> This includes experts such as Dr. Anthony Fauci during the COVID-19 pandemic.

*Doe* also defines involuntary public figures as those “who become public figures through no purposeful action of their own.”<sup>64</sup> Involuntary public figures are significant individuals who are drawn into public controversy.<sup>65</sup> Examples of involuntary public figures may include Captain Sully who landed an airplane on the Hudson River, or an air traffic controller who was assigned partial blame for an airplane crash.<sup>66</sup>

“Defining who is a public figure for purposes of First Amendment protections is a question of federal constitutional law, and therefore the federal courts say on the matter is decisive and binding on state courts.”<sup>67</sup> This means, “state courts cannot remove public-figure status from those who have been deemed public figures by the federal courts, but states can broaden the scope of the classification.”<sup>68</sup> Hence, if definitions would help seek defamation remedies, the federal courts could provide definitions for who is and is not a public figure, and so state courts would have to use these minimum definitions but could expand upon these definitions to increase the level of protection afforded to defamed private figures.

While the courts provide no explicit definition of private figures, the courts implicitly define private figures as individuals who do not meet the public figures categories.<sup>69</sup> The courts’ careful categorizations of public figures in defamation law have brought the law’s attention to the needs of public figures and helped distinguish different groups of people and the protections afforded to them. Yet, simultaneously this attention towards public figures has left private figures out of the spotlight and thus neglected by the law. This sentiment of private figures being the “other” has influenced the law and how it protects or does not protect private individuals when it comes to defamation claims.

---

<sup>61</sup> *Pittsylvania Cty.*, 844 F. Supp. 2d 724.

<sup>62</sup> *Gertz v. Robert Welch*, 418 U.S. 323, 351 (1974).

<sup>63</sup> *Proving Fault: Actual Malice and Negligence*, DIGITAL MEDIA LAW PROJECT, <https://www.dmlp.org/legal-guide/proving-fault-actual-malice-and-negligence#:~:text=A%20limited%2Dpurpose%20public%20figure%20is%20either%3A,reached%20an%20all%2Dencompassing%20level> (last visited Nov. 27, 2020).

<sup>64</sup> *Pittsylvania Cty.*, 844 F. Supp. 2d 724.

<sup>65</sup> *Wells v. Liddy*, 186 F.3d 505 (4th Cir. 1999).

<sup>66</sup> *Dameron v. Wash. Magazine*, 250 U.S. App. D.C. 346, 779 F.2d 736 (1985).

<sup>67</sup> *Proving Fault: Actual Malice and Negligence*, DIGITAL MEDIA LAW PROJECT, <https://www.dmlp.org/legal-guide/proving-fault-actual-malice-and-negligence#:~:text=A%20limited%2Dpurpose%20public%20figure%20is%20either%3A,reached%20an%20all%2Dencompassing%20level> (last visited Nov. 27, 2020).

<sup>68</sup> Eileen Carroll Prager, *Public Figures, Private Figures and Public Interest*, 30 Stan. L. Rev. 157–189, 165 (1977), [www.jstor.org/stable/1228177](http://www.jstor.org/stable/1228177) (last visited Nov. 30, 2020).

<sup>69</sup> *Proving Fault: Actual Malice and Negligence*, DIGITAL MEDIA LAW PROJECT, <https://www.dmlp.org/legal-guide/proving-fault-actual-malice-and-negligence#:~:text=A%20limited%2Dpurpose%20public%20figure%20is%20either%3A,reached%20an%20all%2Dencompassing%20level> (last visited Nov. 27, 2020).

## b. Standards of Proving Fault for the Four Classes of Individuals

The four classes of individuals are afforded different levels of protection under defamation law. These differences are seen in the level of liability that needs to be found, and the standard of evidence that must be provided to find the other party guilty of libel, amongst others.

Public figures are at a lower standard of proving fault under defamation law because a broad sweep of conduct by critics is protected.<sup>70</sup> The ruling of *New York Times Co. v. Sullivan* was that:

“[E]xpression on matters of public interest is protected by the First Amendment. Within that area of protection is commentary about the public actions of individuals. The fact that expression contains falsehoods does not deprive it of protection, because otherwise such expression in the public interest would be deterred by monetary judgments and self-censorship imposed for fear of judgments.”<sup>71</sup>

The Supreme Court then determined that the *Times* rule applied to public officials.

“Candidates for public office were subject to the *Times* rule and comment on their character or past conduct, public or private, insofar as it touches upon their fitness for office, is protected. Thus, a wide range of reporting about both public officials and candidates is protected. Certainly, the conduct of official duties by public officials is subject to the widest scrutiny and criticism. But the Court has held as well that criticism that reflects generally upon an official’s integrity and honesty is protected. Candidates for public office, the Court has said, place their whole lives before the public, and it is difficult to see what criticisms could not be related to their fitness.”<sup>72</sup>

Defamation of public figures is balanced with critics’ First Amendment rights. So libelous statements or images of public figures, even if false, may still be out of reach of a successful defamation suit by public figures. Public figures have to prove “actual malice” to succeed in a defamation suit.<sup>73</sup> Actual malice is present when someone makes the action or statement in controversy “with knowledge that it was false or with reckless disregard of whether it was false or not.”<sup>74</sup>

To prove that an individual is a limited-purpose public figure, the following has to be satisfied: “(1) a public controversy [must] exist, (2) the [individual] played a meaningful role in the controversy, and (3) the allegedly defamatory statement related to the controversy.”<sup>75</sup> “A

---

<sup>70</sup> *Defamation*, CORNELL LEGAL INFORMATION INSTITUTE, <https://www.law.cornell.edu/constitution-conan/amendment-1/defamation> (last visited Nov. 27, 2020).

<sup>71</sup> *Id.*

<sup>72</sup> *Id.*

<sup>73</sup> *Id.*

<sup>74</sup> *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 84 S. Ct. 710 (1964).

<sup>75</sup> *Trivedi, LLC v. Lang*, No. A16-1209, 2017 Minn. App. Unpub. LEXIS 389, at 1 (May 1, 2017).

limited-purpose public-figure plaintiff must also prove that the allegedly defamatory statement was made with actual malice.”<sup>76</sup>

More recently, “involuntary public figure” status has developed in lower court decisions, such as *Dameron v. Washington Magazine*.<sup>77</sup> In *Dameron*, an air traffic controller at the time of a major plane crash was determined to be someone “involved in an event of overriding societal importance.”<sup>78</sup> To prove that a plaintiff is an involuntary public figure the defendant must first “demonstrate to the court that the plaintiff has become a central figure in a significant public controversy and that the allegedly defamatory statement has arisen in the course of discourse regarding the public matter.”<sup>79</sup> To prove that the plaintiff is a central figure in the controversy, “the defendant must put forth evidence that the plaintiff has been the regular focus of media reports on the controversy.”<sup>80</sup> A significant public controversy is “one that touches upon serious issues relating to, for example, community values, historical events, governmental or political activity, arts, education, or public safety.”<sup>81</sup> The court also held that although an involuntary public figure “need not have sought to publicize her views on the relevant controversy, she must have nonetheless assumed the risk of publicity.”<sup>82</sup> This means that the defendant must demonstrate that the plaintiff “has taken some action, or failed to act when action was required, in circumstances in which a reasonable person would understand that publicity would likely inhere.”<sup>83</sup>

The Supreme Court in *Gertz v. Robert Welch* (1974) and in *Time, Inc. v. Firestone* (1976), “provided greater protection for individuals who had not voluntarily placed themselves in the public eye. Unlike public figures, private individuals do not have to prove actual malice to win damages for libel.”<sup>84</sup> To support a defamation claim, in most states a private figure needs to show the defendant was negligent, whereas a public figure needs to show the defendant acted with actual malice.<sup>85</sup> Negligence is a lower standard than “actual malice.”<sup>86</sup> “Some states, however, impose a higher standard on private figures, especially if the statement concerns a matter of public importance.”<sup>87</sup>

---

<sup>76</sup> *Id.*

<sup>77</sup> Gary Bugh, *Public Figures and Officials*, THE FIRST AMENDMENT ENCYCLOPEDIA (2009), <https://www.mtsu.edu/first-amendment/article/1010/public-figures-and-officials>.

<sup>78</sup> *Id.*

<sup>79</sup> *Wells v. Liddy*, 186 F.3d 505 (4th Cir. 1999).

<sup>80</sup> *Id.*

<sup>81</sup> *Id.*

<sup>82</sup> *Id.*

<sup>83</sup> *Id.*

<sup>84</sup> Gary Bugh, *Public Figures and Officials*, THE FIRST AMENDMENT ENCYCLOPEDIA (2009), <https://www.mtsu.edu/first-amendment/article/1010/public-figures-and-officials>.

<sup>85</sup> *Proving Fault: Actual Malice and Negligence*, DIGITAL MEDIA LAW PROJECT, <https://www.dmlp.org/legal-guide/proving-fault-actual-malice-and-negligence#:~:text=A%20limited%2Dpurpose%20public%20figure%20is%20either%3A,Reached%20an%20all%2Dencompassing%20level> (last visited Nov. 27, 2020).

<sup>86</sup> *Id.*

<sup>87</sup> *Id.*

### c. Is Defamation Law Enough?

As for why public figures are protected more than private figures under defamation law and consequently are also protected more than private figures in deepfake laws, could partly be due to the reasoning behind defamation laws. The opinion in *Rosenbloom v. Metromedia* reflects “divergent traditions of thought about the general problem of reconciling the law of defamation with the First Amendment.”<sup>88</sup>

“One approach has been to extend the *New York Times* test to an expanding variety of situations. Another has been to vary the level of constitutional privilege for defamatory falsehood with the status of the person defamed. And a third view would grant to the press and broadcast media absolute immunity from liability for defamation.”<sup>89</sup>

Legal protection under defamation suits was based on the belief that individuals who become public figures lay their entire life before the public and thus need more protection.<sup>90</sup> The reasoning behind this need is flawed because as pointed out in *Rosenbloom*, “the idea that certain “public” figures have voluntarily exposed their entire lives to public inspection, while private individuals have kept theirs carefully shrouded from public view is, at best, a legal fiction.”<sup>91</sup> Private figures just as much as public figures can expose their lives to public inspection, and public figures can shroud their lives from public view if they so choose. “High-profile figures are used [for deepfakes] because their public profiles provides ample source material for AI to learn from, but with the amount of selfies the average person takes in a lifetime and technological advances, perhaps soon anyone could be used as a [deepfake] source.”<sup>92</sup> Such technological advances and ample source materials increase the need for laws protecting all individuals from deepfake defamation, regardless of how much an individual tries to or can successfully shroud their private life from public view.

The conventional belief that disclosure of a public figure’s life may more easily reach a larger audience than a disclosure of a private figure’s life has made it easier for defamation law to protect public figures than private figures.<sup>93</sup> With the advent of technological advances that make sharing one’s personal life with the world possible, it is a fallacy to say that public figures are the only individuals who can easily disclose their lives. Thanks to social media platforms it has never been easier for a private figure’s life to be available to the world. Fortunately, the Supreme Court has moved away from using the exposure of one’s private life to the public as a measure for determining whether a statement was libelous because it afforded insufficient protections to defamed individuals.<sup>94</sup> Regardless of the size of the intended or unintended audience, every person should have the same protections when it comes to deepfakes being made of them.

---

<sup>88</sup> *Gertz*, 418 U.S. 323.

<sup>89</sup> *Id.*

<sup>90</sup> *Rosenbloom v. Metromedia*, 403 U.S. 29, 48 (1971).

<sup>91</sup> *Id.*

<sup>92</sup> Joseph Foley, *10 Deepfake Examples that Terrified and Amused the Internet*, CREATIVE BLOQ (Mar. 23, 2020), <https://www.creativebloq.com/features/deepfake-examples>.

<sup>93</sup> *Id.*

<sup>94</sup> *Id.*

While some people argue that public figures are protected less than private figures when it comes to libel suits, this belief is based on the burden of proof that public figures must prove actual malice to succeed in a libel suit, while private figures must prove negligence to succeed in a libel suit.<sup>95</sup> The focus of the conventional belief has been on the intent of the law protecting private figures more through its development of a lower standard for private individuals, that of negligence. Rather, such a view is too narrow because it does not provide the full context surrounding the protections any individual receives from the law. While the standard of proving negligence is lower than proving actual malice just because a standard is lower does not mean that a private figure is protected more than a public figure. The conventional belief does not consider the impact of the law as to private figures filing defamation suits and winning defamation suits because of this lower standard. A more holistic approach is required to see the full effect of the legal protections for a public or private figure.

Courts' have recognized that the court system "may be a private figure's only avenue for contesting a defamatory statement" and that "the reputations of private figures deserve greater protection than those of public figures."<sup>96</sup> Hence, the courts' recognition is about the disproportionate resources between public and private figures, which is part of a holistic view.

Comparing the resources between a private figure and a public figure, a public figure may have more resources than a private figure to file suit for defamation. In general, for individuals who have financial resources, it is not that difficult to file defamation lawsuits.<sup>97</sup> But most private figures do not have the financial resources to file a defamation lawsuit which creates a financial barrier. Due to this financial barrier, a private figure may be less likely than a public figure to file a defamation lawsuit and as a result, private figures are less likely to obtain defamation damages compared to public figures, and thus while the intent was to protect private figures more, the resulting impact is that private figures are less protected than public figures by defamation law.

Additional ways in which defamation law disproportionately impacts private figures more than public figures is that defamation law assumes that all individuals, regardless of a public figure or private figure status have access to counteract false statements through legal representation and that the level of harm caused to public figures is greater than the harm caused to private figures. Contrary to that assumption, public officials and public figures have "greater access to the channels of effective communication and hence have a more realistic opportunity to counteract false statements than private individuals normally enjoy."<sup>98</sup> Public officials may have better access to legal representation for various reasons, such as the ability to afford legal representation and access to a network of willing attorneys due to the public figure's high-profile status. The public figure's high-profile status combined with their clout may allow them to convince the public that the defamatory statements or images are false. In contrast, a private figure's low-profile status may not garner the same public attention and so a private figure's use of traditional media outlets to correct any defamatory harms may not be effective.

Due to the assumption that defamation harms public figures more than defamation harms private figures, defamation law alone is not enough to provide redress to individuals harmed by

---

<sup>95</sup> *Id.*

<sup>96</sup> *Public Versus Private Figures*, NATIONAL CONSUMER RIGHTS, <https://www.nationwideconsumerrights.com/public-vs-private-figures.html> (last visited Nov. 30, 2020).

<sup>97</sup> Lata Nott, *A Quick Guide to Libel Law*, FREEDOM FORUM INSTITUTE, <https://www.freedomforuminstitute.org/first-amendment-center/primers/libellaw/> (last visited Dec. 5, 2020).

<sup>98</sup> *Marcone v. Penthouse Int'l Magazine for Men*, 754 F.2d 1072 (3d Cir. 1985).

deepfakes. Deepfakes may cause private figures more harm than public figures because deepfakes may cause a private figure to lose their life as they know it, whether that is through work, academic pursuits, personal relationships, or business relationships. While public figures may face the same harms, the degree to which public figures face this harm may be less than the degree to which private figures face this harm because public figures may have the resources to be able to bounce back from deepfake defamation in ways that private figures may not.

Deepfakes may be covered by defamation law, but defamation law alone is not enough because defamation law is tiered in a way that benefits public figures more than it does private figures. Given that deepfakes can be used to damage anyone's reputation the law needs to reflect that by extending protection to all. Defamation law has a long historical development so achieving equal protection in defamation law to address deepfakes would take too long or possibly not happen. Defamation laws do not protect everyone equally, which is an opportunity for deepfake laws to get it right by including the protection of all individuals, both private and public figures. Thus, defamation harm caused by deepfakes is best addressed through deepfake legislation passed on a state by state basis.

#### IV. Deepfake Laws Landscape

In the U.S. there are currently a few pieces of deepfake legislation, in the form of federal and state legislation. Each piece of legislation uses a different method to address the issues deepfakes present today. Unfortunately, deepfake laws are few and limited in scope leaving individuals in many instances to rely on defamation laws. Further, this deepfake legislation while making some advances has also left room for improvement such as protecting private figures equal to public figures.

According to the Committee on Homeland Security and Governmental Affairs of the U.S. Senate:

“[N]ational security experts have identified a series of interim actions Congress ... can take to counter the proliferation of deepfake media content and its adverse effects. Proposed actions to counter deepfake technologies and content include: implementing legislation prohibiting public figures from disseminating deepfake content, ... encouraging social media companies to develop digital signatures, and launching public awareness campaigns about deepfake content.”<sup>99</sup>

Additionally, the Committee on Homeland Security and Governmental Affairs of the U.S. Senate believes “the Federal Government should proactively identify the tools and techniques used by our adversaries to develop deepfake technologies and content, and develop countermeasures and tools to identify and counter deepfake content.”<sup>100</sup>

Given the limitations of defamation law, the recognition by both state and federal governments that there are steps that can be taken now, and that there is a need to take action now, as evidenced by committee reports and legislation already passed demonstrates that effective government legislation is the quickest and the best path forward for handling the issues deepfakes create. One of the quickest paths would be amending current laws.

---

<sup>99</sup> U.S. Senate, *Deepfake Report Act of 2019*, Rep. No. 116-93, at 3 (2019), <https://www.congress.gov/116/crpt/srpt93/CRPT-116srpt93.pdf>.

<sup>100</sup> *Id.*

## A. Federal Legislation

### a. National Defense Authorization Act for Fiscal Year 2020 (NDAA)

On December 20, 2019, the U.S. Congress passed the National Defense Authorization Act (NDAA) for Fiscal Year 2020. Within the NDAA, Section 5709 called the ‘Report On Deepfake Technology, Foreign Weaponization Of Deepfakes, And Related Notifications’ addresses deepfakes.<sup>101</sup> Section 5709 of the Act requires that the DNI provide reports to the congressional intelligence committees about the potential national security impacts of deepfakes, and the actual or potential use of deepfakes by foreign governments to spread disinformation or engage in other malign activities.<sup>102</sup>

It is good that the federal government has passed legislation regarding deepfakes. It demonstrates that deepfakes are a true concern for the country and that the federal government is keeping up with technological advances and threats. It seems that the DNI is tasked with surveying the U.S. landscape to determine where, if anywhere, deepfakes are posing an issue to national security, and surveying the international landscape to determine whether foreign governments are using deepfakes for disinformation or to engage in other malign activities.

While it is great that the federal government has tasked the DNI with tracking the possible negative impacts of deepfakes for national security purposes, the current government directive ignores the issues deepfakes pose at the individual level.

There are three possible actions the federal government could take regarding deepfakes. One, the federal government could take more action on deepfakes affecting individuals. Two, the federal government could take no additional action on deepfakes affecting individuals. Third, given that defamation law is dealt with at the state level the federal government could develop a definition for deepfakes as a guidepost for state deepfake laws, and let states develop their own deepfake laws.

## B. State Legislation

A handful of states have passed deepfake laws, namely California, Texas, and Virginia.

### a. California

California has passed two laws to address deepfakes: AB 602, and AB 730. California had a third deepfake bill, AB 1280, but it was not enacted into law.

#### i. AB 602

On October 3, 2019, the Governor of California signed into law AB 602, also known as “Depiction of Individual Using Digital or Electronic Technology: Sexually Explicit Material:

---

<sup>101</sup> U.S. Congress, *S.1790 - National Defense Authorization Act for Fiscal Year 2020*, (2019-2020), <https://www.congress.gov/bill/116th-congress/senate-bill/1790/text>.

<sup>102</sup> *Id.*

Cause of Action.”<sup>103</sup> AB 602 provides that the depicted individual has a cause of action against a person who:

“[E]ither (1) creates and intentionally discloses sexually explicit material if the person knows or reasonably should have known the depicted individual did not consent to its creation or disclosure or (2) who intentionally discloses sexually explicit material that the person did not create if the person knows the depicted individual did not consent to its creation.”<sup>104</sup>

AB 602 is focused on punishing people who create deepfakes of an individual or individuals engaged in nonconsensual sexual activity.

Disclosure under AB 602 means to “publish, make available, or distribute to the public.”<sup>105</sup> Under AB 602, the person who created the deepfake must have intentionally disclosed the deepfake. Requiring that the perpetrator intentionally disclose the deepfake material is a high standard. If a perpetrator were to accidentally leak such a deepfake, this may not be deemed as “intentional disclosure” of the deepfake, and so the perpetrator may walk away from their harmful action without any legal or monetary punishment. Yet, the individual depicted in the leaked deepfake still has to face the repercussions of that leaked deepfake. While this law attempts to protect the people of California from all walks of life, it focuses only on the use of deepfakes for sexual purposes. This focus on sexual purposes is limiting because deepfakes can defame individuals in other areas of their lives, such as political speech or hate speech they otherwise did not state, or engaging in conduct that they otherwise did not engage in.

AB 602 is one of the more equitable deepfake laws because it aims to protect all individuals regardless of whether the individual is a public figure or a private figure. But AB 602 is too limited in its scope to handle the full breadth of issues deepfakes create.

## ii. AB 730

On October 3, 2019, California passed an amendment to AB 730’s Elections Law, which protects a candidate for elected office.<sup>106</sup> The amendment is known as ‘Deceptive Audio or Visual Media.’<sup>107</sup> This law covers “materially deceptive audio or visual media” such as:

“[A]n image or audio or video recording of a candidate’s appearance, speech, or conduct that has been intentionally manipulated in a manner such that the image or audio or video recording would falsely appear to a reasonable person to be authentic and would cause a reasonable person to have a fundamentally different

---

<sup>103</sup> *AB-602 Depiction of individual using digital or electronic technology: sexually explicit material: cause of action*, CAL. LEGIS. INFO. (2019-2020),

[https://leginfo.legislature.ca.gov/faces/billStatusClient.xhtml?bill\\_id=201920200AB602](https://leginfo.legislature.ca.gov/faces/billStatusClient.xhtml?bill_id=201920200AB602).

<sup>104</sup> *AB-1280 Crimes: deceptive recordings*, CAL. LEGIS. INFO. (2019-2020),

[https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill\\_id=201920200AB1280#:~:text=The%20bill%20would%20criminally%20prohibit,individual%20engaging%20in%20sexual%20conduct.&text=By%20creating%20a%20new%20crime,a%20state%2Dmandated%20local%20program](https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB1280#:~:text=The%20bill%20would%20criminally%20prohibit,individual%20engaging%20in%20sexual%20conduct.&text=By%20creating%20a%20new%20crime,a%20state%2Dmandated%20local%20program).

<sup>105</sup> *AB-602*.

<sup>106</sup> *AB-730 Elections: deceptive audio or visual media*, CAL. LEGIS. INFO. (2019-2020),

[https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill\\_id=201920200AB730](https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB730).

<sup>107</sup> *Id.*

understanding or impression of the expressive content of the image or audio or video recording than that person would have if the person were hearing or seeing the unaltered, original version of the image or audio or video recording.”<sup>108</sup>

A strength of AB 730 is that it provides a broader definition of deepfakes to include audio and visual manipulation. Yet, AB 730 only applies to candidates running for elected office. It does not include candidates for appointed office. It also does not include everyday people who deepfakes can attack or harm.

Given that AB 730 is under a larger umbrella of California’s election law, it seems that this deepfake law is meant to protect only elected political candidates. AB 730 has a more comprehensive definition of deepfakes than any other legislation, federal or state, thus far. AB 730’s definition of deepfakes should also be included in AB 602 and other deepfake statutes because AB 730’s comprehensive definition would help cover more of the methods in which deepfakes can harm an individual’s reputation.

### iii. AB 1280

On February 21, 2019, Assembly Member Grayson introduced AB 1280 to the California Legislature.<sup>109</sup> Unfortunately, this bill was not enacted into law.<sup>110</sup>

AB 1280 provides a cause of action, as well as legal and injunctive remedies, for three groups of people: individuals over 18 years old, individuals under the age of 18, and candidates for elected office.<sup>111</sup> AB 1280’s explicit inclusion of minors is helpful because of the dangers that minors face in today’s digital world. According to a study in the Journal for the American Medical Association, “about 1 in 7 (or 14.8%) of those between the ages of 12 and 17 had sent sexts and approximately 1 in 4 (27.4%) have received them.”<sup>112</sup> This increase in sexting amongst minors demonstrates that minors need protection. While sexting is different from the circulation of deepfakes, sexting can include deepfaked media facilitating the easy circulation of deepfakes in texts.

AB 1280 included the strong and broad definition of deepfakes found in AB 730 and recognized the First Amendment challenges faced by regulating deepfakes.<sup>113</sup> Another strength of AB 1280 is that it included the preparation of deepfakes as a criminal activity. In contrast, AB 602 and AB 730 only criminalized the disclosure of deepfakes, not the preparation of deepfakes. The inclusion of deepfakes preparation as a criminal activity gave AB 1280 greater breadth to help deter the creation of deepfakes.

---

<sup>108</sup> *Id.*

<sup>109</sup> *AB-1280 Crimes: deceptive recordings*, CAL. LEGIS. INFO. (2019-2020), [https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill\\_id=201920200AB1280#:~:text=AB%201280%2C%20as%20amended%2C%20Grayson,or%20soliciting%20goods%20or%20services.](https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB1280#:~:text=AB%201280%2C%20as%20amended%2C%20Grayson,or%20soliciting%20goods%20or%20services.)

<sup>110</sup> *AB-1280 Status*, CAL. LEGIS. INFO. (2019-2020), [https://leginfo.legislature.ca.gov/faces/billStatusClient.xhtml?bill\\_id=201920200AB1280.](https://leginfo.legislature.ca.gov/faces/billStatusClient.xhtml?bill_id=201920200AB1280.)

<sup>111</sup> *Id.*

<sup>112</sup> Bruce Y. Lee, *Here Is How Much Sexting Among Teens Has Increased*, JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION (Sept. 8, 2018 10:22 PM EDT), <https://www.forbes.com/sites/brucelee/2018/09/08/here-is-how-much-sexting-among-teens-has-increased/?sh=1c544afd36fl>.

<sup>113</sup> *AB-1280 Status*.

The punishment imposed under AB 1280 for the preparation, production, and disclosure of deepfakes is a one-year term in county jail. A one-year term in county jail can be a strong deterrent of deepfake creation and dissemination, but it may be harsh, which could explain why AB 1280 was not enacted as law. Another concern is that AB 1280 might criminalize comedic or satirical uses of deepfakes. One such example is “Sassy Justice with Fred Sassy,” a show sprinkled with deepfakes and made by the creators of South Park.<sup>114</sup> If the creators of South Park were subject to California law, they may be subject to the penalties in AB 1280. Yet, perhaps this is what AB 1280(d), the First Amendment section in AB 1280 was meant to address. AB 1280(d) states “No person shall be held liable under this section for any activity protected by the First Amendment to the Constitution of the United States.” Hence, the creators of “Sassy Justice with Fred Sassy” may not be subject to penalties under AB 1280 because their show may be protected as comedic and satirical under the First Amendment.

AB 1280 is more comprehensive than other California deepfake laws, but parts of AB 1280 prevented its passage. Perhaps, the bill did not pass because the penalty of one year in county jail was too harsh of a penalty. Thus, a possible amendment before reintroducing AB 1280 could be to either remove the penalty of one year in county jail or minimize the duration of time in county jail. Alternatively, different punishment methods could be imposed such as a flat fee fine in the range of thousands of dollars, monetary awards to deepfake victims, or a tiered level of punishment.

#### b. Texas

On April 18, 2019, Texas passed its deepfake law, known as SB 751.<sup>115</sup> The Texas legislature defines a deepfake as “an act relating to the creation of a criminal offense for fabricating a deceptive video with intent to influence the outcome of an election.”<sup>116</sup>

Fortunately, Texas has created a deepfake law that criminalizes deepfakes in the context of political harm. Unfortunately, SB 751 limits deepfakes to visual misrepresentations and it does not include audio misrepresentations. SB 751 only includes deepfake videos, but deepfakes can take several forms, including deepfake audio recordings, and deepfake pictures. SB 751 only offers protection to candidates running for elected office, instead of extending that protection to the general public.

SB 751 is limited in scope. It should broaden its definition of deepfakes, expand the setting in which deepfakes can be punished beyond that of political office, and be more inclusive as to the people it protects.

---

<sup>114</sup> Randall Colburn, *The Creators of South Park Launched an Entire Deepfake Studio to Make Sassy Justice*, AV CLUB (Nov. 30, 2020 10:10 AM), <https://news.avclub.com/the-creators-of-south-park-launched-an-entire-deepfake-1845529116>.

<sup>115</sup> *S.B. No. 751*, TEX. CAPITOL (Apr. 8, 2019), <https://capitol.texas.gov/tlodocs/86R/billtext/html/SB00751S.htm>.

<sup>116</sup> *S.B. No. 751*, TEX. CAPITOL (May 22, 2019), <https://capitol.texas.gov/tlodocs/86R/billtext/html/SB00751F.htm>.

### c. Virginia

In 2019, the Virginia legislature passed Virginia Code Annotated § 18.2-386.2, also known as ‘Unlawful Dissemination or Sale of Images of Another; Penalty.’<sup>117</sup> This Virginia law criminalizes the act of an individual creating a deepfake video or still of a person that depicts a person in the nude or a state of undress to expose private parts of the body.<sup>118</sup> Virginia’s deepfake statute defines deepfakes only in the context of revenge porn.

A plus of Virginia’s law is that it includes video and still images of an individual and that the term “individual” can be broadly defined to include public and private figures. Yet, it only criminalizes deepfakes created for revenge porn. So, its scope is limited. While it is commendable that Virginia has created a cause of action and remedy for individuals who are victims of deepfake revenge porn Virginia’s law only penalizes deepfakes in a narrow context. Virginia’s law is missing the inclusion of deepfakes created outside of revenge porn purposes.

Like the recommendations provided for Texas’ deepfake statute, Virginia’s deepfake laws should include scenarios in addition to revenge porn, such as political speech, and hate speech.

### C. Recommendations

Based on the earlier mentioned suggestions, it is recommended that the federal government develop a federal deepfake definition that states could use to create their own deepfake laws. Even if no federal deepfake definition is ultimately provided, states should seek to create their own deepfake laws and define deepfakes.

#### a. Federal Deepfake Definition

There is no uniform definition of deepfakes in either the legislative branch or judicial branch. So, it may be more difficult to pinpoint what is or what is not a deepfake, making the regulation of deepfakes trickier. A lack of a deepfake definition also makes the judicial branch’s job a bit more difficult.<sup>119</sup> Without a set definition of what a deepfake is, it may be difficult for people to make a claim that would provide standing in court.<sup>120</sup> Without a set definition of deepfakes, it also makes it difficult for parties filing suit to explain the claim or harm they are seeking redress for in court.<sup>121</sup> Not only does a lack of definition of deepfakes make it difficult for the court to address deepfake issues in court, but it also makes it difficult for plaintiffs to describe the harm caused to them by deepfakes. Being unable to describe harm, makes it difficult for a court to recognize the claim as one for which a legal remedy can be provided. Hence, at the federal level, it would be helpful if the legislative branch or judicial branch defined deepfakes.

---

<sup>117</sup> § 18.2-386.2. *Unlawful dissemination or sale of images of another; penalty*, CODE OF VA. <https://law.lis.virginia.gov/vacode/title18.2/chapter8/section18.2-386.2/> (last visited Nov. 15, 2020).

<sup>118</sup> *Id.*

<sup>119</sup> *Fariss-Borello v. City of Scottsdale*, 2019 U.S. Dist. LEXIS 155013.

<sup>120</sup> See *Fariss-Borello*, where the court dismissed a second amended complaint because the complaint was difficult to decipher.

<sup>121</sup> See *Fariss-Borello*, where via a second amended complaint the plaintiff was attempting to describe a video or audio that was edited by someone who intended to harm the plaintiff.

Given that defamation law is left for states to develop, the federal government or federal courts can provide a definition of deepfakes that will guide states on modifying their existing deepfake laws or develop their deepfake laws. The best definition of deepfakes from deepfake legislation is from California’s AB 730. Building from California’s AB 730 definition of deepfakes, the federal government can modify the definition as follows to guide the definition in other deepfake statutes:

“[A deepfake is an image or audio or visual manipulation resulting in] an image or audio [recording] or video recording of [an individual’s] appearance, speech, or conduct that has been created or altered in a manner [that it] would falsely appear to a reasonable observer to be [an] authentic [image, speech or conduct of the individual depicted].”<sup>122</sup>

#### b. Five Takeaways for State Deepfake Laws

States considering passing deepfake laws can look towards California, Texas, and Virginia’s deepfake laws for guidance. While the above analysis is specific to the states who have passed deepfake laws, the analysis can prove valuable for states who are considering passing deepfake laws. There are five general takeaways for states who are considering amending their current deepfake laws or creating their own deepfake laws. First, rely on the federal government’s definition of deepfakes, but if a federal definition does not exist, states should create a broad enough definition of deepfakes so that it includes audio and visual manipulation of images. A broad enough definition is necessary to encapsulate as much deepfake misconduct as possible. Second, broaden the scope of liability to include both sexual and non-sexual settings. This is because deepfakes can harm individuals in different aspects of their lives. Third, determine the appropriate standard of liability for the creation of deepfakes (e.g. intent, malicious intent, negligence, or some other level of intent). An inappropriate standard of liability may make passage of a deepfake bill difficult. Fourth, create reasonable punishments for people who do not follow deepfake laws, in other words, the punishment must fit the crime. Fifth, be inclusive by extending deepfake laws to protect everyone (e.g. minors, private figures) not just public figures (e.g. celebrities, and political candidates). States who are interested in creating deepfake laws to protect the people within their jurisdiction should ensure that all people are protected in their deepfake statutes because deepfakes can negatively affect both private figures and public figures.

### CONCLUSION

The Supreme Court has not provided input as to defamation resulting from the Internet.<sup>123</sup> Given the digital age in which we live, where access to audio and visual manipulation is at our fingertips, deepfakes have been brushed aside as just another technological advance, rather than recognized for the harms deepfakes have caused and may cause.<sup>124</sup> It is

---

<sup>122</sup> AB-730.

<sup>123</sup> Shlomit Yanisky-Ravid and Ben Zion Lahav, *Public Interest Vs. Private Lives—Affording Public Figures Privacy In The Digital Era: The Three Principle Filtering Model*, 19 U. of Pa. 976, 1011 (May 2017), <https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=1633&context=jcl>.

<sup>124</sup> *Id.*

important to view deepfakes as a concern for national security, democracy, and defamation instead of diminishing the harms deepfakes can cause. It is also necessary that the law adapts to the digital age we live in. New technology comes along with time, and the law must adapt to the new technology. Deepfakes are just another example of new technology. Now that deepfakes are a way to defame someone, the law should provide redress to all, public figures and private figures alike, who are victims of defamation due to deepfakes. Defamation law is currently insufficient and not targeted enough to address deepfake concerns at the individual level.

Regulation of deepfakes as they affect individuals is important because of the negative consequences that could arise as a result of a deepfake being made of an individual, whether it is defamation as a result of hate speech, political speech, or conduct. There are a wide variety of harms at the individual level including, but not limited to the loss of personal and business relationships, negative effects on academic and professional endeavors, and humiliation.

Federal law has charged the DNI with providing a report on deepfake usage at the national and international levels. In addition to that intel, the federal government can further assist in the fight against deepfakes negatively affecting individuals by providing states with a definition of deepfakes.<sup>125</sup> If guidance as to a definition of deepfakes is not provided by the federal government, states should look towards California's deepfake laws for a more comprehensive definition of deepfakes. Alternatively, states can develop their own definitions of deepfakes.

More importantly, because defamation law disparately protects public figures and private figures, the best approach to address deepfake harms is by states creating deepfake laws. States should make sure their deepfake laws are inclusive of all individuals and protect private figures and public figures equally.

---

<sup>125</sup> James Vincent, *Why we need a better definition of 'deepfake'*, THE VERGE (May 22, 2018, 2:53 PM EDT), <https://www.theverge.com/2018/5/22/17380306/deepfake-definition-ai-manipulation-fake-news>.

## BIBLIOGRAPHY

1. § 18.2-386.2. *Unlawful dissemination or sale of images of another; penalty*, CODE OF VA. <https://law.lis.virginia.gov/vacode/title18.2/chapter8/section18.2-386.2/> (last visited Nov. 15, 2020).
2. *AB-602 Depiction of individual using digital or electronic technology: sexually explicit material: cause of action*, CAL. LEGIS. INFO. (2019-2020), [https://leginfo.legislature.ca.gov/faces/billStatusClient.xhtml?bill\\_id=201920200AB602](https://leginfo.legislature.ca.gov/faces/billStatusClient.xhtml?bill_id=201920200AB602).
3. *AB-730 Elections: deceptive audio or visual media*, CAL. LEGIS. INFO. (2019-2020), [https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill\\_id=201920200AB730](https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB730).
4. *AB-1280 Crimes: deceptive recordings*, CAL. LEGIS. INFO. (2019-2020), [https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill\\_id=201920200AB1280#:~:text=AB%201280%2C%20as%20amended%2C%20Grayson,or%20soliciting%20goods%20or%20services](https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB1280#:~:text=AB%201280%2C%20as%20amended%2C%20Grayson,or%20soliciting%20goods%20or%20services).
5. *AB-1280 Crimes: deceptive recordings*, CAL. LEGIS. INFO. (2019-2020), [https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill\\_id=201920200AB1280#:~:text=The%20bill%20would%20criminally%20prohibit,individual%20engaging%20in%20sexual%20conduct.&text=By%20creating%20a%20new%20crime,a%20state%2Dmandated%20local%20program](https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB1280#:~:text=The%20bill%20would%20criminally%20prohibit,individual%20engaging%20in%20sexual%20conduct.&text=By%20creating%20a%20new%20crime,a%20state%2Dmandated%20local%20program).
6. *AB-1280 Status*, CAL. LEGIS. INFO. (2019-2020), [https://leginfo.legislature.ca.gov/faces/billStatusClient.xhtml?bill\\_id=201920200AB1280](https://leginfo.legislature.ca.gov/faces/billStatusClient.xhtml?bill_id=201920200AB1280).
7. *Another Fake Video of Pelosi Goes Viral*, WASHINGTON POST (Aug. 3, 2020), <https://www.washingtonpost.com/technology/2020/08/03/nancy-pelosi-fake-video-facebook/>.
8. Bruce Y. Lee, *Here Is How Much Sexting Among Teens Has Increased*, JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION (Sept. 8. 2018 10:22 PM EDT), <https://www.forbes.com/sites/brucelee/2018/09/08/here-is-how-much-sexting-among-teens-has-increased/?sh=1c544afd36f1>.
9. *Dameron v. Wash. Magazine*, 250 U.S. App. D.C. 346, 779 F.2d 736 (1985).
10. Dave Johnson, *Audio Deepfakes: Can Anyone Tell If They're Fake?*, HOW TO GEEK (Aug. 3, 2020, 6:40 AM EDT), <https://www.howtogeek.com/682865/audio-deepfakes-can-anyone-tell-if-they-are-fake/>.
11. *Defamation*, CORNELL LEGAL INFORMATION INSTITUTE, <https://www.law.cornell.edu/constitution-conan/amendment-1/defamation> (last visited Nov. 27, 2020).

12. *Defamation Law - Guide to Libel and Slander Law*, HG LEGAL RESOURCES, <https://www.hg.org/defamation.html#:~:text=Libel%20is%20defamatory%20statements%20and,a%20fact%2C%20not%20an%20opinion> (last visited Nov. 30, 2020).
13. *Doe v. Pittsylvania Cty.*, 844 F. Supp. 2d 724 (W.D. Va. 2012).
14. Eileen Carroll Prager, *Public Figures, Private Figures and Public Interest*, 30 Stan. L. Rev. 157–189, 165 (1977), [www.jstor.org/stable/1228177](http://www.jstor.org/stable/1228177) (last visited Nov. 30, 2020).
15. *Fariss-Borello v. City of Scottsdale*, 2019 U.S. Dist. LEXIS 155013.
16. Gary Bugh, *Public Figures and Officials*, THE FIRST AMENDMENT ENCYCLOPEDIA (2009), <https://www.mtsu.edu/first-amendment/article/1010/public-figures-and-officials>.
17. *Gertz v. Robert Welch*, 418 U.S. 323, 351 (1974).
18. Google, *Generative Adversarial Networks Introduction*, (Oct. 8, 2019), <https://developers.google.com/machine-learning/gan>.
19. James Vincent, *Why we need a better definition of 'deepfake'*, THE VERGE (May 22, 2018, 2:53 PM EDT), <https://www.theverge.com/2018/5/22/17380306/deepfake-definition-ai-manipulation-fake-news>.
20. Jessica Ice, *Defamatory Political Deepfakes and the First Amendment*, 70 Case W. Res. L. Review 417, 419 (2019), <https://scholarlycommons.law.case.edu/cgi/viewcontent.cgi?article=4854&context=caselrev>.
21. Joseph Foley, *10 Deepfake Examples that Terrified and Amused the Internet*, CREATIVE BLOQ (Mar. 23, 2020), <https://www.creativebloq.com/features/deepfake-examples>.
22. Lata Nott, *A Quick Guide to Libel Law*, FREEDOM FORUM INSTITUTE, <https://www.freedomforuminstitute.org/first-amendment-center/primers/libellaw/> (last visited Dec. 5, 2020).
23. *Libel vs. Slander: Different Types of Defamation*, NOLO, <https://www.nolo.com/legal-encyclopedia/libel-vs-slander-different-types-defamation.html> (last visited Nov. 30, 2020).
24. Lourdes Vazquez, *Deepfakes And Fake News: Is Regulation of Deepfakes Possible?*, IP BRIEF AT AMERICAN U. WASHINGTON C. OF L. (Jan. 30, 2020), <http://www.ipbrief.net/2020/01/30/deepfakes-and-fake-news-is-regulation-of-deepfakes-possible/>.

25. Luke Dormehl, *These Amazing Audio Deepfakes Showcase Progress of A.I. Speech Synthesis*, DIGITAL TRENDS (June 6, 2020), <https://www.digitaltrends.com/news/best-audio-deepfakes-web/>.
26. *Marcone v. Penthouse Int'l Magazine for Men*, 754 F.2d 1072 (3d Cir. 1985).
27. *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 84 S. Ct. 710 (1964).
28. *Proving Fault: Actual Malice and Negligence*, DIGITAL MEDIA LAW PROJECT, <https://www.dmlp.org/legal-guide/proving-fault-actual-malice-and-negligence#:~:text=A%20limited%2Dpurpose%20public%20figure%20is%20either%3A,reached%20an%20all%2Dencompassing%20level> (last visited Nov. 27, 2020).
29. *Public Versus Private Figures*, NATIONAL CONSUMER RIGHTS, <https://www.nationwideconsumerrights.com/public-vs-private-figures.html> (last visited Nov. 30, 2020).
30. Randall Colburn, *The Creators of South Park Launched an Entire Deepfake Studio to Make Sassy Justice*, AV CLUB (Nov. 30, 2020 10:10 AM), <https://news.avclub.com/the-creators-of-south-park-launched-an-entire-deepfake-1845529116>.
31. Richard L. Hasen, *Deep Fakes, Bots, and Siloed Justices: American Election Law in a Post-Truth World* (July 10, 2019), St. Louis University L. J., 2019, Forthcoming, U.C. Irvine School of L. Research Paper No. 2019-36, Available at SSRN: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3418427](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3418427).
32. *Rosenbloom v. Metromedia*, 403 U.S. 29, 48 (1971).
33. Sally Adee, *What Are Deepfakes and How Are They Created?*, TECH TALK BLOG (April 29, 2020 16:00 GMT), <https://spectrum.ieee.org/tech-talk/computing/software/what-are-deepfakes-how-are-they-created>.
34. *S.B. No. 751*, TEX. CAPITOL (May 22, 2019), <https://capitol.texas.gov/tlodocs/86R/billtext/html/SB00751F.htm>.
35. *S.B. No. 751*, TEX. CAPITOL (Apr. 8, 2019), <https://capitol.texas.gov/tlodocs/86R/billtext/html/SB00751S.htm>.
36. Shlomit Yanisky-Ravid and Ben Zion Lahav, *Public Interest Vs. Private Lives—Affording Public Figures Privacy In The Digital Era: The Three Principle Filtering Model*, 19 U. of Pa. 976, 1011 (May 2017), <https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=1633&context=jcl>.
37. *Trivedi, LLC v. Lang*, No. A16-1209, 2017 Minn. App. Unpub. LEXIS 389, at 1 (May 1, 2017).
38. U.S. Congress, *S.1790 - National Defense Authorization Act for Fiscal Year 2020*, (2019-2020), <https://www.congress.gov/bill/116th-congress/senate-bill/1790/text>.

39. U.S. Const. amend. I.
40. U.S. Senate, *Deepfake Report Act of 2019*, Rep. No. 116-93, at 3 (2019), <https://www.congress.gov/116/crpt/srpt93/CRPT-116srpt93.pdf>.
41. *Wells v. Liddy*, 186 F.3d 505 (4th Cir. 1999).
42. Wilmer Hale, *First Federal Legislation on Deepfakes Signed Into Law*, JD SUPRA (Dec. 24, 2019), <https://www.jdsupra.com/legalnews/first-federal-legislation-on-deepfakes-42346/>.