

SAVING FACE: LESSONS FROM THE DMCA FOR COMBATING DEEFAKE PORNOGRAPHY

Jared de Guzman*

ABSTRACT

Deepfake technology can produce a realistic video of someone doing something they never actually did. Most often, that something is nonconsensual pornography. It is not difficult to imagine the emotional, social, and physical harm that comes with having a lifelike depiction of one's sexual privacy shared with the entire world without consent. Artificial intelligence (AI) is making deepfake pornography easier to create and more realistic, augmenting its accessibility and its capacity for harm. Currently, there are no adequate legal remedies for victims who find themselves in deepfake pornography and there is no mechanism deterring bad actors from creating and distributing it. This Note argues for the creation of federal legislation aimed at combatting deepfake pornography distribution. Specifically, it advocates for a legislative model following and modifying the Digital Millennium Copyright Act's immunity and liability framework.

* J.D. expected 2023, Gonzaga University School of Law. I would like to thank Dean Agnieszka McPeak for sparking my interest in AI policy and for her help in developing this Note. I also want to thank my amazing partner Jenna for her constant support. Finally, a big thank you goes out to all the friends I have made at GU Law.

TABLE OF CONTENTS

INTRODUCTION: A PARALLEL REALITY	111
I. BACKGROUND.....	115
A. <i>What Makes Deepfakes So Dangerous?</i>	115
B. <i>The Lack of Adequate Legal Remedies</i>	117
II. BOTTOM-LINE CONTROL, THE POWER OF VICARIOUS LIABILITY, AND THE NEED FOR FEDERAL LEGISLATION.....	120
A. <i>The DMCA as a Legislative Model for Combatting Deepfake Pornography</i>	123
B. <i>Case Study: Viacom Int'l, Inc. v. YouTube, Inc.</i>	126
III. LESSONS FROM THE DMCA'S LEGISLATIVE FRAMEWORK	130
A. <i>Establishing Strong Foundations</i>	130
B. <i>General Legislative Framework</i>	132
C. <i>Obstacles</i>	135
CONCLUSION	136

INTRODUCTION: A PARALLEL REALITY

We watch movies because of postproduction. Without postproduction, there would be no transitions, no special effects, no background music, and no coherent narrative.¹ The raw footage captured by a camera and the cinematic production that makes it to the screen are two very different things. While we enjoy the escapism and entertainment that the film industry provides, we should be aware that our own reality now exists in a postproduction format on the internet. The postproduction reality is a nearly indistinguishable reflection of our own, but with alterations. Sometimes these alterations are obvious and employed with benevolent intention.²

However, most of the time the postproduction reality aims to deceive by passing altered content off as if it were reality.³ Artificial intelligence (AI) technology has become so advanced that software can instantly generate life-like images of people who never existed.⁴ It can also portray existing people in situations that never transpired. Take for example Emma Gonzales: a Parkland Shooting survivor who, in reality, tore up a bullseye target as part of a gun reform campaign video. In the postproduction reality, she instead tore up the United States Constitution, inciting widespread outrage.⁵

This video was the product of deepfake technology. “Deepfake” is the term used for software that uses AI to realistically superimpose a person’s face and likeness onto someone else’s body.⁶ Because of deepfakes, individuals increasingly find themselves in the postproduction reality without their consent. Most

1. Postproduction is the editing stage after raw footage is recorded. It is a critical step in manipulating the raw footage into the story a filmmaker wants to tell. See *Post-Production Everything You Need to Know*, NASHVILLE FILM INST., <https://www.nfi.edu/post-production/> (last visited June 28, 2022).

2. See Joshua Rothkopf, *Deepfake Technology Enters the Documentary World*, N.Y. TIMES (July 1, 2020), <https://www.nytimes.com/2020/07/01/movies/deepfakes-documentary-welcome-to-chechnya.html> (stating documentary makers in Chechnya used deepfake technology to alter and anonymize the faces of persecuted LGBTQ interviewees).

3. See Bobby Chesney & Danielle Citron, *Deepfakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CAL. L. REV. 1753, 1763 (2019).

4. See, e.g., THIS PERSON DOES NOT EXIST, <https://this-person-does-not-exist.com/en> (last visited June 28, 2022) (website uses AI to instantly produce an entirely new person and face rendered in a realistic photo with a single click).

5. See Chesney & Citron, *supra* note 3, at 1755–56.

6. *Id.* at 1757–58; Christa Laser & Eric Goldman, *Deepfakes, Privacy, and Freedom of Speech*, CLEVE. MARSHALL COLL. L. (June 18, 2021), <http://yourwitness.csulaw.org/uncategorized/deepfakes-privacy-and-freedom-of-speech/>.

of the time, these individuals are women depicted in sexually explicit and degrading situations they were never in.⁷

Rana Ayyub is one of these victims. Ayyub is an accomplished investigative journalist and self-described “anti-establishment” Muslim woman.⁸ After publicly speaking out against members of an Indian nationalist party and its defense of child abusers, a fake Twitter account posing as Ayyub began posting inflammatory tweets. Amidst the raging backlash, Ayyub soon discovered she had been realistically superimposed into a porn video that had gone viral with over 40,000 shares.⁹ Although the content of a deepfake video is not real, the profound impact it has on a victim’s life is. Apart from traumatic feelings of embarrassment, fear, and anger,¹⁰ victims can suffer physical harm¹¹ or receive violent threats from the public.¹² Reputational harm from deepfake pornography can also be devastating.¹³

The software that individuals use to create deepfake content is becoming increasingly accessible. With just a few sample photos and some basic software, a malcontent can insert anyone into the postproduction reality as a pornographic subject.¹⁴ Deepfake technology blurs the line between reality and the

7. See Karen Hao, *Deepfake Porn Is Ruining Women’s Lives*, M.I.T. TECH. REV. (Feb. 12, 2021) [hereinafter Hao I], <https://www.technologyreview.com/2021/02/12/1018222/deepfake-revenge-porn-coming-ban/> (highlighting that 90% of deepfake content is pornographic).

8. Raya Ayyub, *I Was the Victim of a Deepfake Pornography Plot Intended to Silence Me*, HUFFPOST (Nov. 21, 2018), https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316.

9. *Id.*

10. Danielle Keats Citron, *Sexual Privacy*, 128 YALE L.J. 1870, 1875 (2019) (stating sexual privacy invasions cause women to “feel humiliated and ashamed”).

11. See, e.g., Ayyub, *supra* note 8 (“I was sent to the hospital with heart palpitations and anxiety, the doctor gave me medicine. But I was vomiting, my blood pressure shot up, my body had reacted so violently to the stress.”).

12. See, e.g., *id.* (likening the violent threats and nonstop harassment Ayyub experienced to being “exposed to a lynch mob”).

13. See Citron, *supra* note 10, at 1875 (explaining that women “suffer stigmatization in the wake of sexual privacy investigations” and “lose their jobs and have difficulty finding new ones”); see also Danielle S. Van Lier, *The People v. Deepfakes*, L.A. LAW., May 2020, at 16, 18 (stating 80% of employers do online searches of potential job candidates and 70% of employers rejected candidates based on those results); see, e.g., Ayyub, *supra* note 8 (documenting how Ayyub was “devastat[ed]” to the point she “just couldn’t show [her] face” and “just couldn’t see through the humiliation”).

14. See Hao I, *supra* note 7 (discussing how Helen Mort was the victim of a targeted campaign to deepfake her into violent pornography using just a small collection of her photos from social media); see also Citron, *supra* note 10, at 1923 (discussing how an eighteen-year-old Australian girl’s social media photos were used to deepfake her into a series of graphic oral sex videos).

postproduction counterpart that exists online. These two realities are likely to merge even further as society spends more time on the internet.¹⁵ Accountability for perpetrators and a remedy for victims is desperately needed, but nothing adequate currently exists. This is another example of the lethargy of the law as advanced technology speeds forward.

Deepfakes present a major problem, but a familiar pattern. Deepfake technology has rapidly proliferated beyond the law's ability to rectify its social harms. Similarly, at the dawn of the Digital Age, copyright infringement-facilitating technology proliferated beyond the law's ability to rectify its commercial harms.¹⁶ To protect and preserve the rights of copyright holders, Congress enacted several pieces of copyright legislation, including the Digital Millennium Copyright Act (DMCA) in 1998.¹⁷ The DMCA contains two components particularly relevant to the deepfake problem: the Safe Harbor provision¹⁸ and the Notice and Takedown procedures.¹⁹ The Safe Harbor provision provides immunity to digital "service providers" hosting third-party material from incurring liability for contributory copyright infringement if the provider follows the Notice and Takedown procedures.²⁰ The procedures require a service provider to designate an agent to receive statutorily specified infringement notices and expeditiously remove any infringing material.²¹

Together, these provisions create a deterrent framework driven by a host's incentive to avoid vicarious liability: the ever-looming threat of copyright litigation for third-party infringement on the platform incentivizes service providers to qualify for Safe Harbor, causing providers to follow the Notice and Takedown procedures.²² As a result, service providers substantially curtailed copyright infringement instead of facilitating it and allowed the internet to be "largely self-policing due to the interaction between those who send takedown notices and the hosts that receive them."²³ Credited for creating the internet as we know it

15. In 2021 alone, Meta invested \$10 billion in "the metaverse," a virtual reality predicted to be "the next generation of the Internet." Mike Isaac, *Meta Spent \$10 Billion on the Metaverse in 2021, Dragging Down Profit*, N.Y. TIMES (Feb. 2, 2022), <https://www.nytimes.com/2022/02/02/technology/meta-facebook-earnings-metaverse.html>.

16. PETER S. MENELL, MARK A. LEMLEY, ROBERT P. MERGES, & SHYAMKRISHNA BALGANESH, *INTELLECTUAL PROPERTY IN THE NEW TECHNOLOGICAL AGE: 2021*, VOL. II COPYRIGHTS, TRADEMARKS & STATE IP PROTECTIONS 516 (2021).

17. *Id.*; Jeffrey Cobia, Note, *The Digital Millennium Copyright Act Takedown Notice Procedure: Misuses, Abuses, and Shortcomings of the Process*, 10 MINN. J.L. SCI. & TECH. 387, 388 (2009).

18. 17 U.S.C. § 512I(1).

19. *Id.* § 512(c)(2)–(3).

20. *See id.* § 512(k)(1)(A).

21. *Id.* § 512(c)(2)–(3).

22. *See Cobia, supra* note 17, at 388.

23. *Id.*

today,²⁴ the DMCA's success suggests similar federal legislation can be a powerful legal tool against a new social harm caused by rapidly changing technology.

As the postproduction reality continues to merge with our own, we must demand federal action to combat deepfake pornography. Since nearly all fifty states have some sort of criminal statute prohibiting "revenge porn,"²⁵ this signals that the United States values protecting the sexual privacy of individuals. Like revenge porn, deepfake pornography is nonconsensual pornography that violates an individual's privacy, exploits their identity, and reduces them to an object.²⁶ To address this growing issue, this Note argues that Congress should enact federal legislation that not only provides a cause of action to victims of deepfake pornography, but imposes vicarious liability on service providers for hosting the content. However, imposing liability is not enough to stop the harm. This Note proposes that federal legislation should incorporate a similar deterrence framework as the Safe Harbor and Notice and Takedown provisions of the Digital Millennium Copyright Act to best incentivize hosts to quickly remove harmful content. Just as the DMCA did, this Note argues similar federal legislation would enable service providers to substantially curtail the harmful content instead of facilitating it.

To advance this argument, Part I of this Note first provides context on deepfake pornography. Section A explains the underlying technology behind deepfake pornography and illustrates why it is a growing social danger that requires a remedy. Section B analyzes the current legal remedies available and evaluates current strategies for how the law should handle deepfake pornography. Next, Part II demonstrates why the DMCA Notice and Takedown provisions provide an exemplary model for drafting federal legislation to combat deepfake pornography. Part II examines the theory of bottom-line control and the role of vicarious liability in driving it, and why federal legislation is the proper path for implementing vicarious liability of service providers for hosting deepfake pornography. Section A reviews the language of the DMCA to explain its utility as a legislative model. Section B examines *Viacom International, Inc. v. YouTube*,

24. See Katherine Trendacosta, *Reevaluating the DMCA 22 Years Later: Let's Think of the Users*, ELEC. FRONTIER FOUND. (Feb. 12, 2020), <https://www EFF.ORG/deeplinks/2020/02/reevaluating-dmca-22-years-later-lets-think-users> ("The Digital Millennium Copyright Act (DMCA) is one of the most important laws affecting the Internet and technology. Without the DMCA's safe harbors from crippling copyright liability, many of the services on which we rely, big and small, commercial and noncommercial, would not exist. That means Youtube, but also Wikipedia, Etsy, and your neighborhood blog. At the same time, the DMCA has encouraged private censorship and hampered privacy, security, and competition.").

25. Hao I, *supra* note 7; see, e.g., ALA. CODE § 13A-6-240 (2022); CAL. PENAL CODE § 647(j)(4) (West 2022); WASH. REV. CODE § 4.24.795 (2022). "Revenge porn" is the popular term to describe when a consensually distributed sexually explicit image is redistributed without the subject's further consent. See Citron, *supra* note 10, at 1918.

26. Citron, *supra* note 10, at 1921.

Inc.,²⁷ a landmark case that tested the limits of the DMCA's enforceability and spurred institutional change for distributors of copyrighted material.²⁸ Finally, Part III uses the lessons from *Viacom* and other policy sources to argue what successful federal deepfake pornography legislation would need and what obstacles it would need to overcome.

I. BACKGROUND

A. *What Makes Deepfakes So Dangerous?*

There are two principal issues necessitating the creation of federal deepfake pornography legislation: hyperrealism and widespread accessibility. Both are made possible by underlying "deep learning"²⁹ technology called Generative Adversarial Networks (GANs).³⁰ The GAN approach gets its name from its very process: an adversarial competition between two different deep learning networks, the generator network and the discriminator network.³¹ This competition has been likened to a series of rounds between a master counterfeiter, who produces a counterfeit bill, and an expert currency investigator, who analyzes the counterfeit bill along with a genuine bill and then attempts to differentiate between the two.³² Like humans gaining skill and knowledge through the experiences of victory and defeat, the networks marginally improve at the end of each round.³³ As an automated and unsupervised process, many thousands of iterations of this game are played until perfection is achieved.³⁴

When it comes to deepfakes, the generator network trains to create false images of a real person's likeness until the discriminator can no longer determine

27. 676 F.3d 19 (2d Cir. 2012).

28. *See generally id.*

29. "Deep learning is a subset of machine learning," which is capable of interpreting data and developing its own rules to improve at a function, like categorization, without human supervision. *What Is Deep Learning?*, IBM, <https://www.ibm.com/cloud/learn/deep-learning> (last visited June 28, 2022).

30. *See* Chesney & Citron, *supra* note 3, at 1760 ("Growing sophistication of the GAN approach is sure to lead to the production of increasingly convincing deep fakes.").

31. *Id.*

32. This Malicious Life, *GAN and Deep Fakes, Part 1*, CYBERREASON, <https://malicious.life/episode/episode-67/> (last visited Jan. 8, 2023).

33. *Id.*

34. *See* Van Lier, *supra* note 13, at 20 (highlighting that "the competition between generating and detecting fake videos" has been described "[as] analogous to a chess game" (alteration in original) (quoting Siwei Lyu, *Detecting "Deepfake" Videos in the Blink of an Eye*, THE CONVERSATION (Aug. 29, 2018, 6:44 AM), <https://theconversation.com/detecting-deepfake-videos-in-the-blink-of-an-eye-101072>))).

which is real and which is fake.³⁵ With a few source images of a person's face, the GAN approach carries out its process until the networks can convincingly generate and depict the real face in any type of shape or expression.³⁶ With deepfake pornography, the end result of GANs is a face swap. In any real pornographic video, the performer's face can be convincingly replaced with the victim's face, expressions and all.³⁷

Under the hood, creating deepfake pornography is an incredibly complex exploit of AI technology, but because it is an automated process, creation tools are easier to use and more accessible than ever.³⁸ For instance, one app enables users to digitally strip women naked using a single photo, ultimately producing fake nude photos of the victim. The application had 94,464 users in June of 2019 and has likely gained many more since.³⁹ Even popular and innocuous deepfake applications, mostly used for entertainment purposes like face swapping with celebrities or "deepfaking" a co-worker into a comedic situation, can facilitate deepfake pornography creation. One such application, Facemagic, openly advertises its capability for creating deepfake pornography and invites customers to do so on certain adult websites.⁴⁰

A day many have feared has already arrived: in 2021 a website emerged inviting users to "turn anyone into a porn star by using deepfake technology to swap the person's face into an adult video."⁴¹ Considering deepfake technology only appeared in 2017,⁴² we have reached this dangerous point far too quickly. AI is making deepfake pornography easier to create and harder to differentiate

35. See Erik Gerstner, *Face/Off: "DeepFake" Face Swaps and Privacy Laws*, 87 DEF. COUNS. J. 1, 3–4 (2020) (discussing how Fakeapp can train the "deep learning algorithm" after creating datasets).

36. See *id.* at 2–3 ("In its most basic form, all a user needs is a 'base' video and a number of source images of the face of the person being pasted into the video.").

37. See Citron, *supra* note 10, at 1921.

38. See Robert Volkert & Henry Ajder, *Analyzing the Commoditization of Deepfakes*, N.Y.U. J. LEGIS. & PUB. POL'Y QUORUM (Feb. 27, 2020), <https://nyujlpp.org/quorum/volkert-ajder-analyzing-the-commoditization-of-deepfakes/>.

39. The name of the application has been intentionally omitted as to not give it further publicity, however, for a general overview and case study of the app, see HENRY AJDER, GIORGIO PATRINI, FRANCESCO CAVALLI & LAURENCE CULLEN, DEEPTRACE, THE STATE OF DEEPPAKES: LANDSCAPE, THREATS, AND IMPACT 8 (2019), https://regmedia.co.uk/2019/10/08/deepfake_report.pdf.

40. See Samantha Cole, *A Popular Face Swap App Is Advertising Deepfakes on Porn Sites*, VICE (May 10, 2022), <https://www.vice.com/en/article/wxez8y/pornhub-year-in-review-deepfake> (noting that Facemagic, with over one million Google Play downloads, has advertisements stating, "Make AI face porn in a sec. Create now!").

41. Karen Hao, *A Horrifying New AI App Swaps Women into Porn Videos with a Click*, M.I.T. TECH. REV. (Sept. 13, 2021) [hereinafter Hao II], <https://www.technologyreview.com/2021/09/13/1035449/ai-deepfake-app-face-swaps-women-into-porn>.

42. Volkert & Ajder, *supra* note 38.

from reality.⁴³ On this course, we should expect that the issue is only going to get worse if left unchecked.⁴⁴

B. *The Lack of Adequate Legal Remedies*

As the case always seems to be with rapidly evolving technology, the current laws aimed at curtailing socially harmful deepfake content are woefully inadequate.⁴⁵ This is because the United States typically takes a reactive approach to social issues, rather than a proactive one.⁴⁶ A handful of laws have proactively legislated, or at least attempted to anticipate, the worsening deepfake harm,⁴⁷ yet most of them fail for various reasons.

A few attempts at creating federal legislation to combat deepfakes have been made, but so far nothing comes to the aid of deepfake pornography victims. Senators on both sides of the political aisle have introduced bills that attempt to combat deepfake pornography, but neither advanced past introduction.⁴⁸ Even if any

43. See *id.* (“The technology for creating deepfakes is becoming increasingly accessible and commoditized . . . and threats posed by nefarious and criminal uses of deepfakes are becoming increasingly tangible.”).

44. It is also worth noting that while deepfake pornography is the technology’s most common use, it can and has been used for many other social harms such as disrupting the political process, identity theft, and extortion. See Van Lier, *supra* note 13, at 18 (noting deepfake technology was used “to recreate the speech President John F. Kennedy was planning to deliver on the day he was assassinated” and to create videos depicting “behaviors like revenge pornography, sexualized extortion, and voyeurism”).

45. See Julia Griffith, *A Losing Game: The Law Is Struggling to Keep Up with Technology*, J. HIGH TECH. L. (Apr. 12, 2019), <https://sites.suffolk.edu/jhtl/2019/04/12/a-losing-game-the-law-is-struggling-to-keep-up-with-technology/> (“It has been estimated that the law is at least five years behind developing a technology [to combat deepfakes].”).

46. See ROB A. DELEO, *ANTICIPATORY POLICYMAKING: WHEN GOVERNMENT ACTS TO PREVENT PROBLEMS AND WHY IT IS SO DIFFICULT* 3 (2015) (“Government, it seems, is only adept at cleaning up society’s problems . . .”).

47. Bradley Waldstreicher, *Deeply Fake, Deeply Disturbing, Deeply Constitutional: Why the First Amendment Likely Protects the Creation of Pornographic Deepfakes*, 42 CARDOZO L. REV. 729, 738–40 (2021).

48. See, e.g., Deepfake Task Force Act, S. 2559, 117th Cong. (2021) (introduced by Senator Rob Portman (R-OH) attempting to establish the National Deepfake Provenance Task Force “[c]omprised of federal and nonfederal stakeholders” to “address threats” posed by deepfake technology); Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019, H.R. 3230, 116th Cong. (2019) (introduced by Rep. Yvette Clarke (D-NY) attempting to require Deepfake content to be labeled so viewers know the image is altered); Malicious Deep Fake Prohibition Act of 2018, S. 3805, 115th Cong. (2018) (introduced by Senator Ben Sasse (R-NE) attempting to make it illegal to “create, with the intent to distribute, a Deepfake with the intent that the distribution of the Deepfake would facilitate criminal or tortious conduct under Federal, State, local, or Tribal law” and to knowingly distribute the content).

were enacted today, this Note argues that they would be inadequate because they do not deter service providers from removing deepfake content for fear of vicarious liability.

As the deepfake problem worsens, states seem to be legislating with disappointing consistency. Despite the vast majority of harmful deepfake content being nonconsensual pornography,⁴⁹ most deepfake legislation at the state level is too narrow in scope to address that harm. Texas is a clear example of this. The criminal statute only prohibits deepfake political content created with “intent to injure a candidate or influence the result of an election.”⁵⁰ Because of this limitation in scope, Texan deepfake pornography victims are forced to either attempt to fit their situation into the definition of another charge⁵¹ or pursue the disheartening endeavor of tort litigation. Washington, Maryland, and Maine proposed deepfake legislation that narrowly focused on curtailing politically disruptive deepfakes, which did not address the full scope of the problem.⁵²

However, a few states are making positive strides with deepfake laws. California’s and New York’s deepfake laws expressly recognize the harms of non-consensual pornography.⁵³ Both laws create private causes of action to bring against perpetrators rather than imposing criminal penalties.⁵⁴ Rather than creating brand new legislation, Virginia amended its existing criminal revenge porn statute by including “a person whose image was used in creating, adapting, or modifying a videographic or still image with the intent to depict an actual person.”⁵⁵

Some scholars argue there is no need for new legislation to combat deepfake porn because the harm can be addressed by other causes of action, like copyright infringement, or the torts of false light and intentional or negligent infliction of emotional distress (IIED and NIED).⁵⁶ However, this argument crumbles under

49. Van Lier, *supra* note 13, at 18.

50. TEX. ELEC. CODE ANN. § 255.004 (West 2021).

51. See Benson Varghese, *Deepfakes in Texas: What Are They and Are They Illegal?*, VARGHESE SUMMERSETTE (Apr. 17, 2021), <https://versustexas.com/deepfakes/> (listing cyberbullying, extortion, harassment, and revenge porn as examples of possible “other charges that can apply to deepfake creators”).

52. S.B. 6513, 66th Leg., Reg. Sess. (Wash. 2020); H.B. 198, 2020 Gen. Assemb., Reg. Sess. (Md. 2020); S.P. 690, 129th Leg., 2d Reg. Sess. (Me. 2020).

53. CAL. CIV. CODE § 1708.86 (West 2022); N.Y. CIV. RIGHTS LAW § 52-c (McKinney 2022). For an in-depth analysis of the harms of deepfakes and California’s attempt at addressing them, see generally Van Lier, *supra* note 13.

54. See N.Y. CIV. RIGHTS LAW § 52-c; see also CAL. CIV. CODE § 1708.86.

55. VA. CODE ANN. § 18.2-386.2 (2022).

56. See, e.g., David Greene, *We Don’t Need New Laws for Faked Videos, We Already Have Them*, ELECTRONIC FRONTIER FOUND. (Feb. 13, 2018), <https://www EFF.ORG/deep links/2018/02/we-dont-need-new-laws-faked-videos-we-already-have-them>.

a preliminary inspection of each claim.⁵⁷ While a deepfake victim retains the copyright to their photograph, a copyright infringement claim may still fail because the deepfake creator is making transformative “fair use” of the photo.⁵⁸ The tort claims are likely to fail because the victim is unlikely to meet all the required elements. For IIED, victims may struggle to show the producer intended to cause the victim’s distress, that the deepfake caused the distress, and that the distress is severe enough.⁵⁹ For negative infliction of emotional distress (NIED), most states limit the action either to a bystander that suffered emotional distress while witnessing another being harmed, or to an individual that escaped physical harm and consequently suffers emotional harm.⁶⁰ Deepfake victims, however, do not fit well into either category.⁶¹ Additionally, requirements for tort actions can vary drastically depending on the jurisdiction,⁶² while some are not even available in all fifty states.⁶³

Even in the unlikely scenario where a victim has a strong cause of action and lives in a jurisdiction that recognizes it, current basic procedural obstacles render the courtroom path to remedy useless in its current state. First, hauling a perpetrator into the victim’s court would require personal jurisdiction.⁶⁴ To determine if jurisdiction is proper, a victim would have to be able to identify the perpetrator, and most perpetrators act online through platforms that allow anonymity.⁶⁵ Even a victim with the most aggressive and advanced investigative techniques would fail to identify a perpetrator whose IP address is obscured by a Virtual Private Network (VPN) or decentralized browser.⁶⁶ Finally, if somehow the perpetrator

57. See Douglas Harris, *Deepfakes: False Pornography Is Here and the Law Cannot Protect You*, 17 DUKE L. & TECH. REV. 99, 107–21 (2019) (acknowledging potential for each claim in the context of deepfakes but critiquing the feasibility of each one).

58. Harris, *supra* note 57, at 107–10.

59. *Id.* at 111.

60. *Id.* at 113.

61. See *id.* at 113–15 (explaining difficulty of deepfake victims bringing a successful claim under either category).

62. See *id.* at 111–19 (discussing how elements of IIED, NIED, and false light claims vary by jurisdiction).

63. See, e.g., *id.* at 115–16 (noting that a false light claim is only available in thirty-three states).

64. FED. R. CIV. P. 4(k).

65. Elizabeth Caldera, “Reject the Evidence of Your Eyes and Ears”: Deepfakes and the Law of Virtual Replicants, 50 SETON HALL L. REV. 177, 191 (2019).

66. See Rebecca A. Delfino, *Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn’s Next Tragic Act*, 88 FORDHAM L. REV. 887, 898–99 (2019).

was identified, jurisdiction was proper, and the victim won the case, there is always the possibility that the perpetrator would be judgement proof.⁶⁷

Apart from the laundry list of procedural obstacles keeping the courtroom from being a viable avenue for relief, the piecemeal litigation approach does not get to the heart of the problem. Once a video is posted to the internet by someone else, it is difficult to take down, and almost impossible if it disseminates.⁶⁸ Since lawsuits are difficult to bring and even more difficult to win, the current piecemeal litigation approach will do little to deter future perpetrators.

Google's CEO Sundar Pichai likened the benefits and dangers of AI to fire, remarking "it [would] change humanity more profoundly than fire."⁶⁹ It appears he was right. If the issue of deepfake pornography is a growing forest fire, the current legal remedies are a scattered band of volunteers with squirt guns. Justice demands a better line of defense. We need to clear brush and dig lines to contain this fire before it has a chance to grow. Federal legislation that addresses the problem by shifting the burden to those most apt to handle it—the content hosts⁷⁰—escapes the inadequacies of our current remedies.

II. BOTTOM-LINE CONTROL, THE POWER OF VICARIOUS LIABILITY, AND THE NEED FOR FEDERAL LEGISLATION

In a free market society, direct impact on an industry's financial interest is the most effective strategy to transform behavior.⁷¹ Tort liability aimed at deterring certain corporate practices achieves this by shifting costs to those most capable of mitigating risk. It is a simple and proven economic principle: when litigation threatens a company's bottom-line, precautionary measures follow.⁷² This model, utilizing vicarious liability to incentivize corporate change,⁷³ is exactly

67. Creating deepfake pornography does not require much infrastructure, most perpetrators are likely individuals who lack the deep pockets required to provide proper redress to victims.

68. See Delfino, *supra* note 66, at 898–99; Laser & Goldman, *supra* note 6.

69. Harris, *supra* note 57, at 99–100.

70. For the purpose of this Note, "content hosts" refers to websites that host deepfake pornography.

71. See Catherine M. Sharkey, *Institutional Liability for Employees' Intentional Torts: Vicarious Liability as a Quasi-Substitute for Punitive Damages*, 53 VAL. U. L. REV. 1, 32–33 (2018).

72. See Omar Y. Pelled, *The Proportional Internalization Principle in Private Law*, 11 J. LEGAL ANALYSIS 160, 160–62 (2019); see generally Victoria C. Dawson, *Who Is Responsible When You Shop Until You Drop?: An Impact on the Use of the Aggressive Marketing Schemes of "Black Friday" Through Enterprise Liability Concepts*, 50 SANTA CLARA L. REV. 747 (2010) (arguing imposing enterprise liability on major retailers would prevent the aggressive marketing practices that result in injury and death at major sales events).

73. Hereinafter, this is referred to as "bottom-line control."

the type of enforcement mechanism that will make federal legislation targeting deepfake pornography successful.

Vicarious liability is a tort theory in which a tortfeasor's actions are imputed to a secondary party as if the secondary party had committed the tort itself.⁷⁴ Consider the example of rideshare companies. In this example, more rides per hour mean more revenue for the rideshare company and the driver. If the rideshare company cannot be held liable for the negligence of its contracted driver, it will have little incentive to promote risk-averse driving since a faster driver can complete more rides and thus generate more revenue.⁷⁵ However, imputing the driver's negligence to the rideshare company jeopardizes its revenue. Thus, vicarious liability incentivizes the rideshare company to implement and enforce safe driving practices.⁷⁶

In the context of deepfake pornography, content hosts would be the logical secondary party. These are the websites that distribute deepfake pornography that users upload.⁷⁷ Holding content hosts vicariously liable resolves many of the issues with the current, inadequate legal remedies. First, vicarious liability provides an obvious defendant to bring an action against: the content host. The fact that the deepfake creator is anonymous will no longer prevent a victim from bringing suit. Second, hosts are much more likely to have "deep pockets" and be capable of paying higher sums of damages to victims.⁷⁸ This is especially true in the internet porn context, an industry that is conservatively estimated to annually bring in \$15 billion in revenue.⁷⁹ Three of the largest online porn hosts allow deepfake content.⁸⁰ While their revenues are unknown, it can safely be assumed that their ad revenue is astronomical since these companies individually receive

74. See 22 AM. JUR. 2D *Damages* § 606 (2022); see also RESTATEMENT (SECOND) OF TORTS § 909 cmt. b (AM. L. INST. 1979).

75. See Pelled, *supra* note 72, at 164–66.

76. See *id.* at 165.

77. See Caldera, *supra* note 65, at 185; see also BRITT PARIS & JOAN DONOVAN, DATA & SOC'Y, DEEPPAKES AND CHEAPFAKES: THE MANIPULATION OF AUDIO AND VISUAL EVIDENCE 35 (2018), https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_Fakes_Final-1-1.pdf.

78. Benjamin Shmueli, *Legal Pluralism in Tort Law Theory: Balancing Instrumental Theories and Corrective Justice*, 48 U. MICH. J.L. REFORM 745, 789–90 (2015).

79. John Naughton, *The Growth of Internet Porn Tells Us More About Ourselves Than Technology*, THE GUARDIAN (Dec. 30, 2018), <https://www.theguardian.com/commentisfree/2018/dec/30/internet-porn-says-more-about-ourselves-than-technology> (explaining 2018 U.S. revenue estimates for the online porn industry "range[ed] from \$9bn to \$97bn a year" making it "bigger than not only Netflix (\$11.7bn) but also Hollywood as a whole (\$11.1bn) and Viacom (\$13.3bn)").

80. Matt Burgess, *Deepfake Porn Is Now Mainstream. And Major Sites Are Cashing In*, WIRED (Aug. 27, 2020, 6:00 AM), <https://www.wired.co.uk/article/deepfake-porn-websites-videos-law>.

as much internet traffic as non-pornographic internet mainstays such as Wikipedia, Amazon, and Reddit.⁸¹ Third, vicarious liability serves “the goal of optimal deterrence,” because it motivates a large company to invest and engage in practices that will reduce its exposure to future litigation.⁸²

While a small handful of mainstream content hosts have recognized the non-consensual nature of deepfake pornography and banned it to various degrees,⁸³ federal legislation utilizing bottom-line control is the best way to fully address the issue. Even hosts that have banned deepfake content remain unwilling to adequately enforce their own policies.⁸⁴ Apart from that, the most significant perpetrators of deepfake pornography distribution are independent content hosts specifically dedicated to that purpose.⁸⁵

Federal legislation combating deepfakes would create uniform accountability for perpetrators by bringing together three prominent forces: federal law, the courts, and the private sector. This interplay of institutions is necessary to create nationwide change.⁸⁶ Creating a federal law gives plaintiffs a uniform cause of action and provides courts uniform directions regarding jurisdiction and liability of parties.⁸⁷ Guided by proper federal law, courts are better suited to handle emerging technology issues.⁸⁸ Targeted legislation is also easier and less costly to enforce through litigation.⁸⁹ This in turn causes the private sector to internalize the costs of looming litigation and engage and invest in preventative practices.⁹⁰ The private sector is also better suited to bear this cost because it can quickly and efficiently implement technology-based solutions.⁹¹ Simply put, the private sector can go far with some arm twisting. A clear federal law imposing vicarious liability on content hosts would do just that.

81. *See id.*

82. *See Shmueli, supra* note 78, at 790.

83. Samantha Cole, *The Ugly Truth Behind Pornhub's 'Year in Review'*, VICE (Feb. 18, 2020, 5:00 AM), <https://www.vice.com/en/article/wxez8y/pornhub-year-in-review-deepfake> (reporting that Reddit, Discord, and Pornhub all placed bans on deepfake pornography in 2018).

84. *See id.*

85. AJDER ET AL., *supra* note 39, at 6.

86. *See* Van Lier, *supra* note 13, at 20 (“True change will require a multi-pronged approach as the law alone is not equipped to respond to these rapidly advancing technologies.”).

87. *See* Delfino, *supra* note 66, at 927–28.

88. *See* Gerstner, *supra* note 35, at 10.

89. John Blevins, *Uncertainty as Enforcement Mechanism: The New Expansion of Secondary Copyright Liability to Internet Platforms*, 34 CARDOZO L. REV. 1821, 1829 (2013).

90. *See* Gerstner, *supra* note 35, at 13.

91. Van Lier, *supra* note 13, at 20–21 (stating “authentication technologies” and “hashing” are existing AI processes that can be used to limit the dissemination of deepfake pornography).

Deepfake pornography is exponentially becoming more convincing and easier to create.⁹² Without significant deterrence in place, dedicated content hosts will only continue to grow and distribute harmful content and mainstream hosts will have greater difficulty moderating their own policies. Bottom-line control, driven by vicarious liability, is the most effective way to combat deepfake pornography. Federal legislation combatting deepfake pornography using this mechanism would compel mainstream porn companies to police themselves and deter independent dedicated content hosts from existing in the first place.

A. *The DMCA as a Legislative Model for Combatting Deepfake Pornography*

Capitalizing on bottom-line control's efficacy as an enforcement mechanism, the DMCA is an excellent legislative model to follow for crafting federal legislation to combat deepfake pornography distribution. The DMCA imposes severe financial risks for service providers, and as a result, internet platforms are incentivized to alter their business practices.⁹³ The DMCA and its impact on copyright business practices illustrate both the DMCA's viability as a legislative model and expose the statute's weaknesses that can be improved upon in crafting federal deepfake pornography legislation.

A crucial difference to appreciate between the DMCA and potential federal deepfake pornography legislation is the interests they are designed to protect. The DMCA's principal goal is to protect multiple commercial interests.⁹⁴ On one hand, the DMCA protects the commercial interest of copyright holders, on the other, it protects the commercial interests of emerging technologies and companies by providing conditional immunity.⁹⁵ Federal deepfake pornography legislation must be solely aimed at protecting the individual liberties of victims.⁹⁶ Therefore, while using the DMCA as a legislative model, federal deepfake pornography legislation should emphasize imposing liability on distributors and providing remedies for victims.

Much like AI is changing the world we live in today, personal computers and the internet were transforming the landscape of copyright law at the end of the 20th century.⁹⁷ CD-ripping, file sharing, and various online web platforms quickly facilitated an explosion of unauthorized distribution of copyrighted content.⁹⁸ For American copyright law, this was "perhaps the most tumultuous

92. See Volkert & Ajder, *supra* note 38.

93. See Blevins, *supra* note 89, at 1833–34.

94. See Cobia, *supra* note 17, at 388.

95. See Amir Hassanabadi, *Viacom v. YouTube—All Eyes Blind: The Limits of the DMCA in a Web 2.0 World*, 26 BERKELEY TECH. L.J. 405, 411–12 (2011).

96. See Citron, *supra* note 10, at 1896–1901.

97. See Hassanabadi, *supra* note 95, at 409.

98. See *id.*

period of its three hundred year existence.”⁹⁹ The World Intellectual Property Organization (WIPO) realized the law needed to catch up quickly, and passed two treaties in 1996.¹⁰⁰ These treaties were a driving force that spurred Congress to pass the DMCA in 1998.¹⁰¹

Though the DMCA’s interest in protecting intellectual property rights is far from perfectly analogous to protecting the civil liberties of individuals, the DMCA’s Safe Harbor and Notice and Takedown provisions serve as excellent models for federal deepfake pornography legislation. To qualify for Safe Harbor immunity, service providers must adhere to all of the Section 512(c)(1) requirements.¹⁰² A service provider is not liable for copyright infringement if it:

- (A) (i) does not have actual knowledge that the material or an activity using the material on the system or network is infringing; (ii) in the absence of such actual knowledge, is not aware of facts or circumstances from which infringing activity is apparent; or (iii) upon obtaining such knowledge or awareness, acts expeditiously to remove, or disable access to, the material;
- (B) does not receive a financial benefit directly attributable to the infringing activity, in a case in which the service provider has the right and ability to control such activity; and
- (C) upon notification of claimed infringement as described in paragraph (3), responds expeditiously to remove, or disable access to, the material that is claimed to be infringing or to be the subject of infringing activity.¹⁰³

Essentially, the Safe Harbor provision provides protection from liability for distributors of copyright infringing materials if they act diligently and meet the 512(c)(1) conditions. Though the Section 512(c)(1)(B) financial benefit requirement is not relevant,¹⁰⁴ Section 512(c)(1)(A) and Section 512(c)(1)(C) provide two important models for effective federal deepfake pornography legislation.

99. *Id.*

100. *Id.* at 409–10; *see generally* World Intellectual Property Organization (WIPO) Copyright Treaty, *adopted* Dec. 20, 1996, 2186 U.N.T.S. 121; World Intellectual Property Organization: Performances and Phonograms Treaty, *adopted* Dec. 20, 1996, 36 I.L.M. 76.

101. *See* Hassanabadi, *supra* note 95, at 410.

102. *See* 17 U.S.C. § 512(c)(1).

103. *Id.* § 512(c)(1)(A)–(C).

104. Imposing a financial benefit requirement would severely hamper the efficacy of a deepfake porn statute because of the reasons discussed *infra* Section III.A.

First, Section 512(c)(1)(A) imposes a “red flag knowledge” standard.¹⁰⁵ This means that service providers who have actual or apparent knowledge of infringing activity on their platform do not qualify for Safe Harbor immunity.¹⁰⁶ Congress intended this provision to impose an objective and subjective test: service providers must be “subjectively aware of the circumstances relating to the infringement” and that “the infringement would have been apparent to a reasonable person operating under similar circumstances.”¹⁰⁷ Subsequent interpretations of the provision have varied widely.¹⁰⁸ However, the key takeaway for federal deep-fake pornography legislation is imposing a defined standard of knowledge for content hosts.

Second, Section 512(c)(1)(C) imposes a duty to remove infringing content upon notification.¹⁰⁹ The provision compels a reactive response from service providers. If the provider does not fulfill this duty, it may lose its Safe Harbor qualification.¹¹⁰

Section 512(c)(2)-(3), the Notice and Takedown provision, places an additional requirement on service providers to qualify for immunity:

The limitations on liability established in this subsection apply to a service provider *only if* the service provider has designated an agent to receive notifications of claimed infringement described in paragraph (3), by making available through its service, including on its website in a location accessible to the public, and by providing to the Copyright Office, substantially the following information:

- (A) the name, address, phone number, and electronic mail address of the agent.
- (B) other contact information which the Register of Copyrights may deem appropriate.¹¹¹

Furthermore, the provisions establish, in detail, the elements of effective notice and require substantial compliance. Notice must be signed, identify the infringed

105. See *Viacom Int’l, Inc. v. YouTube, Inc.*, 676 F.3d 19, 30–31 (2d Cir. 2012) (referring to Section 512(c)(1)(A) as the “red flag” knowledge provision).

106. *Id.* at 30.

107. Hassanabadi, *supra* note 95, at 414.

108. The red flag knowledge interpretation has ranged from constructive knowledge to knowledge as specific as providing individual URLs. See Blevins, *supra* note 89, at 1838; see generally *In re Aimster Copyright Litig.*, 252 F. Supp. 2d 634, 661 (N.D. Ill. 2002), *aff’d*, 334 F.3d 643 (7th Cir. 2003) (explaining that interpretations of the safe harbor provision may vary substantially).

109. 17 U.S.C. § 512(c)(1)(C).

110. *Id.*

111. *Id.* § 512(c)(2)(A)–(B) (emphasis added).

and the infringing material, provide sufficient contact information of the complaining party, be in good faith, and contain a sworn statement that the notice is accurate.¹¹² Having clear statutory guidelines for how a Notice and Takedown system should function is essential for successful federal deepfake pornography legislation.

Together, these provisions created the process that has allowed the internet to regulate itself for the last two decades. A copyright holder submits statutorily sufficient notice to the service provider's designated agent.¹¹³ The designated agent, incentivized to remove the material if it is infringing in order to qualify for safe harbor, removes the content if it is determined to be infringing.¹¹⁴ This process circumvents the difficulty in protecting copyrighted material from anonymous user-generated content abuse, much of which went unaddressed before the DMCA's enactment.¹¹⁵

Together, these provisions establish standards that must be followed in order to protect service providers from monetary liability. The Safe Harbor provision establishes a standard of knowledge that disqualifies service providers from immunity and imposes a duty to act to remedy violations.¹¹⁶ The Notice and Takedown provision creates a statutory standard for a system to make this process possible.¹¹⁷ These DMCA provisions, despite the drawbacks, serve as excellent models for federal deepfake pornography legislation.

B. *Case Study: Viacom International, Inc. v. YouTube, Inc.*

YouTube, the internet's premier video streaming and sharing platform,¹¹⁸ and its history of DMCA elusion and compliance, illustrates the potency of targeted federal legislation designed to influence and deter certain corporate behaviors. YouTube's transformation from the internet's chief piracy facilitator to one of the strictest moderators of copyrighted content stems from a 2012 copyright infringement lawsuit brought by Viacom and a class of copyright holders.¹¹⁹

112. *Id.* § 512(c)(2)–(3).

113. *Id.*

114. *See* Blevins, *supra* note 89, at 1844–45.

115. *See id.* at 1838–39.

116. *See generally* § 512.

117. *Id.* § 512(c)(1)–(3).

118. Julia Stoll, *Leading U.S. Video Streaming Services 2019, by Monthly Average Users*, STATISTA (Jan. 18, 2022), <https://www.statista.com/statistics/910875/us-most-popular-video-streaming-services-by-monthly-average-users/#:~:text=The%20most%20popular%20video%20streaming,with%20just%20under%2026.5%20million>.

119. *See generally* Viacom Int'l, Inc. v. YouTube, Inc., 676 F.3d 19 (2d Cir. 2012).

YouTube has had issues with copyright infringement since its earliest days.¹²⁰ YouTube is powered by user-generated content (UGC), meaning YouTube serves as a host and distributor of videos, but most of its content is created and uploaded by individual users.¹²¹ The UGC format results in nonstop and essentially infinite video uploads to the platform.¹²² In the mid-2000s, near YouTube's inception and rise to prominence, the platform did very little to curb copyright infringement.¹²³ YouTube was quickly under fire for widespread and repeated unauthorized distribution of copyrighted content, being sued for hundreds of millions by media giants like the Premier League and Mediaset.¹²⁴

Today, however, YouTube (now owned by Google)¹²⁵ has an incredibly aggressive infrastructure for preventing copyright infringement. Google boasts immense success when it comes to antipiracy, removing over 837 million URLs with copyright infringing content in 2017 alone.¹²⁶ Building this robust infrastructure required investments exceeding one hundred million dollars.¹²⁷ When it comes to anti-piracy, YouTube/Google only arrived where they are now because of bottom-line control guided by the DMCA and tort litigation.¹²⁸

*Viacom International, Inc. v. YouTube, Inc.*¹²⁹ began in 2007 when the complaint arrived on the desk of YouTube executives with an audacious prayer: \$1 billion in damages for copyright infringement.¹³⁰ Viacom, a colossal media conglomerate itself, held the copyrights to hundreds of high-value television programs and movies which YouTube was allegedly redistributing on its platform

120. See William L. Hosch, *YouTube*, BRITANNICA, <https://www.britannica.com/topic/YouTube> (Aug. 4, 2022).

121. Lauren G. Gallo, *The (Im)possibility of Standard Technical Measures for UGC Websites*, 34 COLUM. J.L. & ARTS 283, 291–96 (2011).

122. See Hassanabadi, *supra* note 95, at 421 (“Less than a year [after its inception] YouTube was streaming more than thirty million videos a day.”).

123. *Id.*

124. See Michael Kahn, *Premier League Soccer Sues YouTube Over Copyright*, REUTERS (May 4, 2007), <https://www.reuters.com/article/us-youtube-lawsuit/premier-league-soccer-sues-youtube-over-copyright-idUSN0445675320070505>; see also Stephen Lawson, *Italian Media Company Sues YouTube*, NETWORK WORLD (July 30, 2008), <https://www.networkworld.com/article/2274290/127talian-media-company-sues-youtube.html>.

125. GOOGLE, HOW GOOGLE FIGHTS PIRACY 18 (2018) [hereinafter GOOGLE I], https://www.blog.google/documents/27/How_Google_Fights_Piracy_2018.pdf/.

126. See *id.* at 38.

127. See *id.* at 27.

128. See Todd Spangler, *Why the Viacom-YouTube Case Wasn't a Giant Waste of Time*, VARIETY (Mar. 18, 2014, 10:07 AM), <https://variety.com/2014/biz/news/why-the-viacom-youtube-case-i-a-giant-waste-of-time-1201137664/>.

129. 676 F.3d 19 (2d Cir. 2012).

130. See *id.* at 26; Hassanabadi, *supra* note 95, at 405.

via user uploads.¹³¹ Viacom alleged that these videos numbered in the hundreds of thousands and that YouTube was brazenly and uniformly eschewing DMCA-compliant takedown notices issued by copyright holders.¹³² Viacom went as far as to allege that “YouTube implemented a policy of maintaining access to infringing videos unless and until it received a ‘cease and desist’ demand from the copyright owner.”¹³³

Viacom’s argument was pivotal in the mid-2000s copyright polemic because it essentially called into question the enforceability of the DMCA in the internet age. Viacom alleged that, because YouTube had actual knowledge of the infringement but acted with willful blindness and failed to act expeditiously to stop the infringement, it did not qualify for DMCA safe harbor immunity.¹³⁴ However, the trial court did not buy this argument. YouTube was granted summary judgment on the basis that “defendants were entitled to DMCA safe harbor protection primarily because they had insufficient notice of the particular infringements in suit.”¹³⁵ The trial court further reasoned that “item specific knowledge of infringing activity is required for a service provider to have the right and ability to control infringing activity.”¹³⁶

When the Second Circuit heard the case on Viacom’s appeal, the grant of summary judgment was reversed primarily on these two grounds.¹³⁷ The Second Circuit affirmed that knowledge of specific instances of infringement is required to disqualify a service provider from Safe Harbor immunity.¹³⁸ However, internal email communications revealed that YouTube executives were aware of numerous specific instances of infringement via takedown notices, but intentionally and consistently left the content up to promote their nascent platform’s growth.¹³⁹ The Second Circuit also held that common law willful blindness could be applied to demonstrate knowledge of specific instances of infringement under the DMCA, and reversed summary judgment on this ground as well.¹⁴⁰

131. Viacom subsidiaries include MTV, Comedy Central, Nickelodeon, and Paramount Pictures. Hassanabadi, *supra* note 95, at 421.

132. *Id.* at 422.

133. Opening Brief for Plaintiffs-Appellants at 11, *Viacom Int’l, Inc. v. YouTube, Inc.*, 676 F.3d 19 (2d Cir. 2010) (No. 10-3270).

134. Hassanabadi, *supra* note 95, at 423.

135. *Viacom*, 676 F.3d at 26.

136. *Id.*

137. *Id.* at 42.

138. *Id.* at 32.

139. *Id.* at 34 (referring to a takedown notice from CNN, YouTube founder Jawed Karim stated in an email, “the CNN space shuttle clip, I like. We can remove it once we’re bigger and better known, but for now that clip is fine.”).

140. *Id.* at 32.

The Second Circuit's holding in *Viacom* indicated that Safe Harbor immunity appeared to require more substantial compliance with the DMCA than YouTube had bet on.¹⁴¹ Though the remanded case would have been a fact-specific determination for the jury, their decision could have fundamentally changed secondary copyright infringement liability on the internet. However, the remanded case will always be a landmark "what if" because the seven-year suit was settled in 2014 under undisclosed terms.¹⁴²

Still, stepping back from the minutia of what the Second Circuit *Viacom* decision meant specifically for copyright law in the rapidly changing internet landscape, something much bigger is revealed: the lawsuit, by proxy of the statute, drove radical institutional change. The highly successful antipiracy infrastructure utilized today by YouTube/Google is called Content ID,¹⁴³ and this case was undoubtedly a catalyst for its evolution and rapid implementation.

Content ID is Google's proprietary copyright management system. Content ID is highly sophisticated and operates largely unsupervised by human intervention.¹⁴⁴ The system automatically recognizes copyright-protected content uploaded to YouTube and allows rights holders to either "make money from it, leave it up and track viewing statistics; or block it from YouTube altogether."¹⁴⁵

Content ID launched and implemented its monetization and revenue sharing option just months after the initial *Viacom* suit.¹⁴⁶ What's more, Safe Harbor case law following *Viacom* applied the Second Circuit's willful blindness jurisprudence and was much less favorable to defendants in the position YouTube had just been in.¹⁴⁷

141. *Id.* at 41–42.

142. Jonathan Stemple, *Google, Viacom Settle Landmark YouTube Lawsuit*, REUTERS (Mar. 18, 2014, 8:05 AM), <https://www.reuters.com/article/us-google-viacom-lawsuit/google-viacom-settle-landmark-youtube-lawsuit-idUSBREA2H11220140318>.

143. GOOGLE I, *supra* note 125, at 24; *see also How Content ID Works*, GOOGLE (2020) [hereinafter GOOGLE II], <https://support.google.com/youtube/answer/2797370?hl=en>.

144. Joanne E. Gray & Nicolas P. Suzor, *Playing with Machines: Using Machine Learning to Understand Automated Copyright Enforcement at Scale*, BIG DATA & SOC'Y 1, 2 (2020).

145. GOOGLE I, *supra* note 125, at 25.

146. *See id.*; Robert Levine, *Who Really Won the \$1 Billion YouTube-Viacom War?*, VANITY FAIR (Mar. 26, 2014), <https://www.vanityfair.com/news/business/2014/03/youtube-viacom-war>; *see also* Hassanabadi, *supra* note 95, at 426–27 (noting that before formally launching its current iteration Content ID, YouTube used the same system but only protected copyrights of affiliated partners, a practice Viacom likened to "cyber extortion").

147. *See* Nate Raymond, *Ex-MP3tunes Chief Held Liable in Music Copyright Case*, REUTERS (Mar. 19, 2014), <https://www.reuters.com/article/us-mp3tunes-infringement/ex-mp3tunes-chief-held-liable-in-music-copyright-case-idINBRE2129J20140319>; *see also* EMI Christian Music Grp., Inc. v. MP3tunes, LLC, 844 F.3d 79, 88 (2d Cir. 2016).

The private sector already has an abundance of AI-based tools capable of combating the distribution of deepfake pornography.¹⁴⁸ The problem is that the current legal landscape does nothing to push the private sector into implementing these tools or to improve them. Analogize this case study to today's deepfake pornography epidemic: imaginatively replace a few terms like "copyright infringement" with "distribution of nonconsensual digitally altered pornography," and it becomes clear that this legislative framework is fitting. In this altered scenario, the tried-and-true power of bottom-line control, through statutorily enabled litigation, has the potential to profoundly change corporate behavior and deter rogue bad actors in order to combat deepfake pornography.

III. LESSONS FROM THE DMCA'S LEGISLATIVE FRAMEWORK

Rather than replicate specific provisions of the DMCA, federal legislation aimed at combatting deepfake pornography should broadly emulate its legislative framework. This would facilitate achieving the chief goals of federal deepfake pornography legislation. Primarily, federal deepfake pornography legislation should utilize a more tailored Safe Harbor and Notice and Takedown approach to compel companies to invest in their own distribution prevention infrastructure, much like Google did with Content ID.¹⁴⁹ Secondly, federal deepfake pornography legislation should improve upon the DMCA's shortcomings and facilitate remedies for victims through specific drafting that leaves little to judicial interpretation.

A. *Establishing Strong Foundations*

Broadly speaking, the underlying interest protected by federal deepfake pornography legislation must be individual liberties and the downstream harms that result from being a target.¹⁵⁰ Long before the post-production reality came to exist, early advocates of privacy as an inherent right recognized that privacy protects a fundamentally different interest than intellectual property law.¹⁵¹ This right to privacy was conceived and advocated for during the escalating privacy invasions that came with the burgeoning of photographic technology.¹⁵² Likened to the right to be free from assault and imprisonment, privacy is a right to not be

148. Van Lier, *supra* note 13, at 20; Citron, *supra* note 10, at 1955–58.

149. See GOOGLE I, *supra* note 125, at 13, 16.

150. See Matthew B. Kugler & Carly Pace, *Deepfake Privacy: Attitudes and Regulation*, 116 NW. U. L. REV. 624–25 (2021).

151. See, e.g., Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 193–95 (1890) (arguing that the commercial value derived from intellectual property is actually an extension of the individual's greater "right to be let alone").

152. *Id.*

owned or possessed, it is a right to protection from invasion of the “sacred precincts of private and domestic life.”¹⁵³ The capabilities of AI to invade privacy through deepfake alterations go lightyears beyond what early privacy proponents anticipated.¹⁵⁴ Deepfakes enable the hijacking of sexual identity and the obliteration of individual liberties.¹⁵⁵

With a strong foundation focused solely on protecting individual liberties, federal deepfake pornography legislation is much less likely to become entangled in the commercial technicalities that have hampered DMCA enforcement. Within the language of the proposed legislation itself, this would mean there would be no commercial benefit requirement that allows perpetrators to escape liability.¹⁵⁶ More broadly speaking, much of the criticism leveled at the DMCA over the years stems from the practical effects of the Notice and Takedown procedures, namely stifling the creation of lucrative intellectual property and hindering the expansion of profitable companies that host UGC.¹⁵⁷

These concerns vanish when protecting profits is removed from the equation. In fact, the United States has already demonstrated great interest in curtailing the spread of nonconsensual pornography.¹⁵⁸ The United States has gone so far as to allow special categorical exceptions to free speech in prohibiting child exploitation material distribution.¹⁵⁹ Deepfake technology facilitates the creation and distribution of both.¹⁶⁰ Hence, establishing federal deepfake pornography legislation on civil liberty interests is essential.

153. *Id.* at 195, 205–06.

154. Even over a century ago when this notion was conceived, privacy proponents argued individuals should be protected from invasions by any “modern device for recording or reproducing scenes or sounds.” *Id.* at 206.

155. Citron, *supra* note 10, at 1921.

156. See 17 U.S.C. § 512(c)(1)(B) (establishing that a service provider must “not receive a financial benefit directly attributable to the infringing activity, in a case in which the service provider has the right and ability to control such activity”); see also Brett White, *Viacom v. YouTube: A Proving Ground for DMCA Safe Harbors Against Secondary Liability*, 24 S. JOHN’S J. LEGAL COMMENT. 825–28 (2011).

157. See *UMG Recordings, Inc. v. Shelter Cap. Partners LLC*, 718 F.3d 1006, 1024 (9th Cir. 2013); see also Blevins, *supra* note 89, at 1830–31 (noting that Veoh substantially complied with Section 512 of the DMCA but was still forced into bankruptcy after three years of costly litigation); Cobia, *supra* note 17, at 394–97.

158. See Hao I, *supra* note 7.

159. See *New York v. Ferber*, 458 U.S. 747, 773–74 (1982).

160. Kugler & Pace, *supra* note 150, at 624–25; *The Impact of “Deepfake” on Child Safety*, BIOMETRICA, (Feb. 1, 2022), <https://www.biometrica.com/the-impact-of-deepfake-on-child-safety/>.

B. *General Legislative Framework*

With careful refitting, the DMCA's general Section 515(c) framework can be borrowed to create a self-regulatory system that substantially impedes the distribution of deepfake pornography. Though it seems counterintuitive to impose liability, a Safe Harbor provision would be essential to effective enforcement using a bottom-line control model. This is because the opportunity to qualify for immunity is the financial incentive that compels companies to act. Safe Harbor provisions keep companies from internalizing substantial costs.¹⁶¹ As the underlying interest is protecting individual liberties, Safe Harbor for federal deepfake pornography legislation should be much less forgiving than the DMCA.

With this underlying interest, the aim of Safe Harbor would be to impose liability on distributors of deepfake pornography. The Safe Harbor should aim to be expansive in what creates red flag knowledge. This is akin to the approach taken by copyright holders' arguments in recent years, aiming to hold service providers accountable whenever possible.¹⁶² Indeed, a standard like generalized knowledge implies that facts beyond statutorily imposed notices can impose knowledge requisite to disqualify content hosts from safe harbor immunity.¹⁶³ This ultimately causes content hosts to internalize the potential costs of litigation, resulting in aggressive self-regulation and investment in preventative practices.¹⁶⁴ Additionally, imposing a narrower knowledge standard better serves victims because it opens the possibility to disqualify content hosts from liability independently of any takedown notice.¹⁶⁵

What constitutes an adequate takedown notice for deepfake pornography legislation should be straightforward and leave little room for finding inadequacy. Adequate notice has been the subject of much litigation in the DMCA context.¹⁶⁶ Moreover, sending notices should be statutorily required to be a simple and painless process on the complainant's end. This has been an issue with

161. See Gallo, *supra* note 121, at 289 (discussing how UGCs have safe harbor immunity, meaning UGCs do not necessarily need to have policies and measures in place to have safe harbor eligibility, even if they have great incentive to so).

162. See Blevins, *supra* note 89, at 1838 (discussing how content owners want provisions interpreted as broader standards).

163. *UMG Recordings, Inc. v. Shelter Cap. Partners LLC*, 718 F.3d 1006, 1015, 1020 (9th Cir. 2013) (where copyright holders argued knowledge that a UGC video website was capable of facilitating copyright infringement was sufficient to disqualify the website from safe harbor immunity).

164. Blevins, *supra* note 89, at 1838–39.

165. See Hassanabadi, *supra* note 95, at 414–15 (discussing a congressional report noting that any provider wanting to benefit from limited liability must take down materials independently).

166. See Blevins, *supra* note 89, at 1841–42 (discussing the central dispute over how specific notice must span several cases).

the DMCA, since actually contacting a DMCA agent can be a complicated bureaucratic process.¹⁶⁷ Pornographic content hosts should be required to have an easily accessible link for submitting takedown notices. Many content hosts already have something similar for revenge porn.¹⁶⁸ Following a straightforward “flag and request to remove” process should fulfill the statutory notice requirement for federal deepfake pornography legislation. This would ultimately push content hosts to better regulate their practices.

Federal deepfake pornography legislation must be enforced through bottom-line control, meaning legislation must specifically impose vicarious liability on content hosts who fail to comply with Safe Harbor requirements. In *Viacom*, the MPAA recognized the importance of vicarious liability in its brief supporting Viacom, stating “[t]he entire point of vicarious liability is to incentivize the party with the right and ability to supervise sites and facilities to exercise that right to uncover and remedy infringing conduct that otherwise would inure to its financial benefit.”¹⁶⁹

Unlike in the copyright context, imposing vicarious liability on deepfake porn content hosts neither stifles innovation nor unfairly affects the economic rights of certain rights holders. Instead, it prohibits content hosts from committing what many argue, and some have decided, should be a crime.¹⁷⁰ Vicarious liability indeed imposes significant responsibility on content hosts, but it is the standard most appropriate to compel preventive practice from an industry already rife with bad practices as well as to ensure victims can fully recover if a lawsuit is brought.

Finally, to successfully adapt the DMCA’s general legislative framework to combat deepfake pornography, careful drafting is required. To begin, federal legislation must officially and narrowly define the term “deepfake pornography” to preserve its legal meaning and legitimacy.¹⁷¹ Though AI generally produces the most realistic videos and photos, the technology, sources, and processes behind

167. Jonathan Bailey, *5 Ways to Improve DMCA Safe Harbor*, PLAGIARISM TODAY (Aug. 5, 2008), <https://www.plagiarismtoday.com/2008/08/05/5-ways-to-improve-dmca-safe-harbor/>.

168. See *How to Report Revenge Porn on Social Media*, C.A. GOLDBERG, PLLC, <https://www.cagoldberglaw.com/how-to-report-revenge-porn-on-social-media/> (last visited July 18, 2022).

169. Brief for Motion Picture Ass’n of Am. & Indep. Film & Television All. as Amici Curiae Supporting Appellants at 6, *Viacom Int’l, Inc. v. YouTube, Inc.*, 676 F.3d 19 (2d Cir. 2012) (No. 10-3270).

170. See Tamsin Selbie & Craig Williams, *Deepfake Pornography Could Become an ‘Epidemic’, Expert Warns*, BBC NEWS (May 27, 2021), <https://www.bbc.com/news/uk-scotland-57254636>; Delfino, *supra* note 66, at 904–09.

171. Caldera, *supra* note 65, at 197–98 (using the example of “fake news” to illustrate that terms can reach a point where they lose meaning).

creating deepfakes can involve a wide range of processes.¹⁷² Therefore, federal legislation must put forth a simple and all-encompassing definition. There are several academic resources¹⁷³ and enacted statutes¹⁷⁴ that provide excellent guidelines.

Creating causes of action and defining economic penalties are also essential for this legislative framework to function properly. The earliest proponents of privacy rights recognized all cases involving violations of privacy should include an action for damages and an injunction.¹⁷⁵ Thus, federal deepfake pornography legislation must explicitly create a cause of action for victims to bring against distributors of deepfake content. Language describing the cause of action should be based on simple elements with clear definitions.¹⁷⁶ This could be brought against individuals, but also imputed to content hosts through vicarious liability. To substantially incentivize content hosts to monitor their business practices through vicarious liability, federal legislation should specifically enumerate statutory damages.¹⁷⁷

Considered altogether, these adaptations create a legislative framework that can successfully combat the deepfake pornography epidemic. In summation, creating a Safe Harbor with stringent qualifications, and including a low standard of knowledge to disqualify immunity, creates incentives for content hosts to comply with the statute. Requiring a Notice and Takedown procedure that is easy to use and accessible for complainants ensures victims can quickly flag content for removal, and is itself a powerful tool for curbing the spread of deepfake pornography. Failing to follow these procedures would expose content hosts to vicarious liability. As vicariously liable defendants, content hosts would be forced to absorb the steep costs of litigation from the legislation's causes of action and statutorily enumerated damages.

The general effect of this legislative framework is twofold. First, as looming litigation will impose substantial costs to mainstream content hosts, cost

172. *Id.* at 199; PARIS & DONOVAN, *supra* note 77, at 11.

173. *See, e.g.*, Chesney & Citron, *supra* note 3, at 1789; MARY ANNE FRANKS, CYBER C.R. INITIATIVE, DRAFTING AN EFFECTIVE “REVENGE PORN” LAW: A GUIDE FOR LEGISLATORS 6–12 (2021), <https://cybercivilrights.org/wp-content/uploads/2021/10/Guide-for-Legislators-10.21.pdf>.

174. *See, e.g.*, CAL. CIV. CODE § 1708.86(a)(6) (West 2020) (“‘Digitization’ means to realistically depict any of the following: (A) The nude body parts of another human being as the nude body parts of the depicted individual. (B) Computer-generated nude body parts as the nude body parts of the depicted individual. (C) The depicted individual engaging in sexual conduct in which the depicted individual did not engage.”); *see also* Selbie & Williams, *supra* note 170 (noting that Scotland prohibits “any image altered in any way”).

175. Warren & Brandeis, *supra* note 151, at 219.

176. *See* FRANKS, *supra* note 173, at 6–8.

177. *See* Van Lier, *supra* note 13, at 20 (explaining California’s deepfake pornography statute provides statutory damages as high as \$150,000).

internalization will compel content hosts to invest in infrastructure to catch and prevent deepfake pornography from ever disseminating in the first place. Additionally, this legislative framework will cause dedicated independent pornographic content to be litigated out of existence and substantially deter similar platforms from re-emerging.

C. Obstacles

Though two potential obstacles stand in the way of this federal legislative framework, they are likely surmountable. First, some scholars argue that prohibiting deepfake pornography distribution is a violation of freedom of expression protected by the First Amendment,¹⁷⁸ but careful drafting can accommodate free speech concerns. Second, though the current interpretation of Section 230 of the Communications Decency Act (CDA) grants significant protection to content hosts,¹⁷⁹ its power can be diminished.

The constitutionality of deepfake pornography creation is a complicated and unanswered First Amendment question.¹⁸⁰ This Note does not attempt to resolve that question, but rather argues that federal legislation using a vicarious liability model should be adopted to curtail its distribution. Still, with the proposed framework's parallels to the DMCA, some may be concerned that such a framework would invite abusive litigation targeted at suppressing free speech.¹⁸¹ To address this concern, federal deepfake pornography legislation should include a carefully crafted anti-SLAAP provision to deter frivolous lawsuits.¹⁸²

As for Section 230 of the CDA, contextualizing its purpose and history demonstrates it is not impossible to overcome. The statute was originally enacted in 1996 to encourage internet service providers to monitor their own content and ensure the internet developed as a “clean” environment.¹⁸³ Like much of the legislation from that area, it was designed to allow the internet to grow and become

178. See Waldstreicher, *supra* note 47, at 759; see also MATTHEW FEENEY, REGUL. TRANSP. PROJECT, DEEPFAKE LAWS RISK CREATING MORE PROBLEMS THAN THEY SOLVE 6–11 (2021), <https://rtp.fedsoc.org/wp-content/uploads/Paper-Deepfake-Laws-Risk-Creating-More-Problems-Than-They-Solve.pdf>.

179. 47 U.S.C. § 230.

180. See Waldstreicher, *supra* note 47, at 745–47; see also Lindsey Wilkerson, *Still Waters Run Deep(fakes): The Rising Concerns of Deepfake Technology and its Influence on Democracy and the First Amendment*, 86 MO. L. REV. 407, 414–20 (2021).

181. See Waldstreicher, *supra* note 47, at 752–53; see generally *Takedown Hall of Shame*, ELECTRONIC FRONTIER FOUND., <https://www EFF.ORG/takedowns> (last visited July 19, 2022).

182. See Chesney & Citron, *supra* note 3, at 1800.

183. See Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 FORDHAM L. REV. 401, 406 (2017); Chesney & Citron, *supra* note 3, at 1796–97.

the commercial and cultural leviathan it is today.¹⁸⁴ However, the internet has grown beyond the statute's original intent. Interpretation of Section 230 has developed it into an almost impenetrable immunity shield for content hosts. Essentially, Section 230 legally prevents any individual from suing an internet platform for its users' crimes or torts.¹⁸⁵ This "super-immunity" allows content hosts to escape liability when it comes to attacks on sexual privacy in the digital age and provides them no incentive to improve their moderation practices.¹⁸⁶ It essentially gives the content hosts all the power, yet none of the responsibility.¹⁸⁷

Section 230 must be amended before federal deepfake pornography legislation using a vicarious liability can be enacted. Though it seems ambitious, it has been done before. To protect individuals from sexual exploitation, Congress amended Section 230 on two separate occasions in recent years.¹⁸⁸ The two separate bills, aimed at combatting sex trafficking, create exceptions to Section 230 immunity where platforms host advertisements for sex work.¹⁸⁹ In this case, bottom-line control took effect and the private sector responded swiftly. Reddit banned multiple sex-work forums and Craigslist removed its "personals" feature in response to the possible liability.¹⁹⁰ Though Section 230 serves an important role in promoting market competition and protecting free speech,¹⁹¹ it clearly can and should be reasonably amended to protect the individual liberties violated by deepfake pornography.¹⁹²

Enacting federal legislation on cutting-edge issues is an arduous task. However, in the case of federally legislating to fight deepfake pornography, the potential obstacles can be overcome.

CONCLUSION

Deepfake pornography can lead to devastating impacts on nearly every facet of a victim's life. While it already can happen to anyone, the ever-increasing

184. See Citron & Wittes, *supra* note 181, at 407–08.

185. See 47 U.S.C. § 230(c)(1) ("No 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.").

186. Chesney & Citron, *supra* note 3, at 1798.

187. See Rebecca Tushnet, *Power Without Responsibility: Intermediaries and the First Amendment*, 76 GEO. WASH. L. REV. 986, 994–96, 1002–03 (2008).

188. See Waldstreicher, *supra* note 47, at 742.

189. 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>.

190. *Id.*; see Waldstreicher, *supra* note 47, at 743.

191. See FEENEY, *supra* note 178, at 11.

192. See Citron & Wittes, *supra* note 181, at 419–21.

capabilities of AI will allow victimization to become commonplace before long. As we spend more time in front of computers and smart devices, the line between true reality and its postproduction counterpart blurs. We must proactively combat a future where victims are left without remedy and perpetrators go free of consequences.

We have the capability to do so. The law can place the burden on the content hosts—those most apt to curtail the distribution of deepfake pornography—by threatening their bottom-line. Federal legislation can create a framework that imposes vicarious liability on content hosts who do not follow strict statutory procedures, mandate takedown systems that allow victims to quickly address violations, and most importantly, incentivize content hosts to change their infrastructure to quickly remove deepfake content. The DMCA illustrates the success of this model in its ability to change entire industries. We can learn from the DMCA's shortcomings and create federal legislation which aims to protect individual liberties. In doing so, we can successfully combat deepfake pornography's egregious social harms and more securely enter the next generation of the internet.