Fordham Law Review

Volume 88 | Issue 3 Article 2

2019

Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn's Next Tragic Act

Rebecca A. Delfino
Loyola Law School

Recommended Citation

Rebecca A. Delfino, *Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn's Next Tragic Act*, 88 Fordham L. Rev. 887 (2019).
Available at: https://ir.lawnet.fordham.edu/flr/vol88/iss3/2

PORNOGRAPHIC DEEPFAKES: THE CASE FOR FEDERAL CRIMINALIZATION OF REVENGE PORN'S NEXT TRAGIC ACT

*Rebecca A. Delfino**

This could happen to you.

Like millions of people worldwide, you have uploaded digital photographs of yourself to the internet through social media platforms. Your pictures aren't sexually explicit or revealing—they depict your daily life, spending time with friends or taking "selfies" on vacation. But then someone decides they don't like you. Using an app available on any smartphone, this antagonist clips digital images of your face from your innocuous pictures and pastes them seamlessly onto the body of a person engaged in sexually explicit acts. Without your knowledge or consent, you become the "star" of a realistic, pornographic "deepfake."

This hypothetical reflects an emerging phenomenon in sex exploitation cybercrimes—it is the next tragic act in the unauthorized public dissemination of private, sexually explicit photos or videos known as "revenge porn." Is there anything you can do if someone inserts you into a pornographic deepfake image or video against your will? Is it against the law to create, share, and spread falsified pornography on the internet?

At best, the answer to these questions is complicated and uncertain. At worst, the answer is no. Although criminalizing bad acts is the most effective deterrent against bad actors, no federal or state laws currently criminalize the creation or distribution of pornographic deepfakes. And since deepfakes exist in cyberspace, they are not confined to an individual state jurisdiction. This Article is the first to focus on the intersection of the law and pornographic deepfakes and to propose a legislative solution to the harms they unleash. Ultimately, this Article proposes a national response rooted in federal criminal law because everyone, everywhere is a potential deepfake victim—even you.

* Clinical Professor of Law at Loyola Law School, Los Angeles. I am grateful and indebted to Katherine Brosamle, Jennifer Roges, and Kathleen Simers for their research assistance, Loyola Law School for its generous support of faculty scholarship and, as always, my family for their patience, indulgence, and encouragement. This Article is dedicated to my parents, Ken and Rosemary Delfino, who not only taught me to discern fact from fiction at a young age but also inspired me to care about the difference.

INTRODUCTION

> Disguise, I see, thou art a wickedness,
> Wherein the pregnant enemy does much.
>
> —William Shakespeare[1]

We are inundated daily, in both real and cyber spaces, with a barrage of truthful and fake information, news, images, and videos. The law has not kept pace with the problems that result when we cannot discern fact from fiction. Something new has emerged from the dark corners of the internet in recent years: doctored pornographic images and videos featuring one person's face believably mapped onto a body engaged in sexually explicit

[1] TWELFTH NIGHT act 2, sc. 2.

acts, which creators,[2] news commentators,[3] and academics[4] call "deepfakes."[5] Although the images are fictitious, having been created using new artificial intelligence (AI)–assisted technology, they are offered as true images of the individuals—usually women—depicted. Deepfake technology has evolved so quickly that an app designed to create deepfakes is now widely available.[6] The app lowers the technical threshold required to create such images and videos, which will likely make face-swapped, fake porn of public figures and private individuals more prevalent. Now anyone who has appeared in a digital image may "star" in pornography against their will, and currently, the law provides no clear or direct recourse to stop it.

[2] *See* Samantha Cole, *We Are Truly Fucked: Everyone Is Making AI-Generated Fake Porn Now*, VICE: MOTHERBOARD (Jan. 24, 2018, 10:13 AM), https://motherboard.vice.com/en_us/article/bjye8a/reddit-fake-porn-app-daisy-ridley [https://perma.cc/EN85-JSM7] (describing how, in late fall 2017, an anonymous Reddit user posted several porn videos on the internet under the pseudonym "Deepfakes," including a video of actress Daisy Ridley's face superimposed on the body of a porn actress); *see also* Kristen Dold, *Face-Swapping Porn: How a Creepy Internet Trend Could Threaten Democracy*, ROLLING STONE (Apr. 17, 2018, 8:47 PM), https://www.rollingstone.com/culture/features/face-swapping-porn-how-creepy-trend-could-threaten-democracy-w518929 [https://perma.cc/7J2Q-J2QD].

[3] *See, e.g.*, Oscar Schwartz, *You Thought Fake News Was Bad?: Deep Fakes Are Where Truth Goes to Die*, GUARDIAN (Nov. 12, 2018, 5:00 AM), https://www.theguardian.com/technology/2018/nov/12/deep-fakes-fake-news-truth [https://perma.cc/2KBB-FDWH]; *see also All Things Considered: Technologies to Create Fake Audio and Video Are Quickly Evolving*, NPR (Apr. 2, 2018, 4:20 PM), https://www.npr.org/2018/04/02/598916380/technologies-to-create-fake-audio- and-video-are-quickly-evolving [https://perma.cc/V5TM-AUVJ]; BuzzFeed (@BuzzFeed), TWITTER (Apr. 17, 2018, 8:00 AM), https://twitter.com/BuzzFeed/status/986257991799222272 [https://perma.cc/6A5J-RBEG].

[4] *See* Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. (forthcoming 2019), http://dx.doi.org/10.2139/ssrn.3213954 [https://perma.cc/FQG5-RG3Y]. *See generally* Marc Jonathan Blitz, *Lies, Line Drawing and (Deep) Fake News*, 71 OKLA. L. REV. 59 (2018).

[5] Deepfakes take their name from the Reddit user who first introduced sexually explicit doctored videos to the internet. Dold, *supra* note 2. The term is a combination of two ideas: "deep," from the concept of "deep learning," an aspect of artificial intelligence (AI) concerned with emulating the learning approach that human beings use to gain certain types of knowledge; and "fake," from the concept of "fake news," an inaccurate or sensationalistic report created to gain attention, mislead, deceive, or damage a reputation. *See* James Vincent, *Why We Need a Better Definition of 'Deepfake,'* VERGE (May 22, 2018, 2:53 PM), https://www.theverge.com/2018/5/22/17380306/deepfake-definition-ai-manipulation-fake-news [https://perma.cc/Z2VE-56UV]; *see also* Margaret Rouse, *Deepfake (Deep Fake AI)*, WHATIS.COM, https://whatis.techtarget.com/definition/deepfake [https://perma.cc/8UEZ-4NNP] (last visited Nov. 12, 2019). Although the term's origin is clear, its stylization remains an unsettled matter: it alternately appears as one word ("deepfake"), a hyphenated word ("deep-fake"), or two words ("deep fake"). Legal scholars and politicians appear to use the two-word version. *See* Malicious Deep Fake Prohibition Act of 2018, S. 3805, 115th Cong. (2018); Blitz, *supra* note 4; Chesney & Citron, *supra* note 4. The Merriam-Webster blog and at least one online tech dictionary, conversely, present it as one word: "deepfake." *See Words We're Watching: 'Deepfake,'* MERRIAM-WEBSTER (July 31, 2019), https://www.merriam-webster.com/words-at-play/deepfake-slang-definition-examples [https://perma.cc/W8FG-75QU]; *see also* Rouse, *supra*. This Article will use the one-word version because the term is a portmanteau, which is a commonplace construction used to describe technology (e.g., "Pinterest," "sexting," "webinar").

[6] *See* Cole, *supra* note 2.

As deepfake technology becomes more refined and easier to create, the law's inadequacies to protect potential victims have become clear. Consider the following hypothetical: Othello and Desdemona are in a long-term relationship. Othello is a digital photography buff, and everywhere he and Desdemona go—strolling the streets of Venice, sightseeing at the seaport in Cyprus, or picnicking in the garden of his castle—Othello takes photos of Desdemona with his smartphone. Desdemona consents to these images but has never allowed Othello to take nude or sexually explicit photographs of her. After Othello becomes convinced that Desdemona is having an affair with his friend Cassio, Othello and Desdemona break up. In retaliation, Othello uploads all of his digital images of Desdemona onto an app on his smartphone that allows him to superimpose Desdemona's face onto a video of a porn actress's body engaged in pornographic acts. Othello then uploads the deepfake video, along with Desdemona's contact information, to various social media sites online.

When Desdemona contacts law enforcement to have the pornographic deepfake video removed and Othello arrested, they inform her that Othello's conduct is not illegal—he has broken no law. When she asks whether Othello might be charged under the state's new revenge porn statute or another cybercrime statute, they demur and question whether she is a victim of any crime at all, since she did not own the photos and Othello used only her face in the video. Unfortunately, Desdemona's story reflects a new and emerging phenomenon in sex exploitation cybercrimes—it is the next tragic act in the unauthorized public dissemination of private, sexually explicit photos or videos, known as "revenge porn."

Part I of this Article defines and explains the phenomenon of deepfakes and places them in the broader context of the internet. This Part first considers the potential dangers and risks that deepfakes pose to our society and democratic institutions through their connection to fake news and false images involving public figures and officials and then focuses on deepfakes in the context of nonconsensual pornography and the related phenomenon of revenge porn.

Part II explores the complexity of the legal and prudential issues implicated by deepfakes. For example, determining who has actually been victimized by a deepfake can be difficult because each video depicts at least two people—the person whose body is truthfully being represented and the person whose face has been added. Is the proper victim the person whose face is used, the person whose body is used, or both? This Part also discusses the difficulty of identifying, locating, and bringing perpetrators to justice and explores the challenges of removing deepfakes once they have been published on the internet in light of the Communications Decency Act of 1996,[7] which shields websites and content distributors from liability for third-party content.

[7] Pub. L. No. 104-104, tit. V, 110 Stat. 56 (codified as amended in scattered sections of 15, 18, and 47 U.S.C.).

Part III begins by examining available criminal legal remedies. Although prosecutors might bring charges of cyberstalking, criminal threats, unauthorized access to digital files, and child pornography, current federal criminal law does not expressly outlaw deepfakes. This Part also discusses pending federal legislation, including the Malicious Deep Fake Prohibition Act of 2018[8] and the Ending Nonconsensual Online User Graphic Harassment Act of 2017.[9] Part III discusses the criminal laws of four states—California, Texas, Florida, and New York—to demonstrate that even the states with the highest populations, the greatest number of reported cybercrimes, and laws addressing revenge porn have failed to keep pace with harms created by deepfake technology.

Part IV examines the shortcomings and limitations of both existing criminal law and proposed federal legislation, arguing that neither can fully remedy the harms created by nonconsensual deepfake pornography. This Part also considers First Amendment and censorship concerns.

Finally, Part V addresses how to solve the problem of nonconsensual pornographic deepfakes through the enactment of federal criminal law. This Part proposes a legal solution combining existing elements of proposed federal legislation criminalizing revenge porn and deepfakes with new elements, including criminal injunctions and victim restitution. Finally, this Part discusses extralegal solutions that should be simultaneously developed and deployed. These solutions include developing effective technological content verification and improving the public's digital literacy so that each person consuming online content does so with the same concern: "internet content might be real or fake, and I have to figure that out before I make decisions or take actions."

I. THE DEEPFAKE PHENOMENON

A dangerous new technology has emerged on the internet that blurs fact and fiction by allowing users to create deepfakes—doctored images and videos that convincingly map one person's likeness onto another person's body.[10] Although this AI-assisted technology has valid and beneficial uses,[11] it can also be misused,[12] specifically to generate nonconsensual

[8] S. 3805.

[9] S. 2162, 115th Cong. (2017).

[10] Dold, *supra* note 2.

[11] *See* Chesney & Citron, *supra* note 4 (manuscript at 14–16) (discussing the potential "prosocial" applications in education, artistic expression, and self-expression for deepfakes); *see also Rise of the Deepfakes*, WEEK (June 9, 2018), http://theweek.com/articles/777592/rise-deepfakes [https://perma.cc/2P9A-DRWF] ("The technology has also been used to create harmless spoof and parody videos—inserting Reddit cult figure Nicholas Cage into films in which he didn't appear, for example.").

[12] Chesney and Citron have identified a litany of potential exploitative uses for deepfake technology, from harm to individuals (including sabotage, blackmail, and exploitation) to societal harms (including harm to democratic institutions, civil discourse, public safety, and national security). *See* Chesney & Citron, *supra* note 4 (manuscript at 16–29).

pornography.[13] Most deepfake or "face-swapping porn" superimposes images of female celebrities onto the bodies of individuals engaged in sexual acts.[14] Millions of individuals who have posted digital images on the internet are vulnerable to becoming unwitting stars of face-swapped porn via an application that easily generates deepfakes.[15]

A. Development of "Deepfake" Technology

Deepfakes first emerged on the internet in late 2017 when an anonymous Reddit[16] user uploaded realistic pornographic videos featuring celebrities.[17] Following the release of these original face-swapped porn videos, another anonymous Reddit user created and released "FakeApp," a free application enabling users to easily create deepfakes.[18] Before FakeApp's development, the production of realistic doctored videos was an expensive and arduous process confined to Hollywood movie studios.[19] FakeApp's creator achieved the goal of "mak[ing] deepfakes' technology available to people without a technical background or programming experience."[20]

FakeApp works by uploading digital images into a "machine-learning algorithm that's trained itself to stitch one face on top of another,"[21] also

[13] Adam Dodge & Erica Johnstone, *Using Fake Video Technology to Perpetuate Intimate Partner Abuse*, WITHOUT MY CONSENT 4, https://withoutmyconsent.org/perch/resources/2018-04-25deepfakedomesticviolenceadvisory.pdf [https://perma.cc/Q6DS-8G7S] (last visited Nov. 12, 2019); Janko Roettgers, *'Deep Fakes' Will Create Hollywood's Next Sex Tape Scare*, VARIETY (Feb. 2, 2018, 1:04 PM), http://variety.com/2018/digital/news/hollywood-sex-tapes-deepfakes-ai-1202685655/ [https://perma.cc/Q4Q7-SCM5].

[14] *See* Dold, *supra* note 2 (listing Emma Watson, Gal Gadot, Taylor Swift, and Daisy Ridley as some of the targets of deepfakes); David Lee, *Deepfakes Porn Has Serious Consequences*, BBC NEWS (Feb. 3, 2018), https://www.bbc.com/news/technology-42912529 [https://perma.cc/R4TY-LJQK] (discussing how deepfakes have targeted celebrities).

[15] *See generally* Dodge & Johnstone, *supra* note 13 (explaining the likelihood that domestic abusers and cyberstalkers will use deepfakes to harm victims).

[16] Reddit is among the top ten most popular websites in the United States and the top twenty in the world and has branded itself as the "front page of the internet." *See* Jake Widman & Will Nicol, *What Is Reddit?: A Beginner's Guide to the Front Page of the Internet*, DIGITAL TRENDS (May 22, 2019, 6:58 AM), https://www.digitaltrends.com/social-media/what-is-reddit/ [https://perma.cc/6ATF-U9FE]. It is a collection of web forums, also known as "subreddits," where communities of users share news and content or comment on posts about specific topics ranging from the straightforward to the unconventional. *Id.* The name of a subreddit colloquially begins with "r/," mimicking Reddit's URL structure. On r/shakespeare, for example, users discuss William Shakespeare and his work, while on r/birdswitharms, they post pictures of birds with arms. *Id.*

[17] *Rise of the Deepfakes*, *supra* note 11.

[18] Derek Hawkins, *Reddit Bans 'Deepfakes,' Pornography Using the Faces of Celebrities Such as Taylor Swift and Gal Gadot*, WASH. POST (Feb. 8, 2018), https://www.washingtonpost.com/news/morning-mix/wp/2018/02/08/reddit-bans-deepfakes-pornography-using-the-faces-of-celebrities-like-taylor-swift-and-gal-gadot/ [https://perma.cc/2TPE-AYBG].

[19] *See* Kevin Roose, *Here Come the Fake Videos, Too*, N.Y. TIMES (Mar. 4, 2018), https://www.nytimes.com/2018/03/04/technology/fake-videos-deepfakes.html [https://perma.cc/6QA8-78NL]; *see also* Hawkins, *supra* note 18 (explaining that FakeApp "put deepfake technology into a user-friendly package"); Lee, *supra* note 14.

[20] Cole, *supra* note 2.

[21] Dold, *supra* note 2.

known as AI deep learning.[22] This AI-assisted technology analyzes and manipulates images of a person's face and then maps it onto a different person's body in a video.[23] Creating a "face-swapped" video is an easy five-step process.[24]

Further, FakeApp can extract images from Google, Instagram, and Facebook to create face-swapped porn, hence the prevalence of celebrity deepfakes.[25]

B. The Proliferation of Deepfakes

Although digital impersonation technology has most notably been used to produce realistic pornographic videos, it may also be employed to create fake news and false images involving politicians and governmental actors.[26] This AI-assisted technology could be used to create fake videos of politicians accepting bribes, soldiers committing war crimes, presidential candidates engaging in criminal behavior, and emergency officials announcing an impending terrorist attack.[27] All of these examples have the potential to cause harm to our democracy, particularly because the technology used to generate these videos is rapidly advancing.[28] For instance, researchers from the University of Washington generated a photorealistic, lip-synced video of President Barack Obama "by mapping audio from one speech onto an existing video of him talking."[29] The researchers could successfully generate this realistic video because there was "significant video footage available online, in the form of interviews, speeches, newscasts, etc."[30] Additionally, sound engineers used AI technology to "synthesiz[e] 116,777 voice samples—each 0.4 seconds long—from 831 of [President John F.

[22] Dodge & Johnstone, *supra* note 13, at 1.

[23] *Id.*

[24] The process involves the following steps: "1. Download a program like FakeApp. 2. Ensure your computer has a reasonably powerful graphics card . . . . 3. Identify the video to be used. 4. Collect, scrape or take hundreds of photographs . . . to be superimposed into the video. 5. Feed the photos into FakeApp and run the program." *Id.* at 5.

[25] *Id.* at 6; Sarah Rense, *What Are 'Deepfakes,' and Why Are Pornhub and Reddit Banning Them?*, ESQUIRE (Feb. 12, 2018), https://www.esquire.com/lifestyle/sex/a17043863/what-are-deepfakes-celebrity-porn/ [https://perma.cc/3QQS-XUCF].

[26] *See* Dold, *supra* note 2 ("What's most concerning about deepfake tech is that seedy porn is just a preview of what's to come elsewhere."); Lee, *supra* note 14 (claiming that deepfake technology "could down the road be used maliciously to hoax governments and populations, or cause international conflict"); *see also* Chesney & Citron, *supra* note 4 (manuscript at 16–29). A Google search of a person's name can render many high-quality images associated with their identity, which would be ideal for digital misappropriation.

[27] *See* Chesney & Citron, *supra* note 4 (manuscript at 16–29); *see also* John Donovan, *Deepfake Videos Are Getting Scary Good*, HOWSTUFFWORKS (Sept. 5, 2018), https://electronics.howstuffworks.com/future-tech/deepfake-videos-scary-good.htm [https://perma.cc/D6FA-HJA9].

[28] *See Rise of the Deepfakes*, *supra* note 11 (noting that "deep learning" technology used to power deepfakes is improving fast).

[29] *Id.*

[30] Supasorn Suwajanakorn et al., *Synthesizing Obama: Learning Lip Sync from Audio*, ACM TRANSACTIONS ON GRAPHICS, July 2017, at 95:1, 95:1.

Kennedy's] speeches and radio addresses"[31] to create, in his own voice, the speech he would have delivered on November 22, 1963, had he not been assassinated earlier that day. Furthermore, filmmaker Jordan Peele used AI-assisted technology to create a doctored public service announcement by President Obama warning about fake videos.[32] Differentiating real videos and news from doctored videos and images will become increasingly difficult as deepfake technology continues to evolve.

The recent swell of deepfakes highlights the inadequacy of the law and raises complex questions about how to solve the problems deepfakes present.[33] Although websites like Reddit, Pornhub, and Twitter have banned pornographic deepfakes,[34] this AI-assisted technology will likely continue to have an increasing presence in society.[35] With FakeApp and other social media apps, like Snapchat, offering face-morphing technology, more and more individuals will have the unregulated ability to create realistic doctored videos.[36] To protect victims of deepfakes and to prevent the negative societal consequences they cause, the laws need to keep pace with this technology.

C. Deepfakes in the Context of Nonconsensual Pornography

Although deepfakes have primarily targeted celebrities, private individuals can just as easily find their likenesses used in a face-swapped porn video.[37] The proliferation of social media accounts and the public sharing of photographs enable FakeApp users to gather images from any individual's Instagram or Facebook account to create a doctored video.[38] For instance, Reddit users openly discussed how best to create deepfakes featuring friends, teachers, and ex-partners.[39]

Because deepfake technology can be used to create realistic pornographic videos without the consent of the individuals depicted, and since these videos can be broadly distributed on the internet, pornographic deepfakes exist in

[31] Yaron Steinbuch, *Listen to JFK Speak from Beyond the Grave*, N.Y. POST (Mar. 16, 2018, 12:05 PM), https://nypost.com/2018/03/16/jfks-voice-delivers-speech-he-never-gave-day-of-assassination/ [https://perma.cc/TJQ9-6AFQ].

[32] Morgan Gstalter, *'Obama' Voiced by Jordan Peele in PSA Video Warning About Fake Videos*, HILL (Apr. 17, 2018, 12:08 PM), http://thehill.com/blogs/in-the-know/in-the-know/383525-obama-voiced-by-jordan-peele-in-psa-video-warning-about-fake [https://perma.cc/9F4R-466W].

[33] *See* Chesney & Citron, *supra* note 4 (manuscript at 30–58).

[34] Hawkins, *supra* note 18.

[35] *See* Chesney & Citron, *supra* note 4 (manuscript at 2–4, 8–14).

[36] *See* Reid McCarter, *Idiots on the Internet Are Getting Really Good at Splicing Nic Cage's Face onto Every Movie*, AV CLUB (Sept. 4, 2018, 10:30 AM), https://news.avclub.com/idiots-on-the-internet-are-getting-really-good-at-splic-1828799192 [https://perma.cc/5VAP-WQTK]; *see also* Chesney & Citron, *supra* note 4 (manuscript at 8–14).

[37] Dodge & Johnstone, *supra* note 13, at 3.

[38] Lee, *supra* note 14 (explaining that private individuals can be targeted in deepfakes as long as clear images of the individuals are accessible).

[39] Samantha Cole, *People Are Using AI to Create Fake Porn of Their Friends and Classmates*, VICE: MOTHERBOARD (Jan. 26, 2018, 2:00 PM), https://motherboard.vice.com/en_us/article/ev5eba/ai-fake-porn-of-friends-deepfakes [https://perma.cc/8EPJ-GNNA].

the realm of other sexually exploitative cybercrimes such as revenge porn and nonconsensual pornography.[40] Revenge porn is the act of "disclosing a private, sexually explicit image to someone other than the intended audience" without the consent of the involved individuals.[41] It describes the distribution of genuine, as opposed to doctored, sexually explicit photos or videos without the subject's consent.[42] Typically, partners in an intimate relationship consensually share photos or videos that one partner later distributes in retaliation against the other.[43]

Revenge porn materials, however, can also be stolen and used by strangers through hacking or theft of a cell phone or computer.[44] Hackers may post these images and videos online with (or without) the victim's personal identifying information.[45] Revenge porn is also referred to more broadly as "nonconsensual pornography," which includes sexually explicit images or videos recorded without the individual's consent, such as hidden recordings.[46]

Pornographic deepfakes and revenge porn have troubling commonalities. First, like revenge porn, pornographic deepfakes predominately affect women.[47] One journalist described deepfakes as "a way for men to have their full, fantastical way with women's bodies."[48] Second, both revenge porn and pornographic deepfakes involve the nonconsensual distribution of sexually explicit material. Victims of revenge porn do not consent to public dissemination of their sexual images and videos, and likewise, victims of deepfakes do not agree to manipulation of their face onto the body of an individual engaging in sexual acts. Thus, revenge porn and deepfakes both violate individuals' expectations that sexual activity be founded on consent.[49]

[40] *See* Jeff John Roberts, *Fake Porn Videos Are Terrorizing Women. Do We Need a Law to Stop Them?*, FORTUNE (Jan. 15, 2019), http://fortune.com/2019/01/15/deepfakes-law/ [https://perma.cc/TNA3-G7TM].

[41] MARY ANNE FRANKS, DRAFTING AN EFFECTIVE "REVENGE PORN" LAW: A GUIDE FOR LEGISLATORS 2 (2016), https://www.cybercivilrights.org/wp-content/uploads/2016/09/Guide-for-Legislators-9.16.pdf [https://perma.cc/W3KB-Q7YQ]; *see also* Danielle Keats Citron & Mary Anne Franks, *Criminalizing Revenge Porn*, 49 WAKE FOREST L. REV. 345, 346 (2014).

[42] *See generally* Andrew Gilden, *Sex, Death, and Intellectual Property*, 32 HARV. J.L. & TECH. 67 (2018) (describing revenge porn).

[43] Citron & Franks, *supra* note 41, at 346.

[44] *Cyber Exploitation—Law Enforcement FAQs*, OFF. ATT'Y GEN., https://oag.ca.gov/sites/all/files/agweb/pdfs/ce/cyber-exploitation-law-enforcement-faqs.pdf [https://perma.cc/FBX2-55YZ] (last visited Nov. 12, 2019) [hereinafter *Cyber Exploitation FAQs*].

[45] *Id.*

[46] Citron & Franks, *supra* note 41, at 346.

[47] *Id.* at 347–48 ("Our society has a poor track record in addressing harms that take women and girls as their primary targets.").

[48] Samantha Cole, *Deepfakes Were Created as a Way to Own Women's Bodies—We Can't Forget That*, VICE: BROADLY (June 18, 2018, 2:10 PM), https://broadly.vice.com/en_us/article/nekqmd/deepfake-porn-origins-sexism-reddit-v25n2 [https://perma.cc/7FJ4-99T8].

[49] Dodge & Johnstone, *supra* note 13, at 4.

Both revenge porn and pornographic deepfakes can have lasting effects for victims[50] and can cause similar emotional and psychological harms.[51] Notably, revenge porn not only harms a person's reputation and emotional well-being but also poses risks of significant academic, professional, and even physical injury.[52] Physical safety concerns include violence, stalking, or other criminal acts.[53]

Despite the substantial similarities between pornographic deepfakes and revenge porn, they differ in meaningful ways, and these dissimilarities present unique challenges to victims of deepfakes seeking remedies for harm caused. First, deepfakes may constitute benign artistic expression, whereas revenge porn is always a crime. Indeed, the deepfakes that have garnered the most attention involve public figures and celebrities in harmless and humorous juxtapositions—for instance, Steve Buscemi's face imposed on Jennifer Lawrence's body in a video clip in which the actress discusses her favorite "Real Housewife."[54] In such depictions, where the video is unquestionably fake, the viewer is "in" on the joke—thus, even pornographic deepfakes have historically been viewed through the lens of entertainment and artistic expression rather than crime.[55]

Revenge porn, in contrast, depicts private individuals engaged in intimate acts that were intended to remain private and were not recorded for mass dissemination or entertainment. Indeed, scholars' arguments in support of imposing civil and criminal liability for acts of revenge porn have centered on the violation of the victim's right to sexual privacy.[56]

Pornographic deepfakes do not raise these same sexual privacy concerns. Because deepfakes arguably do not depict a person who exists,[57] no individual's privacy is clearly at stake in a deepfake. Nonetheless, in the case

[50] *See Cyber Exploitation FAQs*, *supra* note 44 (quoting the Cyber Civil Rights Initiative's 2014 "End Revenge Porn" survey, which showed that an overwhelming amount of those affected by revenge porn are women); *see also Revenge Porn Statistics*, CYBER C.R. INITIATIVE, https://www.cybercivilrights.org/wp-content/uploads/2014/12/RPStatistics.pdf [https://perma.cc/29NH-7848] (last visited Nov. 12, 2019).

[51] Citron & Franks, *supra* note 41, at 347; Chesney & Citron, *supra* note 4 (manuscript at 16–20).

[52] *See* sources cited *supra* note 51.

[53] *See* sources cited *supra* note 51.

[54] *See* birbfakes, *Jennifer Lawrence-Buscemi on Her Favorite Housewives [Deepfake]*, YOUTUBE (Jan. 14, 2019), https://www.youtube.com/watch?v=r1jng79a5xc [https://perma.cc/VDT2-L54N].

[55] *See* Madeline Buxton, *The Deep Dark World of Fake Porn,* REFINERY29, https://www.refinery29.com/en-gb/2018/02/190383/deepfakes-ai-assisted-fake-porn [https://perma.cc/6NPX-KKQX] (last updated Feb. 7, 2018, 4:00 PM); *see also* Hawkins, *supra* note 18 ("U.S. law offers little recourse to victims of face-swap porn, largely because courts thus far have viewed such material in the same realm as parody or satire, which enjoy strong First Amendment protections.").

[56] *See* Citron & Franks, *supra* note 41, at 350–56.

[57] The deepfakes at issue in this Article—those that are the subject of the proposed statute described in Part V—involve face swapping, where the image of one person's face is believably mapped onto the body of another person engaged in a pornographic act. Other varieties of deepfakes, such as those involving the manipulation of an identifiable person's speech or image to simulate fake conduct or speech, are beyond the scope of this Article.

of a private individual whose face was used to create a pornographic deepfake, the viewer may not realize the video depiction is a fake and may instead assume that the video is a genuine depiction of a real person.

In addition, deepfakes also differ from other forms of nonconsensual pornography and revenge porn because of the nearly incalculable number of potential victims. Possible victims of deepfake pornography include both individuals who consensually appeared in sexually explicit digital images that were later nonconsensually disclosed (i.e., traditional victims of revenge porn) and anyone whose image has been captured digitally.[58]

II. THE PRUDENTIAL AND LEGAL CHALLENGES OF DEEPFAKES

The rapid influx of face-swapped porn videos, as well as expeditious advancements in the technology powering the deepfake phenomenon, raises complex legal and prudential issues. This Part will address the difficulty of identifying the perpetrators and victims of deepfakes and the challenge of convincing internet platforms to remove deepfakes from websites in light of the Communications Decency Act of 1996 (CDA). Given these challenges, there is a need for a criminal law response.

A. Defining and Recognizing Deepfake Victims

This new technology raises the issue of defining and identifying victims of deepfakes, which depict two people: the person whose body is engaging in sexual acts and the person whose face has been transposed onto that body.[59] Although two different people appear in these videos, articles and commentators have generally classified the "victim" as the person whose face was wrongly used.[60] But should the person whose body is used also be classified as a victim if consent was not obtained to use that person's body? In addition, how should the law determine harm when deepfakes do not expose actual intimate details of victims' lives? How can liability be imposed when no real person has been exposed? The person whose face was used likely did not agree to participate in pornography, but consent is less clear as to that person whose body was depicted. Although the actor whose body is featured may have consented to the original pornographic video, they likely never agreed to have another person's face superimposed onto their body. They too have been victimized. Thus, both people depicted in the deepfake should be presumed to be victims.

B. Challenges of Identifying Perpetrators and Seeking Recourse

Deepfakes also raise the complex problem of identifying perpetrators and removing videos after they have been published on the internet. Once a video is posted online, it is difficult to remove it, and, usually, the reputational and

[58] Buxton, *supra* note 55.

[59] *See* Chesney & Citron, *supra* note 4 (manuscript at 9, 16–17).

[60] *See, e.g.*, *id.* (manuscript at 16–20).

emotional damage is already done.[61] Ideally, the law should hold responsible the individual who created and published a deepfake video online. However, it can be challenging to locate or identify the perpetrator who created a deepfake for two reasons. First, many websites, such as Reddit, Twitter, and Pornhub, allow anonymous use; predictably, deepfakes have flourished on these platforms.[62] And second, deepfake creators can use software like Tor to make the IP address associated with a deepfake untraceable.[63]

Because of the elusiveness of perpetrators, victims may consider pursuing action directly involving or against the internet platform where the video was posted. Such recourse implicates section 230 of the CDA,[64] which shields websites from civil and criminal liability under state law for third-party content. Under section 230, "[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."[65] This provision protects content providers from incurring liability as intermediaries for third-party content posted on their websites.[66] In theory, this immunity safeguards free expression on the internet by eliminating the fear of liability for the actions of third-party users.[67] Section 230 allows websites to host videos, reviews, classified ads, and social networking profiles created by hundreds of millions of internet users without being liable for that content.[68] Although section 230 has been lauded as one of the most important laws protecting internet

[61] Megan Farokhmanesh, *Is It Legal to Swap Someone's Face into Porn Without Consent?*, VERGE (Jan. 30, 2018, 2:39 PM), https://www.theverge.com/2018/1/30/16945494/deepfakes-porn-face-swap-legal [https://perma.cc/5HP9-FQM6] ("It's almost impossible to erase a video once it's been published to the internet . . . . If you're looking for the magic wand that can erase that video permanently, it probably doesn't exist." (quoting Eric Goldman, a law professor at Santa Clara University School of Law)).

[62] *See Rise of the Deepfakes*, *supra* note 11 (noting that "deepfakes" were first posted to Reddit by a user that remains anonymous); Roose, *supra* note 19 (noting that the creator of FakeApp has remained anonymous).

[63] *See* Andy Greenberg, *It's About to Get Even Easier to Hide on the Dark Web*, WIRED (Jan. 20, 2017, 7:00 AM), https://www.wired.com/2017/01/get-even-easier-hide-dark-web/ [https://perma.cc/HKB6-C924] (explaining how Tor allows for anonymous internet use).

[64] 47 U.S.C. § 230 (2012). Section 230 states, in relevant part, that

> [n]o provider or user of an interactive computer service shall be held liable on account of . . . any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected.

Id. § 230(c)(2).

[65] *Id.* § 230(c)(1).

[66] *See id.*; *see also* Patricia Spiccia, Note, *The Best Things in Life Are Not Free: Why Immunity Under Section 230 of the Communications Decency Act Should Be Earned and Not Freely Given*, 48 VAL. U. L. REV. 369, 379 (2013).

[67] *Section 230 of the Communications Decency Act*, ELECTRONIC FRONTIER FOUND., https://www.eff.org/issues/cda230 [https://perma.cc/N9CT-XXPY] (last visited Nov. 12, 2019); *see also* Annemarie Bridy, *Remediating Social Media: A Layer-Conscious Approach*, 24 B.U. J. SCI. & TECH. L. 193, 196–97 (2018).

[68] *Section 230 of the Communications Decency Act*, *supra* note 67.

speech,[69] the law also enables internet platforms to host revenge porn—and, now, deepfakes—with impunity.[70]

Section 230 immunity makes it difficult for deepfake victims to sue internet platforms for hosting deepfakes.[71] Victims of revenge porn and deepfakes are unlikely to pierce CDA immunity because "courts give a high degree of deference to website hosts under Section 230."[72] Courts have interpreted the CDA broadly to effect Congress's intent to "encourage free and open communication on the Internet, even if that communication includes potentially harmful speech or conduct."[73]

Indeed, in *Barnes v. Yahoo!, Inc.*,[74] the Ninth Circuit found that section 230 barred a claim "for negligent provision of services" after Yahoo! failed to take down reported sexually explicit images of the plaintiff posted by her ex-boyfriend. *Barnes* "demonstrates the force of Section 230 . . . in barring potential revenge porn cases against website hosts."[75] The same result would certainly follow for a claim based on a deepfake. Section 230 would likely shield websites from liability for hosting deepfakes posted by users. Thus, deepfake victims will be left without recourse to law if they can neither locate the perpetrators of deepfakes because of anonymity nor sue the internet platforms due to section 230. Even if a perpetrator is identified, successful litigation against deepfake creators will not bind content providers.

Despite the protection the CDA affords, many internet platforms have independently banned deepfakes and removed the videos from their platforms.[76] Reddit announced new policies regarding involuntary pornography, which prohibit the dissemination of images or video depicting any person nude or engaged in any act of sexual conduct created or posted without their permission, "including depictions that have been faked."[77]

[69] Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1604 (2018) (noting that the CDA is often called "the law that matters most for speech on the Web").

[70] Samantha Cole, *Targets of Fake Porn Are at the Mercy of Big Platforms*, VICE: MOTHERBOARD (Feb. 5, 2018, 12:44 PM), https://motherboard.vice.com/en_us/article/59kzx3/targets-of-fake-porn-deepfakes-are-at-the-mercy-of-big-platforms [https://perma.cc/6MJF-KHRN].

[71] Rachel Budde Patton, Note, *Taking the Sting Out of Revenge Porn: Using Criminal Statutes to Safeguard Sexual Autonomy in the Digital Age*, 16 GEO. J. GENDER & L. 407, 423–24 (2015).

[72] *Id.* at 423.

[73] *Id.*

[74] 570 F.3d 1096, 1105 (9th Cir. 2009).

[75] Patton, *supra* note 71, at 424.

[76] Natalie Gil, *"Deepfake" Porn: The New Way Women Are Being Exploited Online*, REFINERY29 (June 21, 2018, 7:30 AM), https://www.refinery29.com/en-gb/2018/06/202456/deepfake-pornography [https://perma.cc/J75W-MJJK] (identifying Twitter, Pornhub, Gfycat (a GIF-hosting service), and Discord (a gaming chat platform) as platforms that ban deepfakes).

[77] *Account and Community Restrictions: Do Not Post Involuntary Pornography*, REDDIT, https://www.reddithelp.com/en/categories/rules-reporting/account-and-community-restrictions/do-not-post-involuntary-pornography [https://perma.cc/23EF-SYBD] (last visited Nov. 12, 2019); *see also Account and Community Restrictions: Do Not Post Sexual or Suggestive Content Involving Minors*, REDDIT, https://www.reddithelp.com/en/categories/

Pornhub similarly declared that it considered deepfakes to be nonconsensual pornography that violated the platform's terms of service.[78] Twitter also vowed to suspend all accounts that distribute or produce nonconsensual pornography.[79] More recently, Tumblr announced its new policy banning the "unwanted sexualization or sexual harassment of others."[80]

The challenges of defining the victims, finding the perpetrators, and holding internet platforms responsible are further muddled by the unclear application of current law to deepfakes. The same distortion and anonymity issues involved in deepfakes' creation make it difficult to naturally fit these doctored videos into existing laws, which does not settle the question of who should be held responsible for acts involving deepfakes. Some remedies focus on the creators of the content, others focus on the social networks that fail to monitor the content properly, and some involve both. Notably, platforms like Pornhub have been largely unsuccessful in removing current content and stopping creators from posting new content; this alone demonstrates a clear need for legal solutions for both the networks and creators.[81]

rules-reporting/account-and-community-restrictions/do-not-post-sexual-or-suggestive [https://perma.cc/3F54-P7KX] (last visited Nov. 12, 2019).

[78] Tom McKay, *Pornhub Says Digitally Generated 'Deepfakes' Are Non-Consensual and It Will Remove Them*, GIZMODO (Feb. 6, 2018, 9:30 PM), https://gizmodo.com/pornhub-says-digitally-generated-deepfakes-are-non-cons-1822786071 [https://perma.cc/6EJH-E7HU]; *Terms & Conditions*, PORNHUB, https://www.pornhub.com/information#terms [https://perma.cc/UB7Q-9D78] (last updated May 18, 2018).

[79] *Non-Consensual Nudity Policy*, TWITTER (Mar. 2019), https://help.twitter.com/en/rules-and-policies/intimate-media [https://perma.cc/F36B-G4UB]; James Vincent, *Twitter Is Removing Face-Swapped AI Porn from Its Platform, Too*, VERGE (Feb. 7, 2018, 8:14 AM), https://www.theverge.com/2018/2/7/16984360/twitter-ban-fake-porn-ai-face-swap [https://perma.cc/NE6B-HSXL];

[80] *Tumblr Bans Non-Consensual Creepshots and Deepfake Porn*, BBC NEWS (Aug. 27, 2018), https://www.bbc.com/news/technology-45323287 [https://perma.cc/TXK4-5BP2]; *see also Community Guidelines*, TUMBLR (Dec. 17, 2018), https://www.tumblr.com/policy/en/community [https://perma.cc/P3E5-ULPQ]. Although it is promising that internet platforms have publicly renounced pornographic deepfakes, they remain readily accessible on internet sites such as Reddit and Pornhub. Reddit, Pornhub, Twitter, and Tumblr rely on users to spot and report deepfakes. Rense, *supra* note 25. Deepfakes keep appearing on these internet platforms because "policing every piece of content is virtually impossible." *Id.* Further, perpetrators are turning to alternative internet platforms like Erome and Sendvid to post deepfakes once their content is removed. McKay, *supra* note 78. More concerningly, Naughty America, an adult entertainment company, has begun to monetize deepfakes by allowing users to commission their own deepfake clips. Janko Roettgers, *Naughty America Wants to Monetize Deepfake Porn*, VARIETY (Aug. 20, 2018, 2:30 PM), https://variety.com/2018/digital/news/deepfake-porn-custom-clips-naughty-america-1202910584/ [https://perma.cc/4PLT-BK7V]. Naughty America CEO Andreas Hronopoulos described "customization and personalization as the future" of the porn industry. *Id.*

[81] Bryan Clark, *Pornhub Promised to Ban 'Deepfakes' Videos. And It Failed Miserably.*, TNW (Apr. 18, 2018), https://thenextweb.com/insider/2018/04/19/pornhub-promised-to-ban-deepfakes-videos-and-it-failed-miserably/ [https://perma.cc/GYC8-48AF].

C. The Need for Criminal Law Responses

Each potential deepfake victim faces unique legal hurdles. For example, a victim whose face was used will encounter the issue of whether they can seek redress for exposure of intimate details which are not their intimate details. Conversely, the victim whose body is represented in a deepfake may have trouble proving that their body is significantly identifiable to qualify as a recognizable misrepresentation. And even if both individuals are recognized as victims, they must separately prove the harm caused by the deepfake, which is problematic because the videos are not exposing intimate details of any single victim's life.[82]

Even if harm can be shown, victims identified, and creators located, obtaining and enforcing a remedy can prove challenging. Civil liability offers one legal means to remedy harm and incentivize actors to conform to societal norms of acceptable behavior. Other scholars writing in this area have begun to identify and evaluate the potential civil claims and market responses to address the problem of deepfakes.[83]

Civil liability, however, has limitations and often falls short. For example, many creators of deepfakes may be judgment-proof, and therefore imposing civil liability may not deter their conduct. In addition, victims of deepfakes, especially those whose bodies (rather than their faces) have been used in the depiction, may not want their identities widely disseminated or known; civil plaintiffs must sue in their own names. Second, the high cost of investigating and litigating deepfake cases will limit those who can obtain a civil remedy to only plaintiffs who have the resources to pursue such claims. And even if those claims prove successful against the creator of the deepfake, there is no guarantee that the deepfake will be removed from the internet, given the immunity granted to third-party providers under the CDA.[84]

A criminal law remedy does not face the same challenges as civil remedies. Indeed, criminalizing deepfakes will have a greater deterrent effect on creators. Professors Bobby Chesney and Danielle Citron have observed that "being judgment proof might spare someone from fear of civil suit, but it is no protection from being sent to prison."[85] In addition, law enforcement has resources far beyond the capacity of most litigants, even celebrities, that can be brought to bear in the investigation and prosecution of deepfakes. Scholars who urge the criminalization of revenge porn have successfully deployed these same arguments.[86] In addition, over the last one hundred

[82] Emma Grey Ellis, *People Can Put Your Face on Porn—and the Law Can't Help You*, WIRED (Jan. 26, 2018, 7:00 AM), https://www.wired.com/story/face-swap-porn-legal-limbo/ [https://perma.cc/5VSH-CYRA].

[83] Solutions emerging in civil law and the regulatory arena warrant additional exploration and are beyond the scope of this Article. The possibility of imposing civil liability on the perpetrators of deepfakes and civil remedies to the victims is the subject of thoughtful scholarship by others. *See, e.g.*, Chesney & Citron, *supra* note 4 (manuscript at 31–41).

[84] *See supra* notes 64–68 and accompanying text.

[85] Chesney & Citron, *supra* note 4 (manuscript at 42).

[86] *See* Citron & Franks, *supra* note 41, at 361–65 (discussing the importance of criminalizing revenge porn because civil remedies do not address the harm); *see also* Taylor

years, the criminalization of invasions of personal privacy has gained support in the legal literature[87] and the law—from criminal penalties for disclosure of private financial[88] and health information[89] to identity theft[90] to the Video Voyeurism Prevention Act of 2004[91] that bans intentionally recording or broadcasting an image of another person in a state of undress. Finally, a criminal response to nonconsensual pornographic deepfakes sends a powerful message beyond that conveyed in the context of civil remedies. Criminalizing conduct communicates that society finds the behavior intolerable.[92] Specifically, it conveys the view that the conduct is not trivial and that it is not only hurtful to the individual involved but also harmful and offensive to the community.[93] Given that these deepfakes disproportionately victimize women and girls, a societal response—in the form of criminal punishment—is warranted.

III. THE STATE OF THE LAW

For more than two hundred years, lawmakers have recognized the need for law and our legal institutions to keep pace with technology.[94] The reality, however, does not square with that aspiration.[95] The law lags, struggling to

Linkous, *It's Time for Revenge Porn to Get a Taste of Its Own Medicine: An Argument for the Federal Criminalization of Revenge Porn*, 20 RICH. J.L. & TECH. 14, 36–37 (2014) (urging the federal criminalization of revenge porn).

[87] Almost 130 years ago, Samuel D. Warren and Louis D. Brandeis first argued that criminal laws should punish as a felony the publication of "'any statement concerning the private life or affairs of another, after being requested in writing . . . not to publish such statement' provided that the statement does not concern someone's qualifications for public office or profession or involve a matter of public interest." Citron & Franks, *supra* note 41, at 346 n.11 (quoting Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 219 n.8 (1890)).

[88] *See* 5 U.S.C. § 552(b)(4) (2012).

[89] *See* 42 U.S.C. § 1320d-6 (2012).

[90] *See generally* 18 U.S.C. § 1028 (2012).

[91] 18 U.S.C. § 1801.

[92] Writing about the criminalization of domestic violence, Citron observed that, "Law creates a public set of meanings and shared understandings between the state and the public" and that, "[b]ecause law creates and shapes social mores, it has an important cultural impact." Danielle Keats Citron, *Law's Expressive Value in Combating Cyber Gender Harassment*, 108 MICH. L. REV. 373, 407 (2009).

[93] *See* Citron & Franks, *supra* note 41, at 361–63.

[94] Letter from Thomas Jefferson to Samuel Kercheval (July 12, 1816), *in* 10 THE WRITINGS OF THOMAS JEFFERSON 37, 42–43 (New York, G. P. Putnam's Sons 1899). In 1816, Thomas Jefferson wrote:

> Laws and institutions must go hand in hand with the progress of the human mind. As that becomes more developed, more enlightened, as new discoveries are made, new truths disclosed, and manners and opinions change with the change of circumstances, institutions must advance also, and keep pace with the times.

Id. at 43.

[95] Professor Henry H. Perritt, Jr. presented two ideas that dictate the relationship between law and technology—first, that technological change is a major source of human problems that the law must address and second, that "law lags [behind] technology" because market forces and the common law tradition have required "that the legal system should not predetermine the course of technology application and product development." HENRY H. PERRITT, JR., LAW AND THE INFORMATION SUPERHIGHWAY 2–3 (1996).

keep pace with rapidly evolving technology.[96] The law on deepfakes is no exception.[97] The focus here is on federal and state laws, primarily criminal, that are currently used to punish similar conduct, including the creation and distribution of nonconsensual pornography and revenge porn. Such laws may provide an analogy and analytical framework to deploy in the context of deepfakes.

A. Federal Law

No specific federal law criminalizes deepfakes or revenge porn. Therefore, federal legislators and federal prosecutors continue to grapple with the criminalization of nonconsensual pornography posted on the internet.

1. Existing Federal Criminal Law

Without any law directly on point, prosecutors look to other federal criminal laws relating to cyberexploitation to prosecute individuals who make and distribute deepfakes.[98] These include various provisions under title 18 of the U.S. Code that criminalize communications involving ransom, extortion, and kidnapping threats; cyberharassment; and cyberstalking.[99] Possible useful statutory provisions include those governing threats made through interstate communication; unauthorized access to a computer; and stalking as interstate domestic violence.[100] Each provision will be evaluated in turn.

Under 18 U.S.C. § 875(c), a culpable person can receive a fine, imprisonment for up to two years, or both where he: "(1) knowingly make[s] a communication containing a true threat to injure in interstate commerce or foreign commerce, and (2) intends the communication to be a true threat to injure another or knows that the recipient of the threat would understand it to be a threat."[101] Assuming that uploading a deepfake qualifies as a "communication"[102] under section 875 and further assuming the other

[96] *Id.* "Law lag" is the notion that the law is always falling behind other industries and that technoscientific inventiveness is an "agenda-setting force to which the law responds only by reaction." Sheila Jasanoff, *Making Order: Law and Science in Action*, *in* 3 THE HANDBOOK OF SCIENCE AND TECHNOLOGY STUDIES 761, 768–69 (Edward J. Hackett et al. eds., 2008).

[97] Chris Meserole & Alina Polyakova, *The West Is Ill-Prepared for the Wave of "Deep Fakes" That Artificial Intelligence Could Unleash*, BROOKINGS (May 25, 2018), https://www.brookings.edu/blog/order-from-chaos/2018/05/25/the-west-is-ill-prepared-for-the-wave-of-deep-fakes-that-artificial-intelligence-could-unleash/ [https://perma.cc/NNL4-9F8D].

[98] *Cyber Exploitation FAQs*, *supra* note 44.

[99] Margaret S. Groban, *Intimate Partner Cyberstalking—Terrorizing Intimate Partners with 21st Century Technology*, U.S. ATTORNEYS' BULL., May 2016, at 12, 12–14.

[100] *See generally* 18 U.S.C. §§ 875, 1030, 2261A (2012).

[101] *Cyber Exploitation FAQs*, *supra* note 44 (paraphrasing 18 U.S.C. § 875(c)).

[102] For example, in *United States v. Jeffries*, the Sixth Circuit held that a music video was a threatening communication under § 875(c) because "the statute covers 'any threat,' making no distinction between threats delivered orally (in person, by phone) or in writing (letters,

elements of section 875(c) are met, a prosecutor may be able to charge the creator of a deepfake with a violation of section 875(c).

Charges might also be brought under 18 U.S.C. § 1030 for unauthorized access to a computer—that is, computer hacking.[103] This provision can be used where the nonconsensual pornographic materials were gathered via unauthorized access to the victim's computer or email.[104] Penalties again can include fines, imprisonment, or both[105] and can also increase depending on how a person gains access. For example, if the information is obtained through the use of stolen usernames and passwords, an aggravated identity theft charge with a mandatory two-year sentence can be applied.[106]

Third, federal prosecutors have used the federal cyberstalking statute, 18 U.S.C. § 2261A, to prosecute crimes similar to deepfakes and revenge porn.[107] This provision applies where the conduct "places [a] person in reasonable fear of the death of, or serious bodily injury to . . . (i) that person; (ii) an immediate family member; or (iii) a spouse or intimate partner of that person."[108] The accused must have the intent to "kill, injure, harass, intimidate, or place [the victim] under surveillance" with the same intent and can be punished by a fine, imprisonment, or both.[109] Therefore, if distributing revenge materials is coupled with, or rises to the level of, cyberstalking, then a person can be held criminally liable.[110] In *United States v. Matusiewicz*,[111] prosecutors defeated an overbreadth challenge to 18 U.S.C. § 2261A(2) and in doing so opened the door for this provision to be applied to nonconsensual pornography in cyberstalking instances.[112]

Finally, if the nonconsensual revenge porn material involves minors, a person may face charges under 18 U.S.C. § 2251 for production of child pornography and 18 U.S.C. § 2422(b) for enticement or coercion of a minor.[113] In a study published in 2016, at least 3 percent of revenge porn victims were between the ages of fifteen and seventeen,[114] but research from

emails, faxes), by video or by song, in old-fashioned ways or in the most up-to-date." 692 F.3d 473, 482 (6th Cir. 2012).

[103] *See* 18 U.S.C. § 1030; Groban, *supra* note 99, at 16.

[104] Groban, *supra* note 99, at 16.

[105] *See Cyber Exploitation FAQs*, *supra* note 44.

[106] *See generally* United States v. Barrington, 648 F.3d 1178 (11th Cir. 2011).

[107] 18 U.S.C. § 2261A(1)(A), 2261A(2)(A) (2012); *see* Katlyn M. Brady, *Revenge in Modern Times: The Necessity of a Federal Law Criminalizing Revenge Porn*, 28 HASTINGS WOMEN'S L.J. 3, 12 (2017).

[108] 18 U.S.C. § 2261A(1)(A).

[109] *Id.*; *see Cyber Exploitation FAQs*, *supra* note 44.

[110] According to a survey by EndRevengePorn.org, 49 percent of respondents said they have been harassed or stalked online by users that have seen the content of the revenge porn posted about them. *Revenge Porn Statistics*, *supra* note 50 (showing the results of the survey hosted on EndRevengePorn.org from August 2012 to December 2013, with an important note that the sample was heavily female).

[111] 84 F. Supp. 3d 363 (D. Del. 2015).

[112] *Id.* at 366–68.

[113] 18 U.S.C. §§ 2251, 2422(b).

[114] Amanda Lenhart et al., *Nonconsensual Image Sharing: One in 25 Americans Has Been a Victim of "Revenge Porn,"* DATA & SOC'Y RES. INST. (Dec. 13, 2016),

BBC News shows that revenge porn victims worldwide are as young as eleven years old.[115] Where a victim is underage, these provisions may apply to the prosecution of those responsible for the creation and distribution of these materials.

2. Previously Proposed Federal Criminal Statutes

Legal scholars[116] and internet advocacy organizations and activists[117] have coalesced around the cause of victims of nonconsensual pornography advocating for the adoption of a federal statute expressly criminalizing this conduct.[118] These groups have proposed federal legislation to outlaw revenge porn and deepfakes.

a. Ending Nonconsensual Online User Graphic Harassment Act of 2017: The ENOUGH Act

In 2016, Congresswoman Jackie Speier (D-CA) introduced the Intimate Privacy Protection Act[119] (IPPA), which would criminalize the distribution of nonconsensual pornography.[120] Upon introduction of the bill, Speier reflected that the IPPA was a response to the fact that most revenge porn victims do not have the resources to seek civil remedies, and thus

https://datasociety.net/pubs/oh/Nonconsensual_Image_Sharing_2016.pdf [https://perma.cc/Q6KG-F5P6].

[115] Peter Sherlock, *Revenge Pornography Victims as Young as 11, Investigation Finds*, BBC NEWS (Apr. 27, 2016), https://www.bbc.com/news/uk-england-36054273 [https://perma.cc/S339-LKSR].

[116] *See* Brady, *supra* note 107, at 3; *see also* Meghan Fay, Note, *The Naked Truth: Insufficient Coverage for Revenge Porn Victims at State Law and the Proposed Federal Legislations to Adequately Redress Them*, 59 B.C. L. REV. 1839, 1839 (2018).

[117] CYBER C.R. INITIATIVE, https://www.cybercivilrights.org/ [https://perma.cc/Z5SF-9F6D] (last visited Nov. 12, 2019); *see also* Holly Jacobs, *This Is What It Is Like to Be the Victim of Revenge Porn, and Why We Need to Criminalise It*, INDEPENDENT (Feb. 13, 2015, 4:40 PM), https://www.independent.co.uk/voices/comment/this-is-what-it-is-like-to-be-the-victim-of-revenge-porn-and-why-we-need-to-criminalise-it-10045067.html [https://perma.cc/Y9WV-PW66].

[118] *Advocacy*, CYBER C.R. INITIATIVE, https://www.cybercivilrights.org/advocacy/ [https://perma.cc/TG74-AML3] (last visited Nov. 12, 2019).

[119] Intimate Privacy Protection Act of 2016, H.R. 5896, 114th Cong. (2016).

[120] As introduced, H.R. 5896 provided:

> Whoever knowingly uses the mail, any interactive computer service or electronic communication service or electronic communication system of interstate commerce, or any other facility of interstate or foreign commerce to distribute a visual depiction of a person who is identifiable from the image itself or information displayed in connection with the image and who is engaging in sexually explicit conduct, or of the naked genitals or post-pubescent female nipple of the person, with reckless disregard for the person's lack of consent to the distribution, shall be fined under this title or imprisoned not more than 5 years, or both.

Id. § 1802(a). The legislation provided exceptions to the use of such images by law enforcement, in legal proceedings, and in other cases of bona fide public interest or where the depiction was done voluntarily in a public place or in a lawful commercial setting. It specifically exempted telecommunication and internet service providers, as defined by section 230, from prosecution, unless the provider "promotes or solicits content that it knows to be in violation of this section." *Id.* § 1802(b)(4).

criminalizing revenge porn at the federal level would provide some redress.[121] The bill expired at the end of the 114th Congress.

After that, the bill was revised[122] and introduced in late 2017 in both the Senate and House as the Ending Nonconsensual Online User Graphic Harassment Act of 2017[123] (the "ENOUGH Act"). The Act would have amended title 18 of the U.S. Code to make it

> unlawful to knowingly use any means or facility of interstate or foreign commerce to distribute an intimate visual depiction of an individual—
>
> "(1) with knowledge of or reckless disregard for—
>
> "(A) the lack of consent of the individual to the distribution;
>
> "(B) the reasonable expectation of the individual that the depiction would remain private; and
>
> "(C) harm that the distribution could cause to the individual; and
>
> "(2) without an objectively reasonable belief that such distribution
>
> touches upon a matter of public concern.[124]

The ENOUGH Act defined "intimate visual depiction" as "any visual depiction"

> "(A) of an individual who is reasonably identifiable from the visual depiction itself or information displayed in connection with the visual depiction;
>
> "(B) in which—
>
> "(i) the individual is engaging in sexually explicit conduct; or
>
> "(ii) the naked genitals or post-pubescent female nipple of the individual are visible;

[121] Press Release, Congresswoman Jackie Speier, Congresswoman Speier, Fellow Members of Congress Take on Nonconsensual Pornography, AKA Revenge Porn (July 14, 2016), https://speier.house.gov/media-center/press-releases/congresswoman-speier-fellow-members-congress-take-nonconsensual [https://perma.cc/ZE25-WQST]; *see also* Steven Nelson, *Federal 'Revenge Porn' Bill Will Seek to Shrivel Booming Internet Fad*, U.S. NEWS & WORLD REP. (Mar. 26, 2014, 6:01 PM), https://www.usnews.com/news/articles/2014/03/26/federal-revenge-porn-bill-will-seek-to-shrivel-booming-internet-fad [https://perma.cc/86UZ-LCNQ] (discussing Speier's prior effort to introduce the legislation in 2014).

[122] Critics of the IPPA worried that it could encroach on bona fide journalism and limit journalists' ability to publish photos or videos that were of public interest. Jessica Lahitou, *What Is the ENOUGH Act?: Lawmakers Are Pushing to Criminalize Revenge Porn with a New Bill*, BUSTLE (Nov. 28, 2017), https://www.bustle.com/p/what-is-the-enough-act-lawmakers-are-pushing-to-criminalize-revenge-porn-with-a-new-bill-6337236 [https://perma.cc/3ZXU-HXQD]. The ENOUGH Act addressed critics' concerns by providing that a person who posts images could only be punished if "no reasonable person would consider the shared image to touch on a matter of public concern." *Id.* In a fact sheet provided to Bustle by Senator Kamala Harris's office, the senator clarified that the bill would "narrowly establish criminal liability." *Id.* It also contained a specific carve-out to protect First Amendment rights. *Id.*

[123] S. 2162, 115th Cong. (2017); H.R. 4472, 115th Cong. (2017). This paper will cite primarily to the Senate version of the bill.

[124] S. 2162.

> "(C) in which the content described in subparagraph (B) is not simulated.[125]

Punishment under the Act included a fine, imprisonment of up to five years, or both.[126] The ENOUGH Act provided exceptions for law enforcement, legal proceedings, and "information content providers" as defined by 47 U.S.C. § 230(f) "unless the provider of the communications service intentionally solicits, or knowingly and predominantly distributes, content that the provider of the communications service has actual knowledge is in violation of this section."[127] Although the ENOUGH Act had both Republican and Democratic sponsors, it never made it out of committee and expired at the end of the 115th Congress.[128]

b. The Malicious Deep Fake Prohibition Act of 2018

In late December 2018, Senator Ben Sasse (R-NE) introduced a bill to criminalize the malicious creation and distribution of deepfakes, the Malicious Deep Fake Prohibition Act of 2018[129] (MDFPA). The MDFPA prohibited using any means or facility of interstate commerce to

> "(1) create, with the intent to distribute, a deep fake with the intent that the distribution of the deep fake would facilitate criminal or tortious conduct under Federal, State, local, or Tribal law; or
>
> "(2) distribute an audiovisual record with—
>
> > "(A) actual knowledge that the audiovisual record is a deep fake; and
> >
> > "(B) the intent that the distribution of the audiovisual record would facilitate criminal or tortious conduct under Federal, State, local, or Tribal law.[130]

The MDFPA defines "deep fake" as "an audiovisual record created or altered in a manner that the record would falsely appear to a reasonable observer to be an authentic record of the actual speech or conduct of an individual."[131] It would have applied to an individual deepfake creator who intended to do something illegal (like committing fraud) by posting the deepfake to a platform like Facebook—but it would only implicate Facebook if the platform knew it had distributed a deepfake. The Act proposed as a punishment a fine and up to two years' imprisonment, or—if the deepfake could facilitate violence or disrupt government or an election—up to ten

[125] *Id.*

[126] *Id.*

[127] *Id.*

[128] *See US S2162*, BILL TRACK 50, https://www.billtrack50.com/BillDetail/897787 [https://perma.cc/FX4S-GPK5] (last visited Nov. 12, 2019).

[129] S. 3805, 115th Cong. (2018).

[130] *Id.*

[131] *Id.* Under the MDFPA, "the term 'audiovisual record'—(A) means any audio or visual media in an electronic format; and (B) includes any photograph, motion-picture film, video recording, electronic image, or sound recording." *Id.*

years.[132] Senator Sasse introduced the MDFPA the day before the December 2018 government shutdown; the bill was sent to the Senate Judiciary Committee and expired at the end of 2018.[133] It had no cosponsors.[134]

B. State Law

No state currently criminalizes deepfakes. As in the federal arena, victims of pornographic deepfakes who want to seek redress under state law must look to laws that criminalize related crimes, such as revenge porn. In the absence of federal laws outlawing nonconsensual pornography, victims are left with a "patchwork of state criminal laws [that] is often inadequate."[135] State laws range from prosecuting a revenge porn offense as a felony to punishing the same case as a misdemeanor to lacking felony or misdemeanor charges for the conduct entirely.[136] This Part examines California, Texas, Florida, and New York—the states with the highest number of cybercrime complaints[137] and the largest populations.[138]

1. California

a. Revenge Porn

In 2013, the California State Legislature responded to the revenge porn epidemic by amending the existing criminal statute prohibiting disorderly conduct.[139] Effective October 1, 2013, this amendment to section 647 of the Penal Code of California criminalized nonconsensual pornography by providing that the following persons are guilty of disorderly conduct:

> A person who intentionally distributes the image of the intimate body part or parts of another identifiable person, or an image of the person depicted engaged in an act of sexual intercourse, sodomy, oral copulation, sexual penetration, or an image of masturbation by the person depicted or in which the person depicted participates, under circumstances in which the persons agree or understand that the image shall remain private, the person distributing the image knows or should know that distribution of the image

[132] *Id.*

[133] *See US S3805*, BILL TRACK 50, https://www.billtrack50.com/BillDetail/1000397 [https://perma.cc/2HEQ-4WYU] (last visited Nov. 12, 2019).

[134] *Id.*

[135] *See* Brady, *supra* note 107, at 3.

[136] *See* Jillian Roffer, Note, *Nonconsensual Pornography: An Old Crime Updates Its Software*, 27 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 935, 956–57 (2017) (summarizing various state law approaches to criminalizing revenge porn). According to the Cyber Civil Rights Initiative, forty-one states now have laws criminalizing revenge porn. CYBER C.R. INITIATIVE, *supra* note 117.

[137] *See generally* FED. BUREAU OF INVESTIGATION, 2017 INTERNET CRIME REPORT (2017), https://pdf.ic3.gov/2017_IC3Report.pdf [https://perma.cc/Y2S8-8RRE].

[138] *US States—Ranked by Population 2019*, WORLD POPULATION REV., http://worldpopulationreview.com/states/ [https://perma.cc/FSA7-NGAR] (last visited Nov. 12, 2019).

[139] CAL. PENAL CODE § 647(j)(4)(A) (West 2019).

> will cause serious emotional distress, and the person depicted suffers that distress.[140]

Accordingly, a person intentionally distributing pornographic materials of an identifiable victim without the victim's permission is guilty of disorderly conduct if they know, or reasonably should know, that the action will cause serious emotional distress, and the victim suffers such distress.[141] A person convicted under this statute will be guilty of disorderly conduct, a misdemeanor.[142] Maximum penalties for a violation of this section include up to six months in county jail, a fine of up to $1000, or both.[143] This penalty increases for subsequent violations or if the initial violation includes a victim who was a minor at the time.[144]

b. Cyberstalking

Despite being the first state to criminalize stalking,[145] California does not have a specific cyberstalking statute and rather opts to make amendments to the existing stalking laws to cover cyberstalking cases.[146] Under California Penal Code section 646.9, the general anti-stalking statute, it is unlawful to "willfully, maliciously, and repeatedly follow[] or willfully and maliciously harass[] another person" and to "make[] a credible threat with the intent to place that person in reasonable fear for his or her safety, or the safety of his or her immediate family."[147] A stalking conviction thus requires not only evidence of following or harassment but also evidence of a willful and credible threat.[148] The statute covers threats made "through the use of an electronic communication device," including a cell phone and computer.[149] Stalking is a "wobbler" in California; it may be prosecuted as either a misdemeanor or a felony, depending on the facts of the case.[150]

California passed section 653.2 of its penal code in 2009 to eliminate "an electronic loophole" in section 646.9 and to give "law enforcement the ability to hold accountable those who would prey on victims using electronic means."[151] This narrow statute supplements the anti-stalking law and provides that

[140] *Id.*

[141] *Id.*

[142] *Id.* § 647.

[143] *Id.* § 19.

[144] *Id.* § 647(*l*)(1)–(2).

[145] Tracey B. Carter, *Local, State, and Federal Responses to Stalking: Are Anti-Stalking Laws Effective?*, 22 WM. & MARY J. WOMEN & L. 333, 358 (2016).

[146] PENAL § 646.9.

[147] *Id.* § 646.9(a).

[148] *Id.*

[149] *Id.* § 646.9(g)–(h).

[150] *Id.* § 646.9(a)–(c) (providing that the violation of a restraining order or repeat offenses must be charged as a felony).

[151] Atticus N. Wegman, *Cyberbullying and California's Response*, 47 U.S.F. L. REV. 737, 741 (2013).

> [e]very person who, with intent to place another person in reasonable fear for his or her safety, or the safety of the other person's immediate family, by means of an electronic communication device, and without consent of the other person, and for the purpose of imminently causing that other person unwanted physical contact, injury, or harassment, by a third party, electronically distributes, publishes, e-mails, hyperlinks, or makes available for downloading, personal identifying information, including, but not limited to, a digital image of another person, or an electronic message of a harassing nature about another person, which would be likely to incite or produce that unlawful action, is guilty of a misdemeanor punishable by up to one year in a county jail, by a fine of not more than one thousand dollars ($1,000), or by both that fine and imprisonment.[152]

The inclusion of the phrase "a digital image of another person" could be promising for future prosecutions of revenge porn or deepfake cases.

c. Other Relevant Criminal Laws

Although the California Penal Code includes a chapter about false personation,[153] impersonation—digital or otherwise—is not in and of itself a crime in California.[154] For example, section 528.5 of the California Penal Code makes it unlawful to "knowingly and without consent credibly impersonate[] another actual person through or on an Internet Web site, or by other electronic means, *for purposes of . . . defrauding another person*."[155] This cyberimpersonation crime is a misdemeanor, punishable by up to one year in county jail, a fine of up to $1000, or both.[156] Similarly, California Penal Code section 529 is a more general personation statute that makes it a crime to "falsely personate[] another."[157] Section 529, however, still applies if, while "in that assumed character," the impersonator undertakes enumerated fraudulent acts.[158] Finally, this personation chapter of the California Penal Code criminalizes identity theft.[159] Per section 530.5, identity theft involves the willful taking of another person's "identifying information" and the use of that information for "any unlawful purpose."[160]

[152] PENAL § 653.2(a).

[153] *Id.* § 528.

[154] In addition to the crimes identified here, the Office of the Attorney General catalogues other cybercrimes. *See Cyber Exploitation FAQs*, *supra* note 44.

[155] PENAL § 528.5(a) (emphasis added).

[156] *Id.* § 528.5(d).

[157] *Id.* § 529.

[158] *Id.* § 529(a). Illegal acts include those in which the impersonator: (1) "[b]ecomes bail or surety . . . ," (2) signs or verifies "any written instrument, with intent that [it] . . . be recorded, delivered, or used as true," or (3) subjects the person falsely impersonated to liability or penalty or accrues a benefit. *Id.* False personation is a "wobbler," which is "[a] crime that can be charged as either a felony or a misdemeanor." *Wobbler*, BLACK'S LAW DICTIONARY (11th ed. 2019). As a wobbler, it is punishable by up to one year in county jail, a fine of up to $10,000, or both. PENAL § 529(b).

[159] *Id.* § 530.5.

[160] *Id.* § 530.5(a). Identity theft is a wobbler, chargeable as a misdemeanor or felony depending on the facts of the case. *Id.* § 529(b).

d. Proposed Deepfake Legislation

In an effort to address the problem of deepfakes, the California legislature has quickly introduced bills over the last year that apply criminal and civil penalties to the phenomenon.

The first proposal, Assembly Bill 602, was supported by the Screen Actors Guild.[161] As originally drafted, it would have established misdemeanor liability for a person who creates a deceptive recording with the intent to distribute it while knowing that the recording is likely to deceive a person who views it.[162] Subsequent amendments to Assembly Bill 602 deleted the criminal penalties and replaced them with a civil cause of action for harm resulting from the intentional creation and disclosure of sexually explicit material without consent.[163] As amended, Assembly Bill 602 provides civil remedies, including economic and punitive damages, attorney's fees, and injunctive relief.[164] Assembly Bill 602 exempts from liability any person who discloses the material in the course of reporting unlawful activity; in the course of a legal proceeding; in the course of duty, when the discloser is a member of law enforcement; in relation to a matter of legitimate public concern; in a work of political or newsworthy value; or for the purposes of commentary or criticism.[165] The bill also states that sexually explicit material is not newsworthy simply because the depicted individual is a public figure.[166]

Another bill, Assembly Bill 1280, would define a "deepfake" as a recording that has been created or altered so that it falsely appears to be an authentic record of the actual speech or conduct of the individual depicted in the recording.[167] The bill would have created a new crime when a person creates or distributes, without the subject's consent, a deepfake that depicts a person engaging in sexual conduct.[168]

[161] *See* Press Release, SAG-AFTRA, Action Alert: Support California Bill to End Deepfake Porn (Apr. 29, 2019), https://www.sagaftra.org/action-alert-support-california-bill-end-deepfake-porn [https://perma.cc/WQM9-22XF]; *California Legislation Takes First Steps Against Deepfakes*, SAG-AFTRA (May 3, 2019), https://www.sagaftra.org/california-legislation-takes-first-steps-against-deepfakes [https://perma.cc/Y5QB-GU4J].

[162] *See* Assemb. 602, 2019–2020 Leg., Reg. Sess. (Cal. 2019) (as introduced); *Legislature Grapples with Three Bills to Address "Deepfake" Videos*, CAL. NEWS PUBLISHERS ASS'N (Mar. 29, 2019), https://cnpa.com/legislature-grapples-with-three-bills-to-address-deepfake-videos/ [https://perma.cc/88AS-ULXH].

[163] Cal. Assemb. 602; *see also From a Field of Three, One Deepfake Bill Is Left*, CAL. NEWS PUBLISHERS ASS'N, https://cnpa.com/from-a-field-of-three-one-deepfake-bill-is-left/ [https://perma.cc/4SUV-XYKF] (last visited Nov. 12, 2019). A substantially similar proposal, Senate Bill 564, was held in the legislative committee. *See* S. 564, 2019–2020 Leg., Reg. Sess. (Cal. 2019).

[164] Cal. Assemb. 602.

[165] *Id.*

[166] *Id.*

[167] Assemb. 1280, 2019–2020 Leg., Reg. Sess. (Cal. 2019).

[168] *Id.* Assembly Bill 1280 also would criminally prohibit a person from creating or distributing, without the depicted person's consent, a deepfake with the intent that the deepfake coerce or deceive any voter into voting for or against a candidate or measure in an

By the end of the 2019 legislative session, Assembly Bill 1280 remained in policy committees, while Assembly Bill 602 was approved by the legislature and sent to the governor for his consideration.[169] Each has drawn criticism from First Amendment advocacy groups, including the ACLU of California, the California News Publishers Association, and the California Broadcasters Association.[170]

2. Texas

a. Revenge Porn

In 2015, the Texas Legislature enacted the Relationship Privacy Act of 2015[171] to target revenge porn. Effective September 1, 2015, Texas Penal Code section 21.16 criminalizes the "unlawful disclosure or promotion of intimate visual material."[172] The statute provides that a person is guilty of the offense if: (1) the disclosure or promotion is done intentionally and without the depicted victim's consent, (2) the victim had a "reasonable expectation that the visual material would remain private," (3) the disclosure causes harm to the victim, and (4) the victim's identity is revealed "in any manner" by the disclosure.[173] Distribution of intimate visual material is a "Class A" misdemeanor and is punishable by up to one year in prison, a $4000 fine, or both.[174]

election that is occurring within sixty days. *Id.* Another proposed bill, Assembly Bill 730, would specifically prohibit

> a person, committee, or other entity, within 60 days of an election at which a candidate for elective office will appear on the ballot, from distributing *with actual malice 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.

Assemb. 730, 2019–2020 Leg., Reg. Sess. (Cal. 2019) (emphasis added) (legislative counsel's digest).

[169] *AB-602 Depiction of Individual Using Digital or Electronic Technology: Sexually Explicit Material: Cause of Action*, CAL. LEGIS. INFO., http://leginfo.legislature.ca.gov/faces/billHistoryClient.xhtml?bill_id=201920200AB602 [https://perma.cc/5FA6-Z4HH] (last visited Nov. 12, 2019); *AB-1280 Crimes: Deceptive Recordings*, CAL. LEGIS. INFO., http://leginfo.legislature.ca.gov/faces/billHistoryClient.xhtml?bill_id=201920200AB1280 [https://perma.cc/B4M6-NXNW] (last visited Nov. 12, 2019).

[170] Ben Christopher, *Can California Crack Down on Deepfakes Without Violating the First Amendment?*, KQED NEWS (July 3, 2019), https://www.kqed.org/news/11758994/can-california-crack-down-on-deepfakes-without-violating-the-first-amendment [https://perma.cc/MA3Q-JJ7V]; Andrew Sheeler, *California Is Moving to Ban Deepfakes. What Are They, Anyway?*, SACRAMENTO BEE (July 1, 2019, 9:26 PM), https://www.sacbee.com/news/politics-government/capitol-alert/article232162032.html [https://perma.cc/Y375-9QTK].

[171] 2015 Tex. Gen. Laws 8723–26.

[172] TEX. PENAL CODE ANN. § 21.16 (West 2019).

[173] *Id.* § 21.16(b).

[174] *Id.* §§ 12.21, 21.16.

b. Cyberstalking

Texas has only one general anti-stalking law, which makes no explicit reference to cyberstalking.[175] Rather, it broadly provides that an individual is guilty of stalking if that individual knowingly engages in a pattern of conduct directed at another person (or that person's family) that harasses or threatens (or creates a fear of) bodily injury or death or causes the victim to feel harassed, annoyed, alarmed, abused, tormented, embarrassed, or offended and would cause a reasonable person to similarly suffer.[176]

Stalking is a felony of the third degree, punishable by at least two (and no more than ten) years' imprisonment and a fine of up to $10,000.[177] For subsequent violations, this offense increases to a felony of the second degree, punishable by at least two (and no more than twenty) years' imprisonment and a fine of up to $10,000.[178]

Although section 42.072 does not refer to electronic harassment or cyberstalking, a Texas appellate court has interpreted the statute to include electronic communications.[179] Further, section 42.07, the harassment provision mentioned above, criminalizes "repeated electronic communications [sent] in a manner reasonably likely to harass, annoy, alarm, abuse, torment, embarrass, or offend another."[180] This underlying offense is itself a "Class B" misdemeanor, punishable by up to 180 days in jail, a fine of up to $2000, or both.[181] For subsequent violations, violations involving minors, or violations of restraining orders, this offense increases to a "Class A" misdemeanor, punishable by up to one year in prison, a $4000 fine, or both.[182]

c. Other Relevant Criminal Laws

The Texas Legislature addressed internet harassment by enacting a criminal statute specifically targeting online impersonation in 2009.[183] Section 33.07 criminalizes two types of online impersonation.[184] The first category, a felony in the third degree,[185] makes it unlawful to, "without obtaining the other person's consent and with the intent to harm, defraud, intimidate, or threaten any person," use the name or persona of another to create an online web page or post or send messages through social networking or other internet websites, not including email or message

[175] *Id.* § 42.072.

[176] *Id.* § 42.072(a).

[177] *Id.* §§ 12.34, 42.072(b).

[178] *Id.* §§ 12.33, 42.072(b).

[179] Manuel v. State, 357 S.W.3d 66, 83 (Tex. App. 2011).

[180] PENAL § 42.07(a)(7).

[181] *Id.* §§ 12.22, 42.07(c).

[182] *Id.* §§ 12.21, 42.07(c).

[183] *Id.* § 33.07.

[184] *Id.*

[185] *Id.* §§ 12.34, 33.07(c) (punishable by at least two and no more than ten years' imprisonment and a fine of up to $10,000).

boards.[186] The second category, a "Class A" misdemeanor,[187] makes it unlawful to send an email, instant message, text message, or other message referencing "a name, domain address, phone number, or other item of identifying information belonging to any person" without that person's consent and with the intent of causing the recipient of the message to reasonably believe the other person sent or authorized the message and the intent "to harm or defraud any person."[188] The term "identifying information" is defined in section 32.51 as "information that alone or in conjunction with other information identifies a person."[189] Section 32.51 also criminalizes fraudulent use or possession of identifying information.[190]

3. Florida

a. Revenge Porn

In 2015, the Florida Legislature tackled the revenge porn problem and enacted a statute to criminalize "sexual cyberharassment."[191] Sexual cyberharassment is defined in the statute as "publish[ing] . . . a sexually explicit image of a person that contains or conveys the personal identification information of the depicted person [to a website] without the depicted person's consent . . . for no legitimate purpose, with the intent of causing substantial emotional distress to the depicted person."[192] A person who "willfully and maliciously sexually cyberharasses another person" will be guilty of a misdemeanor of the first degree, punishable by up to one year in prison or a fee of up to $1000.[193] Subsequent offenses are charged as a felony of the third degree, punishable by up to five years' imprisonment or a fee of up to $5000.[194]

b. Cyberstalking

Florida's anti-stalking law has notably contained an explicit cyberstalking provision since July 1, 2004.[195] The statutory definition of cyberstalking is to "engage in a course of conduct to communicate, or to cause to be communicated, words, images, or language by or through the use of electronic mail or electronic communication, directed at a specific person . . . causing substantial emotional distress to that person and serving no

[186] *Id.* § 33.07(a).

[187] *Id.* §§ 12.21, 33.07(c) (punishable by up to one year in prison, a $4000 fine, or both).

[188] *Id.* § 33.07(b).

[189] *Id.* § 32.51(a)(1) (including name and date of birth, "unique biometric data," "unique electronic identification" information, "telecommunication identifying information," and "social security number or other government-issued identification number").

[190] *See generally id.* § 32.51.

[191] FLA. STAT. § 784.049 (2019).

[192] *Id.* § 784.049(2)(c).

[193] *Id.* § 784.049(3)(a); *see also id.* §§ 775.082(4)(a), 775.083(1)(d).

[194] *Id.* § 784.049(3)(b); *see also id.* §§ 775.082(3)(e), 775.083(1)(c).

[195] *Id.* § 784.048.

legitimate purpose."[196] The Florida anti-stalking statute is also unique as it has two categories of stalking—stalking and aggravated stalking.[197] A person who "willfully, maliciously, and repeatedly" cyberstalks another is guilty of stalking, a misdemeanor of the first degree.[198] Stalking is punishable by up to one year in prison or a fee of up to $1000.[199] A person who "willfully, maliciously, and repeatedly" cyberstalks another *and* either (1) makes a credible threat to that person, (2) violates a restraining order or other injunction, or (3) targets a minor under the age of sixteen is guilty of aggravated stalking, a felony of the third degree.[200] Aggravated stalking is punishable by up to five years' imprisonment or a fee of up to $5000.[201]

4. New York

a. Revenge Porn

New York is presently among the twelve states that have not criminalized revenge porn.[202] Senate Bill 642 was the third iteration of the New York legislature's attempt at enacting a criminal ban on nonconsensual pornography.[203] However, any momentum behind this bill stopped after it passed in the Senate on March 19, 2018, and was delivered to the Assembly, where it subsequently expired.[204]

b. Cyberstalking

New York's anti-stalking scheme consists of four categories of offenses: stalking in the fourth degree,[205] the third degree,[206] the second degree,[207] and the first degree.[208] The base-level fourth-degree stalking statute does not

[196] *Id.* § 784.048(d).

[197] *Id.* § 784.048(2)–(5).

[198] *Id.* § 784.048(2).

[199] *Id.* §§ 775.082(4)(a), 775.083(1)(d), 784.048(2).

[200] *Id.* § 784.048(3)–(5).

[201] *Id.* §§ 775.082(3)(e), 775.083(1)(c), 784.048(3)–(5).

[202] Douglas Harris, Student Article, *Deepfakes: False Pornography Is Here and the Law Cannot Protect You*, 17 DUKE L. & TECH. REV. 99, 119 (2019) ("Thirty-eight states and the District of Columbia have nonconsensual pornography laws.").

[203] S. 642, 2017–2018 Leg., Reg. Sess. (N.Y. 2017).

[204] *Senate Bill 642*, N.Y. ST. SENATE, https://www.nysenate.gov/legislation/bills/2017/s642 [https://perma.cc/TZ2F-9ZMU] (last visited Nov. 12, 2019). After the revenge porn bill's demise, proponents pointed fingers at Google and other tech lobbyists. *See* Kirstan Conley & Gabrielle Fonrouge, *Google Kills Revenge Porn Bill*, N.Y. POST (June 21, 2018, 2:31 AM), https://nypost.com/2018/06/21/new-yorks-revenge-porn-bill-dies-after-11th-hour-campaign-by-google/ [https://perma.cc/8YSF-Z8V3]; James Hetherington, *What Happened to New York's Revenge Porn Bill?*, NEWSWEEK (June 21, 2018, 1:46 PM), https://www.newsweek.com/what-happened-new-yorks-revenge-porn-bill-google-989666 [https://perma.cc/7ZHY-C8JD].

[205] N.Y. PENAL LAW § 120.45 (McKinney 2019) (a class B misdemeanor).

[206] *Id.* § 120.50 (a class A misdemeanor).

[207] *Id.* § 120.55 (a class E felony).

[208] *Id.* § 120.60 (a class D felony).

explicitly recognize cyberstalking or online harassment.[209] The higher-degree charges increase in severity based on factors such as repeat offenses, the age of the victim, and the use of weapons.[210]

New York also has a separate criminal statutory scheme specifically focused on criminal harassment.[211] Notably, a person who intentionally "engages in a course of conduct or repeatedly commits acts which alarm or seriously annoy such other person and which serve no legitimate purpose" is guilty of harassment in the second degree.[212] However, harassment in the second degree is only a violation punishable by up to fifteen days' imprisonment and a fine of up to $250.[213] Moreover, the aggravated harassment in the second-degree statute specifically criminalizes harassing electronic communication if a threat against one's physical harm, property, or family is made.[214] As a "class A" misdemeanor, aggravated harassment in the second degree is punishable by up to one year in prison and a fine of up to $1000.[215]

c. Other Relevant Criminal Laws

Similar to the California framework, online impersonation is only criminal under New York Penal Law section 190.25 when it is done in conjunction with one of the specific purposes delineated in the statute.[216] Section 190.25 explicitly provides that a person who "[i]mpersonates another by communication by internet website or electronic means" is guilty of criminal impersonation in the second degree if the impersonation is done with the "intent to obtain a benefit or injure or defraud another."[217] As a "class A" misdemeanor, criminal impersonation in the second degree is punishable by up to one year in prison and a fine of up to $1000.[218]

d. Proposed New Civil Rights

On May 31, 2017, Assembly Bill 8155 was introduced in the New York State Assembly.[219] This bill attempted to address the deepfake problem by creating new civil rights and civil remedies

> [e]stablish[ing] the right of privacy and the right of publicity for both living and deceased individuals; provid[ing] that an individual's persona is the personal property of the individual and is freely transferable and

[209] *Id.* § 120.45 (referencing only telephonic communication and the use of global positioning systems to track a person's location).

[210] *Id.* §§ 120.50, 120.55, 120.60.

[211] *See generally id.* §§ 240.25–240.32.

[212] *Id.* § 240.26(3).

[213] *Id.* §§ 70.15(4), 80.05(4), 240.26.

[214] *Id.* § 240.30(1).

[215] *Id.* §§ 70.15(1), 80.05(1), 240.30.

[216] *See id.* § 190.25.

[217] *Id.* § 190.25(4).

[218] *Id.* §§ 70.15(1), 80.05(1), 190.25.

[219] Assemb. 8155, 2017–2018 Leg., Reg. Sess. (N.Y. 2017).

> descendible; provid[ing] for the registration with the department of state of such rights of a deceased individual; and that the use of a digital replica for purposes of trade within an expressive work shall be a violation.[220]

The bill passed the Assembly and was delivered to the New York State Senate on June 18, 2018.[221] The breadth of the initial law drew strong criticism from entertainment companies, including Disney, NBCUniversal, and the Motion Picture Association of America.[222]

IV. LIMITATIONS TO CURRENT CRIMINAL LAW SOLUTIONS

Although other laws (and proposed laws) might be used in a deepfake prosecution, these laws have limitations. This Part explores shortcomings in the current and proposed laws in the context of deepfakes, including First Amendment concerns.

A. Problems with Current Laws and Proposed Legislation

1. Inadequacies of Current Federal and State Criminal Laws on Cybercrime, Stalking, Criminal Threats, and Harassment

As the revenge porn phenomenon emerged, victims, their advocates, and legal scholars were quick to point out that existing criminal laws—both federal and state—were insufficient to punish the creators and distributors and to remedy the harms victims suffered.[223] These inadequacies persist in the face of pornographic deepfakes, which similarly do not fit within current criminal statutory schemes, even those that criminalize revenge porn. A brief survey of stalking, criminal threats, and harassment statutes illustrates how current laws do not address the problems affiliated with nonconsensual pornography, be it revenge porn or pornographic deepfakes.

The federal cyberstalking statute[224] and the anti-stalking statutes of many states, such as California[225] and Texas,[226] include a specific fear requirement. To be found guilty, the accused must have intentionally or recklessly (depending on the jurisdiction) threatened the victim or acted in a way that

[220] *Assembly Bill A8155B*, N.Y. ST. SENATE, https://www.nysenate.gov/legislation/bills/2017/a8155 [https://perma.cc/Y9PF-8PZR] (last visited Nov. 12, 2019) (summarizing the bill).

[221] *Id.*

[222] Katyanna Quach, *New York State Is Trying to Ban 'Deepfakes' and Hollywood Isn't Happy*, REGISTER (June 12, 2018, 10:22 PM), https://www.theregister.co.uk/2018/06/12/new_york_state_is_trying_to_ban_deepfakes_and_hollywood_isnt_happy/ [https://perma.cc/33HL-8X2P]; *see also* Memorandum from the Motion Picture Ass'n of Am., Inc. (June 8, 2018), https://www.rightofpublicityroadmap.com/sites/default/files/pdfs/mpaa_opposition_to_a8155b.pdf [https://perma.cc/L6DV-TF4H]; Letter from Lisa Pitney, Vice President, Walt Disney Co., to Senator Martin Golden (June 8, 2018), https://www.rightofpublicityroadmap.com/sites/default/files/pdfs/disney_opposition_letters_a8155b.pdf [https://perma.cc/TBL6-V5U5].

[223] *See, e.g.*, Brady, *supra* note 107, at 12; Citron & Franks, *supra* note 41, at 365.

[224] 18 U.S.C. § 2261A (2012).

[225] CAL. PENAL CODE § 646.9 (West 2019).

[226] TEX. PENAL CODE ANN. § 42.072 (West 2019).

made the victim fear bodily harm or death for themselves or their families.[227] Similarly, other statutes include an element requiring proof of the victim's "emotional distress."[228] The laws that proscribe "criminal threats" may also fail to address deepfakes because criminal threat statutes require proof of a communication of a threat to kidnap or injure the victim.[229] Even though deepfakes may injure victims, they do not necessarily constitute a threat of injury. And although many victims of nonconsensual pornography suffer from emotional distress[230] and genuinely fear for their safety, especially where their personal information is also shared,[231] these subjective elements can be difficult, if not impossible, to prove at trial.

Further, these requirements trivialize the harm inherent in nonconsensual pornography. By requiring prosecutors to prove these subjective elements, the message is that the public display of victims' bodies engaged in revealing or sexually explicit behavior without their consent is insufficient, standing alone, to warrant the law's attention.[232] Moreover, many stalking and harassment statutes, such as those of Texas[233] and Florida,[234] require a pattern, course of conduct, or multiple incidents.[235] Thus, a single posting, even if it goes viral, would not qualify as a crime under these legal frameworks.[236]

Impersonation laws similarly fail to address many nonconsensual pornography scenarios. As an initial matter, these laws apply where the depiction at issue was shared or posted in a way that made it seem like the victim was the poster.[237] Clearly, this is not always the case, especially with

[227] *Compare* 18 U.S.C. § 2261A(1)(A) (applying to conduct that "places [a] person in reasonable fear of the death of, or serious bodily injury to (i) that person, (ii) an immediate family member . . . of that person, or (iii) a spouse or intimate partner of that person"), *with* TEX. PENAL CODE ANN. § 42.072 (West 2019) (prohibiting threats of "(A) bodily injury or death for the other person; (B) bodily injury or death for a member of the other person's family or household or for an individual with whom the other person has a dating relationship; or (C) that an offense will be committed against the other person's property").

[228] *See, e.g.*, FLA. STAT. § 784.048(1)(d) (2019).

[229] *See* 18 U.S.C. § 875(c).

[230] Citron & Franks, *supra* note 41, at 351 (reporting that "80% of revenge porn victims experience severe emotional distress and anxiety").

[231] *Id.* at 350–51 ("Nonconsensual pornography raises the risk of offline stalking and physical attack. In a study of 1,244 individuals, over 50% of victims reported that their naked photos appeared next to their full name and social network profile; over 20% of victims reported that their e-mail addresses and telephone numbers appeared next to their naked photos.").

[232] *See id.* at 350–54 (explaining the damage to victims of revenge porn).

[233] TEX. PENAL CODE ANN. § 42.072 (West 2019).

[234] FLA. STAT. § 784.048(1)(d) (2019).

[235] *See* Citron & Franks, *supra* note 41, at 350–54 (describing revenge porn's damage).

[236] *See* Linkous, *supra* note 86, at 24–25.

[237] *See* Allison Greene, Note, *The Ill of Misogyny on the Internet: Why Revenge Porn Needs Federal Criminalization*, 16 COLO. TECH. L.J. 175, 194 (2017) (discussing the intersections between impersonation crimes and revenge porn); *see also* Snehal Desai, Note, *Smile for the Camera: The Revenge Pornography Dilemma, California's Approach, and Its Constitutionality*, 42 HASTINGS CONST. L.Q. 443, 464 (2015) ("[I]f a perpetrator simply posts a picture online without impersonating the victim or giving out her information, he could avoid

regard to deepfakes which gained initial popularity on Reddit with depictions of celebrities in pornographic videos.[238] Further, in situations where there is an element of impersonation, as seen in the state impersonation laws discussed above,[239] online or digital impersonation is not inherently criminal; it is the use of impersonation to perform an unlawful task, such as defrauding a third party or creating liability for the impersonated individual, that is criminalized.[240] These laws will do little to protect victims of nonconsensual pornography because there is rarely a secondary criminal intent behind the posting or sharing of revenge porn or deepfakes.

Identity theft can be a promising option given the frequency with which personal identifying information accompanies nonconsensual pornography posts.[241] The California case of Kevin Bollaert is illustrative. Bollaert created and managed ugotposted.com, a revenge porn website, and its companion page, changemyreputation.com, which charged victims fees of $250 to $350 to have their images removed from the revenge porn page.[242] Bollaert was prosecuted and convicted of identity theft and extortion.[243] Although the website was shut down and Bollaert received an eighteen-year prison sentence, he notably "was not prosecuted for a crime directly related to revenge porn, instead he was charged for his conduct in requiring payment to remove the photograph."[244] Accordingly, although identity theft can be useful in some nonconsensual pornography cases, these laws fail to criminalize the underlying conduct—the *posting* of nonconsensual pornography in the first place.

Similarly, computer crimes, such as unauthorized access or hacking,[245] apply in limited circumstances where the pornographic material or images used in the creation of a pornographic deepfake are obtained via unauthorized access to the victim's computer or email.[246] Again, while prosecution under these laws may result in a desirable outcome, the laws do not focus on the true harm of sharing revenge porn or pornographic deepfakes nor do they have broad applications.

These federal and state laws might be an option to prosecute deepfakes, but they are not ideal—they do not expressly apply to deepfakes. And regardless of the reprehensibility of the conduct, judges might be reluctant to stretch these laws past their plain language. This is illustrated by the Massachusetts Supreme Judicial Court's recent holding that "upskirting" is legal as long as the person being photographed is not nude or partially

prosecution for false personation and unauthorized electronic distribution of personal information.").

[238] *See* Hawkins, *supra* note 18.

[239] CAL. PENAL CODE § 528.5(a) (West 2019); N.Y. PENAL LAW § 190.25 (McKinney 2019); TEX. PENAL CODE ANN. § 33.07 (West 2019).

[240] *See* sources cited *supra* note 239.

[241] *See* Citron & Franks, *supra* note 41, at 351.

[242] People v. Bollaert, 203 Cal. Rptr. 3d 814, 819 (Ct. App. 2016).

[243] *Id.*

[244] Brady, *supra* note 107, at 20–21.

[245] 18 U.S.C. § 1030 (2012).

[246] Groban, *supra* note 99, at 16.

nude.[247] Judges' unwillingness to broaden statutes further highlights the need for specific laws targeting deepfakes.

2. Shortcomings of Actual and Proposed Revenge Porn Laws

Most states have adopted legislation specifically targeting revenge porn to address these insufficiencies.[248] Those laws, often imperfect even in the revenge porn context, are not a fix for pornographic deepfakes.

a. State Statutes

As the criminalization of revenge porn has become increasingly popular across the nation, these new laws have been subject to scrutiny. Three common and particularly salient critiques stand out. First, many revenge porn statutes, including those in California, Texas, and Florida, incorporate an intent requirement.[249] Requiring prosecutors to prove a particular mens rea is difficult.[250] Moreover, it "miss[es] the point"[251]: the wrongdoing is "the disclosure of someone's naked photographs without the person's consent and in violation of their expectation that the image be kept private"; the motivation for doing so is irrelevant.[252] A second issue is the inclusion of a harm requirement; many states, including California, require a showing that the victim suffered actual harm or emotional distress.[253] This element is again difficult to prove in court and it "forces victims to expose even more of their private lives to the public."[254] Finally, the penalties affiliated with many revenge porn statutes are weak.[255]

[247] *See* Commonwealth v. Robertson, 5 N.E.3d 522, 528 (Mass. 2014) (holding that it is legal to secretly photograph underneath a person's clothing when the person is not nude or partially nude); *see also* Haimy Assefa, *Massachusetts Court Says 'Upskirt' Photos Are Legal*, CNN (Mar. 6, 2014), https://www.cnn.com/2014/03/05/us/massachusetts-upskirt-photography/index.html [https://perma.cc/943Q-R2EA].

[248] *46 States + DC + One Territory Now Have Revenge Porn Laws*, CYBER C.R. INITIATIVE, https://www.cybercivilrights.org/revenge-porn-laws/ [https://perma.cc/V276-JCYC] (last visited Nov. 12, 2019).

[249] *See, e.g.*, CAL. PENAL CODE § 647(j)(4) (West 2019); FLA. STAT. § 784.049 (2019); TEX. PENAL CODE ANN. § 21.16 (West 2019).

[250] Diane Bustamante, Comment, *Florida Joins the Fight Against Revenge Porn: Analysis of Florida's New Anti-Revenge Porn Law*, 12 FIU L. REV. 357, 387 (2017) ("Prosecutors have to 'prove beyond a reasonable doubt that the defendant posted the photo to intentionally hurt the victim,' and not because he just thought it was 'profitable,' 'just for fun,' to annoy, or any other reason the offender can come up with." (quoting Haley Fox, *Why Revenge Porn Laws May Not Protect Women*, TAKEPART (Dec. 2, 2014), http://www.takepart.com/article/2014/12/02/revenge-porn-protections/ [https://perma.cc/V6ZZ-Z8V4])).

[251] *Id.*

[252] Citron & Franks, *supra* note 41, at 387.

[253] *See, e.g.*, CAL. PENAL CODE § 647(j)(4) (West 2019).

[254] Christian Nisttáhuz, Comment, *Fifty States of Gray: A Comparative Analysis of "Revenge-Porn" Legislation Throughout the United States and Texas's Relationship Privacy Act*, 50 TEX. TECH L. REV. 333, 358 (2018).

[255] The statutes in California, Florida, and Texas all qualify first-time offenses without any aggravating circumstances as mere misdemeanors, with punishments ranging from just six months to a year in prison and $1000 to $4000 fines. *See* CAL. PENAL CODE § 647(j)(4), (*l*)

b. The ENOUGH Act

Beyond these critiques, revenge porn laws—both current state laws and the proposed federal ENOUGH Act—will not assist in the case of deepfakes for four key reasons.

First, the ENOUGH Act, like the laws enacted by many states including California, Texas, and Florida,[256] requires that there be a victim who (1) is either engaged in sexually explicit conduct or whose intimate body parts are exposed, and (2) is identifiable to some degree in the visual depiction.[257] These are problematic requirements for deepfake cases because victims are not easily identifiable and because revenge porn inherently creates two victims, not one. The victim whose face is shown will likely be more readily identified, but it is the other victim whose intimate body parts are exposed. Both individuals are indeed victims, yet revenge porn laws do not completely cover either.[258]

A second issue is that of consent. The ENOUGH Act and state revenge porn laws require that the sharing or posting be done without the victim's consent.[259] In a typical revenge porn scenario, the victim often initially consents to the creation of the image or visual within the context of a specific relationship. The issue then arises when the image or visual is exposed in another context, one in which the victim did not provide consent.[260] Consent is even more amorphous when it comes to deepfakes. Whose consent is required? The victim whose body is shown or the victim whose face is shown or both? What would consent cover: the creation or distribution of the deepfake? The specific use of their image or likeness? The underlying videos or images used to create the deepfake? The victim whose body is displayed may have consented to the creation and distribution of the original, underlying video or image; however, they most likely did not consent to the superimposition of another person's face onto their body. Similarly, the victim whose face is shown likely consented to the underlying photographs or videos in one context, perhaps as a post on social media, but not for the use in a deepfake.

Third, the ENOUGH Act and the laws of states like Texas and California require that the victim had a reasonable expectation of privacy regarding the depiction.[261] Again, this element is relatively straightforward in a typical revenge porn context where one partner in a relationship shared an intimate image with the other, operating under the assumption that the image would

(West 2019); FLA. STAT. §§ 775.082(4)(a), 775.083(1)(d), 784.049(3)(a) (2019); TEX. PENAL CODE ANN. §§ 12.21, 21.16 (West 2019).

[256] *See, e.g.*, CAL. PENAL CODE § 647(j)(4) (West 2019); FLA. STAT. § 784.049 (2019); TEX. PENAL CODE ANN. § 21.16 (West 2019).

[257] *See* S. 2162, 115th Cong. (2017). The ENOUGH Act also shares similarities with some revenge porn state statutes. As in some states, the ENOUGH Act does not include a harm requirement but does have an intent requirement. *Id.*

[258] *See* Harris, *supra* note 202, at 122.

[259] *See* sources cited *supra* note 256.

[260] *See* Citron & Franks, *supra* note 41, at 345–48.

[261] S. 2162; *see supra* Part III.B.

not be shared outside the relationship. In the deepfake context, this element does not exist and cannot be shown; neither victim would ever have an expectation of privacy in the deepfake.

Fourth, the ENOUGH Act excluded simulated acts from the definition of "intimate visual depiction."[262] This provision directly contradicted the Act's incorporation of the definition of "sexually explicit conduct" set forth in 18 U.S.C. § 2256, which includes simulated conduct.[263] This provision would presumably preclude criminal liability for deepfakes, which are necessarily simulated because they are fake.

In sum, because identification of victims, initial consent, and expectation of privacy are all either implicit or explicit elements of the existing statutes criminalizing revenge porn, and because these elements either are nonexistent (in the case of privacy or consent) or would not be easily proven (in the case of identification of all victims), existing and proposed revenge porn laws cannot be used to prosecute deepfakes. New legislation focused on deepfakes is warranted.

3. Shortcomings of the Malicious Deep Fake Prohibition Act of 2018

Although pornographic deepfakes must be criminalized, the Malicious Deep Fake Prohibition Act of 2018 is an inadequate solution. As an initial matter, the MDFPA is overbroad in many respects. First, the MDFPA's definition of a "deepfake" is extremely broad, including any "audiovisual record created or altered in a manner that the record would falsely appear to a reasonable observer to be an authentic record of the actual speech or conduct of an individual."[264] This definition, without any limitations or modifiers, casts too broad a net; it seemingly encompasses a wide range of media, including legitimate, nonoffensive content like computer-generated imagery in films. To counteract this broad definition, the MDFPA includes an equally overbroad exemption for First Amendment speech.[265] As written, the MDFPA would allow nearly every deepfake to be subject to a parody or satire defense. Moreover, the liability placed on distributors is overly expansive and runs the risk of sparking a reactive sweep of potentially problematic content without giving careful consideration to legitimate content.[266]

[262] S. 2162.

[263] 18 U.S.C. § 2256 (2012) ("'[S]exually explicit conduct' means actual or simulated— (i) sexual intercourse, including genital-genital, oral-genital, anal-genital, or oral-anal, whether between persons of the same or opposite sex; . . . or (v) lascivious exhibition of the genitals or pubic area of any person . . . .").

[264] S. 3805, 115th Cong. (2018).

[265] *Id.* ("No person shall be held liable under this section for any activity protected by the First Amendment to the Constitution of the United States.").

[266] *See* Kaveh Waddell, *Lawmakers Plunge into "Deepfake" War*, AXIOS (Jan. 31, 2019), https://www.axios.com/deepfake-laws-fb5de200-1bfe-4aaf-9c93-19c0ba16d744.html [https://perma.cc/KRH5-ZRWE] ("Danielle Citron, a University of Maryland law professor and co-author of a landmark law article on deepfakes, says the bill places over-broad liability on distributors. She says it could scare platforms into immediately taking down everything that's reported as a deepfake—potentially deleting legitimate posts in the process.").

Further, the MDFPA only criminalizes deepfakes created with the intention that the distribution would "facilitate criminal or tortious conduct."[267] In other words, under the MDFPA, conduct is only criminal if the distribution is done with the intent to violate some other criminal or tort law. By conditioning criminal sanctions for deepfakes on proof that the creator or distributor violated another law, the MDFPA is both under-[268] and overinclusive[269] and adds a layer of complication to the prosecutor's job. At the same time, as illustrated above, the law does not adequately protect victims of nonconsensual pornography. Thus, the MDFPA may fail to criminalize pornographic deepfakes where no secondary criminal or tortious conduct exists. The creation or distribution of a pornographic deepfake, standing alone, deserves criminal punishment without requiring intent to violate another law.

Furthermore, the MDFPA was not written with the goal of protecting pornographic deepfake victims in mind. It focuses on the implications of politicized deepfakes.[270] Election tampering, fake news, and other perceived harms to democratic institutions rightly deserve attention; these concerns, however, also draw more political concern and resources. Conversely, the victims of nonconsensual pornographic, including deepfakes, have not been afforded the same concern or given the same priority.[271] Although pornographic deepfakes involving celebrities first drew attention to this new technology, its potential victims are not just the Scarlett Johanssons of the world.[272] Historically, the law, law enforcement, and even society have

[267] S. 3805.

[268] Under the MDFPA, a creator or distributor of a pornographic deepfake would escape federal criminal liability under the proposed statute where, irrespective of the harm caused by the conduct, it does not violate another law. *Id.* Given the inadequacy of current civil remedies and criminal law to address deepfakes and similar forms of nonconsensual pornography and the challenges of proof where such remedies exist, the likelihood of a successful prosecution under the MDFPA for pornographic deepfakes is slim.

[269] Under the MDFPA, seemingly harmless deepfakes could ensnare their creators and distributors in criminal liability where the conduct violates another statute. *See* Orin S. Kerr, *Should Congress Pass a "Deep Fakes" Law?*, VOLOKH CONSPIRACY (Jan. 31, 2019, 6:05 PM), https://reason.com/volokh/2019/01/31/should-congress-pass-a-deep-fakes-law [https://perma.cc/6M8U-G7MA] ("Consider an example of how the broad language might work with Senator Sasse's bill. Imagine Sally creates a deepfake in which she imposes her own face on a video clip of President Trump at a political rally. Sally thinks her clip is hilarious, as it really looks like Sally is President. The video is so funny that Sally wants to show it to her friends. She decides to throw a party with a live band at which she will hand out copies of her hilarious deepfakes video. The live band is very loud, however. It's so loud that the party is tortious under state law, as it's a private nuisance to her neighbors.").

[270] *See* Ben Sasse, *This New Technology Could Send American Politics into a Tailspin*, WASH. POST (Oct. 19, 2018, 5:26 PM), https://www.washingtonpost.com/opinions/the-real-scary-news-about-deepfakes/2018/10/19/6238c3ce-d176-11e8-83d6-291fcead2ab1_story.html [https://perma.cc/SL33-27K8].

[271] *See* Citron, *supra* note 92, at 402–04.

[272] Sean Hollister, *Scarlett Johansson Slams Deepfakes, Says She Can't Stop the Internet from Pasting Her Face on Porn*, VERGE (Dec. 31, 2018, 5:30 PM), https://www.theverge.com/2018/12/31/18163351/scarlett-johansson-slams-deepfakes-internet-lost-cause [https://perma.cc/SMK2-3H3P].

failed to prioritize and protect the victims of nonconsensual pornography.[273] Thus, just as revenge porn prompted the development of its own set of laws, pornographic deepfakes need to be addressed through a separate criminal statute.

B. First Amendment Concerns

Immediately after celebrity-based pornographic deepfakes emerged in late 2017 and went viral on the internet, legal scholars and journalists raised the alarm that this conduct implicated the First Amendment protections afforded to online content.[274] Speech and expressive conduct reflexively obtain First Amendment free speech protection; the government may not "restrict expression because of its message, its ideas, its subject matter, or its content."[275] Speech that is merely offensive or distasteful is nonetheless protected as a free expression as well.[276]

Similarly, attempts to criminalize revenge porn across the states and at the federal level have been met with First Amendment challenges and concerns.[277] Thus, to survive constitutional challenge, legislation targeting revenge porn has been narrowly tailored to avoid encompassing legitimate and protected, albeit objectionable, speech.[278]

The same concerns appear in the deepfake context.[279] Indeed, the U.S. Supreme Court has explicitly held that false speech is protected under the First Amendment.[280] Further, given the digitalized nature of deepfakes, there is an added layer of concern about parody and satire. These apprehensions are legitimate, and deepfake legislation must recognize the risk posed by content-based restrictions. However, there are ways to consider First Amendment free speech concerns while also safeguarding victims of nonconsensual pornography, such as deepfakes.[281]

[273] *See* Citron, *supra* note 92, at 402–04.

[274] Farokhmanesh, *supra* note 61 ("Getting the content removed could be a possible First Amendment violation. 'All content is presumptively protected by the First Amendment,' [law professor Eric] Goldman says.").

[275] United States v. Alvarez, 567 U.S. 709, 716 (2012) (quoting Ashcroft v. ACLU, 535 U.S. 564, 573 (2002)).

[276] Virginia v. Black, 538 U.S. 343, 358 (2003) ("The hallmark of the protection of free speech is to allow 'free trade in ideas'—even ideas that the overwhelming majority of people might find distasteful or discomforting." (quoting Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting))); Texas v. Johnson, 491 U.S. 397, 414 (1989) ("If there is a bedrock principle underlying the First Amendment, it is that the government may not prohibit the expression of an idea simply because society finds the idea itself offensive or disagreeable.").

[277] *See* Citron & Franks, *supra* note 41, at 374–77.

[278] *Id.*

[279] Chesney & Citron, *supra* note 4 (manuscript at 31–33).

[280] *See Alvarez*, 567 U.S. at 721–22.

[281] Criminal laws are often used to protect privacy, including criminal laws against unauthorized disclosures of private financial or medical information as well as laws against trespass and voyeurism. *See* 18 U.S.C. §§ 1028, 1801 (2012); 42 U.S.C. § 1320d-6 (2012). If criminal laws protecting financial, medical, and other forms of privacy can be compatible with the First Amendment, criminal laws protecting pornographic deepfakes should be as well.

One solution is to include an explicit public interest exception in a statute criminalizing pornographic deepfakes.[282] Several states have included provisions "guard[ing] against the criminalization of disclosures concerning matters of public interest" in their anti–revenge porn laws.[283] Allowing for the distribution of deepfakes or revenge porn that would otherwise be criminalized when it pertains to a legitimate matter of public concern can alleviate First Amendment issues.[284]

Another option is to draft legislation that defines and characterizes deepfakes as unprotected speech—namely, in this context, obscenity.[285] The Supreme Court has firmly established that obscene material is not protected under the First Amendment.[286] The test to determine whether the content is obscene is:

> (a) whether "the average person, applying contemporary community standards" would find that the work, taken as a whole, appeals to the prurient interest; (b) whether the work depicts or describes, in a patently offensive way, sexual conduct specifically defined by the applicable state law; and (c) whether the work, taken as a whole, lacks serious literary, artistic, political, or scientific value.[287]

The Court in *Miller v. California*[288] even provided examples of regulable obscenity, such as "[p]atently offensive representations or descriptions of ultimate sexual acts, normal or perverted, actual or simulated" or "[p]atently offensive representations or descriptions of masturbation, excretory functions, and lewd exhibition of the genitals."[289] Thus, a law targeted at pornographic deepfakes could likely survive a First Amendment challenge if the regulated content is defined within the constraints of obscenity.

V. PROPOSED SOLUTIONS

This Part focuses on potential solutions to the problem of pornographic deepfakes, both legal and extralegal. Deepfakes, like revenge porn, inflict the harm of sexual objectification without consent. They violate a person's "expectation that all aspects of sexual activity should be founded on consent."[290] Thus, victims of pornographic deepfakes, like victims of revenge porn, need concrete solutions for holding creators of content liable, policing distributors, and removing offensive content from the internet. There is a clear need for federal criminal law to provide solutions for victims.

[282] *See* Citron & Franks, *supra* note 41, at 388.
[283] *Id.*
[284] *Id.*
[285] *See id.* at 384–85.
[286] Miller v. California, 413 U.S. 15, 23 (1973).
[287] *Id.* at 24 (citation omitted) (quoting Kois v. Wisconsin, 408 U.S. 229, 230 (1972)).
[288] 413 U.S. 15 (1973).
[289] *Id.* at 25.
[290] *See* Dodge & Johnstone, *supra* note 13, at 4.

A. Criminalization Through Federal Law Rather Than State Law

A federal law criminalizing pornographic deepfakes would provide a strong and effective disincentive to their creation and distribution.[291] The slow, uneven efforts to criminalize revenge porn at the state level over the last decade[292] demonstrate that waiting for the states to outlaw deepfakes will take too long as the technology becomes more sophisticated and more accessible.[293] Federal criminalization of deepfakes is warranted for several reasons. First, internet activities cross jurisdictional boundaries and involve interstate and international communications; because they appear on the internet, deepfakes are not a local crime confined to a single state or local jurisdiction. A deepfake creator may reside in one state, while the creator's victims live in another and the deepfake may be disseminated on the internet. Thus, a pornographic deepfake, like any other internet crime, is by its nature an offense that is beyond the jurisdictional limits of any single state.[294]

Second, because creating pornographic deepfakes is a crime lacking jurisdictional boundaries, criminalization should be national, uniform, and consistent everywhere—the punishment imposed and remedies provided should not depend on the state in which the victims or perpetrators reside. A federal criminal statute would ensure that victims are protected in states that refuse to act or are slow to do so. Moreover, victims may have trouble persuading state law enforcement to help them if, for example, the deepfake creator resides in another jurisdiction, especially one that does not criminalize the activity. In that instance, federal criminal law would ensure that authorities understand that this behavior is against the law and deserves attention.

Third, federal law is necessary because state laws are constrained by section 230 of the CDA, which impedes state actions against website operators who host nonconsensual pornography; the immunity that section 230 provides for internet service providers and other content distributors does not apply to violations of federal criminal law.[295] Fourth, criminalizing

[291] As Mary Anne Franks and Danielle Keats Citron have observed in encouraging the criminalization of revenge porn, "[n]onconsensual pornography's rise is surely related to the fact that malicious actors have little incentive to refrain from such behavior." *See* Citron & Franks, *supra* note 41, at 361.

[292] *See* Linkous, *supra* note 86, at 36–37 (urging the federal criminalization of revenge porn).

[293] *See* Bloomberg, *How Faking Videos Became Easy—and Why That's so Scary*, FORTUNE (Sept. 11, 2018), http://fortune.com/2018/09/11/deep-fakes-obama-video/ [https://perma.cc/64MB-JGG9]; Donie O'Sullivan et al., *When Seeing Is No Longer Believing: Inside the Pentagon's Race Against Deepfake Videos*, CNN BUS., https://www.cnn.com/interactive/2019/01/business/pentagons-race-against-deepfakes/ [https://perma.cc/6ASD-WK3Y] (last visited Nov. 12, 2019).

[294] *See* Robert L. Ullmann & David L. Ferrera, *Crime on the Internet*, BOS. B.J., Nov./Dec. 1998, at 4, 4 (stating that most internet crime involves interstate or international communications); Anne E. Hawley, Comment, *Taking Spam out of Your Cyberspace Diet: Common Law Applied to Bulk Unsolicited Advertising via Electronic Mail*, 66 UMKC L. REV. 381, 385 (1997) (discussing the inadequacy of state legislation in controlling spam because internet activities cross jurisdictional boundaries).

[295] *See* 47 U.S.C. § 230(e) (2012).

pornographic deepfakes as a federal crime brings to bear the greater resources of the federal government, including the prosecutorial power of the Department of Justice and the investigative expertise of the FBI.

Finally, criminalizing deepfakes at the federal level adds gravitas to the situation and shines a spotlight on its harms—it demonstrates that the problem is of national concern and signals the seriousness of the damage to victims. As recent efforts by internet service providers to remove child pornography and police other dangerous or offensive content have shown, criminalizing deepfakes under federal law will encourage search engines and interactive computer services to voluntarily to address the problem.[296]

B. Proposing a Federal Criminal Statute: The Pornographic Deepfake Criminalization Act

To best address the mounting dangers of pornographic deepfakes, legislative action is needed. This section proposes a new federal criminal statute to regulate the creation and distribution of pornographic deepfakes. By blending components of the ENOUGH Act and the MDFPA and adding other necessary elements, this law seeks to combat pornographic deepfakes and protect victims. The proposed statute is set forth first and then analyzed below.

1. Proposed Statutory Text

TITLE.—Pornographic Deepfake Criminalization Act

(a) FINDINGS.—The legislative body finds that—

(1) the technology necessary to create deepfakes is publicly available and requires no technical training, resulting in the quick proliferation of deepfakes;

(2) pornographic deepfakes, like revenge pornography, are a form of nonconsensual pornography;

(3) victims of nonconsensual pornography can suffer serious emotional and psychological harms and can fear for their physical safety; and

(4) nonconsensual pornography disproportionately victimizes women and girls.

(b) PURPOSE.—The purposes of this statute are to—

(1) criminalize the nonconsensual creation and distribution of pornographic deepfakes;

(2) protect victims, and provide victims with adequate remedies; and

(3) hold content creators and service providers accountable.

(c) DEFINITIONS.—In this section—

[296] Alanna Petroff, *Google, Microsoft Move to Block Child Porn*, CNN BUS. (Nov. 18, 2013, 9:10 AM), https://money.cnn.com/2013/11/18/technology/google-microsoft-child-porn/index.html [https://perma.cc/6V7F-JFQM].

(1) AUDIOVISUAL RECORD.—The term "audiovisual record" means—

(A) any audio or visual media in an electronic format; and

(B) any photograph, motion picture, film, video recording, electronic image, or sound recording.

(2) DEEPFAKE.—The term "deepfake" means an audiovisual record created or altered in a manner that the record would falsely appear to a reasonable observer to be an authentic record of the actual speech, conduct, image, or likeness of an individual.

(3) INDIVIDUAL.—The term "individual" refers to either a person whose body is depicted or a person whose face is depicted in the deepfake and the term "individuals" refers to all depicted persons.

(4) PORNOGRAPHIC DEEPFAKE.—The term "pornographic deepfake" means any deepfake (as defined in paragraph (2))—

(A) in which—

(i) either individual is, or is depicted to be, engaging in sexually explicit conduct; or

(ii) the naked genitals or post-pubescent female nipples of any individual are visible;

(B) in an original or modified format, such as with a filter or text overlay.

(5) SEXUALLY EXPLICIT CONDUCT.—The term "sexually explicit conduct" has the meaning given in section 2256(2)(A).

(6) INTERACTIVE COMPUTER SERVICE.—The term 'interactive computer service' has the same meaning given that term in Section 230 of the Communications Act of 1934 (47 U.S.C. § 230).

(7) INFORMATION CONTENT PROVIDER.—The term 'information content provider' has the same meaning given that term in Section 230 of the Communications Act of 1934 (47 U.S.C. § 230).

(d) OFFENSE.—Except as provided in subsection (g), it shall be unlawful to knowingly use any means or facility of interstate or foreign commerce to distribute or create a pornographic deepfake—

(1) with knowledge or reckless disregard for—

(A) the lack of consent of the individual or individuals to the use of their likeness or image in the creation or distribution of the pornographic deepfake; and

(B) the harm that the distribution could cause to the individual or individuals;

(2) without an objectively reasonable belief that such distribution touches upon a matter of public concern.

(e) PENALTY.—Any person who violates subsection (d) shall be—

(1) fined under this title, imprisoned for no more than 5 years, or both;

(2) fined under this title, imprisoned for no more than 10 years, or both, in the case of a second or subsequent violation; or

(3) fined under this title, imprisoned for no more than 10 years, or both, in the case of a violation where any individual depicted in the deepfake is a minor.

(f) REMEDIES.—Remedies shall be in the court's discretion and shall promote the purposes set forth in subsection (b). Potential remedies include, but are not limited to—

(1) issuance of an order to destroy the audiovisual record;

(2) issuance of an order compelling an interactive computer service or information content provider to remove or take down the audiovisual record;

(3) issuance of a temporary or permanent injunction to prevent the further distribution of the deepfake;

(4) reasonable attorney's fees and costs;

(5) damages or restitution to the victims; and

(6) any other relief that the court deems to be proper.

(g) EXCEPTIONS.—

(1) LAW ENFORCEMENT AND OTHER LEGAL PROCEEDINGS.—This section—

(A) does not prohibit any lawful law enforcement, correctional, or intelligence activity;

(B) shall not apply in the case of an individual reporting unlawful activity in good faith; and

(C) shall not apply in the case of a document production or filing associated with a legal proceeding.

(2) SERVICE PROVIDERS.—This section shall not apply to any provider of a communication service with regard to content provided by another information content provider unless the provider of the communications service intentionally solicits, or distributes knowingly or with reckless disregard, content that is in violation of this section.

(h) THREATS AND EXTORTION.—Any person who intentionally threatens to commit an offense under subsection (d), regardless of whether the threat is an act of extortion, shall be punished as provided in subsection (e).

(i) VENUE AND EXTRATERRITORIALITY.—A prosecution under this section may be brought in a district where the defendant or either depicted individual resides or in a district where the pornographic deepfake is distributed or made available. There is extraterritorial federal jurisdiction over an offense under this section if the defendant or depicted individual is a citizen or permanent resident of the United States.

2. Analyzing the Pornographic Deepfake Criminalization Act

This proposed statute seeks to strike a balance between protecting victims, punishing wrongdoers, and protecting freedom of expression. Subsections (a) through (c) set forth the context, purpose, and framework of this law.

Unlike the MDFPA, the new law proposed here is narrowly tailored to address only *pornographic* deepfakes and, in contrast to the ENOUGH Act, it explicitly covers *simulated* nonconsensual pornography.[297] By framing the law within the context of nonconsensual pornography and highlighting the disproportionate harms caused to women, the law's purpose is made clear and the need for regulation is underscored.[298]

Further, the definitions clarify the law's reach and purpose. Seeking to build upon the definition of deepfakes set forth in the MDFPA while avoiding overbreadth problems, this statute adds the modifier "pornographic" to "deepfake."[299] By focusing on deepfakes portraying "sexually explicit conduct,"[300] the scope of the law is further limited. This constraint is crucial to the law's success in protecting the victims of nonconsensual pornography while also reducing the risk of the law's misuse to prosecute harmless deepfakes. The statute also carefully incorporates *both* victims of deepfakes in its definition of affected persons. This protects both individuals whose likeness and image have in some way been manipulated and shared without their consent.[301]

Subsection (d) begins with a federal jurisdictional requirement[302] and sets forth the offense specifically criminalized by the statute—the creation or distribution of a pornographic deepfake with knowledge or reckless disregard for the lack of consent of either victim to the use of their likeness or image in the deepfake and the potential harms caused to the victim. Intentional threats to commit this offense are also included under subsection (h). Accordingly, this offense applies to those who create or distribute pornographic deepfakes and broadens the requisite mens rea to punish those who act with the knowledge or recklessness.[303] An intent requirement is not included lightly. On the one hand, as has been argued in the context of revenge porn, a requirement that the actor disclose or distribute an image with the specific intent to harm or harass may leave victims unprotected and

[297] *See supra* Part IV.A.2.b (noting that pornographic deepfakes do not fit within the framework of the ENOUGH Act); *see also supra* Part IV.A.3 (discussing overbreadth concerns pertaining to the MDFPA).

[298] *See supra* Part I.C (discussing deepfakes in the nonconsensual pornography context and the harms to victims).

[299] *See supra* Part IV.A.3 (discussing overbreadth concerns pertaining to the MDFPA).

[300] 18 U.S.C. § 2256(2)(A) (2012), which is incorporated in this proposed law, defines "sexually explicit conduct" as "actual or simulated—(i) sexual intercourse, including genital-genital, oral-genital, anal-genital, or oral-anal, whether between persons of the same or opposite sex; (ii) bestiality; (iii) masturbation; (iv) sadistic or masochistic abuse; or (v) lascivious exhibition of the genitals or pubic area of any person."

[301] *See supra* Part II.A (defining and recognizing the victims); *see also supra* Part IV.A (highlighting the challenges posed by the "identifiable victim" requirement incorporated in many state revenge porn laws and the ENOUGH Act).

[302] *See supra* Part V.B.1; *see also* Aubrey Burris, Note, *Hell Hath No Fury Like a Woman Porned: Revenge Porn and the Need for a Federal Nonconsensual Pornography Statute*, 66 FLA. L. REV. 2325, 2354 (2014).

[303] *See supra* Part IV.A.2.a (discussing concerns about the ability of prosecutors to successfully prove the requisite mens rea in many state revenge porn statutes).

make the crime unnecessarily difficult to prove.[304] On the other hand, organizations like the ACLU and the Electronic Frontier Foundation have historically opposed statutes that do not include an "intent to harm" provision under the theory that failure to require proof of intent may violate the First Amendment.[305] The proposed statute here, which includes recklessness as an alternative mens rea, is a reasonable compromise. The consent conundrum posed by many revenge porn laws[306] is also alleviated as the proposed law requires knowledge of, or reckless disregard for, the victim's lack of consent to the use of their image in the deepfake. Similarly, there is no actual harm requirement; harm is included in the mens rea requirement of subsection (d)(1)(B).[307]

Moreover, the offense does not punish those who had an "objectively reasonable belief that such distribution touches upon a matter of public concern."[308] This provision is patterned on the ENOUGH Act and seeks to ensure that legitimate First Amendment activity is not criminalized.[309] This provision is also more narrowly drafted than the First Amendment provision included in the MDFPA.[310]

The statute provides for two categorical exceptions in subsection (g)—law enforcement and service providers. The exception for service providers comports with section 230 immunity[311] but is not unlimited. This proposal advocates for imposing liability where "the communications service intentionally solicits, or distributes knowingly or with reckless disregard, content that is in violation of this section."[312] Although it is difficult to hold internet platforms liable, this law encourages providers to self-regulate and put screening mechanisms in place.[313]

The final components of this proposed law are the penalties set forth in subsection (e) and remedies set forth in subsection (f). The base-level penalties are drawn from the proposed ENOUGH Act; a five-year sentence emphasizes the severity of the crime and has been successfully incorporated in other nonconsensual pornography statutes.[314] This proposal additionally adds heightened sentences for repeat offenders or crimes involving minors.

[304] Roffer, *supra* note 136, at 979–80.

[305] *See id.* at 979; *see also* Danny O'Brien & Dia Kayyali, *Facing the Challenge of Online Harassment*, ELECTRONIC FRONTIER FOUND. (Jan. 8, 2015), https://www.eff.org/deeplinks/2015/01/facing-challenge-online-harassment [https://perma.cc/9994-VK52].

[306] *See supra* Part IV.A.2.b.

[307] *See supra* Part IV.A.2.a (discussing challenges associated with the actual harm requirements in many state revenge porn statutes); *see also supra* Part V.B.1.

[308] *See supra* Part V.B.1 (quoting section (d)(2) of the proposed statute).

[309] *See supra* Part IV.B (detailing the First Amendment implications of content-based regulations).

[310] *See* S. 3805, 115th Cong. (2018) ("No person shall be held liable under this section for any activity protected by the First Amendment to the Constitution.").

[311] *See supra* Part II.B (detailing the implications of the CDA for attempts to regulate deepfakes).

[312] *See supra* Part V.B.1 (quoting section (g)(2) of the proposed statute).

[313] Many platforms have already taken steps to self-regulate nonconsensual pornography. *See* Petroff, *supra* note 296.

[314] *See* Citron & Franks, *supra* note 41, at 365–67 (discussing penalties).

Beyond solely penalizing offenders, this proposal includes remedies for victims. Given the serious and often long-lasting harms suffered by victims of nonconsensual pornography, it is imperative that the law assists victims. The statute gives courts discretion to grant remedies that protect and compensate victims and allows for a court order for the destruction of original copies of deepfakes and removal from platforms to protect victims from further harm.[315] Further, restraining orders may be appropriate. The statute also advises reasonable attorney's fees and costs and damages or restitution where appropriate.[316]

C. Supplementing the Pornographic Deepfake Criminalization Act: Extralegal Solutions

In addition to enacting a federal criminal statute, solutions and support should be developed to combat nonconsensual deepfake pornography, including education and training for law enforcement, the public, and the judiciary; support from organizations and advocacy groups; and technological responses. This section highlights each in turn.

1. Awareness, Education, and Training

For a criminal statute such as that proposed here to have its full effect—to deter and punish the criminal conduct—it requires that potential perpetrators, law enforcement, and the judiciary gain a greater awareness and understanding of pornographic deepfakes and their harms. It is important to promote the idea that the potential victims are not only celebrities and public figures but anyone whose image has been digitally captured. The first step is to educate the public on the harms of nonconsensual deepfake pornography through the use of public service announcements on the internet and in traditional media. Information about the crime should be included in education programs in schools as part of their sexual education curricula and on college campuses and in workplaces as a part of sexual harassment and discrimination training.

Law enforcement and the judiciary need training and education as well. Law enforcement has a history of trivializing domestic violence and other forms of sexual violence and harassment of women in both the home and workplace, treating these crimes as a private matter or something that must be tolerated as a part of the everyday work environment.[317] Law enforcement's hands-off attitude persists in its approach to cyberharassment crimes, including revenge porn.[318] Law enforcement and the judiciary must

[315] *See supra* Part V.B.1 (citing section (e) of the proposed statute).

[316] *See supra* Part V.B.1 (citing section (f) of the proposed statute).

[317] In *Hate Crimes in Cyberspace*, Citron describes incidents of the police refusing to help victims of revenge porn or cyberharassment. DANIELLE KEATS CITRON, HATE CRIMES IN CYBERSPACE 20–21, 41, 84–85 (1st ed. 2014). For example, when a victim of online harassment went to the police, the officers did not take her fear seriously, said "[b]oys will be boys," and told her to clean up her online reputation. *Id.* at 41.

[318] *See generally id.*

be better equipped to address and take these crimes seriously; they need sensitivity training, education in the law, and regular updates and training on the rapidly changing technology, computer software, and applications. Law enforcement officers and judges also need technological tools to recognize pornographic deepfakes.

In addition to more effective law enforcement, educating stakeholders about pornographic deepfakes may also result in a decrease in the number of pornographic deepfakes. Such was the case for incidents of domestic violence.[319] To make similar progress, law enforcement, the judiciary, and the public must understand technology and its relationship to deepfakes, as well as the consequences. When attitudes shift to understanding the problem and working toward a solution, the law will have its intended effect.

2. Advocacy and Support Groups

Currently, no advocacy or support groups exclusively focused on assisting victims of pornographic deepfakes exist. Until such groups emerge, victims of pornographic deepfakes should seek assistance, guidance, and support from groups focused on revenge porn, and anti–revenge porn advocates should be encouraged to lend support to victims of deepfakes because both sets of victims suffer similarly.

After becoming a victim of revenge porn in the 2000s, Holly Jacobs[320] started a campaign called "End Revenge Porn," which began as a petition to lawmakers to criminalize the act and then grew into a place for victims to seek advice and information.[321] Eventually, the campaign incorporated the Cyber Civil Rights Initiative (CCRI), which advocates "for technological, social, and legal innovation to fight online abuse" as a 501(c)(3) nonprofit.[322] This victim resource provides everything from a crisis hotline to low-cost or pro bono support and services to those who fall victim to the sharing of these graphic images without their consent.[323] The CCRI provides information on how to report nonconsensual posts and send takedown requests.[324] It also continues to advocate for legislative change and provides model state and federal laws for legislators to consider adopting to criminalize revenge porn.[325]

Other national campaigns, such as "Without My Consent," target online privacy violations and harassment as a whole, with a special focus on

[319] *See* Citron, *supra* note 92, at 409 (detailing the legal development that transformed domestic violence from a private family matter to criminal conduct, resulting in decreased incidents of domestic violence).

[320] *See* Jacobs, *supra* note 117.

[321] *Id.*

[322] *About Us*, CYBER C.R. INITIATIVE, https://www.cybercivilrights.org/welcome/ [https://perma.cc/MPW3-8RE2] (last visited Nov. 12, 2019).

[323] *Id.*

[324] *Id.*

[325] *Id.*

nonconsensual pornography.[326] A powerful tool for victims, Without My Consent provides numerous resources, including hotlines, victim advocacy programs, lawyers, data, blogs, and an in-progress overview of laws in each of the fifty states, as well as a federal overview.[327]

Further, each state may provide resources to victims. California's attorney general, for example, has launched a "Cyber Exploitation" campaign and provided online resources for victims and law enforcement.[328] According to the Office of the Attorney General, "Posting intimate images online without consent undermines privacy, basic civil rights, and public safety."[329] Studies cited by the office show that cyberexploitation, just like sexual harassment, rape, and domestic violence, disproportionately harms women and girls.[330] The Office of the Attorney General has therefore undertaken an "initiative to #EndCyberExploitation [which] focuses on four specific areas: (1) the Attorney General's Task Force; (2) Litigation[;] (3) Legislative Advocacy; and (4) Law Enforcement Training & Education."[331] Other states and the federal government should follow California's lead.

3. Technological Responses

Even as the creation or distribution of pornographic deepfakes are prosecuted under a federal statute, other actions may simultaneously be taken that rely on existing technologies to provide remedies for victims.

a. Takedown Requests

A deepfake victim may not even be aware of the creation of the material or its distribution. Once a person becomes aware of the content, they can begin to send requests to have the material removed, even before the perpetrator is apprehended and prosecuted. The Digital Millennium Copyright Act[332] (DMCA) allows content to be removed by request of the owner of the copyrighted content.[333] This is a widely accepted and standard procedure for website owners and providers.[334] These requests can be made when the victim is the owner of the content and finds it "online without their permission," even if the content has not been copyrighted.[335] The DMCA

[326] WITHOUT MY CONSENT, http://www.withoutmyconsent.org/ [https://perma.cc/6TCB-978S] (last visited Nov. 12, 2019).

[327] *Id.*

[328] *Cyber Exploitation*, OFF. ATT'Y GEN., https://oag.ca.gov/cyberexploitation [https://perma.cc/8B9E-GLJG] (last visited Nov. 12, 2019).

[329] *Id.*

[330] *Id.*

[331] *Id.*

[332] Pub. L. No. 105-304, 112 Stat. 2860 (1998) (codified as amended in scattered sections of the U.S.C.).

[333] *What Is a DMCA Takedown?*, DMCA, http://www.dmca.com/solutions/view.aspx?ID=aa18445c-9d91-44b3-9718-49da3eb208a2&?ref=sol5a32 [https://perma.cc/4BL8-WP5E] (last updated Apr. 9, 2019).

[334] *Id.*

[335] *Id.*

does not, however, provide a solution for a deepfake video, where the victim does not own the copyright. Technically, these videos can infringe on both the rights of the individuals featured in the original video, as well as the people whose faces are being used.[336] This is especially true where the person's face being used is that of a celebrity.[337] In a digital impersonation takedown request, two competing views exist. First, the owner of the content may be only the person who created the original video. This would mean that only the victim whose body is shown could submit a request. Alternatively, the DMCA website provides that a takedown request can be made by anyone who is the subject of a video—a provision arguably broad enough to include all deepfake victims.[338]

Although a takedown request under DMCA may fail to provide complete protection to a deepfake victim, other means exist to request that materials be removed from websites.[339] Facebook, Instagram, Twitter, Reddit, and Tumblr each have their own internal reporting and takedown processes for nonconsensual pornography.[340] In contrast to the DMCA, a victim can report a pornographic deepfake even if they do not have a copyright interest; this report will remove the content from the platform. The CCRI website provides a victim resources section that is complete with takedown or reporting instructions for each of these social media platforms.[341] Victims can also unfollow, unfriend, or block users; however, this would do nothing to help remove the content.[342]

b. Third-Party Monitors

The development of deepfake technology may create a market for what is being called an "immutable authentication trail,"[343] which allows a person to pay for a service (akin to a financial credit monitoring service) to monitor a person's digital data, movement, and images.[344] The service could track a person's physical movement and in-person communications along with their electronic data, communications, navigation history, and internet presence,

[336] Samantha Cole, *Personalized Fake Porn Videos Now for Sale on Reddit*, VICE: MOTHERBOARD (Feb. 6, 2018, 3:00 PM), https://motherboard.vice.com/en_us/article/7x799b/selling-ai-generated-fake-porn-is-probably-a-good-way-to-get-sued [https://perma.cc/P5X7-VASJ].

[337] *Id.*

[338] *Id.*; *see also What Is a DMCA Takedown?*, *supra* note 333.

[339] *See* Dodge & Johnstone, *supra* note 13, at 7.

[340] *See Online Removal Guide*, CYBER C.R. INITIATIVE, https://www.cybercivilrights.org/online-removal/ [https://perma.cc/9BYP-4TC5] (last visited Nov. 12, 2019) (providing a step-by-step guide for submitting a takedown request on various social media platforms, including Facebook, Instagram, Twitter, Reddit, Tumblr, Yahoo!, Google, and Microsoft).

[341] *Id.*

[342] *See* Dodge & Johnstone, *supra* note 13, at 7.

[343] *See* Robert Chesney & Danielle Citron, *Deep Fakes: A Looming Crisis for National Security, Democracy, and Privacy?*, LAWFARE BLOG (Feb. 21, 2018, 10:00 AM), https://www.lawfareblog.com/deep-fakes-looming-crisis-national-security-democracy-and-privacy [https://perma.cc/SBJ2-7B3X].

[344] *Id.*

including social media.[345] In theory, a company would have the power to identify fake information about a client and could work on the client's behalf to remove the identified fake information.[346]

As the technology to create deepfakes is developing at a rapid pace, so too are various ways that technology can detect them. Online and digital content are currently authenticated using machine learning algorithms, watermarks, digital keys, and fingerprint authorizations. It is foreseeable that detection technology will also continue to evolve. The first of these, machine learning algorithms, is already in use today.[347] The algorithm used by Gfycat, for example, can detect when a video is fake by comparing it to the original content.[348] Because the algorithm is not currently able to access videos on private servers or private social media accounts, the deepfake image or video must be posted on the internet for the algorithm to function.[349]

A technological "arms race" is underway to diffuse false information, photos, and videos.[350] The ideal response "to the deepfake threat would be the simultaneous development and diffusion of software capable of rapidly and reliably flagging deep fakes" that can also keep "pace with innovations in deep fake technology."[351] Currently, however, the technology required to monitor and address the problem has not kept up with the software available to create deepfakes.[352] Thus, a federal criminal law remedy is warranted now.

CONCLUSION

Pornographic deepfakes are the latest phenomenon in sex exploitation cybercrimes, which have emerged in the last few years as the internet has proliferated into the daily lives of millions of people. Although many initial pornographic deepfake victims were female celebrities, the rapid advancement and widespread accessibility of AI technology means that anyone who has appeared in a digital image may now "star" in a pornographic deepfake without their consent. Like revenge porn, pornographic deepfakes arrived on the scene before policymakers and

[345] *Id.*

[346] *Id.* However, an idea for this type of service, while well-intentioned, has many flaws. While a company may be able to obtain much of a person's data, it will inherently be incapable of evaluating it all. On top of that, human and/or machine error in detecting what is real or fake will still be a problem. Finally, although perhaps most importantly, are the privacy implications. It is thought that no matter how much good a market like this could bring, that it could "risk[] the 'unraveling of privacy'—that is, the collapse of privacy by social consent regardless of what legal protections for privacy there may be." *Id.* A company that not only has access to but also compiles endless amounts of data on an individual person would be in an astonishing position of power over that individual and would represent an invaluable market and research tool. *Id.*

[347] *See* Dodge & Johnstone, *supra* note 13, at 7.

[348] *Id.*

[349] *Id.*

[350] *See* Chesney & Citron, *supra* note 343.

[351] *Id.*

[352] *Id.*

legislatures could respond. Because no federal or state laws currently criminalize the creation or distribution of pornographic deepfakes, and because this conduct is not limited to a single jurisdiction, a national response rooted in federal criminal law and centered on pornographic deepfakes is required because everyone, everywhere is a potential victim.