School of Law
UNIVERSITY OF GEORGIA

Georgia Law Review

Volume 56 | Number 2 Article 8

2022

The Double-side of Deepfakes: Obstacles and Assets in the Fight Against Child Pornography

Abigail Olson
University of Georgia School of Law

The Double-side of Deepfakes: Obstacles and Assets in the Fight Against Child Pornography

Cover Page Footnote

J.D. Candidate, 2022, University of Georgia School of Law; B.A., 2018, Emory University.

THE DOUBLE-SIDE OF DEEPFAKES: OBSTACLES AND ASSETS IN THE FIGHT AGAINST CHILD PORNOGRAPHY

Abigail Olson[*]

Deepfake technology recently took the internet by storm. Although they can be used for both innocuous and nefarious purposes, deepfakes overwhelmingly depict people who appear to be creating nonconsensual pornography. The rise of deepfake technology must be accounted for in the existing federal legal framework, specifically in cases implicating images of children. While deepfakes' malicious uses ought to be criminalized, exceptions should be made to use deepfake technology as a tool to enforce and deter purveyors of child pornography. This Note explores what the emerging legal framework addressing deepfakes should look like and considers the importance of using the "flipside" of deepfake technology—meaning its potentially safe, beneficial uses—to stop child pornography.

[*] J.D. Candidate, 2022, University of Georgia School of Law; B.A., 2018, Emory University.

TABLE OF CONTENTS

I. INTRODUCTION 867

II. BACKGROUND 870
- A. CHILD PORNOGRAPHY LAW IN THE UNITED STATES 870
- B. THE EMERGENCE OF DEEPFAKE TECHNOLOGY 874

III. ANALYSIS 879
- A. SUGGESTED LEGAL FRAMEWORK TO ADDRESS DEEPFAKES IN THE CONTEXT OF CHILD PORNOGRAPHY 880
- B. EXCEPTIONS TO THE SUGGESTED LEGAL FRAMEWORK 882
  - *1. Allow the Use of Synthetic Pornography Materials for Law Enforcement to Use in Online Stings* 882
  - *2. Allow the Use of Synthetic Child Pornography as a Rehabilitative Tool for Predators* 885

IV. BARRIERS TO IMPLEMENTATION 888

V. CONCLUSION 890

I. INTRODUCTION

Your friend sends you a video. At first glance, it looks like any other scene from *Game of Thrones*. But a few seconds into the scene, something strange happens: the characters' faces change. The video no longer shows a conversation between Sam Tarly and Jon Snow. Instead, you're watching your friend talk to different *Game of Thrones* versions of himself—his face but the actors' bodies. The strangest thing about the scene: it looks realistic.[1]

Anyone with a smartphone can create videos like the one just described, which, like this example alone, can be a creative source of entertainment.[2] These altered videos are called "deepfakes."[3] Simply put, a deepfake is a "fake video or audio recording that look[s] and sound[s] just like the real thing."[4] Deepfakes include any type of falsified video content that appears realistic, and just about anyone can make them.[5] Deepfakes use technology driven by artificial intelligence (AI) to create the fake image, which can

[1] This situation is based on a video made by Allan Xia using the ZAO iOS App. Allan Xia (@AllanXia), TWITTER (Sept. 1, 2019, 4:40 AM), https://twitter.com/AllanXia/status/1168081219768045569. Xia's video is "a neat demonstration of what the app is capable of . . . . [T]he clips were generated in under eight seconds from just a single photograph . . . ." Jon Porter, *Another Convincing Deepfake App Goes Viral Prompting Immediate Privacy Backlash*, VERGE (Sept. 2, 2019, 6:32 AM), https://www.theverge.com/2019/9/2/20844338/zao-deepfake-app-movie-tv-show-face-replace-privacy-policy-concerns.

[2] *See, e.g.*, Jacob Schulz, *The Deepfake iPhone Apps Are Here*, LAWFARE (Apr. 27, 2020, 1:14 PM), https://www.lawfareblog.com/deepfake-iphone-apps-are-here (describing one of the author's quarantine activities: using the smartphone application "Mug Life" to "become late night comedian Seth Meyers").

[3] *Cf.* Zak Doffman, *Chinese Deepfake App ZAO Goes Viral, Privacy of Millions 'At Risk'*, FORBES (Sept. 2, 2019, 4:27 AM), https://www.forbes.com/sites/zakdoffman/2019/09/02/chinese-best-ever-deepfake-app-zao-sparks-huge-faceapp-like-privacy-storm/#214b5d828470 (discussing the privacy implications that resulted from the popularity of the ZAO app).

[4] J.M. Porup, *How and Why Deepfake Videos Work—And What is at Risk*, CSO (Mar. 18, 2021, 2:00 AM), https://www.csoonline.com/article/3293002/deepfake-videos-how-and-why-they-work.html.

[5] *See id.* ("[T]oday anyone can download deepfake software and create convincing fake videos in their spare time.").

appear either extremely doctored or extremely realistic.[6] Deepfake videos are relatively new—the technology needed to create them only became widely available to the public in 2017.[7] Since then, individuals have used deepfakes for an array of purposes: funny videos shared between friends,[8] comedy shows,[9] political videos,[10] and even immersive videos of art and culture.[11] While the these examples demonstrate the versatility of the technology, an overwhelming majority of deepfakes depict nonconsensual pornographic videos.[12] Many of the first deepfakes of this kind

[6] *See* Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753, 1759 (2019) ("This technology often involves the use of a 'neural network' for machine learning. . . . If the network processes a broad array of training examples, it should be able to create increasingly accurate models. It is through this process that neural networks categorize audio, video, or images and generate realistic impersonations or alterations." (footnote omitted)).

[7] *See id.* at 1763 ("Indeed, diffusion has begun for deep-fake technology. The recent wave of attention generated by deep fakes began after a Reddit user posted a tool inserting the faces of celebrities into porn videos." (citing Emma Grey Ellis, *People Can Put Your Face on Porn—And the Law Can't Help You*, WIRED (Jan. 26, 2018), https://www.wired.com/story/face-swap-pom-legal-limbo/)).

[8] *See, e.g.*, Arnold, *10 Crazy Deepfake Apps That Will Make You Question Reality*, DEEPFAKENOW (Apr. 21, 2020), https://deepfakenow.com/10-crazy-deepfake-apps-that-will-make-you-question-reality/ (listing ten of the most popular deepfake apps available for download on iOS and Android); *id.* ("Swapping your own face with that of someone you know is pretty hilarious, and it makes for a memorable picture that can be shared with friends and family.").

[9] One of the most well-known examples of deepfakes in pop culture comes from Bill Hader's appearance on Conan O'Brien's television show. *See, e.g.*, Ctrl Shift Face, *Bill Hader Impersonates Arnold Schwarzenegger [DeepFake]*, YOUTUBE (May 10, 2019), https://www.youtube.com/watch?v=bPhUhypV27w&ab_channel=CtrlShiftFace.

[10] *See* Thomas E. Kadri, *Drawing Trump Naked: Curbing the Right of Publicity to Protect Public Discourse*, 79 MD. L. REV. 899, 957 (2019) (discussing how a deepfake video of Donald Trump saying something that he never actually said spurred "valuable political speech" in Belgium regarding the Paris climate agreement).

[11] *See, e.g.*, Cuseum, *3 Things You Need to Know About AI-Powered "Deep Fakes" in Art & Culture* (Dec. 17, 2019), https://cuseum.com/blog/2019/12/17/3-things-you-need-to-know-about-ai-powered-deep-fakes-in-art-amp-culture (exploring the manner in which "deepfakes are being used to produce art, engage audiences, and provide personalized experiences to visitors in a way that has never been done before").

[12] These videos overwhelmingly target women in acts of nonconsensual violence. *See, e.g.*, Samantha Cole, *This Horrifying App Undresses a Photo of Any Woman with a Single Click*, VICE: MOTHERBOARD (June 26, 2019, 5:48 PM), https://www.vice.com/en/article/kzm59x/deepnude-app-creates-fake-nudes-of-any-woman

featured female celebrities.[13] Now, about 70% of deepfake targets are private individuals.[14]

On the one hand, deepfakes can be a creative and innocuous source of entertainment. On the other hand, deepfakes pose alarming risks in the context of nonconsensual pornography and different obstacles in the context of online child pornography.[15] The harms of deepfakes are apparent in their use as tools to create nonconsensual images depicting abuse. But just as deepfakes can be used to perpetuate abuse, they can also potentially be used to combat that same type of harm. This Note suggests that this flipside of deepfake technology could also serve as a powerful tool in the fight against child pornography.

The relative novelty and prevalence of deepfake technology has caught legal systems on their back foot: technology has far outstripped existing laws on content regulation.[16] The rise of deepfake technology is concerning, and legal frameworks must evolve to account for it. This Note argues that the developing legal frameworks must take special account of pornographic deepfakes of children by enacting new laws and amending current laws criminalizing child pornography and sexual abuse that are applicable to deepfakes. In addition to the development of the law surrounding deepfakes, this Note argues that carveouts should exist for the use of this technology as a tool for law enforcement to catch online predators and to use as a potential rehabilitative tool to divert predators away from continued victimization of children.

("The $50 DeepNude app dispenses with the idea that deepfakes were about anything besides claiming ownership over women's bodies.").

[13] *See* Samantha Cole, *We Are Truly Fucked: Everyone Is Making AI-Generated Fake Porn Now*, VICE: MOTHERBOARD (Jan. 24, 2018, 1:13 PM), https://www.vice.com/en/article/bjye8a/reddit-fake-porn-app-daisy-ridley (depicting the initial deepfake boom in which "several convincing porn videos of celebrities—including Gal Gadot, Maisie Williams, and Taylor Swift" were made using the open-sourced AI and publicly available videos).

[14] Siladitya Ray, *Bot Generated Fake Nudes of over 100,000 Women Without Their Knowledge, Says Report*, FORBES (Oct. 21, 2020, 6:46 AM), https://www.forbes.com/sites/siladityaray/2020/10/20/bot-generated-fake-nudes-of-over-100000-women-without-their-knowledge-says-report/#589757137f6b.

[15] *See infra* Section II. A.

[16] *See* Danielle K. Citron, *Sexual Privacy*, 128 YALE L.J. 1870, 1877 (2019) ("Thanks to networked technologies, sexual privacy can be invaded at scale and from across the globe.").

Part II discusses current child pornography law and the emergence of deepfakes in the context of child sexual abuse. Part III suggests how incorporation of deepfake criminalization could proceed at the federal level and then argues for two exceptions to the criminalization of deepfakes in this context: as a crime-fighting tool to identify and stop purveyors of child pornography and as a deterrent or rehabilitative tool to prevent offenders from reoffending. Part IV considers and responds to counterarguments and potential criticisms of exceptions to the criminalization of synthetic child pornography.

II. BACKGROUND

A. CHILD PORNOGRAPHY LAW IN THE UNITED STATES

Child pornography is an image or video of a child being sexually exploited.[17] Unlike pornography that features consenting legal adults, children depicted in pornography are victims of sexual abuse and exploitation.[18] In child pornography, "[t]he act creating the pornography [is] by force—a rape, for example—and there [is] unequivocally no consent, legally or commonsensically speaking."[19] The sexual abuse and exploitation of children in the United States is not taken lightly, as evidenced by the massive social condemnation and disdain held for perpetrators of this abuse.[20]

[17] 18 U.S.C. § 2252A; *see also Child Pornography*, U.S. DEP'T OF JUST., https://www.justice.gov/criminal-ceos/child-pornography (last updated May 28, 2020) ("Federal law defines child pornography as any visual depiction of sexually explicit conduct involving a minor (persons less than 18 years old. Images of child pornography are also referred to as child sexual abuse images.").

[18] *Id.*

[19] Whitney J. Gregory, Comment, *Honeypots: Not for Winnie the Pooh but for Winnie the Pedo—Law Enforcement's Lawful Use of Technology to Catch Perpetrators and Help Victims of Child Exploitation on the Dark Web*, 26 GEO. MASON L. REV. 259, 263 (2018).

[20] *See id.* at 267 ("Child pornography is deservedly among the most heavily punished offenses in criminal law. Child pornography is a crime federally, internationally, and in all fifty states."); Anthony M. Dillof, *Possession, Child Pornography, and Proportionality: Criminal Liability for Aggregate Harm Offenses*, 44 FLA. ST. U. L. REV. 1331, 1372 (2017) ("[B]ecause the general proliferation of an image of sexual abuse leads to feelings of humiliation, helplessness, and fear of recognition on the part of the image's subject, the practice of trading child porn over the Internet may be considered . . . wrongful and deserving of punishment.").

Furthermore, the abuse and exploitation can cause massive long-term repercussions for victims: "The permanent record of a child's sexual abuse can alter his or her life forever. Many victims of child pornography suffer from feelings of helplessness, fear, humiliation, and lack of control given that their images are available for others to view in perpetuity."[21]

Laws first passed in the 1970s provide the foundation for today's legal response to child pornography.[22] The United States Code dedicates an entire chapter to child sexual exploitation and abuse,[23] which defines child pornography as "any visual depiction, including any photograph, film, video, picture, or computer or computer-generated image or picture, whether made or produced by electronic, mechanical, or other means, of sexually explicit conduct" involving a minor.[24] A "visual depiction" can be any kind of image and can be stored in a number of ways.[25] The Child Pornography Prevention Act of 1996 criminalizes the production, trafficking in, and possession of child pornography.[26] While criminalization of child pornography continues to grow,[27] pushback related to online free speech concerns also persists.[28] The U.S. Supreme Court,

[21] *Child Pornography*, *supra* note 17. Additionally, "[e]xperts and victims agree that victims depicted in child pornography often suffer a lifetime of re-victimization by knowing the images of their sexual abuse are on the Internet forever." *Id.*

[22] *See* Amy Adler, *The Perverse Law of Child Pornography*, 101 COLUM. L. REV. 209, 211–12 (2001) (commenting on "the 'discovery' in the late 1970s of the twin problems of child sexual abuse and child pornography, and the continuation of the problems").

[23] The chapter on child sexual exploitation and abuse is Chapter 110 of Title 18 of the United States Code. 18 U.S.C. §§ 2251–2260A.

[24] 18 U.S.C. § 2256.

[25] *Id.* § 2256(8); *see, e.g.*, Gregory, *supra* note 19, at 268 ("Images may be undeveloped and are not required to be stored in a permanent format.").

[26] 18 U.S.C. § 2252A.

[27] *See, e.g.*, Melissa Hamilton, *The Efficacy of Severe Child Pornography Sentencing: Empirical Validity or Political Rhetoric*?, 22 STAN. L. & POL'Y REV. 545, 545 (2011) ("Congress has repeatedly forced increases in the federal sentencing guidelines using unconventional means for child pornography offenses, notwithstanding opposition from the United States Sentencing Commission.").

[28] *See* Ashcroft v. Free Speech Coal., 535 U.S. 234, 256 (2002) (striking down two provisions in the Child Pornography Prevention Act for being overbroad and infringing on "the freedom to engage in a substantial amount of lawful speech"). *Ashcroft* poses a significant barrier to the implementation of federal criminal law banning deepfakes in the context of child pornography, but as discussed in Part IV, the current situation surrounding online child

however, has found that the First Amendment does not protect child pornography.[29]

Child pornography offenses typically fall under federal jurisdiction when they involve interstate commerce.[30] Interstate commerce includes any "commerce between one State, Territory, Possession, or the District of Columbia."[31] Any person who receives, distributes, reproduces, or advertises material constituting or containing child pornography in a manner that uses or affects interstate commerce violates federal law.[32] Interstate commerce is an expansive umbrella—so expansive that the Internet is considered "an instrumentality of interstate commerce" in U.S. caselaw.[33] For example, federal jurisdiction attaches if the materials used to download or store child pornography traveled through interstate commerce, if a hard copy image of child pornography traveled through interstate or foreign commerce, or if the perpetrator used the Internet to commit a child pornography offense.[34]

Child pornography offenders can also be prosecuted under state law.[35] For example, in 2007 Georgia enacted its own law targeting purveyors of child pornography, called the Computer or Electronic Pornography and Child Exploitation Prevention Act.[36] The language of the statute aligns closely with existing federal law. Both Georgia and federal law criminalize knowingly coercing a minor to

sexual abuse and technology has changed to such an extent that *Ashcroft* may no longer hold water.

[29] *See, e.g.*, United States v. Williams, 553 U.S. 285, 288 (2008) ("We have long held that obscene speech—sexually explicit material that violates fundamental notions of decency—is not protected by the First Amendment." (citation omitted)).

[30] 18 U.S.C. §§ 2251, 2252, 2252A; *see also* Gregory, *supra* note 19, at 269 ("Federal jurisdiction attaches if child pornography activity occurs in interstate or foreign commerce.").

[31] 18 U.S.C. § 10.

[32] *Id.* § 2252A.

[33] United States v. Hornaday, 392 F.3d 1306, 1311 (11th Cir. 2004) (first citing United States v. Pipkins, 378 F.3d 1281, 1295 (11th Cir. 2004); and then citing United States v. Panfil, 338 F.3d 1299, 1300 (11th Cir. 2003)).

[34] Gregory, *supra* note 19, at 269.

[35] *See Citizen's Guide to U.S. Federal Law on Child Pornography*, U.S. DEP'T OF JUST., https://www.justice.gov/criminal-ceos/citizens-guide-us-federal-law-child-pornography (last updated May 28, 2020) ("It is important to note that an offender can be prosecuted under state child pornography laws in addition to, or instead of, federal law.").

[36] O.C.G.A. § 16-12-100.2 (2021).

engage in sexually explicit conduct, possessing sexually explicit depictions of minors, and distributing explicit material depicting minors.[37] The penalties for a federal child pornography conviction are harsher though: the minimum time of imprisonment in Georgia is one year,[38] but the federal minimum is five years.[39] Although state laws may mirror federal legislation, states have a distinct interest in creating state law causes of action to prosecute offenders when federal jurisdiction does not attach to activity related to child pornography.[40]

Despite the societal disapproval and legal ramifications for those who engage in activity relating to child pornography, a large community of people who produce and consume materials depicting the exploitation of children still exists.[41] In 2011, the U.S. Attorney General released a statement noting that the Department of Justice had "seen a historic rise in the distribution of child pornography" and that "[t]ragically, the only place we've seen a *decrease* is in the age of victims."[42] The expansion of the Internet brought with it an explosion of the child pornography market.[43] Purveyors of child pornography have created communities on the "Dark Web," where

[37] *Id.*; 18 U.S.C. § 2252(a).

[38] O.C.G.A. § 16-12-100.2(c)(2).

[39] 18 U.S.C. § 2252(b)(1).

[40] *See* Dillof, *supra* note 20, at 1340 n.67 (noting that in addition to an expansion in federal law, "states have also significantly increased the penalties for possession of child pornography since 2000," indicating an existing state interest in policing this type of behavior); *see also* BHobson, *Concurrent Jurisdiction: Possession of Child Pornography*, ODOM, DAVIS & HOBSON (Apr. 13, 2020), https://www.wendellodom.com/concurrent-jurisdiction-possession-of-child-pornography/ ("The federal government normally saves their resources for egregious cases. For example, the federal government is unlikely to pursue criminal charges for possession of child pornography if the number of images is low.").

[41] *See* Paul Bischoff, *The Rising Tide of Child Abuse Content on Social Media*, COMPARITECH (Jan. 11, 2022), https://www.comparitech.com/blog/vpn-privacy/child-abuse-online-statistics/ (indicating that Facebook flagged 55.6 million instances of child nudity and sexual exploitation in 2021); *Child Sexual Abuse Material*, NAT'L CTR. FOR MISSING & EXPLOITED CHILD., https://www.missingkids.org/theissues/csam (last visited Dec. 16, 2021) (reporting that the National Center for Missing & Exploited Children's Cyber Tipline has received over 82 million reports since its creation in 1998).

[42] Eric Holder Jr., Att'y Gen., Address Before the Nat'l Strategy Conf. on Combating Child Exploitation (May 19, 2011) (emphasis added), https://www.justice.gov/criminal-ceos/child-pornography.

[43] *See Child Pornography*, *supra* note 17 ("The expansion of the Internet and advanced digital technology lies parallel to the explosion of the child pornography market.").

they can anonymously share and participate in child sexual abuse.[44] The Dark Web is an intentionally hidden section of the Internet invisible to common search engines that "requires the use of an anonymizing browser called Tor for access."[45] The anonymous engagement with child pornography materials in an online community takes away the shame and social ramifications that would normally accompany this behavior.[46] Instead, the online digital market emboldens child pornography offenders by encouraging, rather than condemning, this type of behavior. Fortunately, government agencies, such as the Federal Bureau of Investigation (FBI), have an established history of action against perpetrators of child pornography.[47] The FBI uses a wide range of online tools to exploit the anonymity of the Dark Web to investigate, identify, and terminate child pornography websites.[48]

B. THE EMERGENCE OF DEEPFAKE TECHNOLOGY

While hosting and accessing online child pornography sites with anonymous access like Playpen are already illegal,[49] more recent

[44] *See id.* ("Child pornography offenders can also connect on Internet forums and networks to share their interests, desires, and experiences abusing children, in addition to selling, sharing, and trading images.").

[45] Darren Guccione, *What Is The Dark Web? How to Access It and What You'll Find,* CSO: SPOTLIGHT (July 1, 2021, 2:00 AM), https://www.csoonline.com/article/3249765/what-is-the-dark-web-how-to-access-it-and-what-youll-find.html.

[46] *See* Daniel Armagh, *A Safety Net for the Internet: Protecting Our Children*, 5 JUV. JUST. J. (1998) https://ojjdp.ojp.gov/sites/g/files/xyckuh176/files/jjjournal/jjjournal1598/net.html ("Because of its anonymity, rapid transmission, and unsupervised nature, the Internet has become the venue of choice for predators who transmit and receive child pornography.").

[47] *See, e.g.*, Gregory, *supra* note 19, at 262 (discussing the evolution of FBI tactics in tracking predators).

[48] *See Crimes Against Children/Online Predators*, FED. BUREAU OF INVESTIGATION, https://www.fbi.gov/investigate/violent-crime/cac (last visited June 2, 2022) (outlining the FBI's approach to addressing crimes against children, including child pornography); *see also* Gregory, *supra* note 19, at 262 ("What was once intercepted via mail by Postal Inspectors is now a global game of virtual hide-and-seek—an undertaking involving foreign and domestic law enforcement agencies, cryptocurrency, and technology named after food. Illustrative is the FBI's takedown of Playpen.").

[49] *See* Daniel Masakyan & Elfin Noce, *Recent "Playpen" Cases, The Onion Router (TOR) Network, and Privacy Considerations for Hidden IP Addresses*, *in* CLOUD COMPUTING LEGAL DESKBOOK § 8:7, Westlaw (database updated October 2021) ("Playpen [was] a known child pornography website that operate[d] as a 'hidden service' to maintain user anonymity.").

technological innovations might not fit into existing laws.[50] The emergence of synthetic child pornography—which is one term for computer-generated images of child pornography that is basically indistinguishable from traditionally generated images—poses one such issue.[51] Judicial interpretations can differ over whether synthetic or altered images fall outside the realm of child pornography, and thus, over whether they are entitled to First Amendment protections.[52] Despite this potential for disagreement, a majority of courts have held that synthetic child pornography is covered by the Child Pornography Prevention Act.[53] The specific acts criminalized on the state level relating to synthetic child pornography materials remain jurisdictionally dependent.[54]

A technological boom of synthetic porn videos—"deepfakes" —emerged in 2017, with catastrophic results.[55] Deepfakes are "videos that have been manipulated to make it look like the subject is

[50] *See* Lori J. Parker, *Validity, Construction, and Application of Federal Enactments Proscribing Obscenity and Child Pornography or Access Thereto on the Internet*, 7 A.L.R. FED. 2d 1 (2005) ("As noted by the U.S. Court of Appeals for the Armed Forces, '[n]ew technologies create interesting challenges to long established legal concepts,' and the Internet represents new technology over which a body of existing law regulating . . . child pornography has been superimposed.").

[51] Synthetic child pornography falls under existing federal laws on child pornography. *See Citizen's Guide to U.S. Federal Law on Child Pornography*, *supra* note 35 (indicating that child pornography definitions include "synthetic" or "digital or computer generated images indistinguishable from an actual minor, and images created, adapted, or modified, but appear to depict an identifiable, actual minor").

[52] *See* Carissa Byrne Hessick, *The Expansion of Child Pornography Law*, 21 NEW CRIM. L. REV. 321, 328 (2018) (noting that "reasonable judges can disagree over whether sexting images and morphed images are entitled to First Amendment protection").

[53] *See id.* at 327 ("Courts have decided that morphed computer images also qualify as child pornography and are not entitled to First Amendment protection."); *see, e.g.*, United States v. Hotaling, 634 F.3d 725, 730 (2d Cir. 2011) (holding that the defendant's possession of fake or synthetic images qualified as child pornography because they were "not mere records of the defendant's fantasies, but child pornography that implicate[d] actual minors and [was] primed for entry into the distribution chain"); Doe v. Boland, 630 F.3d 491, 497 (6th Cir. 2011) ("Once Boland modified the images of the minors, he crossed the line between possessing lawful images and violating the statute.").

[54] Chesney & Citron, *supra* note 6, at 1802 ("In certain jurisdictions, creators of deep fakes could also face charges for criminal defamation if they posted videos knowing that they were fake or if they were reckless as to their truth or falsity.").

[55] *See supra* note 7 and accompanying text.

realistically saying or doing something they didn't."[56] Technology like the Sweetie avatar, a computer animated child used to catch online predators, has also caused an explosion in "the practice of producing AI-assisted fake porn."[57] Sweetie is a virtual AI run by an international organization attempting to unmask sexual predators, as discussed further in the following Section, but Sweetie's technology is available to a host of other parties, who often do not have the same altruistic concerns that Sweetie's creators do.[58] Deepfake videos are similar to Sweetie in that they depict a "false" child on a screen, but due to technological developments, deepfake technology is available to even more people who lack the good intentions that drove Sweetie's creation, including to many people who use deepfakes maliciously.[59] In this scenario, Sweetie demonstrates a flipside to deepfake technology: while the actual Sweetie avatar is used to find predators, the very same technology can be used to superimpose a child's face onto another person's body, or vice-versa.

Deepfakes are particularly concerning in the context of child sexual abuse because they can "be used to produce new online child sexual abuse material from already existing material."[60] Theoretically, creators of deepfake materials could create more images of children being abused, or even "produce material using images of children who have not been subjected to actual sexual assault."[61] Because deepfake technology can superimpose a person's face onto another's body in a video, a creator could take images of a

[56] *See* Benjamin Goggin, *From Porn to 'Game of Thrones': How Deepfakes and Realistic-Looking Fake Videos Hit It Big*, BUS. INSIDER (June 23, 2019, 10:45 AM), https://www.businessinsider.com/deepfakes-explained-the-rise-of-fake-realistic-videos-online-2019-6 ("Deepfakes also have the potential to differ in quality from previous efforts to superimpose faces onto other bodies. A good deepfake, created by AI that has been trained on hours of footage, has been specifically generated for its context, with seamless mouth and head movements and appropriate coloration.").

[57] Cole, *supra* note 13.

[58] *See id.* (discussing the prominence of Reddit deepfake communities that are dedicated to creating and sharing successful pornographic deepfake videos of celebrities and other women).

[59] Christian Berg, *The Main Challenge Is Victim Identification*, NETCLEAN (May 14, 2019), https://www.netclean.com/2019/05/14/christian-berg-the-main-challenge-is-victim-identification/.

[60] *Id.*

[61] *Id.*

child from any online site—Facebook, for example—and put that child's face onto an image or video depicting another child's sexual abuse.[62] Deepfake technology can even be used to create an image of a completely fake or unreal person.[63] The emergence of deepfakes within child pornography demonstrates that while the technology used to fight online child exploitation can be effective and useful, that same technology can sometimes be used to proliferate exploitation.[64]

The ability to hold perpetrators legally accountable for abusing deepfake technology currently falls short. As Professor Danielle Citron put it,

> No federal criminal law covers the practice, though a smattering of state statutes might apply. . . . The most-recognized privacy torts—intrusion on seclusion and public disclosure of private fact—provide no redress for deep-fake sex videos. Creation and dissemination of videos generated from publicly available images would not amount to an intrusion on seclusion since no private space or activity is intruded upon. Nor would deep-fake sex videos amount to a public disclosure of private fact since no truthful, private facts are revealed.[65]

The notion that the law is well behind the technology is an unfortunate, yet accurate, identification of the problem posed by

[62] *Id.*

[63] The creation of material depicting entirely invented people poses questions regarding expenditure of resources on tracking down supposed victims in images that are entirely fake. For an example of a website that offers images of "people" who are not real, see THIS PERSON DOES NOT EXIST, https://www.thispersondoesnotexist.com/ (last visited Dec. 16, 2021).

[64] For further discussion on the potential dangers of deepfake technology on child sexual abuse, see Hosting Journalist Ed. Team, *Ransomware, DDoS Attacks and Child Abuse Among Key Cybercrime Threats, Says Europol*, HOSTING JOURNALIST, (Oct. 11, 2019), https://hostingjournalist.com/cybersecurity/ransomware-ddos-attacks-and-child-abuse-among-key-cybercrime-threats-says-europol/.

[65] Citron, *supra* note 16, at 1939. Although it is outside the scope of this Note, Professor Citron also raises a convincing argument for civil tort remedies for victims whose likenesses appear in deepfake videos and other violations of sexual privacy. *Id.* at 1949.

deepfake technology.[66] Many deepfake videos are harmful, making the technology generally unfit for general public use. The sexual abuse and cruelty that the majority of deepfake videos enable are testimony to that alone.[67] The law surrounding deepfakes is, for the most part, nonexistent.[68] Deepfakes need to be regulated, monitored, and controlled.[69] With the current scramble to address deepfakes in the law will come a wide array of new regulations. So far, only three states have passed four general deepfake laws.[70] Of those laws, two concern deepfake technology in electoral interference,[71] while the other two address synthetic pornography.[72] In 2019, Maryland enacted a criminal law that encapsulates child pornography deepfakes specifically.[73] The statute "expanded the possession of child pornography . . . to include computer-generated images that are indistinguishable from an actual child under the

[66] *See* Aasha Shaik, *Deepfake Pornography: Beyond Defamation Law*, YALE CYBER LEADERSHIP F. (July 20, 2021), https://cyber.forum.yale.edu/s/Deepfake-Pornography-Beyond-Defamation-Law.pdf. ("Deepfakes are yet another example of technology growing exponentially faster than our laws, leaving people already at greater risk of harm without legal protection.").

[67] *Cf.* Robert Chesney, Danielle Citron & Hany Farid, *All's Clear for Deepfakes: Think Again*, LAWFARE (May 11, 2020, 4:19 PM), https://www.lawfareblog.com/alls-clear-deepfakes-think-again ("[O]f the approximately 15,000 deepfake videos appearing online, 96 percent involve deepfake sex videos; and 99 percent of those involve women's faces being inserted into porn without consent.").

[68] A close but insufficient parallel is nonconsensual porn law. Nonconsensual porn laws often cannot apply to deepfake videos given the lack of body being exposed. Even if a real person's face is depicted in the media, their body is not shown in these images or videos, and their "victory in criminal court thus depends on the [p]roducers' intent and judges' interpretations of key phrases like 'privacy,' 'intimate areas,' and 'depiction.'" Douglas Harris, *Deepfakes: False Pornography Is Here and The Law Cannot Protect You*, 17 DUKE L. & TECH. REV. 99, 123 (2019).

[69] Deepfake videos have the potential to alter our way of life in a number of different ways. Not only could they potentially wreak havoc as tools to create realistic revenge porn, but they also "could be used to create fake videos of politicians accepting bribes, soldiers committing war crimes, presidential candidates engaging in criminal behavior, and emergency officials announcing an impending terrorist attack." Rebecca A. Delfino, *Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn's Next Tragic Act*, 88 FORDHAM L. REV. 887, 894 (2019).

[70] VA. CODE ANN. § 18.2-386.2 (West 2019); CAL. ELEC. CODE § 20010 (West 2020); CAL. CIV. CODE § 1708.86 (West 2020); TEX. ELEC. CODE ANN. § 255.004 (West 2019).

[71] *See* CAL. ELEC. CODE § 20010 (West 2019); TEX. ELEC. CODE ANN. § 255.004 (West 2019).

[72] *See* VA. CODE ANN. § 18.2-386.2 (West 2019); CAL. CIV. CODE § 1708.86 (West 2020).

[73] MD. CODE ANN., CRIM. LAW § 11-208 (West 2021).

age of sixteen."[74] Other states are rushing to introduce new laws that address deepfakes.[75] Additionally, a number of federal bills to regulate deepfakes are currently on the table.[76] The surge in new laws will also need to account for the creation of deepfaked child pornography, not just political deepfakes and deepfaked pornography depicting adults.

III. ANALYSIS

The potential for abuse due to the current lack of redress and regulation in the law is concerning. Technological advances in online abuse must be heavily monitored and may be put to good use. But technology on the side of law enforcement and other government agencies is useless without laws to guide the criminalization of online exploitation. New laws addressing deepfakes must fully account for crimes concerning sexual exploitation, abuse, and violence. The convergence of child pornography laws and the ongoing development of emerging deepfake law is critical to the fight against online child abuse. To properly address online child exploitation, developing deepfake law needs to intertwine with existing child pornography law. This Note identifies some mechanisms that emerging laws should include to ensure that deepfakes and synthetic AI technology are both sufficiently regulated and able to serve as tools to fight, rather than enable, the problem of online child pornography.

[74] *In re* S.K., 215 A.3d 300, 315 n.22 (Md. 2019).

[75] Massachusetts and New York have introduced bills in their respective legislatures to address deepfakes: the Massachusetts bill would make it illegal to use a deepfake in conjunction with other crimes, and New York's bill concerns an individual's right to their digital likeness. *See* David Ruiz, *Deepfakes Laws and Proposals Flood US*, MALWAREBYTES LABS (Jan. 23, 2020), https://blog.malwarebytes.com/artificial-intelligence/2020/01/deepfakes-laws-and-proposals-flood-us/.

[76] Four federal bills of note: Identifying Outputs of Generative Adversarial Networks Act, H.R. 4355, 116th Cong. (1st Sess. 2019); Deepfakes Report Act of 2019, H.R. 3600, 116th Cong. (1st Sess. 2019); S. 1348, 116th Cong. (1st Sess. 2019); Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability (DEEP FAKES) Act, H.R. 3230, 116th Cong. (1st Sess. 2019).

A. SUGGESTED LEGAL FRAMEWORK TO ADDRESS DEEPFAKES IN THE CONTEXT OF CHILD PORNOGRAPHY

Victims of pornographic deepfakes depicting children need and deserve legal frameworks for holding the creators, distributors, and consumers of this material liable. Additionally, it is important to police and remove explicit material depicting children from the Internet. Given the strong federal approach to child pornography law,[77] and the piecemeal compilation of state-level deepfake statutes,[78] it makes sense for the legal framework surrounding deepfaked child pornography to exist mainly on the federal level.[79]

A marriage between developing deepfake laws and existing child pornography laws should lead to a legal framework that criminalizes pornographic deepfakes depicting children. Virginia serves as a model,[80] both domestically and internationally. Virginia criminalized the sharing and creation of deepfakes, setting an example for changes to be made on the federal level.[81] In doing so, Virginia amended its nonconsensual porn laws "to include realistic fake videos and photos, including computer-generated 'deepfakes.'"[82] While the Virginia statute directly applies to

[77] Given the online nature of child pornography currently, it is difficult to imagine a situation in which federal jurisdiction would not attach. *See supra* Section II.A.

[78] *See supra* note 75; *see also* Lvxiao Chen, *Deepfake is Here. What Should We Do?*, JIPEL BLOG (Feb. 14, 2020), https://blog.jipel.law.nyu.edu/2020/02/deepfake-is-here-what-should-we-do/ (discussing how federal regulation is "a more promising solution" to the problem posed by deepfakes because state law addressing deepfakes would face jurisdictional issues, incomplete and narrow regulation, and limited state resources).

[79] This topic has already been identified an important area for intervention by legal scholars in discussion surrounding the differing benefits of civil and criminal laws in response to deepfakes. *See, e.g.*, Delfino, *supra* note 69, at 902 (discussing the promise of civil remedies as one legal response to deepfakes but indicating that criminal law would help in ways that civil laws would not).

[80] *See* Adi Robertson, *Virginia's 'Revenge Porn' Laws Now Officially Cover Deepfakes*, VERGE (July 1, 2019, 4:49 PM), https://www.theverge.com/2019/7/1/20677800/virginia-revenge-porn-deepfakes-nonconsensual-photos-videos-ban-goes-into-effect (explaining that the Virginian amendment now covers "deepfakes").

[81] *See Virginia Bans 'Deepfakes' and 'Deepnudes' Pornography*, BBC (July 2, 2019), https://www.bbc.com/news/technology-48839758 ("Virginia has become one of the first places to outlaw the sharing of computer-generated pornography known as deepfakes.").

[82] Robertson, *supra* note 80; VA. CODE ANN. § 18.2-386.2 (West 2019).

nonconsensual pornography (known as revenge porn),[83] the differences between that and criminalization of computer-generated images of children are not so substantial that state practices cannot be used as a touchstone for amendments to federal law to include deepfakes. Amendments to 18 U.S.C. §§ 2251–2260A, the chapter on child pornography and sexual abuse, could be amended to mirror language from Virginia's statute to ensure that deepfaked videos fall under the federal statutory scheme for criminalization of child pornography.[84]

This is not to say that federal amendments should be the sole response to deepfakes, but this manner of intervention is critical.[85] There are already bills in circulation to address deepfake regulation on the federal level.[86] But those bills do not address deepfakes, nor their potential for misuse by way of creating child pornography.[87] The hole left by the existing bills and laws surrounding deepfakes, and child pornography specifically, can be filled by amending

[83] *See* Abrar Al-Heeti, *Sharing Deepfake Revenge Porn Is Now a Crime in Virginia*, CNET (July 1, 2019, 3:08 PM), https://www.cnet.com/news/sharing-deepfake-revenge-porn-is-now-a-misdemeanor-in-virginia/ (discussing the text of the Virginia law and what the new law could represent moving forward).

[84] Virginia's new statutory language added "videographic or still image[s] created by any means whatsoever." VA. CODE ANN. § 18.2-386.2 (West 2019).

[85] For a detailed proposal concerning "a new federal criminal statute to regulate the creation and distribution of pornographic deepfakes," see Delfino, *supra* note 69, at 928.

[86] *See id.* at 927–28 (detailing five reasons for federal law to address deepfakes: (1) it "would provide a strong and effective disincentive to their creation and distribution"; (2) "pornographic deepfakes is a crime lacking jurisdictional boundaries"; (3) "state laws are constrained by section 230 of the [Communications Decency Act], which impedes state actions against website operators who host nonconsensual pornography"; (4) "criminalizing pornographic deepfakes as a federal crime brings to bear the greater resources of the federal government"; and (5) such criminalization "adds gravitas to the situation and shines a spotlight on its harms").

[87] The bills do not even do a good job addressing the most common abuse currently being perpetuated by deepfake technology because they erroneously identify political deepfakes as the big threat when, as of September 2019, 96% of deepfakes and synthetic media were nonconsensual sexual images. HENRY AJDER, GIORGIO PATRINI, FRANCESCO CAVALLI & LAURENCE CULLEN, THE STATE OF DEEPFAKES: LANDSCAPE, THREATS, AND IMPACT 1 (2019). Only half of the current state laws address pornographic deepfakes; the others target political deepfakes. Korey Clark, *'Deepfakes' Emerging Issue in State Legislatures*, LEXISNEXIS, https://www.lexisnexis.com/en-us/products/state-net/news/2021/06/04/Deepfakes-Emerging-Issue-in-State-Legislatures.page (last visited Feb. 2, 2022).

existing child pornography laws to encompass synthetically created or altered sexual images of children.

B. EXCEPTIONS TO THE SUGGESTED LEGAL FRAMEWORK

Accompanying the need for a legal framework that strictly regulates and outlaws deepfaked child pornography is a need for limited exceptions to these laws. These exceptions should enable government agencies to use this technology as a tool to fight the spread of child pornography online. Deepfake technology is an invaluable asset for online stings that shut down abuse,[88] as well as for rehabilitative services that keep offenders from victimizing real children.

1. Allow the Use of Synthetic Pornography Materials as Tools for Law Enforcement in Online Stings. One suggested exception to the criminalization of synthetic child pornography lies in the use of synthetic materials for law enforcement to use in online stings. As mentioned previously, government agencies like the FBI aggressively pursue and remove online sites and users who share and create child pornography.[89] A recent example is the FBI's Operation Pacifier, which resulted in the complete takedown of Playpen, the largest online child pornography website.[90] Operation Pacifier, an online sting operation, utilized what is known as a "honeypot" attack strategy to target and eliminate Playpen.[91] Originally designed to catch hackers and scammers, a honeypot "is

[88] *See* Graeme R. Newman, OFF. OF CMTY. ORIENTED POLICING SERVS., U.S. DEP'T OF JUST., STING OPERATIONS 3 (2007) https://cops.usdoj.gov/RIC/Publications/cops-p134-pub.pdf (defining general sting operations as a police-created opportunity to commit a crime targeted towards a particular offender or group that is monitored by an undercover officer and culminates in an arrest).

[89] *See, e.g.*, United States v. Adams, No. 6:16-cr-11-Orl-40GJK, 2016 WL 4212079, at *1 (M.D. Fla. 2016) (concerning one of many prosecutions of child pornography offenders caught in the FBI takedown of Playpen).

[90] *See* Press Release, U.S. Dep't of Justice, Florida Man Sentenced to Prison for Engaging in a Child Exploitation Enterprise (May 1, 2017), https://www.justice.gov/opa/pr/florida-man-sentenced-prison-engaging-child-exploitation-enterprise (reporting the arrest and sentencing of the creator of Playpen and the additional 350 U.S.-based arrests following the operation); Gregory*,* *supra* note 19, at 261 ("[C]elebrated by some, maligned by others—is the FBI's Operation Pacifier, the total takedown of the world's largest child pornography website, Playpen.").

[91] Gregory*, supra* note 19, at 261.

a designated area within a computer system or network that has been designed specifically with the expectation that it will be attacked by unauthorized users, whether internal or external to the organization operating the honeypot."[92] A honeypot that specifically targets perpetrators of child pornography is a site that "purport[s] to contain child pornography but in fact [is] set up by police and designed to capture the IP address or credit card details of visitors trying to download images."[93] Honeypots lure perpetrators in and then provide law enforcement with information to capture those who take the bait.[94] After a honeypot attracts a target, a code deploys to grab the user's IP address and identifiers.[95]

Another well-known online attack strategy is an animated internet avatar called "Sweetie." Sweetie is a virtual girl animation used to engage with predators first in chat rooms and then in video calls with the goal of identifying online predators who sought to take advantage of an underage girl.[96] Much like deepfakes, Sweetie is an AI technology that purports to, but does not, represent a real person. Once enough data on identifying information and the details of exploitation requested by predators is collected, Sweetie's operators submit the data to Interpol.[97] Sweetie remains an example of a technological innovation successfully working to catch and take down online predators,[98] potentially serving as a model for future government-operated stings.

The above-mentioned examples indicate that technology that would be dangerous in the wrong hands can assist in monitoring and stopping online child sexual exploitation. Deepfake technology is not much different from Sweetie and can be used with honeypot

[92] Ian Walden & Anne Flanagan, *Honeypots: A Sticky Legal Landscape?*, 29 RUTGERS COMPUT. & TECH. L.J. 317, 318 (2003).

[93] RICHARD WORTLEY & STEPHEN SMALLBONE, INTERNET CHILD PORNOGRAPHY: CAUSES, INVESTIGATION, AND PREVENTION 60 (Graeme R. Newman ed., 2012).

[94] Gregory, *supra* note 19, at 279.

[95] *Id.*

[96] *See id.* at 279–80; *see also Sweetie, Our Weapon Against Child Webcam Sex,* TERRE DES HOMMES, https://www.terredeshommes.nl/en/programs/sweetie (last visited Oct. 28, 2021) ("Sweetie was a ten-year-old virtual Filipino girl. . . . When men started talking with her in a sexually suggestive way, she engaged back. All the information from their exchanges got stored and used to warn, track down or even arrest and convict perpetrators.").

[97] Gregory, *supra* note 19, at 280.

[98] *Id.*

technology to detect, track, and eventually locate online predators. The use of deepfaked or completely synthetic materials in online stings like Operation Pacifier would be a dramatic improvement in these law enforcement efforts.

This new technology would prevent the ongoing victimization of children, which was a marked criticism of the FBI's Operation Pacifier in *United States v. Anzalone*.[99] After discovering Playpen's IP address on the Tor Network, the FBI hacked Playpen and gained control over the site.[100] To catch predators who were using the site, the FBI left Playpen operational for twelve days, tracking the IP addresses of users who frequented the site during that time.[101] The sting eventually led to the complete shutdown of Playpen and over 135 criminal prosecutions against people who accessed the site.[102] Critics of Operation Playpen argued that the FBI's actions in the sting "amounted to outrageous government conduct that violated . . . due process."[103] In *Anzalone*, the defendant claimed that the FBI itself participated in the distribution and perpetuation of child pornography to every person who frequented Playpen while the FBI controlled it.[104] The court dismissed the defendant's claims that the FBI's conduct violated his due process rights, but the related point raised about the danger of the FBI's conduct in operating a child pornography site remains relevant.[105]

[99] *See* United States v. Anzalone, 923 F.3d 1, 6 (1st Cir. 2019) (dismissing the defendant's claim that the FBI's decision to operate Playpen for two weeks was outrageous government conduct that violated his due process rights); Maureen Weidman, Comment, *Jurisdiction, the Internet, and the Good Faith Exception: Controversy over the Government's Use of Network Investigative Techniques*, 122 DICKINSON L. REV. 967, 973 (2018) (describing the background of Operation Pacifier and mentioning that "the government chose to keep the website operational from February 20 to March 4, 2015").

[100] *See* Weidman, *supra* note 99, at 972 ("Based on the information provided by the foreign law enforcement agency and information from its own investigation, the FBI determined the location of the website's operator and . . . then took control of the website's server . . . .").

[101] *See id.* at 973 ("Due to difficulties in obtaining criminals' IP addresses, the government chose to keep the website operational from February 20 to March 4, 2015.").

[102] *See* Kaleigh E. Aucoin, *The Spider's Parlor: Government Malware on the Dark Web*, 69 HASTINGS L.J. 1433, 1449 ("Operation Pacifier . . . led to an estimated collection of IP addresses ranging somewhere in the thousands, at least 350 domestic arrests, and over 135 cases nationwide . . . ." (footnote omitted)).

[103] *Anzalone*, 923 F.3d at 5.

[104] *Id.*

[105] *Id.* at 6.

More specifically, this claim raises a serious criticism of online child pornography stings, which could be *completely circumvented* by the use of synthetic materials in future takedowns of child pornography offenders. Rather than contributing to the ongoing victimization of sexually abused children in the process of trying to help them, government agencies can instead use images that do not victimize real children in the process of catching predators. The use of deepfake and synthetic AI technology in cases like Operation Pacifier would help further government and victim interests simultaneously: it would cut off systems of abuse while also continuing to find, catch, and stop online predators and child pornography sites.

2. Allow the Use of Synthetic Child Pornography as a Rehabilitative Tool for Predators. AI technology also stands to be useful in areas other than government attack strategies. In addition to serving as a tool to catch and stop the spread of online child pornography, this technology could also serve as a rehabilitative tool to prevent people who seek out and engage with child pornography from victimizing real children. This is not an entirely new concept. The original Sweetie program is being further developed into "Sweetie 2.0," which will function not only as a way to catch predators but also as an attempt to divert them from child pornography.[106] Sweetie 2.0 operates as a form of intervention for people who seek out explicit content of children to provide them with access to resources to help divert them from exploiting children and instead help them focus on controlling those urges,[107]

[106] *See* Tilburg University, *Sweetie 2.0 Software Tackles Online Child Sex Abuse*, PHYS.ORG (Mar. 15, 2015), https://phys.org/news/2015-03-sweetie-software-tackles-online-child.html ("[Sweetie 2.0 will] continue the Sweetie [1.0] project. . . . [A]n advanced software system to combat webcam sex with children across the world will be developed. It will help law enforcement agencies to recognize and/or deter millions of potential perpetrators.").

[107] *Id.* ("Pedophiles who operate with different e-mail accounts can be traced by means of the 'catch-recatch method.' It works as follows: person X is spotted in chatroom Y and is given a warning with information on the crimes committed, their consequences under criminal law, and tips on how to get help. If X later visits chatroom Z with a different e-mail address, he will still be recognized as X. A second and last warning will be issued. If X is caught a third time, the information will be passed on to the police.").

demonstrating the potential for AI programs to serve as deterrence tools in combatting child pornography.[108]

Current understandings of online predators' motivations also support the argument that synthetic AI has a deterrent role. There have been many studies of pedophilia in attempts to better understand how to prevent the victimization of children.[109] Pedophilia is technically a mental disorder; those who have it are sexually interested in prepubescent children for an extended period of time."[110] Many people who are sexually or romantically interested in children never act on their urges,[111] but the boom in online photos and videos depicting images of child sexual abuse indicates that

[108] The Sweetie 3.0 project, also known as #Sweetie 24/7, began in 2019 as an even more integrated form of identifying suspected online predators. *See Sweetie, Our Weapon Against Child Webcam Sex*, TERRE DES HOMMES https://www.terredeshommes.nl/en/programs/sweetie (last visited Mar. 24, 2022) (explaining that Sweetie 3.0 enables "track[ing of] suspects across all kinds of social media and online platforms" and for "undercover employees [to] use it to make contact with intermediaries who provide children for sexual abuse," allowing operators to "locate where the children are being exploited").

[109] *See, e.g.*, Gillian Tenbergen et al., *The Neurobiology and Psychology of Pedophilia: Recent Advances and Challenges*, 9 FRONTIERS IN HUM. NEUROSCIENCE 1, 1 (2015) (discussing neuroscientific inquiries into pedophilia); Kristen Jordan, Tamara Sheila Nadine Wild, Peter Fromberger, Isabel Müller & Jürgen Leo Müller, *Are There Any Biomarkers for Pedophilia and Sexual Child Abuse? A Review*, 10 FRONTIERS IN HUM. NEUROSCIENCE 1, 3, 14 (2020) (summarizing a recent review of studies and research on pedophilia conducted to learn more about preventing the victimization of children, concluding that "further work remains to be done").

[110] Ryan C.W. Hall & Richard C.W. Hall, *A Profile of Pedophilia: Definition, Characteristics of Offenders, Recidivism, Treatment Outcomes, and Forensic Outcomes*, 82 MAYO CLINIC PROC. 457, 457 (2007); *see also* Benedict Carey, *Preying on Children: The Emerging Psychology of Pedophiles*, N.Y. TIMES (Sept. 29, 2019), https://www.nytimes.com/2019/09/29/us/pedophiles-online-sex-abuse.html (defining those with pedophilia as those who "fantasize[ ] about, [are] sexually aroused by, or experience[ ] sexual urges toward prepubescent children (generally <13 years) for a period of at least 6 months").

[111] *See* Carey, *supra* note 110 ("There's a subgroup [of pedophiles] out there, . . . and they are quite convinced that they do not want real-life sex with children." (quoting Dr. Fred Berlin)); Catherine Burns, *The Young Paedophiles Who Say They Don't Abuse Children*, BBC (Sept. 11, 2017), https://www.bbc.com/news/uk-41213657 (discussing "anti-contact" pedophiles who do not want to victimize any children).

there are plenty who do seek abusive content.[112] In doing so, they victimize the people depicted in that media with every click, download, and message sent.[113] There are a number of different treatment methods for pedophilia.[114] Treatment of pedophiles is extremely challenging, though, partially because the individuals participating in treatment must be willing to engage in treatment and avoid offending again.[115] Should individuals reoffend while engaging in treatment, the deterrent quality of the treatment is effectively destroyed, as each new offense further victimizes a child. Additionally, "few pedophiles voluntarily seek treatment," which significantly reduces the effectiveness of existing treatment methods in terms of stopping child sexual abuse and exploitation, both online and offline.[116] The majority of these treatments focus "on stopping further offenses against children rather than altering the pedophile's sexual orientation towards children."[117] While government initiatives using deepfakes would work as a way of identifying pedophiles and enforcing existing child pornography laws,[118] implementing deepfakes as a rehabilitation tool after finding offenders could also play a role in stopping the circulation of real images of children.

A significant gap exists between current interventions and treatment methods for people interested in child pornography and actually stopping people from committing offenses that sexually abuse and harm children. The introduction of AI technology to deter people with pedophilic tendencies could serve to attract more people who are afflicted with pedophilia to come forward to seek treatment,

[112] *See* NETCLEAN, COVID-19 IMPACT 2020: A REPORT ABOUT CHILD SEXUAL ABUSE CRIME 22 (2020), https://www.netclean.com/wp-content/uploads/2021/01/NetCleanReport_COVID19Impact2020_spreads1200-1.pdf (observing that "online child sexual abuse activity has increased," specifically during the COVID-19 pandemic).

[113] *See Child Pornography*, *supra* note 17 ("Child pornography is a form of child sexual exploitation, and each image graphically memorializes the sexual abuse of that child.").

[114] *See* Hall & Hall, *supra* note 110, at 466 (listing known treatments for pedophilia).

[115] *See id.* at 467 ("The published rates of recidivism are in the range of 10% to 50% for pedophiles . . . .").

[116] *See id.* at 460 ("It is difficult to estimate the true prevalence of pedophilia because few pedophiles voluntarily seek treatment and because most of the available data are based on individuals who have become involved with the legal system.").

[117] *Id.* at 465.

[118] *See supra* Section III.A.*1*.

as it is a much less invasive form of intervention and would prevent the exploitation of children by depicting images of people who do not exist.

IV. BARRIERS TO IMPLEMENTATION

While the federal criminalization of pornographic deepfakes seems like a natural extension of existing law, First Amendment concerns must be overcome before this Note's proposal could be enacted.[119] A successful First Amendment challenge was raised in *Ashcroft v. Free Speech Coalition*.[120] *Ashcroft* concerned a pre-enforcement First Amendment challenge to certain provisions of the Child Pornography Prevention Act of 1996—specifically, the Act's definition of child pornography: (1) "'visual depiction[s], including any . . . computer-generated image or picture,' that 'is or appears to be, of a minor engaging in sexually explicit conduct,'"[121] and (2) "any sexually explicit image . . . 'that conveys the impression' it depicts 'a minor engaging in sexually explicit content.'"[122] These provisions hinged on the concept of virtual (or synthetic) child pornography as something that fell within child pornography, and therefore, within the Act's content-based ban on such material.[123] The U.S. Supreme Court found that the Act banned speech that "records no crime and creates no victims by its production,"[124] and determined that the illegality of child pornography stemmed from the harm its creation, distribution, and consumption inflicts on the children implicated in such issues.[125] Additionally, the Court stated that the Act was not specific enough in its language, and was thus "overbroad and unconstitutional."[126]

After *Ashcroft*, Congress enacted a law, called the PROTECT Act of 2003, to criminalize pornographic images of children, whether or

[119] *See supra* note 27 and accompanying text.

[120] 535 U.S. 234 (2002).

[121] *Id.* at 241 (quoting 18 U.S.C. § 2256(8)(D)(1)).

[122] *Id.* (quoting 18 U.S.C. § 2256(8)(D)(2)).

[123] *Id.*

[124] *Id.* at 250.

[125] *Id.* at 250–53 (citing New York v. Ferber 458 U.S. 747, 759 (1982)).

[126] *Id.* at 256.

not the child in the image exists.[127] Congress also amended the definition of child pornography contained in 18 U.S.C. § 2256(8) to include "computer or computer-generated image[s] or picture[s], whether made or produced by electronic, mechanical, or other means, of sexually explicit conduct."[128] While it appears that in the wake of *Ashcroft* the law has done its best to adapt to potential First Amendment issues that can arise from these regulations, those protections are based on the notion that "child pornography always harms the child when it is made, sold, shared, and possessed, and the government's interest in preventing such is of the highest order."[129]

While fake images of children lack the direct harm that images of real children cause, there is enough societal harm present in the unregulated dissemination of these types of images that they should still fall within the scope of government regulation. Theoretically, the creation and dissemination of completely fake child pornography would not inflict the harms associated with images depicting real children because there is no actual child to victimize. People accessing or creating this type of material could couple this notion with the argument that the fake material does not look realistic enough to fall under the statutory language of 18 U.S.C. § 2256(8) and claim that they deserve First Amendment protections. Although this Note argues for government and rehabilitative exceptions to an outright ban on deepfaked child pornography, there is a critical distinction between the argument set forward in this Note and that of the potential predator. Specifically, there is still, in fact, a significant harm perpetuated by anonymous users freely using and accessing this type of content that does not appear in the severely limited exceptions argued herein. The exceptions mentioned in this Note would be limited to narrow circumstances under the control of law enforcement agencies and rehabilitative programs.[130]

[127] *See* 18 U.S.C. § 1466A(a) (criminalizing "a visual depiction of any kind [of a minor], including a drawing, cartoon, sculpture, or painting").

[128] *Id.* § 2256(8).

[129] Gregory, *supra* note 19, at 275.

[130] The rehabilitative programs would be conducted with some form of scrutiny and monitoring system in place to ensure that the technology was being used effectively to actually curtail predators from engaging in their uncontrollable urges.

Moreover, the type of harm that would be caused by widespread, unrestricted access to this type of content would have the opposite effect of the strict regulations and restrictions. The type of harm that widely available access to deepfakes causes comes in two parts: when the viewers believe that the content that they are viewing is real and when the viewers know that it is fake.[131] When the deepfake is believed to be real, the unregulated and unmonitored distribution of it would actively enable and encourage the viewer to continue to seek out content like it in the future.[132] This could lead viewers (unknowingly) to another "fake" image of a nonexistent child or to a real image of a real child; should this happen, the abuse and victimization of children shown in child pornography would merely continue, having been enabled yet again in an anonymous, unmonitored context. Furthermore, when the viewer knows that the media is fake and seeks it out of their own accord as a part of their own exploration into the realm of child pornography, they sexualize and fetishize children in an uncontrolled setting where they could continue to perpetuate other forms of abuse. It is important to flag that the exceptions argued for in this Note would need to be hyper-monitored so as to avoid the harms of child pornography altogether, rather than merely stopping them after they begin.

As indicated by the surge in reports of online exploitation in the past years,[133] there is a serious problem of online child pornography. To allow it to go unregulated and unsupervised would be a disservice to all child pornography victims.

V. CONCLUSION

We are witnessing a new frontier of technological innovation with the development of deepfake technology. While the abuse of synthetic child pornography has yet to contribute to the ongoing child pornography crisis, Congress should take a forward-looking

[131] ENDTAB, DEEPFAKES: A VICTIM RESOURCE GUIDE (2019), https://static1.squarespace.com/static/58b8cb1846c3c4543ab7b863/t/5dcb1ede16bdca031cccb4ff/1573592805299/EndTAB+Deepfake+Victim+Guide+1.0.pdf.

[132] Additionally, if the video contained the face of a real child, the harm caused there would be akin to the harms caused by nonconsensual pornography and cause serious trauma to the victim.

[133] *See supra* note 41 and accompanying text.

approach and amend its child pornography laws before this abuse occurs. The peak of deepfaked child pornography is yet to come. Given our current knowledge about deepfake technology and how poorly prepared our legal framework is to account for it, it is best to seriously consider the implications of this emerging technology now, before it is overwhelmingly in the hands of bad actors. Federal laws criminalizing the use of deepfakes for child pornography purposes should be enacted, with specific exceptions laid out so that law enforcement and government agencies can utilize the flipside of this potentially harmful technology to attempt to alleviate, rather than exacerbate, the ongoing issue of child pornography.