Arizona Law Journal of Emerging Technologies

Volume 4 *Article 6*

03-2021

LET'S TALK ABOUT (FAKE) SEX BABY: A DEEP DIVE INTO THE DISTRIBUTIVE HARMS OF DEEPFAKE PORNOGRAPHY

Shelby Akerley
Michigan State College of Law

2020-2021 **https://azlawjet.com** **Vol. 4, Art. 6**

Arizona Law Journal of Emerging Technologies

LET'S TALK ABOUT (FAKE) SEX BABY: A DEEP DIVE INTO THE DISTRIBUTIVE HARMS OF DEEPFAKE PORNOGRAPHY

Shelby Akerley, JD Candidate

Table of Contents

LET'S TALK ABOUT (FAKE) SEX BABY: A DEEP DIVE INTO THE DISTRIBUTIVE HARMS OF DEEPFAKE PORNOGRAPHY

Shelby Akerley

I. Introduction

At seventeen years old, Noelle Martin's career as a porn star had taken off.[1] To this day, you can find her image on page after page of dozens of pornographic sites.[2] A simple Google search of Noelle's name will bring up hundreds of pictures of her in various sexual positions, videos featuring Noelle performing oral sex and being ejaculated on, and the cover of two adult films, one titled "Treat Me Like a Whore."[3] Noelle's name, home address, and other personal information appear next to nearly all of this content.[4] She was, by any estimation, making a popular name for herself in the pornographic sector.[5] But, there was a problem – by the time this graphic content

[1] Cara Curtis, *Deepfakes Are Being Weaponized to Silence Women – but This Woman Is Fighting Back*, NEXT WEB (Oct. 5, 2018) https://thenextweb.com/code-word/2018/10/05/deepfakes-are-being-weaponized-to-silence-women-but-this-woman-is-fighting-back/ ("Noelle Martin's battle with deepfake pornography started six years ago [when] [a]nonymous predators stole non-sexual images of her from social media and posted them onto porn sites and threads.").

[2] Ruby Harris, *How it Feels to Find Your Face Photoshopped Onto Internet Porn*, VICE (Apr. 17, 2019) https://www.vice.com/en_au/article/gy4p47/how-it-feels-to-find-your-face-photoshopped-onto-internet-porn (stating "[Noelle Martin] instantly found page after page of search results with dozens of pornographic sites").

[3] Curtis, *supra* note 1 ("The situation escalated [and] 'moved to doctoring images of [Noelle] into graphic pornography, on the cover of pornographic DVDs, to fake images of [Noelle] being ejaculated on . . . they then doctored [Noelle] into pornographic videos performing oral sex and having sexual intercourse.'"). *See also* Ally Foster, *Teen's Google Search Reveals Sickening Online Secret About Herself*, AU NEWS (June 30, 2018) https://www.news.com.au/technology/online/security/teens-google-search-reveals-sickening-online-secret-about-herself/news-story/ee9d26010989c4b9a5c6333013ebbef2 (explaining that Noelle Martin had been put on the cover of two adult movies, one bearing the title "Treat Me Like A Whore," with additional text on the cover stating, "[s]he'll do things your wife won't!").

[4] *See* Foster, *supra* note 3 ("What is even more terrifying than these disgusting images is many were accompanied by identifying information like Ms. Martin's name, where she lived and studied, along with highly graphic comments.").

[5] Kirsti Melville, *The Insidious Rise of Deepfake Porn Videos – and One Woman Who Won't be Silenced,* ABC AUSTRL. (Aug. 29, 2019) https://www.abc.net.au/news/2019-08-30/deepfake-revenge-porn-noelle-martin-story-of-image-based-abuse/11437774 ("[N]ew, increasingly explicit images were spreading like wildfire . . . referencing her name, what she was studying and where she lived.").

was first discovered by Noelle, she was an eighteen-year-old university student studying law who had never voluntarily participated in pornography.[6]
Noelle's nightmare began when she decided to perform a Google search of herself.[7] Expecting to find the typical mundane information available in a self-search, Noelle was horrified to instead find that complete strangers had stolen non-sexual images from her social media accounts, manipulated them into pornographic images and videos, and posted them on the internet for the world to see.[8] While Noelle did not know it at the time, this simple Google search would spark a six-year journey through the darkest corners of the internet.[9] This journey would include rampant slut shaming by the public, ineffective legal remedies by her home country of Australia, and blackmail by internet service providers.[10] Consequently, Noelle Martin's journey also catalyzed a movement fighting to demonstrate the irreversible harm created when the line between reality and internet fiction becomes indistinguishable.[11]

The ever-blurring line between reality and fiction has been presented by writers, philosophers, and scientists for years.[12] In his dystopian novel, 1984, George Orwell cautioned that "reality exists in the human mind and nowhere else."[13] Unfortunately, the advent of increasingly powerful technology in modern society has created a muddled philosophical realm, causing us to question whether reality exists at all.[14]

---

[6] *See* Harris, *supra* note 2 (stating Martin was eighteen years old and at university when she first discovered anonymous predators had taken images from her social media accounts and posted them on pornographic sites).
[7] *Id.* (stating Noelle Martin's traumatizing journey started when, at eighteen, Martin "decided to google herself, just for kicks, as everyone does").
[8] *Id.* (stating Noelle Martin's understanding is that the perpetrator is not anyone she knew, but rather, "some strangers somewhere who had seen images of [her] at an event and fetishized [her]").
[9] *See, e.g.*, Curtis, *supra* note 1 (stating it has been six years since the first deepfake of Martin); Jake Sturmer, *Noelle Martin Fights to Have Harmless Selfie Removed from 'Parasite' Porn Sites*, ABC NEWS AUSTRL. (Oct. 12, 2016) https://www.abc.net.au/news/2016-10-12/womans-fight-to-have-harmless-selfie-removed-from-porn-site/7924948?pfmredir=sm ("Noelle Martin was just 17 years old when predators stole a 'selfie' she posted on her Facebook feed and plastered it over porn websites around the world . . . [n]ow aged 22, Ms. Martin is finishing a law degree and said she felt violated by the continuing use of her image.").
[10] *See* Sturmer, *supra* note 9 ("[W]hen [Noelle] asked one site to remove her photo, the webmaster tried to blackmail her. He asked for nude images of her for his own private collection in exchange for removing her photos from the public site.").
[11] *Id.* ("[Noelle] is concerned the online images make her appear as if she has voluntarily contributed to the sites. 'They literally can ruin a girl's life by what they're doing,' she said.").
[12] Sarah Gretter et al., *Walking the Line between Reality and Fiction in Online Spaces: Understanding the Effects of Narrative Transportation*, 9 J. MEDIA LITERACY EDUC. 1 ("From cave paintings to today's social networking sites, human beings have enjoyed sharing their experiences of the world and of their social interactions through narrative forms . . . [however,] the Internet can sometimes blur the boundaries between reality and fiction, rendering the distinction between factual and fictional information more difficult.").
[13] George Orwell, 1984 205 (1989).
[14] *See, e.g.*, Melville, *supra* note 5 ("'[Y]ou're likely to believe what you see if it confirms what you already believe, even if you later find out it's false. Unfortunately, it doesn't have to be that sophisticated to convince

The distortion and manipulation of images, piles of unverified information consistently pumped into the digital social sphere, and increasing awareness of "fake news" outlets have undoubtedly affected the human perception of reality and challenged society's definition of the world in which it exists.[15] However, scholars opine the newest, most unnerving addition to this technology and perhaps one most poised to shatter our discernment of reality and fiction is involuntary synthetic imagery – more commonly known as deepfakes.[16]

Deepfakes combine existing images and video using machine-learning technology to create a new video.[17] The result is an ultrarealistic video that shows an individual performing acts in which they never actually participated.[18] In other words, a deepfake is a forged video that represents a situation that has never occurred by manipulating pre-existing pictures or videos.[19] This technology has found a new home in a dark but well explored space of the internet: pornography.[20]

---

people of what they want to believe,' Dr Franks says. 'Having fact-checking resources at your disposal doesn't matter if people don't care about the facts. And that's going to become an even greater problem if people are looking at what they believe to be real.'"); Samantha Cole, *We Are Truly Fucked: Everyone Is Making AI-Generated Fake Porn Now*, VICE (Jan. 24, 2018) https://www.vice.com/en_us/article/bjye8a/reddit-fake-porn-app-daisy-ridley ("The combination of powerful, open-source neural network research, our rapidly eroding ability to discern truth from fake news, and the way we spread news through social media has set us up for serious consequences.").

[15] *See, e.g.*, J.M. Porup, *How and Why Deepfake Videos Work – and What is at Risk*, CSO (Apr. 10, 2019) https://www.csoonline.com/article/3293002/deepfake-videos-how-and-why-they-work.html (stating "[h]uman beings seek out information that supports what they want to believe and ignore the rest. Hacking that human tendency gives malicious actors a lot of power. We see this already with disinformation (so-called 'fake news') that creates deliberate falsehoods that then spread under the guise of truth"); Gretter, *supra* note 12, at 1 ("Recent contentions about 'fake news' and misinformation online has shed light on the critical need for media literacy at a global scale . . . the line between facts and fiction can often become blurry in these online spaces and being able to distinguish between reality and fantasy can have important consequences in the lives of young Internet users.").

[16] Jeremy Hsu, *Can AI Detect Deepfakes To Help Ensure Integrity of U.S. 2020 Elections?*, IEEE SPECTRUM (Feb. 28, 2019) https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/will-deepfakes-detection-be-ready-for-2020 (referring to deepfakes as "involuntary synthetic . . . imagery").

[17] Elizabeth Caldera, "*Reject the Evidence of Your Eyes and Ears:*" *Deepfakes and the Law of Virtual Replicants*, 50 SETON HALL L. REV. 177, 178-79 ("Combining the words 'deep learning' and 'fake,' a deepfake is a 'hyper-realistic digital falsification of images, video, and audio.' Put simply, a deepfake is a forged video; it depicts something that has never happened by manipulating previously existing video footage or pictures.").

[18] *See id.* at 181 ("By utilizing previously existing images and videos, the technology creates a generated video that nevertheless looks authentic.").

[19] *See id.* at 182 ("[T]he result is a video that both looks and sounds like the figure in the video, but that in actuality is fabrication.").

[20] Danielle Keats Citron, *Sexual Privacy*, 128 YALE L.J. 1870, 1921 (2019) ("Machine-learning technologies are being used to create 'deep-fake' sex videos where people's faces and voices are inserted into real pornography. Deep-fake technology enables the creation of impersonations out of digital whole cloth. The end result is realistic-looking video or audio that is increasingly difficult to debunk.").

As technology advances, the hyper realism of deepfake pornography and its ability to be mass distributed on the internet increases.[21] The distribution of deepfake pornography causes irreparable harm to women and stifles their free speech rights.[22] Existing laws are no longer viable solutions to prevent the distribution of deepfake pornography.[23] The best method to truly prevent the distribution of deepfake pornography is to hold Internet Service Providers (ISPs) liable for their part in the passive distribution of deepfake pornography.[24] While Section 230 of the Communications Decency Act grants ISPs extensive immunity, this protection should be tailored to strike a balance between the First Amendment rights of site members and the harms associated with the distribution of deepfake pornography.[25] This tailoring should involve a federal law criminalizing the distribution of deepfake pornography, and an amendment to Section 230 that would carve out exceptions barring immunity for content such as deepfake pornography.[26]

This Note explores the distribution of deepfake pornography and the complicated legal web it has created – pitting internet free speech, existing law, and human dignity against each other.[27] Part I provides a brief overview of the machine-learning technology used to create deepfake pornography.[28] Part II discusses the current legal

[21] Jillian Roffer, *Nonconsensual Pornography: An Old Crime Updates Its Software*, 27 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 950 (2017) ("The harm experienced by victims of nonconsensual pornography is exacerbated by the unique nature of the internet (including social media) because it facilitates an exponential growth in publication.").
[22] Rana Ayyub, *I Was the Victim of a Deepfake Porn Plot Intended to Silence Me*, HUFFINGTON POST (Nov. 21, 2018) https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316 (describing Rana Ayyub's experiences as a victim of deepfake pornography and the stifling of her free speech experienced as a result).
[23] Douglas Harris, *Deepfakes: False Pornography is Here and the Law Cannot Protect You*, 17 DUKE L. & TECH. REV. 99, 102 (2019) ("Unfortunately, as with many new technologies, the law is unequipped to handle these impending issues.").
[24] Dalisi Otero, *Confronting Nonconsensual Pornography with Federal Criminalization and a "Notice-and-Takedown" Provision*, 70 U. MIAMI L. REV. 585, 600 (2016) ("Amending § 230 to exclude nonconsensual pornography websites from the benefit of immunity would allow nonconsensual pornography victims to have some sort of leverage to pressure revenge porn website operators to remove damaging material or save Internet protocol addresses of users who post the material in the first place.").
[25] *See* Citron, *supra* note 20, at 1931 ("Congress gave platforms a broad liability shield for user-generated content in the form of section 230 of the Communications Decency Act.").
[26] *See infra* Part IV (advocating for the enactment of a federal criminal law prohibiting the distribution of deepfake pornography and an amendment to Section 230 excluding ISP immunity for distribution of deepfake pornography).
[27] *See, e.g.*, *infra* Part II (discussing internet free speech and existing law contributing to the distribution of deepfakes); *infra* Part III (discussing the harms and loss of dignity experienced by victims of deepfake pornography).
[28] *See infra* Part I (discussing machine-learning, its benefits, and its migration from altruistic purposes to the creation of deepfake pornography).

environment that protects the distribution of deepfake pornography.[29] Particularly, Part II explains the relationship between free speech, pornography, obscenity, Section 230, and lack of federal regulations in the context of deepfake pornography.[30] Part III explores the impact that distribution of deepfake pornography has on victims.[31] Part IV presents the argument that it is the distribution of deepfake pornography that causes harm and analyzes why existing law does little to prevent distribution.[32] Part IV offers the solution that ISPs, as the mechanisms for distribution, should held liable for the distribution of deepfake pornography.[33] Part IV advocates for a combined solution of a federal criminal law prohibiting the distribution of deepfakes and an amendment to Section 230 to exclude immunity for ISPs who knowingly distribute deepfake pornography.[34] Such a solution eradicates the distribution of deepfake pornography without censoring lawful and socially valuable deepfakes, such as parodies and satire.[35]

II. The Advent of Deepfake Pornography

Technology has come a long way since the advent of the digital camera, desktop computer, and cell phone of the last fifty years.[36] Now, nearly everyone has access to a computer capable of using the popular program "Photoshop" to edit photos on a computer.[37] Similarly, nearly anyone can use his or her smart phone to open an app called Snapchat and watch a photo of themselves morph into that of a monkey, dog,

[29] *See infra* Part II (discussing the various legal mechanisms contributing to the current legality of distribution of deepake pornography on the internet).
[30] *See infra* Part II (discussing First Amendment protections of pornography, obscenity as an exception to free speech rights, the immunity afforded to ISP's under Section 230, and the lack of a federal criminal law prohibiting the distribution of deepfake pornography).
[31] *See infra* Part III (discussing the harmful effects of distribution of deepfake pornography and the disproportionate amount of deepfake pornographies featuring women).
[32] *See infra* Part IV (discussing the harms distribution, as opposed to creation, of deepfake pornography poses).
[33] *See infra* Part IV (discussing ISP's role as the mechanism for distribution of deepfake pornography).
[34] *See infra* Part IV (advocating for a new federal criminal law prohibiting the distribution of deepfake pornography and an amendment to Section 230 excluding immunity for ISP's who knowingly allow the distribution of deepfake pornography on their servers).
[35]Rebecca A. Delfino, *Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn's Next Tragic Act*, 88 FORDHAM L. REV. 887, 925 (2019) ("[G]iven the digitalized nature of deepfakes, there is an added layer of concern about parody and satire.").
[36]Daniel Gutierrez, *The Birth of Modern Technology – 50 Years Ago to Now a Look at How Far We've Come*, INSIDE BIG DATA (Aug. 27, 2018) https://insidebigdata.com/2018/08/27/birth-modern-technology-50-years-ago-now-look-far-weve-come/ ("Much of the technology we take for granted today stems from 50 years ago . . . [t]echnology has come a long way over the last 50 years, but many of the complex and high-performing technologies we have now have their roots in the unprecedented changes of the late 1960s.").
[37]*In Celebration of Photoshop World Conference & Expo 2015*, BUSINESS.COM (Aug. 11, 2015) https://www.business.com/articles/photoshop-reigns-supreme-how-the-software-has-maintained-market-dominance/ ("[O]n its 20th anniversary, Photoshop had more than 10 million users worldwide.").

or cat.[38] This same application allows you to swap your face with the face of your friends and can turn any "selfie" of the user into a photo of the user depicted as the opposite sex.[39] Each of these new day-to-day normalcies implicate some sort of machine-learning technology, which paved the way for the creation of deepfake pornography.[40]

a. Understanding Machine-Learning Technology

"Machine learning" is an application of artificial intelligence which allows a computer to learn the way the human mind does – from experience.[41] A computer can learn and improve from the data that runs through it with the addition of machine-learning technology.[42] The more data the program collects and analyzes, the more the computer using machine-learning will learn, and the "smarter" the computer becomes.[43]

The potential benefits of machine-learning technology are enormous.[44] Besides offering day-to-day comforts like customizing your Netflix recommendation list or improving your Google searches, machine-learning technology is being used to enhance data security, detect cancer, prevent fraud, and improve the safety of smart

[38] J. Clement, *Daily Active Users of Snapchat 2014-2019*, STATISTA (Oct. 23, 2019) https://www.statista.com/statistics/545967/snapchat-app-dau/ ("As of the third quarter of 2019, photo and video sharing app Snapchat had 210 million daily active users worldwide.").
[39] Cammy Harbison, *Snapchat's New Gender Swap Filter Will Make You Question Your Identity: How to Get the Male to Female Filter*, NEWSWEEK (May 14, 2019) https://www.newsweek.com/snapchat-gender-swap-filter-how-get-girl-boy-change-male-female-how-use-not-1425014 (describing Snapchat's new gender swap filter).
[40] Dyani Sabin, *How Snapchat Uses A.I. to Make a Custom "Discover" Tab*, INVERSE (May 11, 2017) https://www.inverse.com/article/31497-snapchat-ai-newsfeed ("In Snap's first earnings call on Wednesday, Snap CEO Evan Spiegel mentioned that Snap uses machine learning to determine which stories to show you in your newsfeed.").
[41] *See e.g.*, Daniel Faggella, *What is Machine Learning?*, EMERJ (Oct. 23, 2019) https://emerj.com/ai-glossary-terms/what-is-machine-learning/ ("Machine learning research is part of research on artificial intelligence, seeking to provide knowledge to computers through data, observations and interacting with the world. That acquired knowledge allows computers to correctly generalize to new settings."); Trefis Team, *Reasons Why Google's Latest AI-TensorFlow is Open Sourced*, FORBES (Dec. 1, 2015) https://www.forbes.com/sites/greatspeculations/2015/12/01/reasons-why-googles-latest-ai-tensorflow-is-open-sourced/#46b09461765b ("Machine learning is a subset of the intellectual domain of Artificial Intelligence, which comprises the study of various sorts of intelligent, self-learning machines.").
[42] *See id.* ("Machine learning is the science of getting computers to act without being explicitly programmed").
[43] *See id*. ("The fundamental goal of machine learning algorithms is to generalize beyond the training samples i.e. successfully interpret data that it has never 'seen' before").
[44] *See id*. ("Machines that learn are useful to humans because, with all of their processing power, they're able to more quickly highlight or find patterns in big (or other) data that would have otherwise been missed by human beings. Machine learning is a tool that can be used to enhance humans' abilities to solve problems and make informed inferences on a wide range of problems, from helping diagnose diseases to coming up with solutions for global climate change").

cars.[45] Because machine-learning systems are poised to make such altruistic societal advances, Google made headlines in 2015 when it released its latest machine-learning system, TensorFlow, to the public.[46] At the time, Google's CEO Sundar Pichai stated that this kind of technology is more profound than electricity or fire.[47] In the same breath, however, Pichai made the sobering observation that fire has the potential to kill.[48]

b. The First Deepfakes

The humble beginnings of deepfakes start in the form of "shallow" fakes, which are real videos that have been slightly manipulated – either by changing speed or modifying pitch – to produce a new video.[49] A well-known example of a "shallow" fake is the 2019 video of House Speaker Nancy Pelosi speaking at a Center for American Progress event.[50] The shallow fake was slightly manipulated to make Pelosi

[45] *See* Bernard Marr, *The Top 10 AI And Machine Learning Use Cases Everyone Should Know About*, FORBES (Sept. 30,2016) https://www.forbes.com/sites/bernardmarr/2016/09/30/what-are-the-top-10-use-cases-for-machine-learning-and-ai/#599f11d194c9 (stating that machine learning can "look for patterns in how data in the cloud is accessed, and report abnormalities that could predict security breaches," "process more information and spot more patterns than their human counterparts, [in fact], [o]ne study used computer assisted diagnosis to review the early mammography scans of women who later developed breast cancer, and the computer spotted 52% of the cancers as much as a year before the women were officially diagnosed," "precisely distinguish between legitimate and fraudulent transactions between buyers and sellers," "analyze your Netflix or Amazon activity and compare it to the millions of other users to determine what you might like to buy or binge watch next," "learn from [search engine] mistakes and deliver a better result next time," and "offer real time advice about traffic and road conditions" in smart cars).

[46] *See* Trefis Team, *supra* note 41 ("Google recently announced that it was open-sourcing its latest machine learning engine engine called TensorFlow").

[47] *See, e.g.*, Theodore Schleifer, *Google CEO Sundar Pichai Says AI is More Profound than Electricity and Fire*, VOX (Jan. 19, 2018) https://www.vox.com/2018/1/19/16911180/sundar-pichai-google-fire-electricity-ai (explaining Google CEO Sundar Pichai stated "AI is one of the most important things that humanity is working on. It's more profound than . . . electricity or fire").

[48] *See* Catherine Clifford, *Google CEO: A.I. is More Important Than Fire of Electricity*, CNBC (Feb. 1, 2018) https://www.cnbc.com/2018/02/01/google-ceo-sundar-pichai-ai-is-more-important-than-fire-electricity.html ("'Well, it kills people, too,' Pichai says of fire. 'We have learned to harness fire for the benefits of humanity but we had to overcome its downsides too. So my point is, AI is really important, but we have to be concerned about it.'").

[49] *See* Kalev Leetaru, *The Real Danger Today is Shallow Fakes and Selective Editing Not Deepfakes*, FORBES (Aug. 26,2019) https://www.forbes.com/sites/kalevleetaru/2019/08/26/the-real-danger-today-is-shallow-fakes-and-selective-editing-not-deep-fakes/#62097fb94ea0 ("[T]raditionally manipulated videos like the Nancy Pelosi speech . . . deliberately [slow] down or [speed] up a video to portray the subject in a misleading manner. Such editing does not actually alter the contents of the video in any way. Instead, by merely reframing how the viewer sees it, such actions can ascribe new meaning to a previously innocuous video.").

[50] *See* Drew Harwell, *Faked Pelosi Videos, Slowed to Make her Appear Drunk, Spread across Social Media*, WASHINGTON POST (May 24, 2019) https://www.washingtonpost.com/technology/2019/05/23/faked-pelosi-videos-slowed-make-her-appear-drunk-spread-across-social-media/ ("The video of Pelosi's onstage speech Wednesday at a Center for American Progress event, in which she said President Trump's refusal to cooperate with congressional investigations was tantamount to a 'coverup,' was subtly edited to make her voice sound garbled and warped. It was then circulated widely across Twitter, YouTube and Facebook").

appear drunk and slur her words.[51] Although the video was not made using machine-learning technology, as with deepfakes, this widely shared video was one of the first instances which caused Congress to pause and consider the impact this kind of technology has on the falsification of information.[52] This fear of misinformation via manipulated videos eventually spurred Congress to introduce two Acts regulating deepfakes: the Malicious Deep Fake Prohibition Act and the Deepfake Accountability Act.[53] However, both Acts aimed to end deepfakes in the political realm, ignoring the ramifications of deepfakes in the pornographic sector entirely.[54]

When Google released TensorFlow to the public, "shallow" fakes were quickly replaced by deepfakes that could seamlessly place the face of one person on the body of another and accurately track that body's movements and expressions, creating a convincingly realistic video.[55] The dystopian nature of this technology was showcased in a public service announcement in which American actor and filmmaker Jordan Peele used Fakeapp, a machine-learning program, to ventriloquize former president Barack Obama.[56] Among other things, Peele used this technology to make it appear as though Obama was calling former president Donald Trump "a total and complete dipshit."[57] Individuals also began using machine-learning technology to create

[51] *See id.* ("Analyses of the distorted Center for American Progress video by Washington Post journalists and outside researchers indicate that the video has been slowed to about 75 percent of its original speed. To possibly correct for how that speed change would deepen her tone, the video also appears to have been altered to modify her pitch, to more closely resemble the sound of her natural speech").
[52] *See id.* ("One version, posted by the conservative Facebook page Politics WatchDog, had been viewed more than 2 million times by Thursday night, been shared more than 45,000 times, and garnered 23,000 comments with users calling her 'drunk' and 'a babbling mess.'").
[53] *See* Nina Brown, *Congress Wants to Solve Deepfakes by 2020*, SLATE (July 15, 2019) https://slate.com/technology/2019/07/congress-deepfake-regulation-230-2020.html ("The Malicious Deep Fake Prohibition Act, for example, would make it a federal crime to create or distribute a deepfake when doing so would facilitate illegal conduct . . . The Deepfakes Accountability Act would require mandatory watermarks and clear labeling on all deepfakes.").
[54] *See* Joseph Cox, *Most Deepfakes Are Used for Creating Non-Consensual Porn, Not Fake News*, VICE (Oct. 7, 2019) https://www.vice.com/en_us/article/7x57v9/most-deepfakes-are-porn-harassment-not-fake-news ("While media, politicians, and technologists panic over the risk of deepfakes impacting elections, a new study has found that the vast, vast majority of deepfakes are pornographic in nature.").
[55] *See* Harris, *supra* note 23, at 101 ("[O]ther open-source tools like the DownAlbum and Instagram Scraper easily allow individuals to create a faceset . . . [and] [b]rowser based applications employing facial recognition software enable users to upload a photo of the person they want in the fake video, and the website outputs the most comparable adult performer").
[56] *See* James Vincent, *Watch Jordan Peele Use AI to Make Barack Obama Deliver a PSA About Fake News*, THE VERGE (Apr. 17, 2018) https://www.theverge.com/tldr/2018/4/17/17247334/ai-fake-news-video-barack-obama-jordan-peele-buzzfeed ("The video was made by Peele's production company using a combination of old and new technology: Adobe After Effects and the AI face-swapping tool FakeApp. The latter is the most prominent example of how AI can facilitate the creation of photorealistic fake videos").
[57] *See id.* ("Using some of the latest AI techniques, Peele ventriloquizes Barack Obama, having him voice his opinion on Black Panther ('Killmonger was right') and call President Donald Trump 'a total and complete dipshit.'").

humorous deepfakes, including inserting actor Nicholas Cage into a variety of movies such as Indiana Jones, James Bond, and The Terminator. [58]

c. The Transition to Deepfake Pornography

As machine-learning programs developed and became more prevalent in the public sphere, it became easy for anyone with a computer to create a deepfake.[59] Unsurprisingly, it was at this point that machine-learning technology moved from technological altruism, such as cancer detection and cybersecurity, to a more illicit use.[60] The term deepfake was first coined in 2018 when a user by the name of "deepfake" on the social content platform Reddit began using machine-learning technology to insert the faces of celebrities into existing pornographic videos.[61] These videos were published on a subreddit which amassed more than 100,000 followers and exploited celebrities such as Gal Gadot, Scarlett Johannsson, Taylor Swift, and Emma Watson.[62] Celebrities were easy for "deepfakes" to exploit due to the large

[58] *See, e.g.*, *Deepfake Video Of Nic Cage Replacing Other Actors' Faces In Iconic Movie Scenes*, GEEKOLOGIE (Oct. 29, 2018) https://geekologie.com/2018/10/deepfake-video-of-nic-cage-replacing-oth.php (describing the new phenomena of inserting Nicolas Cage into a variety of movies via deepfake technology); Russell Spivak, *"Deepfakes": The Newest Way to Commit One of the Oldest Crimes*, 3 GEO. L. TECH. REV. 339, 346-48 (2019) ("Deepfakers have also focused on generating celebrity videos outside the adult film industry. In a slightly more good-natured use of the technology, the deepfake community turned to one particular movie star for comedic relief — Nicolas Cage. Cage's face was superimposed onto Harrison Ford's Indiana Jones in Raiders of the Lost Ark and onto Amy Adams' Lois Lane in Man of Steel. In one particularly humorous and meta deepfake, Cage's face was superimposed onto Andy Samberg's face in a Saturday Night Live sketch in which Samberg was impersonating Cage").

[59] *See* Harris, *supra* note 23, at 101 ("A Hollywood production budget is not necessary to create deepfakes from home. All one needs is a computer, a decent graphics card, the FakeApp program (which uses the open-source software Google released), hundreds of pictures of the desired person (known as a 'faceset'), and a few hours of time").

[60] *See id.* at 100 ("Although . . . machine-learning tools can be used in beneficial ways, like discovering new planets, they can also be used for deviant purposes").

[61] *See* Alex Hern, *AI Used to Face-swap Hollywood Stars into Pornography Films*, THE GUARDIAN (Jan. 25, 2018) https://www.theguardian.com/technology/2018/jan/25/ai-face-swap-pornography-emma-watson-scarlett-johansson-taylor-swift-daisy-ridley-sophie-turner-maisie-williams ("Advanced machine learning technology is being used to create fake pornography featuring real actors and pop stars, pasting their faces over existing performers in explicit movie . . . [a] community on the social news site Reddit has spent months creating and sharing the images, which were initially made by a solo hobbyist who went by the name 'deepfake' . . . [t]he creation of face-swapped pornography rapidly scaled up in late December, when another Reddit user (going by the name 'deepfaceapp') released a desktop app designed to let consumers create their own clips").

[62] *See Definition of Subreddit*, LEXICO https://www.lexico.com/en/definition/subreddit (Last Visited Feb. 17, 2020) (defining a subreddit as "[a] forum dedicated to a specific topic on the website Reddit"); *see, e.g.*, Citron, *supra* note 20, at 1921-22 ("A Subreddit (since closed) featured deep-fake sex videos of female celebrities, amassing more than 100,000 users. One such video featured Gal Gadot having sex with her stepbrother - but of course Gadot never made the video. Deep-fake sex videos have also featured the likenesses of Scarlett Johansson, Taylor Swift, and Maisie Williams."); Cole, *supra* note 14 ("In one such hyper realistic video, actress Gal Gadot is depicted as having sexual intercourse with her stepbrother . . . [f]akes posted in the subreddit have already been pitched as real on other websites; a deepfake of Emma

amounts of downloadable images and videos of celebrities that exist on the internet.[63] Reddit removed the "deepfakes" page from Reddit in 2018, spurring another Reddit user to create an app called Fakeapp, which was specifically designed to allow users without a computer science background to create deepfake pornography.[64] According to the creator, anyone who downloads Fakeapp can create deepfake pornography using only one or two high quality videos of the face of the person they wish to portray in a pornography.[65] This advancement allowed users to bypass using Tensorflow's machine-learning technology and create a deepfake pornography of any individual they desired using tools that were free, readily available, and easy to learn.[66]

As machine-learning programs became more advanced and users began sharing tips, deepfake pornography dramatically shifted from a choppy, unrecognizable video of a person to a seamless and believable portrayal of sex on film.[67] It was this pervasive and easily accessible technology that led to the transition from deepfake pornography

Watson taking a shower was reuploaded by CelebJihad—a celebrity porn site that regularly posts hacked celebrity nudes—as a 'never-before-seen video above is from my private collection,' and appears to feature Emma Watson fully nude and flaunting her naked sex organs while showering with another girl").

[63] *See* Gregory Barber, *Deepfakes Are Getting Better, But They're Still Easy to Spot*, WIRED (May 26, 2019) https://www.wired.com/story/deepfakes-getting-better-theyre-easy-spot/ ("Celebrities are the easiest targets, with ample public imagery that can be used to train deepfake algorithms; it's relatively easy to make a high-fidelity video of Donald Trump, for example, who appears on TV day and night and at all angles").

[64] *See* Spivak, *supra* note 58, at 345 ("Additionally, another Redditor, wanting to break down barriers to entry for this technology even further, 'created an app specifically designed to allow users without a computer science background to create AI-assisted fake porn' . . . [t]he app[,] appropriately titled 'FakeApp[,]'opened the door to even more deepfake creators").

[65] *See* Cole, *supra* note 14 ("[S]ome users are also creating videos that show the far-reaching implications of a technology that allows anyone with sufficient raw footage to work with to convincingly place any face in any video . . . anyone who can download and run FakeApp can create one of these videos with only one or two high-quality videos of the faces they want to fake").

[66] *See id.* ("The subreddit's wiki states that FakeApp is "a community-developed desktop app to run the deepfakes algorithm without installing Python, Tensorflow, etc."); *see, e.g.*, Katayanna Quach, *FYI: There's Now An AI App that Generates Convincing Fake Smut Vids Using Celebs' Faces*, THE REGISTER https://www.theregister.co.uk/2018/01/25/ai_fake_skin_flicks/ ("You don't need to be an AI whiz to wield this, since you don't need to mess around with much coding. 'This app is intended to allow users to move through the full deep-fake creation pipeline—creating training data, training a model, and creating fakes with that model—without the need to install Python and other dependencies or parse code,' according to [Fakeapp's] documentation."); Cole, *supra* note 14 (explaining that the creator of fakeapp stated, "I think the current version of the app is a good start, but I hope to streamline it even more in the coming days and weeks . . . [e]ventually, I want to improve it to the point where prospective users can simply select a video on their computer, download a neural network correlated to a certain face from a publicly available library, and swap the video with a different face with the press of one button").

[67] *See, e.g.*, Quach, *supra* note 66 ("There are countless threads of people posting their results, or asking for advice on how to make the faces less blurry or how to align the eyebrows."); Harris, *supra* note 23, at 101 ("The now-banned Deepfakes subreddit and a now-closed Discord chatroom were hotbeds for users to exchange tips on producing deepfake porn videos of each other's crushes and ex-significant others").

featuring celebrities to deepfakes portraying private individuals like Noelle Martin.[68] Creating a deepfake pornography of an unsuspecting classmate, coworker, or ex-girlfriend became as simple as collecting a few photos and videos from social media, running them through a machine-learning software that matches their physique to that of a similar pornstar, and using the Fakeapp program to seamlessly blend the unsuspecting individuals face and voice onto the body of that pornstar.[69] If time is an issue, multiple platforms exist that sell services offering the creation of deepfake pornography, often for less than one U.S. dollar. [70] The time to create these deepfake pornographies can range from a few hours up to a week, but rapidly developing technology continues to decrease the amount of time necessary.[71]

[68] *See id.* (explaining that the Fakeapp creator stated "we should expect such code to emerge in public as powerful AI technology becomes more and more accessible – and this tech will be used for good, and bad . . . [u]ltimately, for better or worse, it's impossible to stop anyone from doing what they want with this tool, and unfortunately that means some will abuse it"); *see, e.g.*, Spivak, *supra* note 58, at 348 ("Deepfakes are also used to superimpose an average member of the public onto a celebrity's body."); Citron, *supra* note 20, at 1922 ("Ex-intimates have seized upon the deep-fake trend. As one Reddit user asked, 'I want to make a porn video with my ex-girlfriend. But I don't have any high-quality video with her, but I have lots of good photos.' A Discord user explained that he made a 'pretty good' video of a girl he went to high school with, using around 380 photos scraped from her Instagram and Facebook accounts ").

[69] *See* Harris, *supra* note 23, at 101 ("Other open-source tools like the DownAlbum and Instagram Scraper easily allow individuals to download all images on a person's social media account to create a faceset. . . . to make a seamless deepfake, the producer needs to find a body that matches the unwary victim's face. Finding the ideal body has also become quasi-automated. Browser-based applications employing facial recognition software enable users to upload a photo of the person they want in the fake video, and the website outputs the most comparable adult performer ").

[70] *See* Laurie Chen, *China's Deepfake Celebrity Porn Culture Stirs Debate About Artificial Intelligence Use*, YAHOO FINANCE (July 20, 2019) https://finance.yahoo.com/news/chinas-deepfake-celebrity-porn-culture-093000173.html ("One discussion named 'Face swap + female celebrity's surname' on Baidu Tieba, one of China's most popular internet forums, offered to sell face-swapped porn for as little as 4 yuan (less than 60 US cents) per video, while a package of 700 videos cost 158 yuan, the report said. Sellers also listed 'customised face-swap' porn videos on the second-hand e-commerce app Xianyu, with prices starting from 20 yuan per minute of footage. The listing said buyers could provide celebrity or personal photos, according to the report").

[71] *See* Samantha Cole, *This Program Makes It Even Easier to Make Deepfakes*, VICE (Aug. 19, 2019) https://www.vice.com/en_us/article/kz4amx/fsgan-program-makes-it-even-easier-to-make-deepfakes ("A new method for making deepfakes creates realistic face-swapped videos in real-time, no lengthy training needed. Unlike previous approaches to making deepfakes—algorithmically-generated videos that make it seem like someone is doing or saying something they didn't in real life—this method works on any two people without any specific training on their faces. Most of the deepfakes that are shared online are created by feeding an algorithm hundreds or thousands of images of a specific face. The algorithm 'trains' on that specific face so it can swap it into the target video. This can take hours or days even with access to expensive hardware, and even longer with consumer-grade PC components. A program that doesn't need to be trained on each new target is another leap forward in making realistic deepfakes quicker and easier to create").

d. The Difference Between Deepfake Pornography and Revenge Porn

Because revenge porn and deepfake pornography parallel in many ways, it is important to clarify their crucial differences.[72] Revenge pornography is defined as the online posting of sexually explicit images of a person without their consent, and is often used as a form of harassment against that person.[73] Deepfake pornography is defined as a pornographic video that applies the nonsexual image of an individual onto the body of an existing porn star with the purpose of creating a new, sexually explicit video.[74] The first important distinction between the two is the issue of consent.[75]

The case of *Novak v. Simpson* highlights the typical scenario involving revenge porn.[76] In *Novak*, a college student sent naked photos to her then-boyfriend.[77] The college student sent the photos to her then-boyfriend consensually but did not consent to him sharing the photos with anyone else.[78] The college student was horrified to find that her then-boyfriend had posted her images online.[79] In the case of deepfake

[72] *Compare Action Sheet on Revenge Porn*, MCO LAW https://www.mcolaw.com/white-papers-research/action-sheet-on-revenge-porn (last visited Feb. 17, 2020) (providing that some "statistics on revenge porn include: 1 out of 10 ex-partners has threatened to post naked images of their exes online. 60% of them carry out the threat. 80-90% of revenge porn victims are women. 93% of victims suffer significant emotional distress. 49% of victims are harassed or stalked online by someone who saw the material. 3,000 pornography websites feature a "revenge porn" genre.); *with* Cox, supra note 54 ("The company found a total of 14,678 deepfake videos online. According to Giorgio Patrini, CEO and chief scientist at Deeptrace, the company then examined the gender of targets in videos from five deepfake porn sites (7,144 videos) and 14 YouTube channels (under 500 videos). Videos on four top, dedicated deepfake pornography websites also had over 134 million views, and all but 1 percent of the subjects featured in deepfake pornography videos were female actors . . . [t]he new report adds, 'Deepfake pornography is a phenomenon that exclusively targets and harms women'").

[73] *See, e.g.*, *Revenge Porn*, MERRIAM-WEBSTER https://www.merriam-webster.com/dictionary/revenge%20porn (last visited Jan. 22, 2020) (defining revenge porn as "sexually explicit images of a person posted online without that person's consent especially as a form of revenge or harassment"); Aine Cain, Former Rep. Katie Hill Says the Wave of Harassment She Faced After Alleged Revenge Porn Leak Left Her Contemplating Suicide, Business Insider (Dec. 7, 2019) https://www.businessinsider.com/katie-hill-fallout-revenge-porn-leak-2019-12 ("Since her resignation, Hill has advocated on behalf of victims of revenge porn, an increasingly prevalent issue where perpetrators publish intimate photos of their victims").

[74] *See* Spivak, *supra* note 58, at 339 (defining deepfake pornography).

[75] *See* Delfino, *supra* note 35, at 897 (describing the differences between deepfake pornography and revenge porn).

[76] *See* Complaint 19, Novak v. Simpson, No. 6:18-cv-00922 (M.D. Fla. July 13, 2018) (describing the events leading up to the victim filing suit against the alleged perpetrator of revenge porn).

[77] *See id.* at 23, 25-26 (describing the relationship between the victim and the alleged perpetrator of revenge porn).

[78] *See id.* (describing the actions of the alleged perpetrator, who was accused of sharing sexually explicit materials of the victim on a Facebook group without her permission).

[79] *See id.* (explaining that the victim's then-boyfriend had posted the images to a Facebook page titled "Dog Pound," where members of his fraternity could view the photos as well as share videos and images of their

pornography, as experienced by Noelle, a stranger stole nonsexual images from Noelle's Facebook profile and manipulated them without her permission to falsely depict her as participating in pornography.[80] In the former, the explicit images were initially shared with consent.[81] In the latter, the victim never consensually shared explicit photos or videos with the perpetrator.[82]

The second important distinction between revenge porn and deepfake pornography is the medium used and its importance in the language of laws criminalizing revenge porn.[83] Many instances of revenge porn involve still photographs, and the language of state laws criminalizing revenge porn reflect this.[84] As such, although there is a plethora of state laws criminalizing revenge porn, deepfake pornography, which encompasses videos as opposed to still photographs, does not always fall within the parameters of revenge porn laws.[85]

---

own sexual encounters with women). S*ee also* Citron, *supra* note 20, at 1918-19 ("[C]onsider the experience of Holly Jacobs. Jacobs shared sexually explicit images and videos with her boyfriend. The images and videos were for their eyes only. After their break-up, her ex betrayed her trust, posting the photos and videos on hundreds of revenge-porn sites, porn sites, and adult-finder sites. He also sent her nude photos to her boss.").

[80] *See supra* Part I (discussing the circumstances under which Noelle Martin came to be depicted in deepfake pornography).

[81] *See* Novak *supra* note 75, at 23, 25-26 (describing the initial consensual sexual relationship between the victim and the alleged perpetrator).

[82] *See* Sturmer, *supra* note 9 ("Noelle Martin was just 17 years old when predators stole a 'selfie' she posted on her Facebook feed and plastered it over porn websites around the world. Other harmless social photos of the young woman were also copied and pasted onto porn sites.").

[83] *See* Harris, *supra* note 23, at 120, 122 (explaining that revenge porn laws "feature certain phrases that are inapplicable to deepfakes . . . [s]ixteen states use the phrase 'intimate parts' or 'intimate areas' which are typically defined in statutes as the unclothed genital areas. Some states also use 'engaged in a sexual act' and 'state of nudity.' Given that personal deepfakes often superimpose a Victim's head on the body of another, the central question is: can these terms apply to a Victim whose actual body parts are not visible? Or, can 'intimate body parts' refer to areas that are not the Victim's? The answer to this question may lie in the usage of 'depiction.' Like North Carolina, thirty-one states use the term "depicted person's intimate parts. However, depiction is not usually defined in these statutes.").

[84] *See, e.g.*, N.C. GEN. STAT. § 14-190.5A (defining revenge porn as applied to North Carolina); Harris, *supra* note 23, at 120-21 ("For example, North Carolina's revenge porn statute contains 'typical language' of a nonconsensual pornography statute: (1) The person knowingly discloses an image of another person with the intent to do either of the following: a. Coerce, harass, intimidate, demean, humiliate, or cause financial loss to the depicted person. b. Cause others to coerce, harass, intimidate, demean, humiliate, or cause financial loss to the depicted person. (2) The depicted person is identifiable from the disclosed image itself or information offered in connection with the image. (3) The depicted person's intimate parts are exposed or the depicted person is engaged in sexual conduct in the disclosed image. (4) The person discloses the image without the affirmative consent of the depicted person. (5) The person discloses the image under circumstances such that the person knew or should have known that the depicted person had a reasonable expectation of privacy.").

[85] *See* Harris, *supra* note 23, at 120 ("Statutory language varies widely by jurisdiction with many states targeting 'revenge porn" cases where former sexual partners post sexually explicit photos or videos of a person online to cause distress or embarrassment. These laws feature certain phrases that are both applicable and inapplicable to personal deepfakes.").

Therefore, a victim of revenge porn may pursue a successful criminal claim against a perpetrator, halting distribution and setting a standard of intolerance for such distribution.[86] Victims of deepfake pornography, on the other hand, are subject to legal inconsistencies which allow the continued online distribution of nonconsensual, falsified videos of themselves.[87]

III. The Legal Mechanisms Affecting Distribution of Deepfake Pornography

Deepfake pornography is poised to become a problem for courts, individuals, and internet service providers as advancements in technology and its increasing affordability result in a surge of deepfake pornography appearing on all corners of the internet.[88] Deepfake pornography can now be found in expected places, like Pornhub, as well as large yet less assuming platforms, such as Twitter.[89] As the number of deepfake pornographies continue to rise, so too do questions regarding the legality of the distribution of deepfake pornography, particularly in light of free speech rights, internet service provider immunity, and federal regulation.[90]

a. The First Amendment: Protection of Pornography

The First Amendment's protection of free speech is the starting point for the discussion of legal mechanisms affecting the distribution of deepfake pornography.[91] While First Amendment protections do not extend to obscenity, pornography is

[86] *See id.* at 119-20 ("Thirty-eight states and the District of Columbia have nonconsensual pornography laws.").
[87] *See id.* at 128 ("Tort doctrines and revenge porn statutes were not intended to tackle the consequences of a technology that transforms a person's sexual fantasy into reality.").
[88] *See* Cole, *supra* note 14 ("The practice of producing AI-assisted fake porn has exploded. More people are creating fake celebrity porn using machine learning, and the results have become increasingly convincing . . . [a]ll the tools one needs to make these videos are free, readily available, and accompanied with instructions that walk novices through the process.").
[89] *See* Harris, *supra* note 23, at 101-02 ("Some websites have taken marginal steps to ensure that deepfakes are not being created with the photos of non-consenting individuals. Reddit has banned the deepfakes subreddit that had a hundred thousand members. Discord has shut down two servers where the chats centered on deepfakes, and has banned several users. Pornhub and Twitter have also banned deepfake videos. However, the websites hosting these videos are shielded by a 1996 statute, the Communications Decency Act, which immunizes them from being legally responsible for user-generated content. The webpages are not incentivized to take swift action to fight these uploads, and many videos are still online.").
[90] *See* Citron, *supra* note 20, at 1936 ("First Amendment objections from perpetrators are most likely to arise in cases involving the nonconsensual disclosure of real or deep-fake nude images or sex video.").
[91] *See* Delfino, *supra* note 35, at 925 ("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.").

generally not considered obscene under the Supreme Court's test for obscenity.[92] However, the issue of whether deepfake pornography can be categorized as obscene has never been answered by courts.[93] Therefore, deepfake pornography currently remains protected speech.[94]

i. The Importance of Free Speech Rights

The right to the freedom of speech is a hallmark of the fundamental rights bestowed to citizens of the United States under the First Amendment to its Constitution.[95] First Amendment rights are lauded as essential human rights which allow progress, change in a free society, and the right to express opinion, art, and passion.[96] While the United States is known for its proclivity to uphold free speech protection, those protections are limited and aim to strike a balance between fundamental rights and preserving a dignified society.[97] Perhaps the most restricted free speech on the internet is obscenity.[98]

---

[92] *See* David Hudson, *Pornography and Obscenity*, FREEDOM FORUM INST. (July 2009) https://www.freedomforuminstitute.org/first-amendment-center/topics/freedom-of-speech-2/adult-entertainment/pornography-obscenity/ ("There are two types of pornography that receive no First Amendment protection — obscenity and child pornography. The First Amendment generally protects pornography that does not fall into one of these two categories.").

[93] *See* Harris, *supra* note 23, at 106-07 ("[N]o court has ruled on the constitutionality of banning personal deepfakes.").

[94] *See id.* at 119 ("[C]ourts must decide how to balance free speech rights and the harm that personal deepfakes can cause.").

[95] *See* U.S. Const. amend. I ("Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.").

[96] *See* Danielle Citron, *Cyber Civil Rights*, 89 B.U. L. REV. 61, 97, 98 (2009) ("One of free speech's most important functions is promoting individual autonomy. This view urges that people be free to choose their own path. Free speech facilitates self-mastery, allowing people to author their own narratives. Commentators characterize respect for autonomy of speech and thought as necessary for legitimate government. For some, freedom from any form of coercion is paramount for autonomy and dignity.").

[97] *See, e.g.*, Elise Gabrielle Sweeney, *Freedom of Speech: Protections and Limitations*, 5 GEO. J. GENDER & L. 77 (2004) ("Although the First Amendment provides broad protections to ensure an individual's right to freedom of speech, these protections are not absolute."); Kathleen Ann Ruane, *Freedom of Speech and Press: Exceptions to the First Amendment*, CONG. RESEARCH SERV. (Sept. 8, 2014) https://fas.org/sgp/crs/misc/95-815.pdf ("Even speech that enjoys the most extensive First Amendment protection may be subject to 'regulations of the time, place, and manner of expression which are content-neutral, are narrowly tailored to serve a significant government interest, and leave open ample alternative channels of communication.'").

[98] *See* C. Richard Martin, *Censorship in Cyberspace*, 34 HOUS. L. REV. 45 ("Censorship of 'obscene' and 'indecent' material in cyberspace raises several interesting issues.").

ii. Limitations to Free Speech: Obscenity and the Miller Test

While obscenity is not protected by the First Amendment, courts have difficulty articulating why.[99] The inability to form this conclusion likely stems from the trouble of defining obscenity in the first place.[100] Courts agree that the interest in limiting free expression in obscenity cases is that such protections should not apply to speech that is without any redeeming societal value.[101] But what does it mean for speech to be valueless? Perhaps Justice Stewart said it best when asked to define obscenity in *Jacobellis v. Ohio* – "I know it when I see it."[102]

The generalized approach of Justice Stewart was not always the case.[103] From 1879 to 1930, courts used the *Hicklin* test to determine obscenity.[104] This test hinged on whether the material had a tendency to corrupt the minds of those who were open to immoral influences.[105] The *Hicklin* test was ultimately replaced in 1933 by *United States v. One Book Called Ulyssses*, in which a federal district court defined obscenity as material that tended to lead to sexually impure thoughts.[106] In 1959, *Roth v. United*

[99] *Obscenity*, CORNELL LAW SCHOOL https://www.law.cornell.edu/wex/obscenity (last visited Jan. 24, 2020) ("The Supreme Court has repeatedly grappled with problematic elements of the Miller test for obscenity. However, to date, no standard has replaced it.").
[100] *See id.* ("A comprehensive, legal definition of obscenity has been difficult to establish.").
[101] *See* Modern Concept of Obscenity, 5 A.L.R. 3d 1158, ("Questions concerning the concept of obscenity are of great importance in view of the fact that Congress has passed 20 obscenity laws between 1942 and 1956, and there are similar laws in force in practically all the states and supported by international agreements of over 50 nations. All the authorities agree that obscene matters are not protected by constitutional guaranties of free speech and press.").
[102] Jacobellis v. Ohio, 378 U.S. 184, 197 (1964) ("I know it when I see it, and the motion picture involved in this case is not that.") (Stewart, J., concurring).
[103] *See* Joey Senat, *An Overview of How Courts Have Defined Obscenity*, UNIV. OF N.C. SEMINAR ON INTERNET POLICY AND FUTURE INITIATIVES ("From 1879 until the early 1930s, American courts followed the Hicklin test.").
[104] *See id.* (describing the previously applied Hicklin test); *See also Anthony Comstock's "Chastity" Laws*, PBS https://www.pbs.org/wgbh/americanexperience/features/pill-anthony-comstocks-chastity-laws/ (last accessed Jan. 24, 2020) ("In 1872 Comstock set off for Washington with an anti-obscenity bill, including a ban on contraceptives, that he had drafted himself. On March 3, 1873, Congress passed the new law, later known as the Comstock Act. The statute defined contraceptives as obscene and illicit, making it a federal offense to disseminate birth control through the mail or across state lines.").
[105] *See* Senat, *supra* note 103 ("Under the Hicklin test, judges considered a work to be obscene if any portion of the material had a tendency 'to deprave or corrupt those whose minds are open to such immoral influences, and into whose hands a publication of this sort may fall.'").
[106] *See, e.g.*, Senat, *supra* note 103 ("In 1933, the Hicklin test was toppled in United States v. One Book Called Ulysses, in which a federal district judge decided to allow James Joyce's 'Ulysses' to be imported and sold in America. Judge John M. Woolsey focused on the literary value of the entire work and its effect on a person with average sex instincts. He defined obscene as 'tending to stir the sex impulses or to lead to sexually impure and lustful thoughts."); *U.S. v. One Book Entitled Ulysses by James Joyce*, 72 F.2d 705, 706 (1934) (explaining that on appeal, the court noted that "[t]hough the depiction happily is not of the 'stream of consciousness' of all men and perhaps of only those of a morbid type, it seems to be sincere, truthful, relevant to the subject, and executed with real art").

States entered as the first constitutional challenge to obscenity law.[107] *Roth* affirmed the view that obscenity was not protected speech and ruled the appropriate test for obscenity was whether the average person, applying community standards, would find that the work appealed to prurient interest.[108] Seven years later, in *Jacobellis v. Ohio*, the Supreme Court held that under the First Amendment, criminal laws in the area of obscenity violations were limited to "hard core" pornography, but was unable to deliver a definitive definition of the term.[109]

The current obscenity standard stems from a three-prong test developed by the Supreme Court in *Miller v. California*. [110] In this 1973 case, Marvin Miller violated a statute prohibiting the distribution of obscene material when he mailed advertisements involving sexually explicit material.[111] The Supreme Court ruled the material could be found obscene when the average person would find the work to: 1) appeal to prurient interest based on contemporary community standards, 2) depict sexual conduct in a patently offensive way, and 3) lack serious literary, artistic, political, or scientific value.[112] Under the *Miller* test, material must satisfy all three prongs before it can be considered obscene and outside the bounds of First Amendment protection.[113] The *Miller* test served to balance the competing interests of First

[107] *See, e.g.*, Senat, *supra* note 103 (explaining that "Roth was the first constitutional challenge to obscenity law" and that "before *Roth*, obscenity cases did not implicate the First Amendment because it was understood that obscenity was prima facie unprotected by freedom of expression"); SUSAN DWYER, THE PROBLEM OF PORNOGRAPHY (1995) (explaining that the court in Roth "ruled that both state and federal provisions concerning obscenity were constitutional").

[108] *See* Roth v. United States, 354 U.S. 476, 487 (1957) ("Obscene material is material which deals with sex in a manner appealing to prurient interest.").

[109] *See Jacobellis*, *supra* note 102 (providing Justice Stewart's opinion that states "I have reached the conclusion, which I think is confirmed at least by negative implication in the Court's decisions since Roth and Alberts, that under the First and Fourteenth Amendments criminal laws in this area are constitutionally limited to hard-core pornography. I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description; and perhaps I could never succeed in intelligibly doing so").

[110] *See* Miller v. California, 413 U.S. 15, 19-20 (1973) ("It is in this context that we are called on to define the standards which must be used to identify obscene material that a State may regulate without infringing on the First Amendment as applicable to the States through the Fourteenth Amendment.").

[111] *See id.* at 18 ("This case involves the application of a State's criminal obscenity statute to a situation in which sexually explicit materials have been thrust by aggressive sales action upon unwilling recipients who had in no way indicated any desire to receive such materials.").

[112] *See id.* at 24 ("The basic guidelines for the trier of fact must be: (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. If a state obscenity law is thus limited, First Amendment values are adequately protected by ultimate independent appellate review of constitutional claims when necessary.").

[113] *See* Senat, *supra* note 103 ("Material must meet all three parts if it is to be ruled obscene and outside of First Amendment protection.").

Amendment protections and the State's interest in protecting citizens from exposure to obscene pornographic materials.[114]

Although *Miller* sought to clarify obscenity, its resulting test was not without gaps.[115] Particularly, scholars debate the implications of balancing community and national standards.[116] The first two prongs of the Miller test involve community standards, while the third is held to the standard of a reasonable person of the United States as a whole.[117] The idea behind this standard is to protect works that may have societal value, despite being considered obscene in a certain community.[118] Scholars argue the first Miller prong does not define the relevant community to be used in its application.[119] Further, questions arise regarding the lack of cohesiveness of the second Miller prong, which leaves the issue of obscenity up to each state.[120] Finally, the Miller test has been subject to strict scrutiny as the digital age has emerged.[121] The advent of the internet has caused even more confusion when applying the community standards of the first two prongs, which permits a different standard of obscenity in New York than Mississippi.[122] In *United States v. Kilbride*, the State Court of Appeals

[114] *See* Spivak, *supra* note 58, at 360 ("The Miller standard . . . was an accommodation between the State's interests in protecting the 'sensibilities of unwilling recipients' from exposure to pornographic material and the dangers of censorship inherent in unabashedly content-based laws.").
[115] *See* Bradley Shafer, *Sex, Lies, and Videotape: In Critique of the Miller Test of Obscenity*, 70 MICH. B.J. 1038, 1041 (1991) (describing the confusing nature of Miller's three-part test).
[116] See E. Morgan Laird, The Internet and the Fall of the Miller Obscenity Standard: Reexamining the Problem of Applying Local Community Standards in Light of a Recent Circuit Split, 52 SANTA CLARA L. REV. 1503, 1514 (2012) ("In explaining the test, the Court rejected the proposition that 'contemporary community standards' should be a national standard, but rather held that the community in which the material was found should judge the material.").
[117] *See* Pope v. Illinois, 481 U.S. 497, 500-01 (1987) ("The proper inquiry is . . . whether a reasonable person would find [serious literary, artistic, political, or scientific] value in the material, taken as a whole.").
[118] *See* Shafer, *supra* note 115, at 1041 ("Interestingly, the [Miller] Court concluded that serious value could not be evaluated pursuant to community standards since the value of material does not vary from community to community.").
[119] *See* Robin Whitehead, *"Carnal Knowledge" is the Key: A Discussion of How Non-Geographic Miller Standards Apply to the Internet*, 10 NEXUS 49, 51 (2005) ("In Miller, the Court approved of the instruction of a "statewide" community standard but did not mandate any precise geographic yardstick.").
[120] *See* Spivak, *supra* note *58*, at 360 ("For example, the District of Columbia has determined that under the District's statute barring obscenity, materials depicting or live performances of oral sex are per se obscene, meaning the Government need not proffer any evidence of national community standards. Similarly, the Court of Appeals of South Carolina has determined '[n]ude dancing per se is not illegal.'"). *See also Miller*, 413 U.S. at 32 (1973) ("It is neither realistic nor constitutionally sound to read the First Amendment as requiring that the people of Maine or Mississippi accept public depiction of conduct found tolerable in Las Vegas, or New York City.").
[121] *See* Sarah Kagan, *Obscenity on the Internet: Nationalizing the Standard to Protect Individual Rights*, 38 HASTINGS CONST. L.Q. 223, 242 (2010) ("The problem we encounter today is due in part to the fact that the Court in the time of Miller could not envision the amorphous and viral nature of the Internet.").
[122] *See id.* at 241 ("The notion of a designated community seems antiquated in the digital era, particularly when material can be produced in one part of the world or nation and then, in the blink of an eye, be sold and transferred automatically, from one site to another, finally ending up in a third locale.").

for the Ninth Circuit ruled that a national community standard should be implemented when evaluating obscenity on the internet, but such a ruling has yet to be solidified at a national level.[123]

iii. The legality of Pornography Under the Miller Test

The legality of pornography and its distribution on the internet has had just as tumultuous a relationship with the courts as the definition of obscenity.[124] Importantly, the way in which both interact paves the way for First Amendment protection of pornography.[125] Under the *Miller* obscenity test, pornography can be protected by the First Amendment.[126] A common argument regarding First Amendment protection of pornography is its failure to satisfy the third Miller prong, which asks if the work has any societal value.[127] In many cases, pornography can be considered artistic expression due to its use of actors.[128] Further, pornography has been argued to provide other societal benefits such as therapeutic purposes involving sexual impotence, or educational purposes involving sexuality and sexual intercourse.[129] Therefore, the

[123] United States v. Kilbride, 584 F.3d 1240, 1250 (9th Cir. 2009) (describing the Ninth Circuit's holding that courts should apply a national community standard when evaluating the obscenity of online speech).
[124] *Pornography*, US LEGAL (last visited Feb. 17, 2020), https://internetlaw.uslegal.com/pornography/. ("Internet Pornography is a battlefield in U.S. law.").
[125] *See, e.g., id.* ("Internet pornography is a battlefield in U.S. law. Since the explosion of public interest in the Net in the 1990s, the public, lawmakers, and the courts have argued over how to control online porn."); *Why is Pornography Legal and Prostitution is Not*, HG.ORG (last visited Feb. 17, 2020), https://www.hg.org/legal-articles/why-is-pornography-legal-and-prostitution-is-not-31164. ("Pornography has had a contentious relationship with the law since the middle of the Twentieth Century.").
[126] *See* Shafer, *supra* note 115, at 1043 ("Sex and obscenity are simply not synonymous. Yet, even the Supreme Court has noted that the two are separated only by a 'dim and uncertain line.' In addition, courts agree that, merely because materials are erotic, sexually explicit, or even 'hard core,' those characteristics by themselves do not necessarily render the materials 'obscene' or mean that they are anything other than fully protected expression under the First Amendment.").
[127] *See Miller*, 413 U.S. at 24-25 (describing the third prong of the Miller test).
[128] *See Why is Pornography Legal and Prostitution is Not*, *supra* note 125 ("[A] string of cases find that porn performances actually constitute acting subject to the artistic expression protections of the First Amendment to the United States Constitution.").
[129] *See, e.g.*, Jeneanne Orlowski, *Beyond Gratification: The Benefits of Pornography and the Demedicalization of Female Sexuality*, 8 MOD. AM. 53, 54, 64 (2012) ("Proponents for the protection of pornography argue that pornography can be a release of sexual tension that contributes to a decrease in sexual violence . . . [p]ornography has given women an outlet to express themselves, a form of literature to educate themselves, and a tool with which to communicate their feelings and lack of fulfillment."); Emily Rothman, *Domestic Violence – What's Porn Got to Do With It?*, B.U. SCH. PUB. HEALTH (Oct. 20, 2015), https://www.bu.edu/sph/news/articles/2015/viewpoint-domestic-violence-whats-porn-got-to-do-with-it/. ("The results of multiple studies suggest that pornography can help some individuals realize or negotiate their sexual identity, improve couples' sexual satisfaction, provide helpful information about the mechanics of sex, promote safer sex practices, and improve sexual response for individuals suffering from dysfunction disorders.").

First Amendment protects all pornography so long as it is not considered obscene.[130] However, even obscene pornography can legally be possessed in the privacy of one's home, a distinction the Supreme Court made clear in the 1969 case *Stanley v. Georgia*.[131]

The legality of pornography and the internet's ability to mass distribute copious amounts of it led to difficulties policing pornography when applying the *Miller* test.[132] If the *Miller* test were to be applied to the internet, it's community standards provisions would be based on the most conservative community in the United States, drastically restricting free speech.[133] Attempts then to limit obscene pornography on the internet moved from judicial to legislative action.[134]

b. Section 230: Protection of Internet Service Providers

In 1996, Congress enacted the Communication Decency Act (CDA). [135] This Act aimed to change the internet in two significant ways.[136] First, the CDA attempted to regulate indecency and obscenity on the internet.[137] Second, the CDA included a provision holding Internet Service Providers (ISPs) immune from liability for the

[130] *See* Orlowski, *supra* note 129, at 53 ("Under the First Amendment, there is a presumptive protection of all speech. In order for sexual speech to fall beyond that protection, there must essentially be a showing that the speech is obscene.").
[131] *See* Stanley v. Georgia, 394 U.S. 557, 568 (1969) ("We hold that the First and Fourteenth Amendments prohibit making mere private possession of obscene material a crime.").
[132] *See* Stephanie Morrow, *How is Obscenity Regulated on the Internet*, LEGAL ZOOM (last visited Feb. 17, 2020), https://www.legalzoom.com/articles/how-is-obscenity-regulated-on-the-internet ("The Internet has made the law of obscenity much more convoluted. Federal obscenity laws apply to interstate and foreign issues, such as distribution; intrastate issues are mostly governed by state law. Today, materials considered 'obscene' can be sent from a computer in California to someone across the U.S. as fast as a click of a button. The question is: What state governs the issue of obscenity when the Internet can reach multiple areas? Interestingly, the Miller Test is based on what is offensive in a certain 'community,' not the United States as a whole. For example, what's offensive to someone from New York City may differ from what offends a person in Topeka, Kansas. But, the Miller Test's basis of 'community' becomes blurred with the advent of the Internet; a state can define a community as the state as a whole, a county, a city or another geographic area.").
[133] *See id.* ("The geographic area of the Internet, however, is nonexistent, and geographic boundaries are essential to the 'community' definition for the Supreme Court's Miller Test.").
[134] *See Pornography*, US LEGAL (last visited Feb. 17, 2020) ("Seeking to control Internet porn, Congress first passed legislation in 1996. The Communications Decency Act (CDA) criminalized the dissemination over computer networks of obscene or indecent material to children.").
[135] *See* Mary Leary, *The Indecency and Injustice of Section 230 of the Communications Decency Act*, 41 HARV. J.L. & PUB. POL'Y 553, 554 (2018) ("Passed in 1996, the CDA was an attempt by Congress to accommodate competing values and facilitate an uncertain but promising future digital world.").
[136] *See id.* at 559 ("Section 230 was a component of a broader effort to limit access to explicit material through the Internet. The CDA intended to limit such access and was attached to Title V of the Telecommunications Act of 1996.").
[137] *See id.* at 558-59 ("Congress acknowledged and expressed concern about the potential of the Internet to spread or expose children to obscene material.").

postings of third parties.[138] This provision, known as Section 230, is the only surviving portion of the CDA and is heralded as the communication law that established the modern internet.[139] However, as illicit materials including obscenity and deepfake pornography continue to multiply on the internet, questions arise as to whether ISP immunity promotes free speech or encourages the distribution of unprotected speech.[140]

i. Policing Porn: The CDA and Section 230

The CDA was created in response to heightened fear that children would be exposed to obscene pornography as the internet moved from its infancy into a more accessible, widely used medium.[141] Language of the statute prohibited knowingly disseminating obscene materials to children and encouraged telecommunication companies to block explicit content from reaching impressionable minors.[142] As part of its mission, the CDA included Section 230, which served two purposes within the statute. [143] First, Section 230 allowed ISPs to police content on servers by permitting the removal or restriction of material the ISP deemed lewd, harassing, obscene, violent, or otherwise objectionable.[144] Second, Section 230 sought to encourage free speech by holding

[138] *See id.* at 559 ("Section 230 was added to the CDA to protect tech companies.").
[139] *See CDA 230: The Most Important Law Protecting Internet Speech*, EFF.org (last visited Feb. 17, 2020), https://www.eff.org/issues/cda230 ("This legal and policy framework has allowed for YouTube and Vimeo users to upload their own videos, Amazon and Yelp to offer countless user reviews, craigslist to host classified ads, and Facebook and Twitter to offer social networking to hundreds of millions of Internet users.").
[140] *See* Haley Halverson, *Ending Immunity of Internet-Facilitated Commercial Sexual Exploitation Through Amending the Communications Decency Act*, 21 NO. 12 J. INTERNET L. 3, 5 (2018) ("Legal interpretation of section 230 of the CDA sparked controversy from its inception and laid the groundwork for on-going debates regarding the tension between unfettered free speech and online safety.").
[141] *See id.* at 4 ("The CDA was passed when public use of the Internet was growing in leaps and bounds; and not surprisingly, it contained sections that relate to the Internet, including section 223, which attempted to regulate sexually explicit content on the Internet.").
[142] *See* Leary, *supra* note 135, at 558-59 ("[C]ongress recogniz[ed] a concern about online exploitation . . . [t]he CDA prohibited the knowing dissemination of obscene material to children, and sought to incentivize telecommunication companies to participate in blocking explicit material from reaching children.").
[143] *See* Julio Sharp-Wasserman, *Section 230(c)(1) of the Communications Act and the Common Law of Defamation: A Convergence Thesis*, 20 COLUM. SCI. & TECH. L. REV. 195, 197 (2018) ("Section 230(c)(1) of the CDA states, '[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.' Section 230(f) defines 'interactive computer service' as 'any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server;' and defines 'information content provider' as 'any person or entity that is responsible, in whole or in part, for the creation or development of information'").
[144] *See* Leary, *supra* note 135, at 561 ("Congress . . . sought to address two goals with § 230. First, consistent with the CDA's effort to protect children from access to obscene or explicit materials, Congress sought to 'encourage telecommunications and information service providers to deploy new technologies and policies' to block or filter offensive material.").

ISPs harmless from liability for the postings of third parties.[145] As a result, a third party who posts a Facebook status could be held liable for its defamatory language, whereas Facebook, as an ISP, could not.[146]

The CDA became law on February 8, 1996 and was met with severe backlash from the American Civil Liberties Union (ACLU) and the online community.[147] The ACLU argued the portion of the law regulating obscenity was too broad and, therefore, too constricting of free speech rights.[148] In protest, several websites chose to black out their webpages, and the ACLU joined several civil liberties organizations in a suit against the overbroad provisions.[149] The case eventually reached the Supreme Court in *Reno v. American Civil Liberties Union*, where the anti-indecency sections of the Act were struck down in the name of protecting First Amendment rights.[150] However, Section 230, which promoted free speech by holding ISPs immune from liability for the posting of third parties, survived.[151]

ii. Section 230: Shifting Liability

Section 230 was created with the intent to protect internet speech by reducing overzealous policing by ISPs who would otherwise be held liable for content posted by third parties.[152] However, as the internet progressed and its content grew exponentially, scholars began to question whether the immunity provisions of Section

[145] *See id.* ("On the other hand, it did not want companies to over-screen, as Congress recognized the desire for the Internet to reach its full potential as "a forum for a true diversity of political discourse, unique opportunities for cultural development, and myriad of avenues for intellectual activity.").
[146] *See* Wasserman, *supra* note 143, at 197 ("[F]or instance, while one could hold a YouTube video uploader (an information content provider) liable for defamation, one could not hold YouTube (an interactive computer service) liable as a 'publisher or speaker' of that video, because, under CDA § 230(c)(1) the video is 'information provided by another information content provider.'").
[147] *See CDA 230 The Most Important Law Protecting Internet Speech*, EFF (last visited Feb. 17, 2020), https://www.eff.org/issues/cda230/legislative-history ("With Section 230 in the bill, the Telecommunications Act was signed into law on February 8, 1996. That same day, the ACLU filed a legal challenge for a temporary restraining order on the bill's indecency provisions. The online community was outraged by the passage of the bill.").
[148] *See id.* ("EFF decried CDA's overly broad language.").
[149] *See id.* ("Several sites chose to black out their websites in protest."). *See also* Leary, *supra* note 135, at 562 ("In Reno v. ACLU, the Supreme Court struck down as vague some of the more controversial criminal provisions of the CDA, such as the prohibition on the transmission of 'indecent material.' However, § 230 was not challenged, and this protection remains effective law to this day.").
[150] *See id.* ("The ACLU's case, which several civil liberties organizations like the EFF as well as other industry groups joined, reached the Supreme Court. On June 26, 1997, in a 9-0 decision, the Supreme Court applied the First Amendment by striking down the anti-indecency sections of the CDA.").
[151] *See id.* ("Section 230, the amendment that promoted free speech, survived.").
[152] *See Fair Housing Council of San Francisco v. Roommates*, 521 F.3d 1157, 1163 (9th Cir. 2008) (Congress enacted § 230 to remove the disincentives to self-regulation).

230 still made sense.[153] Proponents argue that the expansion of the internet is an even more compelling reason to uphold Section 230 because an ISP is unable to review every posting by a third party.[154] This argument was the basis for the holding in *Zeran v. American Online, Inc*, in which the Fourth Circuit Court of Appeals ruled that Section 230 created federal immunity against any cause of action that would hold ISPs liable for information posted by third parties.[155] Zeran's interpretation was followed by many future court decisions and cemented Section 230 as a check against an impermissible restriction of free speech rights.[156]

Opponents of Section 230, however, argue that the advancement of the internet warrants more responsibility from ISPs because at the time Section 230 was enacted, the internet was in its infancy.[157] In 1996, artificial intelligence was nearly nonexistent, standards of obscenity were drastically different, and monolithic ISPs were not the norm.[158] Now, the world is at the fingertips of any individual. Artificial intelligence has exploded, constant exposure to sexual materials has changed standards of obscenity, and large ISPs are pervasive.[159] There now exists a fine line between constitutionally protected and valuable speech as opposed to speech that is obscene, illicit, or otherwise valueless – and there is little way to police it.[160] While this line becomes finer, opponents of Section 230 argue that the immunity afforded

[153] Halverson, *supra* note 140, at 5 ("Legal interpretation of section 230 of the CDA sparked controversy from its inception and laid the groundwork for on-going debates regarding the tension between unfettered free speech and online safety.").
[154] *Id.* at 6 ("On the one hand, some scholars embrace court decisions that interpret section 230 of the CDA to effectively instill immutable immunity regarding third-party postings.").
[155] *See* Zeran v. American Online, Inc., 129 F.3d 327, 331 (4th Cir. 1997) ("§ 230 forbids the imposition of publisher liability on a service provider for the exercise of its editorial and self-regulatory functions.").
[156] Halverson, *supra* note 140, at 6 (describing *Zeran v. American Online, Inc.* as a "hallmark case [that] paved the way for many other similar court decisions").
[157] Citron, *supra* note 20, at 1934 ("We find ourselves in a very different moment now than we were in five or ten years ago, let alone twenty years ago when § 230 was passed.")
[158] Leary, *supra* note 135, at 556 ("In 1996 the Internet was in its infancy and Congress was struggling with the implications of its development. The Internet of 1996 is unrecognizable today. That 'new' 'dial up' Internet engine connected people through a novel and experimental 'bulletin board' through which events could be organized. Newspapers were just considering having an online presence. 'Google' was not a verb, and online research was described as 'tough for the amateur researcher.' Congressional debate discussed floppy disk drives, usenet groups, and message boards over telephone lines. In this climate, Congress could not have imagined what the Internet would look like two decades into the twenty-first century.").
[159] Halverson, *supra* note 140, at 5-6 ("It is important to recognize that the CDA came on the scene in the mid-1990s, nearly simultaneous with the widespread use of the Internet. Although the Internet evolved over time, it was not until around 1993 that laymen usership spiked.").
[160] Harris, *supra* note 23, at 102 ("Unfortunately, as with many new technologies, the law is unequipped to handle these impending issues. Courts must answer questions like: should state tort doctrines or involuntary porn statutes be interpreted to encompass fictitious fabricated videos? Does Congress need to pass a law to handle these types of cases? Or, does the First Amendment completely immunize the publication and creation of deepfakes as a form of protected speech?").

to ISP creates no incentive to remove harmful materials, even though ISP are granted broad deference to do so.[161] This argument was raised in *Blumenthal v. Drudge* when a federal district judge questioned whether granting immunity to ISP as an incentive to police the internet for obscenity and other offensive material was logical when self-policing by ISPs was rarely attempted if at all.[162]

As the internet technically advances and is a platform for more illicit content, the resulting struggle to distinguish when or whether an ISP should step in to prevent the distribution of such content remains constant, particularly in the case of deepfake pornography.[163] This is because the technology used to create deepfake pornography is also used to create constitutionally protected speech. For example, deepfakes can be used to communicate humor, such as placing friends' faces in popular movies, or facilitate discussion in the form of politically satirical videos.[164] Deepfakes can be used for innocent purposes, as in the case of one deepfake creator who used the software to place his wife on the body of Anne Hathaway in an interview with David Letterman.[165] Even more problematic is the fact that deepfakes can also be used to create protected pornography, such as recreations of intimate movie scenes by consenting adults.[166] A conflict therefore arises as to the extent to which ISPs can or should police deepfake pornography.[167]

[161] Halverson, *supra* note 140, at 6 ("On the other hand, some scholars question the broad interpretation of section 230, arguing for some level of further responsibility for interactive computer services.").
[162] *See* Blumenthal v. Drudge, 992 F. Supp 44, 51-52 (D.D.C. 1998) (describing Judge Paul Friedman's disagreement with the common interpretation of Section 230 because immunity is granted to ISP "even where the self-policing is unsuccessful or not even attempted").
[163] Halverson, *supra* note 140, at 7 ("This broad interpretation of section 230 has increasingly apparent negative consequences 'as more and more criminal activity migrates to the Internet, and the online intermediaries that knowingly host such activity are held immune from traditional modes of checking such lawlessness.'").
[164] Caldera, *supra* note 17, at 179 ("Examples of deepfakes range from the silly to the sinister. Some of the lighter applications of deepfakes include videos putting Nicholas Cage into famous scenes from movies such as Raiders of the Lost Ark or videos of a Wall Street Journal reporter performing Bruno Mars's dance moves.").
[165] Spivak, *supra* note 58, at 348 ("Deepfakes are also used to superimpose an average member of the public onto a celebrity's body. As one blogger wrote, 'we can leverage these celebrities for other things, such as inserting your friends and family into blockbuster movies and shows!' That blogger then turned his wife's likeness[,] on the body of Anne Hathaway[,] into an interviewee opposite David Letterman and a film star opposite Steve Carrell. In his words: I personally think it's fun, can be innocent, and even makes for a nice surprise/gift . . . [n]ow you can put your best friend into his favorite movie: have her dance with Patrick Swayze and have the time of her life, or have an alien burst out of his stomach.")
[166] Harris, *supra* note 23, at 105 ("If the deepfake does not violate community standards (e.g., a non-graphic pornographic deepfake) or has some artistic value (e.g., a deepfake featuring a unique blend of colors) then a state or federal law prohibiting deepfakes would be unconstitutional.").
[167] Halverson, *supra* note 140, at 6 (describing the conflicting views of proponents and opponents of Section 230 immunity for ISP).

Because of the conflicts in determining what should be policed by an ISP, a middle ground between opponents and proponents of Section 230 advocates specific exceptions to ISP immunity through legislative amendments.[168] Since its inception, however, Section 230 has only been amended once with the implementation of the Stop Enabling Sex Traffickers Act (SESTA), which excludes ISP immunity from the enforcement of federal or state sex trafficking laws.[169] Besides amendments like SESTA, Section 230 does not contain a force of law which compels an ISP to remove harmful content.[170] Without a compelling force of law, deepfake pornography can be distributed on the internet and platforms like Pornhub have no incentive to ensure the pornography is "real" or respond to removal requests from individuals like Noelle Martin.[171]

c. Lack of Federal Regulation: Protection of Distribution of Deepfake Pornography

No existing federal criminal law prohibits the broad umbrella of nonconsensual pornography, let alone the specific sector of deepfake pornography.[172] However, legislative attempts to federally regulate nonconsensual pornography and deepfake pornography have begun to gain traction.[173] In 2016, Congresswomen Jackie Spier proposed the Intimate Privacy Protection Act (IPPA) in an attempt to deliver justice to victims of revenge porn who were unprotected by state criminal laws and unable to afford civil suits.[174] IPPA makes it a crime to distribute sexually intimate images despite having knowledge that the victim did not consent to their distribution.[175] However, the bill expired at the culmination of the 114th Congress.[176] Following IPPA in 2017 was the Ending Nonconsensual Online User Graphic Harassment (ENOUGH)

[168] *Id.* at 13 (describing the necessity of legal deterrents to Section 230, such as the FOSTA-SESTA amendment which holds ISP's liable for facilitating sex-trafficking).
[169] *See* Mark Sullivan, *The 1996 Law That Made the Web is in the Crosshairs*, FAST COMPANY (Nov. 29, 2018) https://www.fastcompany.com/90273352/maybe-its-time-to-take-away-the-outdated-loophole-that-big-tech-exploits (stating Section 230 doesn't "contain the force of law to compel" ISP to remove content).
[170] *Id.* ("There's no language saying, 'Get that garbage off your site within 24 hours or else!' So, for the tech companies, it remains a largely PR and public policy issue, not something that directly affects their bottom line. The big platform companies have been doing just enough content takedowns and bad-actor ejections to keep new regulations at bay.").
[171] Harris, *supra* note 23, at 105 ("The publication of all pornographic . . . deepfakes cannot be deemed obscene under the Miller test.")
[172] Delfino, *supra* note 35, at 904 ("No specific federal law criminalizes deepfakes or revenge porn.").
[173] *Id.* (describing the proposed criminal federal laws used to punish nonconsensual pornography).
[174] *Id.* at 906 ("In 2016, Congresswoman Jackie Speier (D-CA) introduced the Intimate Privacy Protection Act (IPPA) . . . IPPA was a response to the fact that most revenge porn victims do not have the resources to seek civil remedies.").
[175] *Id.* (discussing IPPA's criminalization of the distribution of deepfake pornography).
[176] *Id.* at 907 ("The bill expired at the end of the 114th Congress.").

Act.[177] The ENOUGH Act was a revised version of IPPA and criminalized the distribution of nonconsensual pornography done with knowledge or reckless disregard for both consent and harm caused by distribution.[178] Despite bipartisan support for the Act, it was never brought to fruition and expired with the 115th Congress.[179] The first federal bill criminalizing deepfakes was introduced in 2018.[180] Known as the Malicious Deep Fake Prohibition (MDFP) Act, the bill prohibited the use of interstate commerce to either create, with the intent to distribute, a deepfake with the intent that such distribution would facilitate illegal conduct or distribute an audiovisual record despite knowing such record was a deepfake.[181] The bill was introduced just days before the December 2018 government shut down and expired as a result.[182]

Absent a federal law prohibiting the distribution of deepfake pornography, victims like Noelle Martin must look to state criminal and civil law to achieve redress.[183] However, while many states enforce some type of law prohibiting revenge porn, very few states specifically prohibit deepfake pornography.[184] California is one of few states which has enacted deepfake pornography legislation.[185] In 2019, AB-602 was passed by the California State Senate.[186] The bill created a private right of action

[177] Jessica Magaldi, *Revenge Porn: The Name Doesn't Do Nonconsensual Pornography Justice and the Remedies Don't Offer the Victim's Enough Justice*, 98 OR. L. REV. 197, 226 (2020) ("In November 2017, another federal nonconsensual pornography law was proposed contemporaneously in the Senate and the House.This bill is known as the ENOUGH Act.").
[178] *Id.* (describing the criminalization of distribution of nonconsensual pornography under the ENOUGH Act).
[179] *Id.* ("The bill was referred to the Senate Committee on the Judiciary in November 2017 . . . [i]t has not progressed to the floor.").
[180] Delfino, *supra* note 35, at 908 ("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.").
[181] *Id.* ("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.").
[182] *Id.* at 909 ("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.").
[183] *Id.* ("In the absence of federal laws outlawing nonconsensual pornography, victims are left with a "patchwork of state criminal laws [that] is often inadequate.").
[184] *Id.* ("[V]ictims of pornographic deepfakes who want to seek redress under state law must look to laws that criminalize related crimes, such as revenge porn.")
[185] *Id.* at 912 ("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.")
[186] Kamran Salour et al., If Signed by Governor, California Bill AB-602 Will Provide Private Right of Action for Victims of Sexually Explicit 'Deepfakes,' BAKER HOSTETLER (Sept. 26, 2019) https://www.dataprivacymonitor.com/state-legislation/if-signed-by-governor-california-bill-ab-602-will-provide-private-right-of-action-for-victims-of-sexually-explicit-deepfakes/ ("AB-602, [was] passed by the California State Senate on September 12, 2019.").

against a person who creates and knowingly discloses sexually explicit material despite knowing the individual depicted in the material did not consent or against a person who does not create the material but intentionally distributes it despite having knowledge of a lack of consent.[187] If a state lacks laws prohibiting the distribution of deepfakes pornography, or if that state's revenge porn statute's language excludes the unique medium of deepfake pornography, victims like Noelle must attempt to fit the unique circumstances of deepfake pornography within the walls of state criminal and civil laws.[188] However, it is unclear whether either applies to deepfake pornography or will prevent the rampant dissemination of the falsified sexual videos.[189]

Accordingly, the protection of pornography on the internet, lack of incentive for an ISP to police deepfake pornography, and absence of a federal law prohibiting the distribution of deepfake pornography all contribute to the current legal environment, which allows the distribution of deepfake pornography to multiply on the internet.[190] As such, victims like Noelle Martin are left with few options to stop the online proliferation of their forged sex lives and the irreparable personal harm that will result.[191]

IV. The Distribution of Deepfake Pornography and Its Disproportionate Effect on Women

Over the last decade, several studies have concluded that a significant correlation exists between pornography and violent behavior and attitudes towards women.[192]

[187] *Id.* (stating AB-602 will "create a private right of action against persons who create or disclose another's sexually explicit content through use of 'deepfake' technology. Specifically, the cause of action may be brought against a person who creates and intentionally discloses sexually explicit material where the person knows, or reasonably should know, that such creation or disclosure was not consented to by the depicted individual, or where such person did not create but intentionally discloses such material knowing that the depicted individual did not consent to its creation").

[188] Delfino, *supra* note 35, at 901 ("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.").

[189] *Id.* (describing the differences between the focus of various deepfake remedies and how the lack of cohesiveness contributes to ineffective recourse for victims).

[190] *Id.* at 898 (explaining the legal challenges of deepfakes, including the lack of a criminal solution and the difficulties of convincing internet platforms to remove deepfake pornography).

[191] *Id.* at 918 (describing the "shortcoming and limitations" of current solutions to deepfake pornography and arguing that "neither can fully remedy the harms created by nonconsensual deepfake pornography").

[192] Kiel Brown, *How Pornography Impacts Violence Against Women and Child Sex Abuse*, FOCUS FOR HEALTH , https://www.focusforhealth.org/how-pornography-impacts-violence-against-women-and-child-sex-abuse/ (last visited Aug. 6, 2019) ("According to a 2010 study that analyzed 304 scenes from best-selling pornography videos, almost 90% of scenes contained physical aggression, while nearly 50% contained verbal aggression, primarily in the form of name-calling. Targets of these displays of aggression were overwhelmingly women and either showed pleasure or neutrality in response to the aggression. Some studies

These studies concluded that the prevalence of violence in nearly all pornography has an influence on non-conscious and conscious stimuli, meaning those who watch violent pornography are being conditioned to be desensitized not only to the pornography itself but the violent behavior toward women in general.[193] Further, studies find the vast majority of pornography includes misguided and harmful behavior-teaching and permission-giving experiences.[194] These experiences include the absence of consent, the encouragement to induce violence upon women, and the pleasure or nonchalance women exhibit in response to that violence.[195] Even worse, pornography creates a scene in which viewers can place themselves, and often takes place in familiar settings, such as work, school, social events, even medical appointments.[196] In these scenes, consent is disregarded, and female actors treated violently respond with pleasure.[197]

This constant flux of familiarity and aggression as well as the lack of consent makes the ability to distinguish reality from fictional pornography increasingly difficult and is thought to be a reason for increasing manifestations of violence against real women.[198] This correlation is increasingly true as pornography is viewed at a younger

---

that have shown nearly 90% of pornography depicts violence while other studies have placed the prevalence at only 2%. One of the most disturbing facts about the prevalence of violence in porn is that nobody can agree on what they consider to be violent content. What can be proven rather definitively is the association between pornography use in general and violence against women.").

[193] *Id.* ("A meta-analysis published in Aggressive Behavior confirmed this link, and went on further to state that there was a significant correlation between sexually violent pornography and attitudes supporting violence against women. This correlation supports findings suggesting that increased pornography use has an influence on non-conscious responses to stimuli, meaning that we are both consciously and unconsciously being conditioned by pornography in a negative way. Consuming any content on a consistent basis has a way of altering our perceptions about that content, and pornography is no different. If one were to watch violent pornography frequently, it would desensitize them to both pornography and violence, specifically towards women.").

[194] Patrick Hough, *The Social Costs of Pornography: A Statement of Findings and Recommendations*, J. OF THE WEATHERSPOON INST. (Mar. 23, 2010), (finding that "much of the social harm associated with pornography consumption seems to spring from its psychological nature as an intense behavior-teaching and permission-giving experience within the highly effective teaching context of sexual arousal, where actions are demonstrated, repeated, encouraged and/or proscribed via information-rich images.").

[195] Brown, *supra* note 192 ("If most porn was teaching respect, consent, and healthy sexual expectations, this conscious and unconscious conditioning might be a boon. As it stands, porn serves to build and reinforce dangerous perceptions.").

[196] *Id.* ("Many production studios run series sexualizing women at work, at school, and even at the doctor, often with female participants shown as either enjoying or being victimized by violent sexual acts.").

[197] *Id.* ("Targets of these displays of aggression were overwhelmingly women and either showed pleasure or neutrality in response to the aggression.").

[198] *11 To 14-Year-Olds Want To Mimic Sex Acts Shown In Porn, Survey Find*s, FIGHT THE NEW DRUG (Dec. 7, 2020) https://fightthenewdrug.org/massive-study-reveals-what-kids-are-watching-learning-from-online-porn/ ("The National Society for the Prevention of Cruelty to Children (NSPCC) conducted a survey of more than 1,000 children aged 11-16, and found that at least half had been exposed to online porn . . . [o]ne of the

age than ever before.[199] Studies revealed that boys who view pornography at a young age are more prone to sexual aggression and coercion.[200] Conversely, females who view pornography at a young age are prone to tolerate emotional and sexual abuse.[201] Further, the internet has increased accessibility to pornography and, therefore, consumption.[202] As tens of thousands of online sexual images replace the occasional glance at the Playboy magazines of the past, pornography is now consumed at higher rates than ever before.[203] A 2015 study confirmed that over thirty-five percent of men viewed pornography at least once a year.[204] That same study found that by the age of fourteen, sixty six percent of males had viewed pornography in the last year.[205] Along with the increase in the consumption of pornography is the increase in murder, rape, sexual assault, and sexual harassment of women.[206]

most unsettling findings was that over half of the boys (53%) believed that the porn they had seen was realistic. They believed that what they saw in porn was an accurate depiction of sex and sexuality.").

[199] *What's the Average Age of a Child's First Exposure to Porn?* FIGHT THE NEW DRUG (Nov. 23, 2020) https://fightthenewdrug.org/real-average-age-of-first-exposure/ ("An estimated 93% of young men under the age of 18 have seen porn.").

[200] Rostad et al., *The Association Between Exposure to Violent Pornography and Teen Dating Violence in Grade 10 High School Students*, 48 ARCH. SEX. BEHAV. 2137-47 ("The confluence model suggests that for boys and men who are high in both hostile masculinity (e.g., domineering attitude toward women) and sexual promiscuity (i.e., engaging in impersonal sex acts), pornography consumption may intensify the risk of sexual violence).)".

[201] Brown, *supra* note 192 ("[T]he age at which males are first exposed to porn shapes their sexual behavior and tendency to seek power over women. Conversely, adolescent girls are more likely to tolerate emotional, physical, and sexual abuse as a result.").

[202] Hough, *supra* note 194 ("With the arrival of the internet age, people of all ages, genders, and classes now have an almost unlimited access to pornographic content that is tailored to every acquired taste and fantasy. The material's immediate accessibility is enhanced by seemingly endless development of more vivid, more realistic digital media.").

[203] *What's the Average Age of a Child's First Exposure to Porn? supra* note 199 ("[W]hat is clear is that exposure to porn is happening earlier than it ever used to, and it's more hardcore and accessible than it ever used to be.").

[204] Rothman, *supra* note 129 ("Approximately 36 percent of US men age 18 years old and older . . . view pornography at least once per year.").

[205] *Id.* ("By the time US youth are 14 years old, 66 percent of males . . . have viewed print, film, or internet pornography at least once in the past year, either on purpose or accidentally.").

[206]Wasiq Agha, *How Porn is Contributing to the Rape Culture*, BINGE DAILY (Oct. 14, 2019) https://www.bingedaily.in/how-porn-promotes-rape ("The Michigan State Police Department found that pornography is used or imitated in 41% of the sex crimes they have investigated"); *Report of the Attorney General's Commission on Pornography: Section 5.2.1 Sexually Violent Material*, Berkman Center for Internet and Society, https://cyber.harvard.edu/vaw00/module5.html (last visited Feb. 17, 2020) ("The evidence also strongly supports the conclusion that substantial exposure to violent sexually explicit material leads to a greater acceptance of the 'rape myth' in its broader sense – that women enjoy being coerced into sexual activity, that they enjoy being physically hurt in sexual context, and that as a result a man who forces himself on a woman sexually is in fact merely acceding to the 'real' wishes of the woman, regardless of the extent to which she seems to be resisting.").

Deepfake pornography is no different in the danger it poses to women, and yet is entirely unique in the way that it cultivates danger[207] Deepfakes use existing pornography, with all of its portrayals of harm against women, and throw in a heightened non-consensual factor.[208] While legal pornographies depict sexual violence against women, they do so through the use of consensual actors.[209] Although fantasy pornography and reality are increasingly difficult to separate, viewers of pornography are aware the films participants are acting out roles and that pornographies are not the bedroom films of real people's lives.[210] This is not the case with deepfake pornography, which is created using digitally manipulated images and videos that make a nonconsenting individual look as though they are participating in hard core pornographies.[211] The result is that deepfakes makes it appear as though real people with real names and identities, not actors, are participating in the violent and sexually degrading acts.[212] This creates many problems when considering almost all deepfake pornographies involve women.[213]

Consequently, deepfake pornography is an increasingly prevalent phenomenon that exclusively targets real, identifiable, and nonconsenting women.[214] The women who

[207] Dave Lee, *Deepfake Porn Has Serious Consequences*, BBC (Feb. 3, 2018) https://www.bbc.com/news/technology-42912529 ("As these tools have become more powerful and easier to use, it has enabled the transfer of sexual fantasies from people's imaginations to the internet. It flies past not only the boundaries of human decency, but also our sense of believing what we see and hear.").
[208] Citron, *supra* note 20, at 1921 ("Much like nonconsensual pornography, deep-fake sex videos exercise dominion over people's sexuality, exhibiting it to others without consent.").
[209] Cory Silverberg, *This is How Real the Sex You See in Pornography is*, MY DOMAINE (Nov. 19, 2019) https://www.mydomaine.com/how-real-is-the-sex-you-see-in-pornography-2982550 ("However, there's one thing that everyone should agree on when it comes to pornography: It's recorded and—to some extent—staged. Like other kinds of movies, porn bears about as much similarity to our real sex lives as a romantic comedy does to our daily lives.").
[210] Id.
[211] Citron, *supra* note 20, at 1921 ("Machine-learning technologies are being used to create "deep-fake" sex videos--where people's faces and voices are inserted into real pornography. Deep-fake technology enables the creation of impersonations out of digital whole cloth. The end result is realistic-looking video or audio that is increasingly difficult to debunk.").
[212] *Id.* ("Deep-fake sex videos are different from the nonconsensual disclosure of intimate images because they do not actually depict a victim's naked body. Yet even though deep-fake sex videos do not depict featured individuals' actual genitals, breasts, buttocks, and anuses, they hijack people's sexual and intimate identities.").
[213] Ivan Mehta, *A New Study Says Nearly 96% of Deepfake Videos are Porn*, NEXT WEB (Oct. 7, 2019) https://thenextweb.com/apps/2019/10/07/a-new-study-says-nearly-96-of-deepfake-videos-are-porn/ ("A study from Deeptrace, a Netherland based cybersecurity company, has published a new report stating 96 percent of deepfake videos online are porn, and they received over 134 million views. What important to note is that all porn videos feature female subjects.").
[214] *Id.* ("Danielle Citron, Professor of Law, Boston University, and author of Hate Crimes in Cyberspace, told the company that deepfakes are being used as a weapon against women, [stating] '[d]eepfake technology is being weaponized against women by inserting their faces into porn. It is terrifying, embarrassing, demeaning,

become victims of deepfake pornography suffer damaging consequences when their images are falsified in a sexual way and distributed publicly.[215] Current victims of deepfake pornography, ranging from celebrities to investigative journalists to ordinary individuals like Noelle Martin, suffer from constant fear of re-exposure, humiliation, blackmail, and stifling of their own First Amendment rights[216] They report feelings of shame, an inability to return to previous lifestyles, and increasing sexual aggression from males.[217]

V. Distribution by ISP's Cause the Harm of Deepfake Pornography and Necessitates Changes to the Legal Landscape

Despite the serious harm inflicted on women by the distribution of deepfake pornography, a solution has yet to be found to prevent its online dissemination.[218] First, absent a court holding that identifies deepfake pornography as obscene, the doctored, nonconsensual video is arguably protected speech.[219] Second, without any legal or civil means to halt distribution, victims of deepfake pornography will be forced to watch as their fake sexual depiction is irretrievably dispersed throughout the internet.[220] Further, without any incentive for an ISP to remove deepfake pornography, victims will be left hopelessly submitting removal petitions as their fabricated sex lives multiply in cyberspace.[221]

---

and silencing. Deepfake sex videos say to individuals that their bodies are not their own and can make it difficult to stay online, get or keep a job, and feel safe.'").

[215] Victoria Turk, *Deepfakes Are Already Breaking Democracy. Just Ask Any Women*, WIRED (Nov. 18, 2019) https://www.wired.co.uk/article/deepfakes-pornography ("[D]eepfake pornography is used as a tool to humiliate, demean and silence women.").

[216] *Id.* (describing the effects that the distribution of deepfake pornography and nonconsensual pornography have had on investigative reporter Rana Ayyub and Democrat Representative Katie Hill).

[217] *Id.* ("As law professor Danielle Citron has pointed out, it's not difficult to imagine a woman attempting to run for political office in the future and finding herself the target of a deepfake video intended to undermine her to her supporters, or shame her so much that she feels compelled to give up – even though the content is entirely fake.").

[218] Delfino, *supra* note 35, at 904 ("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.")

[219] Spivak, *supra* note 58, at 360 ("[W]hether deepfakes, or computer-generated pornography, are obscene is not easily answered.").

[220] Magaldi, *supra* note 177, at 225 ("The absence of nonconsensual pornography criminal statutes in four states and the variance of the conduct legally prescribed by the forty-six different state nonconsensual pornography criminal statutes leaves substantial nonconsensual pornography conduct unrestricted.").

[221] *Id.* at 209 ("The Communications Decency Act deprives victims of a bona fide remedy by providing a sort of immunity to ISPs for nonconsensual pornography posted through or on their internet services.").

a. Creation: The Red Herring of Distribution

While the law's focus has largely been on the creation of deepfake pornography, it is actually the distribution that produces harmful effects on society and women.[222] A man who creates a deepfake pornography of his classmate and uses it in the privacy of his own home steps into unethical but protected territory.[223] Conversely, a man who creates a deepfake pornography of a classmate and distributes it to the world does not merely step, but leaps into territory that is unethical, unprotected, and irreparably harmful to both the society he lives in and the individual he has victimized.[224] History has narrowed the issue of deepfake pornography to reflect a presumption that it is distribution, not creation, of explicit materials which has given courts reason to stifle First Amendment protections.[225] In 1857, the court in *Regina v. Hicks* found the distribution of an anti-Catholic pamphlet obscene due to fear the material would fall into the wrong hands.[226] In 1873, Anthony Comstock successfully persuaded Congress to pass the Comstock Act which, among other things, defined contraceptives as obscene and made it a federal offense to distribute them through mail or across

[222] Roffer, *supra* note 21, at 950-51 ("The harm experienced by victims of nonconsensual pornography is exacerbated by the unique nature of the Internet (including social media) because it facilitates an exponential growth in publication. Mary Anne Franks, who a leader in the fight against nonconsensual pornography, outlined four reasons why cyber harassment can be more damaging than real-life harassment: (1) the veil of anonymity, (2) amplification, (3) permanence, and (4) virtual captivity and publicity.").
[223] Henry Cohen, *Obscenity and Indecency: Constitutional Principle and Federal Statutes*, Congressional Research Service (2013) ("The Supreme Court has allowed one exception to the rule that obscenity, as defined by *Miller*, is not protected under the First Amendment. In *Stanley v. Georgia*, the Court held that 'mere private possession of obscene material' is protected. The Court wrote: Whatever may be the justifications for other statutes regulating obscenity, we do not think they reach into the privacy of one's own home. If the First Amendment means anything, it means that a State has no business telling a man, sitting alone in his house, what books he may read or what films he may watch.").
[224] Roffer, *supra* note 21, at 948 ("A recent Cyber Civil Rights Initiative ('CCRI') survey revealed that ninety-three percent of victims suffered 'significant emotional distress' and forty-two percent 'sought out psychological services.' In addition, because personal information like names and contact information frequently accompany the images posted online, victims are at a higher risk of stalking and physical attacks. The same CCRI survey reported that fifty-nine percent of victims had their full name posted and forty-nine percent had their social network information or a 'screenshot' of their social network profile included. By including such personal information, perpetrators essentially invite others to contact the victim, instilling fear in the victim of additional contact or confrontation from others, both online and offline.").
[225] *See, e.g.*, Regina v. Hicklin, L.R. 3 Q.B. 360 (1868) (discussing the fear that obscene material would fall into the wrong hands); Roth v. U.S., 354 U.S. 476 (1957) (discussing the distribution of sexually explicit circulars); *Miller v. California*, 413 U.S. 15 (1973) (discussing the distribution of sexually explicit mail campaign materials).
[226] See Gretchen Brooke Gould, Obscenity and pornography: A historical look at the American Library Association, the Commission on Obscenity and Pornography, and the Supreme Court (2010) ("Lord Cockburn, the judge in the case, stated that material was considered obscene . . . 'whether the tendency of the matter charged as obscenity is to deprave and corrupt those whose minds are open to such immoral influences, and into whose hands a publication of this sort may fall.'").

state lines.[227] In 1957, the Supreme Court held in *Roth v. United States* that the distribution, not the creation, of sexually explicit circulars violated federal obscenity statutes.[228] The current test within the Supreme Court for obscenity was developed in 1973 by *Miller v. California* after a mass mail campaign advertising the sale of sexually explicit materials was found to violate a California statute prohibiting the distribution, not creation, of obscene materials.[229]

The history of policing obscenity makes it possible to identify the harm caused by distributing deepfake pornography, but there are modern harms experienced by victims that support this theory as well.[230] The serious consequences of the distribution of deepfake pornography has already been established in the case of Noelle Martin, who described the public dissemination of the forged videos as humiliating and the proliferation as beyond her control.[231] When Noelle attempted to fight the distribution of the deepfake pornography by distributing her own narrative of their invalidity, the abuse magnified.[232] Noelle was doxxed, slut shamed, and blackmailed.[233]

---

[227] *See id.* ("In 1873, Anthony Comstock, a private citizen with a great deal of influence, succeeded in persuading Congress to pass an 'Act for the Suppression of Trade in, and Circulation of, Obscene Literature and Articles of lmmoral Use (17 Stat. 598 (1873)),' more popularly known as the Comstock Act.").

[228] *See* Roth v. U.S., 354 U.S. at 493 ("We therefore hold that the federal obscenity statute punishing the use of the mails for obscene material is a proper exercise of the postal power delegated to Congress by Art I, s 8, cl. 7.").

[229] *See* Miller v. California, 413 U.S. at 24 ("Appellant was convicted of mailing unsolicited sexually explicit material in violation of a California statute. . .").

[230] *See* Roffer, *supra* note 21, at 951 ("For example, one victim said: 'I am victimized every time someone types my name into the computer.' Another victim stated: 'It just makes me feel like a piece of meat [that is] being passed around for a profit.' When an image is posted online, the viewer gains complete control of the image and the amount of time they spend viewing the image. Moreover, the images 'often dominate Internet searches for victims' names' and are 'easily accessible to everyone a victim knows.' With the click of a button, the image can be shared again and again as it continues to be seen by users further down the chain. The journey an image takes online can be summarized as 'unchartered, unpredictable, and uncontrollable.'").

[231] *See, e.g.*, Melville, *supra* note 5 ("'The helplessness and powerlessness of tackling something that was proliferating beyond my control … I didn't cope,' [said Martin]."); Harris, *supra* note 2 ("It just was a never-ending battle as well because the more I'd try and get the sites deleted the more sites were popping up, and the more people had been seeing the photos, and the more it was just getting out of my control. It was proliferating to the point where I would never, ever, even to this day, be able to fully get the pictures deleted.").

[232] *See* Melville, *supra* note 5 ("As soon as I started speaking out I got a different kind of abuse, which was all the trolling and hate comments. 'She's fat,' 'she's a whore,' 'she's a slut,' 'she's attention-seeking,' 'look at the way she dresses,' Noelle says. 'I was blamed for the conduct of the perpetrators.'").

[233] *See Definion of Dox*, MERIAM WEBSTER https://www.merriam-webster.com/dictionary/dox (last visited Feb. 17, 2020) (defining dox as "publicly identify[ing] or publish[ing] private information about (someone) especially as a form of punishment or revenge"). *See also* Harris, *supra* note 2 ("One webmaster told me he would only delete the site if I sent him nude photos of myself within 24 hours.").

Similarly, investigative journalist Rana Ayyub experienced firsthand the ramifications caused by the distribution of deepfake pornography.[234] After an eight-year-old Kashmiri girl had been raped, Rana spoke out against her home country of India and it's practice of sweeping child sexual abuse under the rug.[235] A misinformation campaign quickly took form, culminating in a deepfake pornography portraying Rana that was widely circulated on Whatsapp.[236] Rana had a visceral reaction to the video, which was shared more than 40,000 times.[237] In addition to the physical reaction, Rana was doxxed, forced to delete her social media accounts, subjected to male harassment, ignored by police, and stifled.[238] The ordeal caused Rana to self-censor and has effected the credibility she needs to perform her job as an investigative reporter.[239]

In a related matter, Democrat Representative Katie Hill stepped down from her congressional position in 2019 after a conservative news site published nude images of Katie without her consent.[240] Katie, a rising star within her party who had just

[234] *See* Rana Ayyub, *I Was the Victim of a Deepfake Porn Plot Intended to Silence Me*, HUFFINGTON POST (Nov. 21, 2018) https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316 ("[T]he effects have stayed with me. From the day the video was published, I have not been the same person.").

[235] *See id.* ("An eight-year-old Kashmiri girl had been raped and there was outrage across the country. The nationalist Bharatiya Janata Party (BJP) was marching to support the accused. I had been invited to speak on the BBC and Al Jazeera about how India was bringing shame on itself by protecting child sex abusers.").

[236] *See id.* ("It started with a misinformation campaign to discredit me as an investigative journalist. Then my face was edited into a porn video.").

[237] *See id.* ("I started throwing up. I just didn't know what to do. In a country like India, I knew this was a big deal. I didn't know how to react, I just started crying . . . [t]hen, the fanpage of the BJP's leader shared the video and the whole thing snowballed. The video was shared 40,000 more times.").

[238] *See id.* ("[T]hey doxxed me. Another tweet was circulated on social media with a screenshot of the video and my number alongside, saying 'Hi, this is my number and I'm available here'. People started sending me WhatsApp messages asking me for my rates for sex . . . I deleted my Facebook, I just couldn't take it. But on Instagram, under every single one of my posts, the comments were filling with screenshots of the video. . . [e]very other person was harassing me with comments like 'I never knew you had such a stunning body' . . . [w]hen we went to the station, the police wouldn't file a report. The people who were sharing this video were political and the officers weren't prepared to take on the powerful. There were about six men in the police station, they started watching the video in front of me. You could see the smirks on their faces . . . [f]rom the day the video was published, I have not been the same person. I used to be very opinionated, now I'm much more cautious about what I post online.").

[239] *See id.* ("I've self-censored quite a bit out of necessity. Now I don't post anything on Facebook. I'm constantly thinking what if someone does something to me again . . . [w]hen I exposed a scandal around a high-profile murder investigation, people started putting photoshopped images of me online in sexualised positions. When my book was published, a police officer wrote on social media I had been sleeping with my sources and using unethical methods to get information . . . [i]t was devastating. I just couldn't show my face. You can call yourself a journalist, you can call yourself a feminist but in that moment I just couldn't see through the humiliation. It had exposed me to a lynch mob in India. People were thinking they could now do whatever they wanted to me.").

[240] *See* Jessica Bennett, *The Complicated Case of Katie Hill*, N.Y. TIMES (Nov. 1, 2019) https://www.nytimes.com/2019/11/01/us/katie-hill-photos-relationship.html ("'I am leaving because of a

usurped a Republican incumbent, was the victim of revenge porn when her nude images were posted online and distributed through various ISP's.[241] The distribution, while not involving deepfake pornography, involved irreparable harm to Katie including her resignation after just ten months in office and a pervasive feeling of fear which she attributed to the cyber exploitation that too commonly targets women.[242] Similar to Noelle and Rana, the distribution of Katie's sexually explicit images caused shame, oppression, and censorship of a woman poised to contribute meaningful change in the world.[243]

b. Distribution: The Law and Its Limitations

While the forged, nonconsensual, and explicit nature of deepfake pornography arguably warrants classification as obscenity, this distinction is yet to be made by courts.[244] Currently, no federal criminal law exists that provides redress for victims or halts the distribution of deepfake pornography.[245] Further, even if a comprehensive federal regulation did exist, it would likely only be construed against the third-party poster of the deepfake pornography.[246] Section 230 would still shield ISP's, the true mechanisms of distribution, from liability.[247] Victims of deepfake pornography are

misogynistic culture that gleefully consumed my naked pictures, capitalized on my sexuality and enabled my abusive ex to continue that abuse, this time with the entire country watching.'").

[241] *See* Marsha Gessen, *The Terrorization of Katie Hill*, THE NEW YORKER (Nov. 5, 2019) https://www.newyorker.com/news/our-columnists/the-terrorization-of-katie-hill (stating Hill "served just ten months in office, after unseating a Republican incumbent. She resigned after the right-wing Web site RedState, and later the British tabloid the *Daily Mail*, published intimate photos of her with a female former campaign staffer").

[242] *See* Maureen Down, *Now Comes the Naked Truth*, N.Y. TIMES (Nov. 2, 2019) https://www.nytimes.com/2019/11/02/opinion/sunday/katie-hill-resigns-millennials-boomers.html (stating Hill's resignation stemmed from "[t]he forces of revenge by a bitter, jealous man, cyber exploitation and sexual shaming that target our gender, and a large segment of society that fears and hates powerful women").

[243] *See* Bennett, *supra* note 240 ("Duncan Hunter, a Republican congressman from California, allegedly had 'intimate relationships' with staffers and faces criminal charges for using campaign funds to pay for dates with them. President Trump has more than a dozen allegations of sexual misconduct against him. Both men remain in office. Katie Hill, a Democratic congresswoman who was once considered a rising star in her party, meanwhile, resigned this week after her own affair was revealed in sensational fashion: when her nude photographs were published by a conservative website, Red State, whose main authors were later revealed to have worked for the Republican congressman she ousted a year ago.").

[244] *See* Spivak, *supra* note 58, at 361 ("[D]eepfakes are not on their face obscene speech.").

[245] *See* Delfino, *supra* note 35, at 922 ("[R]evenge porn laws[,] both current state laws and the proposed federal ENOUGH Act[,] will not assist in the case of deepfakes.").

[246] *See* Magaldi, *supra* note 177, at 209 ("Section 230 . . . provides that an ISP that simply serves as a digital bulletin board is not liable for content created, developed, or posted on or through the ISP's site, unless the ISP somehow curated the content.").

[247] *See id.* (explaining that Section 230 "deprives victims of a bona fide remedy by providing a sort of immunity to ISPs for nonconsensual pornography posted through or on their internet services").

thus left to rely on a piecemealing of related remedies in hopes of obtaining justice.[248] However, such remedies as applied to deepfake pornography yield inconsistent and ineffective results.[249] It is unsurprising then, that when Noelle Martin found deepfake videos of herself being ejaculated on, she was confronted by the same harsh reality many private victims of deepfake pornography will undoubtedly face: the law could not help her.[250]

ii. The First Amendment's Failure to Prevent Distribution of Deepfake Pornography

An ideal first line of defense to prevent the distribution of deepfake pornography would be a judicial holding that deepfake pornography is obscene.[251] Under the Miller test, this seems feasible on its face.[252] The first prong of Miller - which asks whether the material appeals to a prurient interest - could certainly be found in the case of deepfake pornography, whose illicit depictions of nonconsenting women appeal to an unhealthy, erotic, or degrading interest in nudity or sex.[253] Further, the second prong of Miller, which asks whether the work is patently offensive, easily applies to deepfake pornography.[254] The patently offensive element's focus is not whether the material depicts an offensive act, but whether the act is depicted in a patently offensive way.[255] Deepfake pornography, by using machine learning to collect hundreds of images of a woman for the purpose of placing her face in a pornography without her

[248] *See id.* (discussing the shortcomings of the current legal mechanisms in place to prevent the distribution of deepfake pornography).
[249] *See* Roffer, *supra* note 21, at 957 (discussing the differing approaches to criminalizing nonconsensual pornography between states).
[250] *See* Citron, *supra* note 20, at 1923 ("Martin went to law enforcement and was told that nothing could be done.").
[251] *See* Danielle Keats Citron et al., *Criminalizing Revenge Porn*, 49 WAKE FOREST L. REV. 345, 384 (2014) ("Although the Court's obscenity doctrine has developed along different lines with distinct justifications, nonconsensual pornography can be seen as part of obscenity's long tradition of proscription.").
[252] *See id.* ("Noted First Amendment scholar Eugene Volokh argues that sexually intimate images of individuals disclosed without consent belong to the category of 'obscenity,' which the Supreme Court has determined does not receive First Amendment protection.").
[253] *See* Whitehead, *supra* note 119, at 50 ("Obscene material appeals to a 'prurient interest' when it is directed or makes its appeal to an unhealthy or abnormally lustful or erotic interest, or to a lascivious or degrading interest, or to a shameful or morbid interest in nudity, sex, or excretion.").
[254] *See* Citron, *supra* note 251, at 385 ("Disclosing pictures and videos that expose an individual's genitals or reveal an individual engaging in a sexual act without that individual's consent could qualify as a 'patently offensive representation' of sexual conduct.").
[255] *See* Whitehead, *supra* note 119, at 50 (explaining that "is not on whether the matter depicts offensive sex acts, but rather, on whether the sex acts are depicted in a patently offensive manner").

consent, would certainly be found patently offensive by the average person.[256] Finally, because deepfake pornography is falsified imagery, nonconsensual, and illicit, it is highly unlikely that it could ever be classified as having some sort of societal value whether it be scientific, literary, artistic, or political.[257]

However, the Miller test has limitations in defining deepfake pornography as obscene.[258] Difficulty arises in the community and national standards used in the Miller test when applied to the internet.[259] With no inherent community standard to apply to the internet, which crosses all state borders, the issue of whether deepfake pornography is obscene may vary from state to state.[260] More importantly, an obscenity remedy serves as a band-aid solution because it will hold third party posters liable but will do little to stop distribution.[261] While third party posters may initiate distribution, it is the continued passive distribution by ISP's protected by Section 230 immunity that causes the real harm of deepfake pornography.[262] A more likely solution would involve the creation of a federal criminal law prohibiting the distribution of deepfake pornography.[263] However, this federal statute must be narrow enough in scope that it does not offend the First Amendment requirement that the government may not restrict expression because of its message, ideas, subject matter, or content.[264]

---

[256] *See id.* ("[W]hen determining whether visual or written works are obscene, juries must avoid applying community standards . . . [i]nstead, they must determine the judgment that would be made by a hypothetical 'average person' applying the standards of the community as a whole.").
[257] *See* Citron, *supra* note 251, at 385 (arguing that deepfake pornography offers no "serious literary, artistic, political, or scientific value").
[258] *See* Spivak, *supra* note 58, at 358 (explaining that deepfake regulation hinges on whether deepfakes fall into the obscenity exception to free speech, but "whether it in fact does so, however, is dubious at best").
[259] *See* Whitehead, *supra* note 119, at 51 ("The Court has faced the issues of obscenity and indecency in several different mediums: print, broadcast television and radio, cable television, telephone, and most recently, the Internet. In doing so, it has applied differing standards of First Amendment protection.").
[260] *See id.* ("[E]ven with the uniform standard of applying community standards (or even if a 'national' standard were employed), the types of material that may be found obscene may vary from trial to trial.").
[261] *See* Delfino, *supra* note 35, at 900 (quoting Rachel Budde Patton, *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 (2015)) ("Section 230 immunity makes it difficult for deepfake victims to sue internet platforms for hosting deepfakes. 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.'").
[262] *See* Roffer, *supra* note 21, at 950 ("The harm experienced by victims of nonconsensual pornography is exacerbated by the unique nature of the Internet . . . because it facilitates an exponential growth in publication.").
[263] *See* Delfino, *supra* note 35, at 906 ("Legal scholars . . . have coalesced around the cause of victims of nonconsensual pornography advocating for the adoption of a federal statute expressly criminalizing this conduct. These groups have proposed federal legislation to outlaw revenge porn and deepfakes.").
[264] *See* Police Dep't of Chi. v. Mosley, 408 U.S. 92, 95 (1972) (explaining that under the First amendment, a government has "no power to restrict expression because of its message, its ideas, its subject matter, or its content").

ii. Limitations of Existing Law on Distribution of Deepfake Pornography

A federal criminal law prohibiting the distribution of deepfake pornography may be the best solution because it creates a cohesive remedy currently lacking in state laws and validates the serious consequences experienced by women who have been victimized by deepfake pornography.[265] However, proposed legislation has not passed and it has significant limitations.[266] The current absence of a federal criminal law prohibiting the distribution of deepfake pornography limits the options for legal redress of a victim to state law, which is unlikely to yield success.[267]

A. Limitations of State Law

In the United States, state criminal and civil laws do little to allow redress for victims of deepfake pornography and to stop distribution.[268] Without comprehensive federal regulation, victims must attempt to fit the specific circumstances of the deepfake within the confines of existing state criminal and civil laws.[269] In doing so, victims will be subject to a legal process whose application may vary wildly from one court to the next.[270]

1) Criminal Law Limitations: The Judgement Proof Perpetrator

From a criminal law standpoint, there are a variety of ways in which the law falls short in preventing the distribution of deepfake pornography.[271] First, it is unlikely that the victim will ever be able to identify a perpetrator.[272] Creators of a deepfake pornography often go to extensive lengths to hide and disguise their digital

[265] *See* Otero, *supra* note 24, at 602 ("To provide the greatest possible deterrence and the most useful remedies for victims, nonconsensual pornography should be criminalized at the federal level . . . .").
[266] *See* Delfino, *supra* note 35, at 906-10 (describing previously proposed federal laws).
[267] *See* Roffer, *supra* note 21, at 955 ("The existing approaches to protecting victims of nonconsensual pornography are extremely inconsistent and inadequate.").
[268] *See* Harris, *supra* note 23, at 107-20 (describing current private legal recourses available to victims of deepfake pornography and their limitations).
[269] *See* Delfino, *supra* note 35, at 902-04 (describing the need for a criminal law addressing deepfake pornography and lack of redress available through current state law).
[270] *See* Harris, *supra* note 23, at 118-19 (describing the current remedies available to victims of deepfake pornography and their inconsistencies as well as inadequacies).
[271] *See* Citron, *supra* note 20, at 1939 ("Deep-fake sex videos are another area where current law falls short. No federal criminal law covers the practice, though a smattering of state statutes might apply . . . .").
[272] *See* Ruby Harris, *supra* note 2 ("[T]he big difficulty is the resources it requires to actually go and try and find the perpetrators.").

footprint.[273] In fact, to this day Noelle Martin has no idea who distributed the deepfake pornography using her images.[274] Second, the creators of deepfakes are often living in other countries and nearly impossible to track down.[275] Third, because most laws dealing with revenge porn or related topics categorize the offense as a misdemeanor, there is little incentive for law enforcement to find the creator of the deepfake, let alone criminally prosecute.[276] Finally, jurisdictional problems arise because many states have differing approaches to revenge porn laws.[277] The end result is the lack of consistency and cohesiveness of various states makes it difficult for victims of deepfake pornography to bring charges against the perpetrators.[278]

2) Civil Law Limitations: Loopholes in Existing Laws

Victims of deepfake pornography are equally unlikely to receive justice through civil means.[279] Tort claims involving emotional distress, such as intentional infliction of emotional distress (IIED) or negligent infliction of emotional distress (NIED) are not viable options for legal redress.[280] To prevail on an IIED claim, the victim of deepfake pornography must prove the creator intended to cause the victim severe emotional distress.[281] This intention is difficult to prove, however, because deepfake pornography is generally created for the pleasure of the creator rather than the

[273] *See id.* ("Depending on how savvy they are, they often use encrypted emails, fake emails, and fake names. They'll use VPNS, so they'll mask where they're actually living. They could be anywhere around the world and they'll be able to cover their tracks.").
[274] *See id.* (explaining that Noelle Martin "had no idea who was doing this to [her]").
[275] *See id.* (explaining that the perpetrators of Noelle Martins deepfake pornography were difficult to identify because "[t]he sites were hosted overseas, and the perpetrators were most likely overseas too.").
[276] *See* Roffer, *supra* note 21, at 938 ("As the law currently stands, thirty-five states take different approaches to criminalizing revenge porn. Some states have tried to use existing statutes while others have drafted new ones. Some states classify the crime as a misdemeanor or a felony, while others classify it as a sexual offense or an invasion of privacy. These inconsistencies have led to unpredictable results among the states.").
[277] *See id.* at 956-57 ("As of April 7, 2017, thirty-five states and the District of Columbia have laws criminalizing revenge porn, although all states treat the crime differently. Various statute titles include: 'Disorderly [C]onduct,' 'Sexual [C]yberharassment,' Disclosure of [P]rivate [I]mages,' and 'Non-[C]onsensual [D]issemination of [P]rivate [S]exual [I]mages.' Depending on the state, the crime may be classified as a felony or a misdemeanor; for some states, it depends on the presence of certain factors. In sum, of the thirty-six jurisdictions that criminalize nonconsensual pornography, twenty-five define it as a misdemeanor and eight define it as a felony. Three states do not denote such offenses as either felonies or misdemeanors.").
[278] *See id.* at 954 ("[T]here is no consistent approach, which has led to variable results among the states.").
[279] *See id.* at 955 ("[T]here is extensive literature on current civil remedies and their respective issues that render the options insufficient.").
[280] *See* Harris, *supra* note 23, at 111 ("[T]he Victim may attempt to bring a bevy of tort claims [which] . . . are subject to their own flaws and limitations.").
[281] *See id.* (quoting RESTATEMENT (SECOND) OF TORTS § 46 (AM. LAW INST. 1965)) ("State doctrines vary, but most require a showing that (1) the Producer intended to (2) cause the Victim severe emotional distress (3) by extreme and outrageous conduct and (4) the Victim suffered severe emotional distress as a result of the extreme and outrageous conduct.").

humiliation of the victim.[282] In many cases, the creator will share the deepfake without anticipation that the victim will ever see it themselves.[283] To prevail on a NIED claim, the victim would likely be required to show some connection to the physical incident that caused the emotional harm.[284] In the case of deepfake pornography, the victim is never actually touched and is not, at least by definition of the law, perceiving an actual traumatizing event.[285] It is therefore unlikely that courts will construe the NIED doctrine to apply to videos that merely appear real.[286]

False light claims are a potential avenue for justice but they also pose hurdles.[287] The tort of false light attaches liability when a perpetrator gives publicity to another individual that places the individual in a false light that is highly offensive to a reasonable person.[288] Difficulties arise in proving that the false portrayal was highly offensive and whether the deepfake pornography is actually considered to be seen by the public.[289] States also have differing stances on the amount of people that must have viewed the content before it can be established as "publicity."[290]

[282] *See id.* at 112 ("The majority of the Producers who share a video online with friends or the general public will likely not know that any emotional distress is imminent because they do not expect that the Victim will watch the video or that the Victim will even learn of its existence. This high standard will prevent many Victims from succeeding on this cause of action when they stumble upon the video online or are made aware of the video by a third party.").
[283] *See id.* ("IIED claims, thus, appear to be limited to instances where the Producer intentionally sends the deepfake to the Victim or informs her of its circulation on the internet.").
[284] *See id.* at 113 ("Six states require a plaintiff to show that there was a physical impact as an effect of the negligent act, and over a dozen states require the plaintiff show that she was in the zone of danger during the negligent conduct.").
[285]*See id.* at 114 ("NIED cases highlight how many states require some physical incident that caused some real emotional harm—with or without accompanying physical symptoms. With personal deepfakes, the Victim never came close to being physically touched by something harmful or perceiving a traumatizing event that actually happened.").
[286] *See id.* ("[I]t is unlikely that courts will begin to construe the NIED doctrine in a way that encapsulates fake videos that merely appear real. Even when personal deepfakes become so indistinguishable from videos of real events, Victims will generally know that they are not actually appearing in the videos and the portrayed act never occurred.").
[287] *See id.* at 118 ("Moreover, false light invasion of privacy claims will not cover most Victims, unless the public at large becomes aware of the personal deepfake's presence.").
[288] *See id.* at 116 ("With personal deepfakes, the success of a false light claim also depends on the specific context of the deepfake and its publication as well as the state where the claim was brought.").
[289] *See id.* (quoting RESTATEMENT (SECOND) OF TORTS § 652E (AM. LAW INST. 1977)) ("An individual may be liable when 'giv[ing] publicity to a matter concerning another that places the other before the public in a false light.' The portrayal must be highly offensive to a reasonable person, and the actor must have had knowledge or acted in reckless disregard in publicizing this person in a false light . . . [w]ith personal deepfakes, the success of a false light claim also depends on the specific context of the deepfake and its publication as well as the state where the claim was brought.").
[290] *See id.* at 116-17 (quoting *Solano v. Playgirl, Inc.*, 292 F.3d 1078, 1082 (9th Cir. 2002)) ("For instance, in California, 'the information [must be] understood by one or more persons to whom it was disclosed as stating or implying something highly offensive'. . . California's standard of only requiring one individual to view the publication is much more victim-friendly than the Restatement's requirement of being viewed by the public at

Appropriation claims, which occur when a person's name, likeness, or image is used without permission, also lack traction in the context of redress for victims of deepfake pornography.[291] Most jurisdictions limit this tort to apply solely where a person's image is being used for commercial purposes.[292] However, few creators of a deepfake pornography are making money off of their offensive handiwork.[293] Defamation claims too lack the consistency necessary to ensure each victim of a deepfake pornography receives adequate justice.[294] Defamation claims rest upon the idea that perpetrators should be held liable for statements that harm the reputation of another.[295] However, defamation claims examine not just the content, but the context of the defamatory material.[296] Therefore, the use of a disclaimer or caption can diminish the strength of a defamation claim if the creator of the deepfake pornography expressly states that the video is a fictitious account.[297] Courts are unwilling to swing the First Amendment pendulum away from content that does not intrude on a victim's rights.[298] When a deepfake pornography is prefaced with a disclaimer professing its fabricated nature, it may be unlikely that a court would find that a victim's rights have been intruded to the extent that suppression of free speech rights is warranted.[299]

Given the difficulty victims have had in achieving justice through state law, Congress began drafting bills to regulate deepfake pornography.[300] However, the majority of

large. Twenty states have followed the Restatement. Personal deepfake Victims in these states will likely have to wait until a substantial population of people watch the video to bring a false light cause of action.").

[291] *See* Spivak, *supra* note 58, at 381 ("The tort of wrongful appropriation requires that the defendant appropriate the plaintiff's likeness to his own *use* or benefit.").

[292] *See id.* ("Usually, such use or benefit is attributed to a commercial or financial benefit. Though opponents may rebut that they are not benefitting commercially. . ..").

[293] *See id.* at 383 ("[C]ourts may be reluctant to recognize a deepfaker's personal use and enjoyment of a fabricated video, even if it is disseminated on the Internet for others' personal, analogous use and enjoyment. Without any promise of monetary value, personal deepfakes are likely insufficient to satisfy the elements of appropriation.").

[294] *See id.* at 368 ("Defamation is one means of civil recourse for pursuing deepfakers.").

[295] *See id.* ("A defamatory statement is defined as a communication that tends to harm the reputation of another as to lower him in the estimation of the community or to deter third persons from associating or dealing with him.").

[296] *See id.* at 373 ("Content is not the only factor considered, however; the video's context (e.g. its caption) play a role in defamation analysis.").

[297] *See id*. ("[I]f the deepfaker is quite clear about the fact that the video is fabricated or fantastical, he or she has a stronger defense that the video does not inflict the same [defamatory] harm on the video's subject.").

[298] Harris, *supra* note 23, at 118-19 ("A discussion of First Amendment rights will remain in the background for these claims, and the courts must decide how to balance free speech rights and the harm that personal deepfakes can cause.").

[299] *See id.* at 117 (discussing similar concerns in the context of false light claims, stating "victims of personal deepfakes will not be able to recover when that video discloses its fabrication to viewers.").

[300] Delfino, *supra* note 35, at 921, 927 ("Most states have adopted legislation specifically targeting revenge porn to address these insufficiencies. Those laws, often imperfect even in the revenge porn context, are not a

these bills have expired at the federal level.[301] Further, even if the bills had passed, each had significant limitations that would have failed to halt distribution.[302]

B. Limitations of Proposed Legislation on Distribution of Deepfake Pornography

Although no federal statute exists criminalizing deepfake pornography, various nonconsensual pornography and deepfake bills have been introduced at the federal level, and a state bill regulating deepfake pornography has been passed by the California Senate.[303] However, none of these options would effectively halt the distribution of deepfake pornography.[304]

1) AB-602: Depiction of Individual Using Digital or Electronic Technology

AB-602 is a California bill that provides recourse for a depicted individual against a person who discloses sexually explicit material with knowledge that the depicted

fix for pornographic deepfakes . . . [t]he slow, uneven efforts to criminalize revenge porn at the state level over the last decade demonstrate that waiting for the states to outlaw deepfakes will take too long as the technology becomes more sophisticated and more accessible . . . [therefore] Federal criminalization of deepfakes is warranted.").

[300] *See id.* at 907-910 (explaining that all three federal proposals regulating deepfakes have expired in congress).

[300] *See id.* at 922-26 (describing the limitations of proposed federal statutes regulating deepfakes).

[300] *See id.* (describing current proposals regulating nonconsensual and deepfake pornography).

[300] *See id.* at 922 ("[C]urrent state laws and the proposed federal ENOUGH Act will not assist in the case of deepfakes.").

[300] Assemb. B. 602, 2019-2020 Leg., Reg. Sess. (Cal. 2019) (as introduced) ("This bill would provide that a depicted individual, as defined, has a cause of action against a person who either (1) creates and intentionally discloses sexually explicit material if the person knows or reasonably should have known the depicted individual did not consent to its creation or disclosure or (2) who intentionally discloses sexually explicit material that the person did not create if the person knows the depicted individual did not consent to its creation.").

[300] *See id.* ("'Depicted individual' means an individual who appears, as a result of digitization, to be giving a performance they did not actually perform or to be performing in an altered depiction.").

[300] *See id.* ("'Disclose means to publish, make available, or distribute to the public.").

[300] Salour, *supra* note 186 ("AB-602 is limited in notable ways."). fix for pornographic deepfakes . . . [t]he slow, uneven efforts to criminalize revenge porn at the state level over the last decade demonstrate that waiting for the states to outlaw deepfakes will take too long as the technology becomes more sophisticated and more accessible . . . [therefore] Federal criminalization of deepfakes is warranted.").

[301] *See id.* at 907-910 (explaining that all three federal proposals regulating deepfakes have expired in congress).

[302] *See id.* at 922-26 (describing the limitations of proposed federal statutes regulating deepfakes).

[303] *See id.* (describing current proposals regulating nonconsensual and deepfake pornography).

[304] *See id.* at 922 ("[C]urrent state laws and the proposed federal ENOUGH Act will not assist in the case of deepfakes.").

individual did not consent to its creation or disclosure.[305] "Depicted individual" is defined as an individual who, as a result of digitization, appears to be giving a performance that never occurred.[306] Under the Act, "disclose" can be interpreted as distribution to the public.[307] However, while AB-602 is arguably the most comprehensive legislation targeting the distribution of deepfake pornography currently available, significant limitations still exist.[308] First, as a state civil action, AB-602 will be preempted by Section 230 and unable to hold an ISP liable for distribution.[309] Second, the Act protects only the individual whose face is portrayed in the deepfake pornography, leaving the individual whose body is depicted without recourse.[310] While private individuals whose faces were pasted onto existing porn stars' bodies would be able to achieve justice, the porn star whose body was depicted outside the realm of their agreed work would be left without recourse.[311] Third, AB-602's definition of consent is an agreement limited to plain language, posing problems for individuals such as existing sex workers who may want to consensually enter into complex legal agreements in the future.[312] Finally, the Act requires a potential plaintiff to prove the defendant did not know the Plaintiff did not consent – a double negative which is likely to yield difficulties for victims and courts.[313]

[305] Assemb. B. 602, 2019-2020 Leg., Reg. Sess. (Cal. 2019) (as introduced) ("This bill would provide that a depicted individual, as defined, has a cause of action against a person who either (1) creates and intentionally discloses sexually explicit material if the person knows or reasonably should have known the depicted individual did not consent to its creation or disclosure or (2) who intentionally discloses sexually explicit material that the person did not create if the person knows the depicted individual did not consent to its creation.").
[306] *See id.* ("'Depicted individual' means an individual who appears, as a result of digitization, to be giving a performance they did not actually perform or to be performing in an altered depiction.").
[307] *See id.* ("'Disclose means to publish, make available, or distribute to the public.").
[308] Salour, *supra* note 186 ("AB-602 is limited in notable ways.").
[309] *See id.* ("[T]he legislation is likely preempted by the federal Communications Decency Act, 47 U.S.C. § 230.").
[310] *See id.* ("[T]he bill seems to protect only persons whose faces are superimposed on another's body but not the person (i.e., the body) shown to be engaging in the sexually explicit conduct.").
[311] Delfino, *supra* note 35, at 898 ("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.").
[312] Salour, *supra* note 186 ("[T] the bill defines 'consent' as 'an agreement written in plain language signed knowingly and voluntarily by the depicted individual that includes a general description of the sexually explicit material and the audiovisual work in which it will be incorporated.' But what does 'plain language' mean exactly if an individual enters a complex legal agreement – does the complexity of such a contract render the 'consent' invalid?").
[313] *See id.* ("Finally, how is a prospective plaintiff to prove that a defendant 'knew' plaintiff did not, in fact, consent? Proving a negative is difficult and for this reason creates uncertainty about AB-602's potential as an effective remuneration tool for plaintiffs.").

2) Intimate Privacy Protection Act of 2016

IPPA advocates for federal criminalization when an individual distributes visual depictions of a person's intimate parts or sexually explicit conduct with reckless disregard for a lack of consent to the distribution.[314] Importantly absent from IPPA is the common requirement of an intent to harass.[315] This is particularly relevant as the line between revenge porn and deepfake pornography continues to blur in legislation.[316] However, a pivotal distinction between the two is that deepfake pornography's harms stem from privacy violations and distribution, whereas revenge porn, true to its name, largely encompasses harassment.[317] Further, by including a mens rea requirement of reckless disregard, IPPA narrows its scope by limiting liability only in cases where distribution occurs recklessly as opposed to unintentionally, increasing its chances of complying with First Amendment protections.[318] The bill is not without its flaws, however, as the bill's exceptions contain vague, undefined language, particularly in the context of consent, which would invariably lead to broad and possibly unconstitutional application.[319] This may have contributed to the fact that, while promising, IPPA was never enacted.[320]

[314] Megan Fay, 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, 1861 (2018) ("Specifically, the proposed bill reads, in part: 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.").
[315] Intimate Privacy Protection Act, H.R. 5896, 114th Cong. (2016) (lacking language regarding an intent to harass).
[316] Fay, *supra* note 314, at 1870 ("The 2016 IPPA bifurcates the requisite *mens rea* for lack of consent. Prosecutors must prove that an individual perpetrator intentionally disseminated nonconsensual pornography with 'reckless disregard' for the subject's lack of consent. The *mens rea* for ISPs, however, is higher and requires an 'intentional' solicitation of involuntary pornography.").
[317] Delfino, *supra* note 35, at 895-96 ("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 the realm of other sexually exploitative cybercrimes such as revenge porn and nonconsensual pornography.").
[318] Fay, *supra* note 314, at 1870 ("Ascribing a higher *mens rea* in order to impose liability on ISPs shields websites from criminal liability for negligently publishing nonconsensual pornography. In turn, websites will not be incentivized to over-censor user-uploaded content because they are not liable for third-party content.").
[319] *See id.* (describing limitations of the IPPA).
[320] Delfino, *supra* note 35, at 907 (stating the IPPA expired at the end of the 114th Congress).

3) Ending Nonconsensual Online User Graphis Harassment Act of 2017

The ENOUGH Act is a revised version of IPPA that created a federal crime for knowingly distributing an intimate visual depiction of a person with reckless disregard for lack of consent, the victim's privacy, or harm caused by the distribution and without the belief that distribution involves a matter of public concern.[321] ENOUGH differentiates from IPPA in that its focus leans heavily on distribution.[322] While this is promising in the context of regulating deepfake pornography, the bill fails to include simulated or forged acts in its definition of intimate visual depiction and may be interpreted to preclude criminal liability for deepfake pornography as a result.[323] Further, the exceptions detailed in ENOUGH hold ISP's immune from liability unless the ISP had actual knowledge that it was distributing content in violation of the bill, but actual knowledge is not defined within the bill.[324] Without a concrete definition, ISP's may resort to over or under-policing their content, depending on which way actual knowledge is interpreted.[325]

4) Malicious Deep Fake Prohibition Act of 2018

The MDFP Act was the first federal bill that aimed to criminalize the creation and distribution of deepfakes.[326] The proposed bill holds individuals liable for a federal felony when they create a deepfake with the intent to distribute despite knowing the

[321] Ending Nonconsensual Online User Graphic Harassment Act, H.R. 4472, 115th Cong. (2017) ("Except as provided in subsection (d), it shall be 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.").
[322] *See id.* (acknowledging distribution as part of the offense).
[323] *See id.* ("The term 'intimate visual depiction' means any visual depiction (as that term is defined in section 2256(5))—(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; (C) in which the content described in subparagraph (B) is not simulated; and (D) in original or modified format, such as with a filter or text overlay.").
[324] *See id.* ("This section shall not apply to any provider of a communications service with regard to content provided by another information content provider 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.").
[325] Delfino, *supra* note 35, at 922 (discussing the ambiguities of the ENOUGH Act).
[326] Brown, *supra* note 53 (describing the Malicious Deepfake Prohibition Act as "[t]he first federal bill targeted at deepfakes").

distribution will violate federal, state, local, or tribal laws.[327] Alternatively, individuals can also be criminally prosecuted if they distribute an audiovisual record with actual knowledge that it is a deepfake and intent that distribution would facilitate criminal conduct.[328] The MDFP Act's strength and weakness stem from its definition of deepfake, because while the MDFPA includes the definition of a deepfake within its writing, the definition is overly broad, and could reasonably be interpreted to include any altered video.[329] A federal law criminalizing every altered video, including protected content such as parodies, would certainly be violative of free speech.[330] This ubiquity is exacerbated when viewing the bill in its entirety, which imposes harsher penalties for violations that affect the administration of an election or facilitate violence.[331] By casting such a wide net, the MDFP opens itself up to First Amendment violations.[332]

Because each proposed or enacted legislation regulating deepfake pornography includes both benefits and limitations, a new federal criminal statute prohibiting the distribution of deepfake pornography should combine their strengths to create a narrow, clear-cut statute that can impose penalties without violating the freedoms contained in the First Amendment.[333]

[327] Malicious Deepfake Prohibition Act, S.3805, 115th Cong. (2018) ("It shall be unlawful to, using any means or facility of interstate or foreign commerce—(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.").

[328] *See id.* ("It shall be unlawful to, using any means or facility of interstate or foreign commerce— (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.").

[329] Delfino, *supra* note 35, at 923 ("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.'").

[330] *See id.* (explaining how the broad language of the Malicious Deepfake Prohibition Act opens itself up to First Amendment scrutiny).

[331] *Id. at 924* ("[T]he MDFPA was not written with the goal of protecting pornographic deepfake victims in mind. It focuses on the implications of politicized deepfakes").

[332] See i*d.* at 925("[A]ttempts to criminalize revenge porn across the states and at the federal level have been met with First Amendment challenges and concerns. Thus, to survive constitutional challenge, legislation targeting revenge porn has been narrowly tailored to avoid encompassing legitimate and protected, albeit objectionable, speech. The same concerns appear in the deepfake context").

[333] *Id.* at 927 ("A federal law criminalizing pornographic deepfakes would provide a strong and effective disincentive to their creation and distribution").

iii. Section 230 Limitations

A federal law criminalizing the distribution of deepfake pornography, while seemingly plausible, may still fall short due to the protections for ISPs provided by Section 230.[334] When a federal criminal law is at issue, the immunity provided by Section 230 can no longer be applied to ISPs.[335] However, as previously seen in federal criminal laws prohibiting online sex trafficking and the consequent enactment of the Stop Enabling Sex Traffickers Act, Section 230 may continue to be construed in such a way as to hold ISPs harmless despite the presence of a federal criminal law.[336]

On April 11, 2018 the Fight Online Sex Trafficking Act (FOSTA) and Stop Enabling Sex Traffickers Act (SESTA) were combined and signed into law.[337] The new law (SESTA) creates an exception to Section 230 which holds ISPs responsible if third parties are found to be posting advertisements construed as sex trafficking on ISPs' servers.[338] In doing so, SESTA carved out an exception to Section 230 excluding immunity for civil claims resulting from sex-trafficking.[339]

The newly-minted law was passed to curb the rise of sex trafficking, which has greatly increased due to the Internet's ability to mass distribute sex trafficking advertisements.[340] While this purpose was met with support from various advocacy

[334] *Id.* at 935 ("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").
[335] *Id.* at 927 ("[T]he immunity that section 230 provides for internet service providers and other content distributors does not apply to violations of federal criminal law").
[336] Aja Romano, *A New Law Intended to Curb Sex Trafficking Threatens the Future of the Internet as We Know it*, VOX (July 2, 2018) https://www.vox.com/culture/2018/4/13/17172762/fosta-sesta-backpage-230-internet-freedom ("[P]revious attempts by authorities to hold [ISP's] responsible for illegal content on its website have failed due to Section 230's dictum that websites aren't liable for content posted by their users").
[337] I*d.* ("President Trump signed a set of controversial laws enabling state and federal authorities to pursue websites that host sex trafficking ads in the Oval Office on April 11, 2018").
[338] I*d.* ("President Trump signed into law a set of controversial bills intended to make it easier to cut down on illegal sex trafficking online. Both bills — the House bill known as FOSTA, the Fight Online Sex Trafficking Act, and the Senate bill, SESTA, the Stop Enabling Sex Traffickers Act — have been hailed by advocates as a victory for sex trafficking victims").
[339] *See* FOSTA-SESTA, H.R. 1865, 115th Cong. (2017-2018) (explaining that Section 230 does not limit federal civil claims for conduct that constitutes sex trafficking or a federal criminal charge for conduct that constitutes sex trafficking, or a state criminal charge for conduct that promotes or facilitates prostitution. The bill defines "participation in a venture," to mean "knowingly assisting, supporting, or facilitating a sex-trafficking violation").
[340] Trafficking for Sexual Exploitation, EQUALITY NO https://www.equalitynow.org/trafficking (last accessed Feb. 5, 2021) ("[T]he National Center for Missing & Exploited Children directly correlated a five-year 846% increase in child sex trafficking reports to the growing use of the internet to sell children for sex").

groups, others criticized SESTA arguing its language was too broad and restricted freedom of speech.[341] Most notable in this debate is SESTA's definition of "participation in a venture," under 18 U.S.C. § 1591, as "knowingly assisting, supporting, or facilitating" sex trafficking.[342] Critics voice apprehension that this language is too vague and will cause ISPs to either over-police, stifling free speech in the process, or under-police so as not to be held liable for "knowing" anything.[343] Examples of over-policing are already being seen.[344] Craigslist has removed its personals section, for example, and sites dedicated to providing safety for sex workers have migrated to off-shore hosting services.[345]

The Electronic Frontier Foundation, one of SESTA's harshest critics, filed a lawsuit challenging the law on the grounds that its broad language is constitutionally defective and stifles free speech.[346] While the suit was initially dismissed for lack of standing, the District of Columbia Court of Appeals reversed in January 2020, finding merit in the argument that the overbroad language of SESTA may have impacted the plaintiff's freedom of speech.[347] Therefore, the possibility exists that SESTA incorrectly weighed the interests of society in holding ISPs liable for the facilitation of online sex trafficking against society's interest in free speech rights.[348]

---

[341] Romano, *supra* note 334 ("[M]any activists and internet freedom advocates have charged FOSTA-SESTA with threatening free speech").
[342] *Id.* ("The bill's language penalizes any websites that 'promote or facilitate prostitution,' and allows authorities to pursue websites for 'knowingly assisting, facilitating, or supporting sex trafficking.'").
[343] *Id.* ("[W]ebsites will have to decide whether to over-police their platforms for potential prostitution advertisements or to under-police them so they can maintain a know-nothing stance, which would likely be a very tricky claim to prove in court").
[344] *Id.* ("The bill's language penalizes any websites that 'promote or facilitate prostitution,' and allows authorities to pursue websites for 'knowingly assisting, facilitating, or supporting sex trafficking,' which is vague enough to threaten everything from certain cryptocurrencies to porn videos to sites for perfectly legal escort services").
[345] *See id.* (describing the impact FOSTA-SESTA had on Craigslist and sex worker verification sites and explaining that "[i]n the immediate aftermath of SESTA's passage on March 21, 2018, numerous websites took action to censor or ban parts of their platforms in response — not because those parts of the sites actually were promoting ads for prostitutes, but because policing them against the outside possibility that they might was just too hard").
[346] *See* Jon Fingas, *Court Reinstates Lawsuit Challenging Online Sex Trafficking Law*, ENGADGET (Jan. 26, 2020) https://www.engadget.com/2020/01/26/court-reinstates-lawsuit-challenging-online-sex-trafficking-law/ (describing the lawsuit filed by the EFF against the allegedly unconstitutional provisions of FOSTA-SESTA).
[347] *Id.* ("While a judge had previously tossed the lawsuit on the grounds that plaintiffs Alex Andrews and Eric Koszyk didn't face a credible threat of prosecution, the appeals court disagreed. It determined that Andrews faced a real threat due to her sex worker support site, while FOSTA may have harmed Koszyk by denying him the ability to offer therapeutic massages anywhere on Craigslist").
[348] David Greene, *EFF Sues to Invalidate FOSTA, an Unconstitutional Internet Censorship Law*, EFF (June 28, 2018) https://www.eff.org/deeplinks/2018/06/eff-sues-invalidate-fosta-unconstitutional-internet-censorship-law ("As a result of these hugely increased risks of liability, many platforms for online speech have shuttered or restructured").

A similar amendment to SESTA would be required to ensure ISPs are held liable for the distribution of deepfake pornography in addition to a federal criminal law.[349] While critics of SESTA argue the law violates the First Amendment due to its stifling effect on more than sex traffickers – such as forums dedicated to fostering safe and consensual sex work – ISPs positions as mechanism for the distribution of deepfake pornography warrants a narrow exclusion to Section 230 immunity that imposes liability on ISPs without violating free speech rights.[350] Further, without the ability to ensure deepfake pornography is not distributed, women's own free speech rights will continue to be stifled.[351]

a. Holding ISPs Accountable: A Federal Criminal Law and Amendment to Section 230

ISPs are responsible for the harmful distribution of deepfake pornography.[352] To prevent the distribution of future deepfake pornography and halt the dissemination of existing deepfake pornography, a two-pronged approach combining a federal criminal law and an amendment to Section 230 is necessary to hold ISPs liable.[353] However, the language of both must be narrow in scope to ensure ISPs do not resort to over-policing constitutionally valid speech in an effort to remain compliant.[354] This solution relies on these components working in a narrowly - construed tandem to prevent violations to crucial free speech protections.[355]

[349] Halverson, *supra* note 140, at 7 ("The courts have a long history of interpreting section 230 of the CDA to immunize Web sites from criminal responsibility for behavior that would be criminal if committed offline").
[350] *Id.* ("Given the exploitive and harmful effects of commercial sexual exploitation, it is vital the law works to break the chain of supply and demand. Unfortunately, section 230 of the CDA has long set a precedent against holding online facilitators of commercial sexual exploitation accountable").
[351] Turk, *supra* note 215 ("[P]ornographic deepfakes of women do threaten the integrity of our democracy . . . deepfake pornography is used as a tool to humiliate, demean and silence women. When women's voices are silenced simply because they are women; when women are humiliated because they are women; when women are subjugated because they are women – these all pose a threat to democracy").
[352] Halverson, *supra* note 140, at 13 ("For the women . . . who are being sexually exploited . . . it is vital to hold Web sites accountable for their role in this chain of crime . . . blanket immunity in the face of human suffering cannot be tolerated").
[353] *See* Delfino, *supra* note 35, at 928-37 (advocating for a similar two-pronged approach involving a federal law regulating deepfakes and extra-legal measures such as take-down requests).
[354] *Id.* at 925 ("[T]o survive constitutional challenge, legislation targeting revenge porn [must be] narrowly tailored to avoid encompassing legitimate and protected, albeit objectionable, speech").
[355] *Id.* at 933 ("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").

ii. Proposed Legislation Prohibiting the Distribution of Deepfake Pornography

The harms associated with the distribution of deepfake pornography make enacting a federal law prohibiting such conduct necessary.[356] Through comprehensive reform of the IPPA, ENOUGH, and MDFP Acts, the following proposed bill attempts to stop distribution of deepfake pornography without compromising free expression on the Internet.[357]

A. Criminalization of Deepfake Pornography Distribution Act

A BILL
To amend title 18, United States Code, to prohibit the distribution of nonconsensual deepfake pornography
SEC. 1. SHORT TITLE
This act may be cited as the "Criminalization of Deepfake Pornography Distribution Act of 2020"
SEC. 2. FINDINGS - The legislative body finds:
(1) deepfake pornography is a growing phenomenon that allows a person to use artificial intelligence to replace the face of an existing pornographic actor with the face of a private individual to create a hyper realistic portrayal of that individual performing sexual acts, despite those acts never occurring;
(2) deepfake pornography almost exclusively targets women who have not consented to their image being used in the forged sexual acts;
(3) the nonconsensual distribution of deepfake pornography results in irreparable harm to women; and
(4) as the technology used to create deepfake pornography becomes more advanced, accessible, and affordable, the distribution of deepfake pornography on the internet will continue to rise.[358]
SEC. 3. PURPOSE - This statute is designed with the purpose of:

[356] *Id. at 928* ("To best address the mounting dangers of pornographic deepfakes, legislative action is needed").
[357] *See id.* (describing the ability to create a narrow, comprehensive criminal law regulating deepfakes by "blending components of the ENOUGH Act and the MDFPA").
[358] *E.g.*, Delfino, *supra* note 35, at 892-93 ("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 [through] AI-assisted technology . . . specifically to generate nonconsensual pornography"); *id.* at 903 ("[D]eepfakes disproportionately victimize women and girls"); See Otero, *supra* note 24, at 593 (describing the harms associated with the distribution of deepfake pornography, including the "destruction of an individual's personal dignity and professional reputation, as well as their mental and emotional health"); Harris, *supra* note 23, at 128 ("[T]he technology is only going to improve. It takes hours to make a deepfake. Soon, it will take seconds, and the product will be indistinguishable from real videos").

(1) criminalizing the nonconsensual distribution of deepfake pornography on the Internet;
(2) protecting victims of deepfake pornography and implementing potential methods of redress; and
(3) attaching liability to any party who knowingly distributes nonconsensual deepfake pornography.[359]

SEC. 4. CERTAIN ACTIVITIES RELATING TO INTIMATE VISUAL DEPICTIONS: DISTRIBUTION OF DEEPFAKE PORNOGRAPHY

(a) IN GENERAL. —Chapter 88 of title 18, United States Code, is amended by adding at the end the following:

§ 1802. CERTAIN ACTIVITIES RELATED TO INTIMATE VISUAL DEPICTIONS: DISTRIBUTION OF DEEPFAKE PORNOGRAPHY

(a) Definitions – In this section:

(1) Internet Service Provider – The term 'internet service provider' has the same meaning given to the term 'interactive computer service' in Section 230 of the Communications Act of 1934 (47 U.S.C. § 230).[360]

(2) Deepfake Pornography – The term 'deepfake pornography' means a video created or altered in a manner that would falsely appear to a reasonable observer to be an authentic record of the actual speech, conduct, image, or likeness of an individual

(A) in which

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

(ii) the naked genitals or post-pubescent female nipples of any individual are visible.[361]

(3) Individual – the term 'individual' can either refer to the person whose body is depicted in the deepfake pornography or whose face is depicted in the deepfake pornography.[362]

(4) Distribute – The term 'distribute' as used within this section means enabling access to deepfake pornography, whether through direct disclosure or passive dissemination.[363]

---

[359] *See* Delfino, *supra* note 35 at 921, 930 (discussing a similar proposal that "seeks to strike a balance between protecting victims, punishing wrongdoers, and protecting freedom of expression").
[360] 47 U.S.C. § 230(f)(2) ("The term 'interactive computer service' means any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server, including specifically a service or system that provides access to the Internet and such systems operated or services offered by libraries or educational institutions").
[361] *See* Delfino, *supra* note 35, at 929 (discussing a similar proposal in which deepfake and pornographic deepfake are separately defined but include similar language such 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, conduct, image, or likeness of an individual" and "the naked genitals or post-pubescent female nipples of any individual are visible").
[362] *See id.* (discussing a similar proposal which defines individual as "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").
[363] *See* Assemb. 602, 2019-2020 Leg., Reg. Sess. (Cal. 2019) (as introduced) (discussing a similar proposal in which disclosure is defined as "publish, make available, or distribute to the public").

(5) Knowledge – the term knowledge means direct and clear awareness of a fact or condition sufficient to create notice. In this section, the existence of a complaint filed by a victim asserting the nonconsensual distribution of deepfake pornography portraying her image sufficiently constitutes knowledge and will be construed as notice to an Internet Service Provider. [364]

(6) Consent – the term 'consent' means a written agreement, in contractual or plain language, signed knowingly and voluntarily by an individual that includes a general description of the sexually explicit material.[365]

(7) Information Content Provider – The term 'information content provider' has the meaning given that term in section 230(f) of the Communications Act of 1934 (47 U.S.C. § 230(f)).[366]

(8) Sexually Explicit Conduct – The term 'sexually explicit conduct' has the meaning given that term in 18 U.S.C § 2256(2)(B). [367]

(b) Offense – Except as provided in subsection (e), it shall be a federal crime to use any internet service provider, electronic communication service, electronic communication system of internet commerce, or any other facility of interstate or foreign commerce to distribute deepfake pornography of an individual:

(1) with knowledge of or reckless disregard for:

(A) the lack of consent of the individual(s) to the distribution; and

(B) the resulting harm distribution could cause the individual(s).

(2) without an objective reasonable belief that distribution touches upon a matter of public concern.[368]

(c) Penalty – Any party who violates subsection (a) shall be:

(1) fined under this title, imprisoned for not more than 3 years, or both, in initial offenses; or

(2) fined under this title, imprisoned for not more than 5 years, or both, in any subsequent case; or

---

[364] *See* Delfino, *supra* note 35, at 932 (discussing a similar proposal in which "the exception for service providers comports with section 230 immunity but is not unlimited [through] imposing liability where 'the communications service intentionally solicits, or distributes knowingly or with reckless disregard, content that is in violation of this section'").

[365] *See* Assemb. 602, 2019-2020 Leg., Reg. Sess. (Cal. 2019) (as introduced) (providing a broader definition of consent as an agreement "written in plain language").

[366] 47 U.S.C. § 230(f)(3) ("The term "information content provider" means any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service.").

[367] 18 U.S.C. § 2256(2)(B) ("[S]exually explicit conduct means—(i)graphic sexual intercourse, including genital-genital, oral-genital, anal-genital, or oral-anal, whether between persons of the same or opposite sex, or lascivious simulated sexual intercourse where the genitals, breast, or pubic area of any person is exhibited; (ii)graphic or lascivious simulated; (I)bestiality; (II)masturbation; or (III)sadistic or masochistic abuse; or (iii)graphic or simulated lascivious exhibition of the anus, genitals, or pubic area of any person.").

[368] *See* Delfino, *supra* note 35, at 930-32 (describing a similar proposal which "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").

(3) fined under this title, imprisoned for not more than 10 years, or both, in any case, initial or subsequent, of a violation where:

(A) the individual depicted in the deepfake pornography is a minor; or

(B) the deepfake pornography facilitates violence against a particular individual.[369]

(d) Remedies – Remedies shall by provided at the discretion of the court, promote the purpose of preventing the further distribution of the deepfake pornography, and may include the following non-exhaustive list of options:

(1) temporary or permanent injunction;

(2) order compelling an Internet Service Provider to respond to a victim's removal request;

(3) damages or restitution; and

(4) any other relief deemed proper by the court.[370]

(e) 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.[371]

(2) Service Providers – This section shall not apply to

(A) any voluntary action of any internet service provider taken in good faith to restrict access to or distribution of deepfake pornography; or

(B) any action taken by an internet service provider to make information available to information content providers regarding the technical means to restrict access to deepfake pornography.[372]

(f) Threats and Extortion – Any party who intentionally threatens to commit an offense under subsection (b), regardless of whether the threat is an act of extortion, shall be punished as provided in subsection (c).[373]

(g) Venue and Extraterritoriality – A prosecution under this section may be brought in a district where the defendant or individual(s) resides or in a district where the pornographic deepfake is distributed or made available. There is extraterritorial

[369] *See id.* at 930-932 (describing a similar proposal which draws from the ENOUGH Act to impose similar penalties, such as "a five-year sentence [which] emphasizes the severity of the crime and has been successfully incorporated in other nonconsensual pornography statutes").

[370] *See id.* at 930-933 (describing a similar proposal which "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").

[371] *Id.* at 932.

[372] *Id.*

[373] Id.

federal jurisdiction over an offense under this section if the defendant or individual(s) is a citizen or permanent resident of the United States.[374]

B. Analysis of Proposed Legislation

The proposed bill is an attempt to protect victims, hold distributers of deepfake pornography accountable, and protect free speech on the internet.[375] The bill seeks to achieve this purpose by narrowing the previously broad legislation that attempted to regulate and prohibit deepfakes.[376] First, the bill explicitly addresses deepfake pornography as opposed to the umbrella term of deepfakes in general.[377] Second, the bill's focus is not on creation of deepfake pornography but on distribution.[378] The bill does not seek to regulate creation, and acknowledges the protected privacy of an individual's home cannot be policed in the context of deepfake pornography.[379] Rather, the bill seeks to eliminate online distribution through making it criminal to knowingly distribute nonconsensual deepfake pornography.[380] Second, the bill includes an intent requirement in that it punishes those who distribute deepfake pornography knowingly or recklessly.[381] Historically, proposed bills are more likely to pass constitutional muster when an intent element is provided, as it heightens the degree of proof needed to loosen First Amendment protections.[382] Third, unlike the IPPA, the bill does not contain any language comporting with Section 230, thereby continuing to hold ISP's immune from liability, as is later addressed in the bill amending Section 230 to hold ISP's liable for distribution of deepfake pornography.[383] Finally, this bill specifically lays out penalties and remedies whose purpose is to halt distribution.[384] For example, penalties are heightened for each subsequent violation, and remedies for victims include injunctions as well as removal orders.[385]

[374] Id.
[375] *Id.* at 930.
[376] *Id.* at 931.
[377] Id.
[378] *Id.* at 928-31.
[379] *Id.* at 931.
[380] Id.
[381] Id.
[382] *Id.* at 932.
[383] *Id.* at 908.
[384] Id.
[385] *Id.*

ii. Proposed Amendment to Section 230

ISPs are the mechanisms for distribution of deepfake pornography, but Section 230 may still protect an ISP from liability for distribution despite a federal criminal law prohibiting it.[386] Such was the case with online sex-trafficking, resulting in the SESTA amendment.[387] Based on the similarity in purpose between SESTA and deepfake pornography, proposed legislation amending Section 230 to hold ISP's liable for the distribution of deepfake pornography should be modeled after SESTA, but include much narrower provisions to avoid censoring free speech.[388]

B. Stop Online Distribution of Deepfake Pornography Act

A Bill To amend the Communications Act of 1934 to clarify that section 230 of such act does not prohibit the enforcement against providers or users of interactive computer services of Federal and State criminal and civil law relating to the distribution of deepfake pornography.[389]

Sec. 1. Short Title

This act may be cited as the "Stop Online Distribution of Deepfake Pornography Act"

Sec. 2. Findings

It is the sense of Congress that:

(1) Section 230 of the Communications Act of 1934 (47 U.S.C. 230, commonly known as the Communications Decency Act of 1996") was never intended to provide legal protection to websites that knowingly distribute deepfake pornography;

(2) Websites that distribute deepfake pornography have been reckless by allowing the nonconsensual, sexually explicit videos of victims to circulate on the internet and have done nothing to prevent their distribution; and

(3) Clarification of such section is warranted to ensure that such section does not provide such protection to such websites.[390]

SEC. 3. Ensuring ability to enforce Federal and State Criminal And Civil law relating to the distribution of deepfake pornography.

(a) IN GENERAL. Section 230 of the Communications Act of 1934 (47 U.S.C. 230) is amended

(1) in subsection (b)

(A) in paragraph (4), by striking "and" at the end;

(B) in paragraph (5), by striking the period at the end and inserting "; and"; and

[386] Otero, *supra* note 24, at 607.
[387] Romano, *supra* note 336.
[388] Greene, *supra* note 348.
[389] FOSTA-SESTA, H.R. 1865, 115th Cong. (2017-2018).
[390] Wasserman, *supra* note 143, at 195.

(C) by adding paragraph (6), to include: "(6) to promote the balance of the protected interests of Interactive Computer Services, Internet Users, and victims of deepfake pornography."[391]

(2) in subsection (e), by adding at the end of the following:

"(6) NO EFFECT ON DEEPFAKE PORNOGRAPHY LAWS – Nothing in this section (other than subsection (c)(2)(A)) shall be construed to impair or limit –

(A) any claim in an action brought under section 1802 of title 18, United States Code, if the conduct underlying the claim constitutes a violation of section 1802; and

(B) any charge in a criminal prosecution brought under state law if the conduct underlying the charge constitutes a violation of section 1802 of title 18, United States Code.[392]

(b) EFFECTIVE DATE – The amendments made by this section shall take effect on the date of enactment of this Act, and the amendment made by subsection (a)(2) shall apply regardless of whether the conduct allegedly occurred, or is alleged to have occurred before, on, or after such date of enactment.[393]

SEC. 4. ACTIONS BY STATE ATTORNEYS GENERAL.

(a) IN GENERAL.—Section 1802 of title 18, United States Code, is amended by adding at the end the following:

"(h) In any case in which the attorney general of a State has reason to believe that an interest of the residents of that State has been or is threatened or adversely affected by any person who violates section 1802, the attorney general of the State, as parens patriae, may bring a civil action against such person on behalf of the residents of the State in an appropriate district court of the United States to obtain appropriate relief."[394]

B. Analysis of Proposed Amendment

This proposed bill seeks to ensure that ISPs will not remain immune from criminalization under proposed federal legislation 18 U.S.C. 1802 prohibiting the distribution of deepfake pornography.[395] The proposed language is narrower than the previous SESTA bill as it does not seek to hold ISP's liable for an umbrella term such as sex-trafficking or nonconsensual pornography, but rather a specific, defined, and easily detected form of sexual exploitation – deepfake pornography.[396] Further, the

[391] 47 U.S.C. § 230(b).
[392] FOSTA-SESTA, H.R. 1865, 115th Cong. (2017-2018).
[393] Id.
[394] Id.
[395] Id.
[396] Id.

bill holds an ISP accountable for conduct underlying a violation of Section 1802, meaning the ISP must be aware that the deepfake pornography of a victim is being distributed on its server.[397] Because Section 1802 specifies that a complaint filed by the victim with the ISP regarding removal of the deepfake pornography constitutes knowledge, an ISP will not be held liable unless it ignores the complaints of victims.[398] Therefore, the proposed amendment to Section 230 clarifies that liability can be brought against an ISP for distribution of deepfake pornography, but narrows the circumstances in which liability is found; therefore the amendment remains constitutionally compliant.[399]

The following amendments to Section 230 as well as a federal criminal law prohibiting the distribution of deepfake pornography must be used in conjunction to ensure ISPs, as the mechanisms for distribution, take responsibility for their historically passive approach to regulating the sexual exploitation of women.[400]

VI. Conclusion

As technology continues to advance, deepfakes will not only become more prevalent, but more seamless as well, further blurring the line between reality and fiction.[401] The distribution of deepfake pornography has the potential to cause severe harm to women who have not consented to the use of their image in pornography.[402] Current legal remedies serve as a band aid solution instead of a preventative measure and do not adequately protect victims of deepfake pornography.[403] Accordingly, ISPs allowing the distribution of deepfake pornography should be held accountable.[404] While the implementation of Section 230 has afforded ISPs the necessary immunity to foster First Amendment protections, this immunity has proven to be detrimental when applied to the distribution of deepfake pornography.[405] Therefore, a combination of the implementation of a new federal law criminalizing the distribution of deepfake pornography and an amendment to Section 230 excluding immunity of ISPs for distribution of deepfake pornography is required.[406] This solution does not intrude on

[397] Delfino, *supra* note 35, at 928-31.
[398] *Id.*
[399] *Id.* at 931.
[400] *Id.* at 932.
[401] Harris, *supra* note 2, at 102.
[402] Otero, *supra* note 24, at 592.
[403] *Id.* at 593.
[404] *Id.* at 594.
[405] *Id.* at 598.
[406] *Id.* at 598-99.

First Amendment protections, as it does not penalize the creation of deepfake pornography in the privacy of one's home and narrows criminalization of distribution to deepfakes that are both pornographic and nonconsensual in nature, leaving deepfakes outside these categories protected.[407] This solution balances the protection of victims, immunity of ISPs, and freedom of speech on the internet.[408]

[407] Stanley v. Georgia, 394 U.S. 557, 567-68 (1969); Delfino, *supra* note 35, at 931.
[408] Citron, *supra* note 20, at 390.