

Deepfakes, Deeper Issues: Moderating Explicit Deepfake Content Through a Federal Right of Publicity Exception

Nako Caternor*

TABLE OF CONTENTS

I.	INTRODUCTION	307
II.	BACKGROUND.....	308
	<i>A. Deepfakes</i>	308
	<i>B. The Dark Side: Exploitation to Create Explicit Content</i>	309
	<i>C. A New Nightmare: False Explicit Images</i>	310
	1. Use for Sextortion	310
	2. Use for Revenge, Humiliation, & Degradation of Reputation	311
	3. Effects on Employment, Mental Health, and Personal Dignity.....	312
III.	CURRENT LEGAL LANDSCAPE.....	313
	<i>A. Traditional Torts</i>	314
	<i>B. State Legislation</i>	315
	<i>C. Shortcomings of Available Legal Remedies</i>	316
IV.	ISSUES WITH CONTENT MODERATION.....	317

* J.D., May 2026, The George Washington University Law School; B.S. 2020 American Politics, The United States Military Academy-West Point. Praise God for what He has done! I am incredibly thankful to Professor Dawn Carla Nunziato whose encouraging mentorship during the topic selection and writing process was vital in bringing this Note into fruition (and advancing my curiosity for First Amendment jurisprudence). I also thank the Federal Communications Law Journal Editorial Board for their dedication and diligence to prepare this piece for publication. A special shout out to my amazing friends, especially the gals who love a cute brunch or a chill time spent together. Last but certainly not least, to my family—your abundant love, jokes, mom's delicious cooking, and encouragement have supported me through these years!

A.	<i>Communications Decency Act</i>	317
B.	<i>Safe Harbor Provision</i>	319
C.	<i>Broad Interpretation of Section 230</i>	320
V.	ANALYSIS	322
A.	<i>Right of Publicity Law</i>	322
1.	Circuit Split on CDA Section 230	323
2.	Right of Publicity is an Intellectual Property	324
B.	<i>A (Narrow) Federal Right of Publicity</i>	325
C.	<i>Digital Millennium Copyright Act</i>	327
1.	Incorporating the Right of Publicity into DMCA Through Amendment	328
2.	Challenges to Right of Publicity Framework	329
VI.	CONCLUSION.....	330

I. INTRODUCTION

In late January 2024, the social media platform X was flooded with sexually explicit images of Grammy-award winning singer and songwriter Taylor Swift.¹ These graphic images were later identified as sexually explicit images generated by artificial intelligence (AI).² Although disturbing, one such image shared by an X account garnered 47 million views before the account was suspended.³ Fans of Taylor Swift quickly mobilized and flooded X with positive images and videos of Swift under the hashtag #ProtectTaylorSwift in an ardent effort to make the disturbing images harder to find.⁴ This disturbing use of AI is unfortunately not new; rather it is a more recent example of digital sexual exploitation that has affected many people, including actress Scarlett Johansson six years prior in 2018.⁵ But make no mistake, this abusive use of technology does not exclusively affect celebrities; it can and has been used to create false sexual images of even non-celebrities using just their face.⁶

This Note argues for a federal right of publicity automatically granted upon birth that allows a person to control the use of their image or likeness, exclusively in the context of explicit deepfake content. Incorporating this right into the Digital Millennium Copyright Act would give victims a mechanism to remove this content online. As an intellectual property right, the right of publicity falls out of the safe harbor provision of Section 230 of the Communications Decency Act, allowing victims an avenue to hold platforms liable for this content if they refuse to remove it.⁷ This Note will examine a brief history about deepfakes and how they have been disproportionately used to create sexual content that impacts women more significantly than men. This Note also explores the damaging effects of explicit deepfake content and the current legal remedies for victims. This Note will then discuss the lack of mechanisms to force content takedown on platforms and the significant barrier Section 230 of the Communications Decency Act poses for content moderation. This Note finally discusses how

1. See Jess Weatherbed, *Trolls have flooded X with graphic Taylor Swift AI fakes*, THE VERGE (Jan. 25, 2015, at 11:04 EST), <https://www.theverge.com/2024/1/25/24050334/x-twitter-taylor-swift-ai-fake-images-trending> [<https://perma.cc/62EH-4RMM>].

2. See Katie Conger & John Yoon, *Explicit Deepfake Images of Taylor Swift Elude Safeguards and Swamp Social Media*, N.Y. TIMES (Jan. 25, 2024), <https://www.nytimes.com/2024/01/26/arts/music/taylor-swift-ai-fake-images.html>.

3. See *id.*

4. See *id.*; see also *Taylor Swift deepfakes spread online, sparking outrage*, CBS NEWS (Jan. 26, 2024, at 21:32 EST), <https://www.cbsnews.com/news/taylor-swift-deepfakes-online-outrage-artificial-intelligence> [<https://perma.cc/B3SV-GQQH>].

5. See Ben Beaumont-Thomas, *Taylor Swift deepfake pornography sparks renewed calls for US legislation*, THE GUARDIAN (Jan. 26, 2024, at 08:44 EST), <https://www.theguardian.com/music/2024/jan/26/taylor-swift-deepfake-pornography-sparks-renewed-calls-for-us-legislation> [<https://perma.cc/66KT-5DPE>].

6. See Jake Sturmer & Allison Bradley, *Noelle Martin fights to have harmless selfie removed from 'parasite' porn sites*, ABC NEWS (Oct. 12, 2016, at 07:45 EST), <https://www.abc.net.au/news/2016-10-12/womans-fight-to-have-harmless-selfie-removed-from-porn-site/7924948> [<https://perma.cc/RDR3-A62Q>].

7. See discussion *infra* Section V.A.2.

the proposed solution solves the issue of content moderation that existing remedies are unable to address and potential challenges that may be brought up against the solution.

II. BACKGROUND

A. Deepfakes

A deepfake is a type of fabricated media “where a person in an image or video is swapped with another person’s likeness.”⁸ This term originated from a Reddit user called “deepfakes” in 2017.⁹ The term deepfake is a portmanteau of deep learning, which is the type of AI used to create this content, and fake.¹⁰ Deep learning is the underlying process in which technology uses large sets of data to solve problems and is often used to manipulate and create false content using images of real people.¹¹ From this, a computer will generate content from existing images that the computer has been trained on.¹² Deep learning uses artificial neural networks to mimic the complex decision-making of the human brain.¹³ The network model takes in data for processing to recognize, classify, and describe content within the data and over time becomes more accurate.¹⁴

One common type of framework is a generative adversarial network (GAN).¹⁵ Within a GAN, there are two networks that compete against the other: the generator tries to produce synthetic data that looks real and the discriminator tries to classify the data as either real or synthetic.¹⁶ The generator uses feedback from the discriminator after each round to adjust its parameters and increase the accuracy of the synthetic data.¹⁷ This competition

8. See Meredith Somers, *Deepfakes, explained*, MIT SLOAN (July 21, 2020), <https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained> [https://perma.cc/J2S5-384F].

9. See *id.*

10. See James Vincent, *Why we need a better definition of ‘deepfake’*, THE VERGE (May 22, 2018, at 14:53 EDT), <https://www.theverge.com/2018/5/22/17380306/deepfake-definition-ai-manipulation-fake-news> [https://perma.cc/A3N5-GLQA].

11. See Dave Johnson & Alexander Johnson, *What are deepfakes? How fake AI-powered audio and video warps our perception of reality*, BUS. INSIDER (June 15, 2023, at 10:58 EDT), <https://www.businessinsider.com/guides/tech/what-is-deepfake> [https://perma.cc/TC3K-BHWE].

12. *Id.*

13. See *What is deep learning?*, IBM (June 17, 2024), <https://www.ibm.com/topics/deep-learning> [https://perma.cc/5D87-GBBP].

14. *Id.*

15. See Scott Robinson, Kinza Yasar & Sarah Lewis, *What is a generative adversarial network (GAN)?*, TECHTARGET (Oct. 18, 2024), <https://www.techtarget.com/searchenterpriseai/definition/generative-adversarial-network-gan> [https://perma.cc/AEY5-HKRV].

16. *Id.*

17. *Id.* See also Shipra Garg, *What is a GAN? – Generative Adversarial Networks Guide*, SOLULAB (Oct. 17, 2024), <https://www.solulab.com/generative-adversarial-network/> [https://perma.cc/M3NC-K742].

continues until the discriminator is unable to distinguish the synthetic data from the generator and real samples.¹⁸

Because the goal of this framework is to generate synthetic data that is almost impossible to distinguish as fake, explicit deepfake content will only become harder to separate from authentic images.¹⁹ One early example of how effective deepfake technology is in creating highly realistic content was a video of President Obama appearing to call President Trump derogatory words.²⁰ This video went viral and brought attention to the potential misuse of deepfake technology to mislead people.²¹ Deepfake content has significantly increased over the past several years, with a 550% increase in deepfake videos from 2019 to 2023.²² Deepfake content will likely continue to increase at an exponential rate as the technology becomes more accessible.²³

B. The Dark Side: Exploitation to Create Explicit Content

The Reddit user “deepfake” came into prominence in 2017 because he had created a space on the website to share explicit deepfake videos that used technology to swap faces on the videos.²⁴ Deepfake pornography makes up 98% of all deepfake videos.²⁵ The unfortunate reality is that deepfake pornography disproportionately affects women more than men.²⁶ In 2023, 99% of the individuals whose image was used for deepfake pornography were women.²⁷ Entire computer applications have been developed solely for the purpose to artificially “undress” and expose the nude body of women’s images.²⁸ These applications take images of innocent women and artificially generate nude images of the women without their consent.²⁹ One such website called DeepNude was launched June 23, 2019 and allowed users to generate images of an “undressed” woman in 30 seconds for free but required a \$50

18. See, *What is a GAN?*, AWS, <https://aws.amazon.com/what-is/gan/> [<https://perma.cc/G5FP-8HUX>].

19. See Don Philmlee, *Practice Innovations: Seeing is no longer believing — the rise of deepfakes*, THOMSON REUTERS (July 18, 2023), <https://www.thomsonreuters.com/en-us/posts/technology/practice-innovations-deepfakes/> [<https://perma.cc/DF3A-62PW>].

20. See Kaylee Fagan, *A viral video that appeared to show Obama calling Trump a 'dips--' shows a disturbing new trend called 'deepfakes'*, BUS. INSIDER (Apr. 17, 2018, at 16:48 EDT), <https://www.businessinsider.com/obama-deepfake-video-insulting-trump-2018-4> [<https://perma.cc/BUD8-8SPT>].

21. See *id.*

22. See *2023 State of Deepfakes*, SEC. HERO, <https://www.securityhero.io/state-of-deepfakes/#deepfake-porn-survey> [<https://perma.cc/N8GS-EEVY>].

23. See Rick Hutchinson, *Deepfakes Are the New Frontier Of Cyberattacks: 5 Steps to Fight Back*, FORBES (Jan. 6, 2025, at 08:45 EST), <https://www.forbes.com/councils/forbestechcouncil/2025/01/06/deepfakes-are-the-new-frontier-of-cyberattacks-5-steps-to-fight-back/> [<https://perma.cc/5XEE-4ZJK>].

24. See Somers, *supra* note 8.

25. See *id.*

26. See *2023 State of Deepfakes*, *supra* note 22.

27. *Id.*

28. See Henry Ajder et al., *The State of Deepfakes: Landscape, Threats, and Impact*, DEEPTRACE, 8 (2019) [<https://perma.cc/EC3B-JDL7>].

29. See *id.*

fee to remove the watermark.³⁰ Of significance is that this application could only “undress” women; there was no similar algorithm to “undress” men.³¹ Additionally, this website received over 545,000 visits and over 95,000 active users within seven days of launching; a majority of these visits likely occurred within 24 hours of launching which caused the website to go offline.³² On Telegram, the messaging application, there are at least 50 different bots that create similar explicit images and videos.³³ These bots combined have over 4 million monthly users using this technology to generate explicit content.³⁴ All a user needs to do is upload a photo of a woman and they will receive a fake nude image or video within minutes.³⁵ These applications reveal how easy it is now to create explicit deepfake content and how high the appetite is to create this type of content.

C. *A New Nightmare: False Explicit Images*

1. Use for Sextortion

The FBI released a public service announcement in 2023 highlighting concerns and the dangerous uses of deepfake content.³⁶ One of the harmful uses that they had observed an increase in was sextortion.³⁷ These victims reported that they were blackmailed with explicit deepfake images or videos created from their photos posted online or on social media.³⁸ The offenders would then demand either payment or for victims to send real sexual content, threatening to share the deepfake content with family or friends if the demand was not met.³⁹ Lauren Book, a Florida state senator, experienced this when she was sent explicit images of herself that included distorted and fake images.⁴⁰ The offender threatened to release the photos to the public and ruin her political career if she did not pay them \$5000.⁴¹ She began to cooperate with law enforcement, who were able to locate the offender under the pretext

30. *Id.*

31. *Id.*

32. *Id.*

33. Sammi Caramela, ‘Nudify’ Deepfake Bots on Telegram Are Up to 4 Million Monthly Users, VICE (Oct. 16, 2024, at 09:38 EDT), <https://www.vice.com/en/article/nudify-deepfake-bots-telegram/> [https://perma.cc/B36Q-P8BT].

34. *See id.*

35. *See id.*

36. *Malicious Actors Manipulating Photos and Videos to Create Explicit Content and Sextortion Schemes*, FBI (June 5, 2023), <https://www.ic3.gov/PSA/2023/PSA230605> [https://perma.cc/9QZ2-X8GV].

37. *See id.*

38. *See id.*

39. *See id.*

40. *Man enters no contest plea in Florida Sen. Lauren Book extortion case*, ASSOCIATED PRESS (June 16, 2022), <https://www.wtsp.com/article/news/regional/florida/no-contest-plea-lauren-book-extortion/67-a7f993c9-bf33-4b86-af6c-4e2f96d6b7d5> [https://perma.cc/VF4X-HS6E].

41. *Id.*

of paying the ransom.⁴² As a brave politician with the resources to hire an attorney and trust to seek help, Lauren Book was able to successfully bring her extorter to justice; this unfortunately may not be the case for many others with similar threats.⁴³

2. Use for Revenge, Humiliation, & Degradation of Reputation

Explicit deepfake content can also be utilized to create false non-consensual intimate images of former partners.⁴⁴ No longer do former partners need actual nude images to ruin a person's reputation: they can simply create it themselves. Explicit deepfake content has also been used as a tool to humiliate and degrade the reputation of women.⁴⁵ Investigative journalist Rana Ayyub became a victim of a barrage of explicit deepfake content after speaking out against a political party in India protecting a child sex abuser.⁴⁶ A video was generated placing her face on a pornographic video and circulated on messaging platform Whatsapp.⁴⁷ Ayyub experienced a large volume of harassment on Facebook, Twitter, and Instagram and later her phone after her number was released alongside the content.⁴⁸ Her video was additionally shared over 40,000 times from a fan page of the political party leader she had spoken out against.⁴⁹ This attack came as a way to discredit and attack Ayyub's character by humiliating her and ruining her reputation on a national scale.⁵⁰

42. See *id.*; see *For-Profit Companies Charging Sextortion Victims for Assistance and Using Deceptive Tactics to Elicit Payments*, FBI (Apr. 7, 2023), <https://www.ic3.gov/PSA/2023/PSA230407> [<https://perma.cc/LSP8-QQKS>] (some sextortion victims have sought services from companies that advertise providing "assistance." Victims are charged high fees and some reporting shows that some of these companies are involved in the sextortion activity).

43. See *Florida senator extorted over nude photos is taking legislative action to protect others*, ASSOCIATED PRESS (Jan. 25, 2022, at 08:42 EST), <https://www.foxnews.com/us/florida-senator-extorted-nude-photos-protect-others> [<https://perma.cc/S6U5-AYLV>] [hereinafter *Florida senator extorted*].

44. See Tim Mak & Dina Temple-Raston, *Where Are The Deepfakes In This Presidential Election?*, NPR (Oct. 1, 2020, at 05:05 EDT), <https://www.npr.org/2020/10/01/918223033/where-are-the-deepfakes-in-this-presidential-election/> [<https://perma.cc/K2VY-JVY6>].

45. See Shona Moreau & Chloe Rourke, *Fake porn causes real harm to women*, POL'Y OPTIONS (Feb. 8, 2024), <https://policyoptions.irpp.org/magazines/february-2024/fake-porn-harm/> [<https://perma.cc/9XPT-TYTH>].

46. See Rana Ayyub, *I Was The Victim Of A Deepfake Porn Plot Intended To Silence Me*, HUFFINGTON POST UK, (Nov. 21, 2018, at 08:11 GMT), https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316 [<https://perma.cc/ZK3F-FXDL>].

47. *Id.*

48. *Id.*

49. *Id.*

50. See *id.*

3. Effects on Employment, Mental Health, and Personal Dignity

Employers will often make hiring decisions based on what they may find online about a potential candidate.⁵¹ A study of over 2,300 hiring managers and human resources professionals found almost 70% of employers use search engines to conduct research on a potential candidate.⁵² More than 50% of employers found content on a candidate's social media that caused them to not hire them.⁵³ Of the employers who passed on a hire, 39% did not hire because they had found inappropriate or provocative content of the candidate.⁵⁴ An explicit deepfake image or video could potentially ruin a person's chance of securing employment.⁵⁵ An employer may have concern or hesitation about a candidate's character or abilities.⁵⁶ And for people who are already employed, explicit deepfake content can cost them their job.⁵⁷ Although a worker or candidate's personal sexual life should not impact their eligibility for hire, many employers still will hesitate based on the concern that their reputation would be risky for their business.⁵⁸

Explicit deepfake content can have damaging effects to women "by causing psychological trauma and feelings of humiliation, fear, embarrassment, and shame."⁵⁹ Explicit deepfake content can also have devastating psychological effects on victims.⁶⁰ Some victims have reported suffering from anxiety and panic attacks from the false intimate exposure of their body.⁶¹ One young girl experienced a fake nude image of herself

51. See *Number of Employers Using Social Media to Screen Candidates at All-Time High, Finds Latest CareerBuilder Study*, PR NEWSWIRE (June 15, 2017, at 03:02 EDT), <https://www.prnewswire.com/news-releases/number-of-employers-using-social-media-to-screen-candidates-at-all-time-high-finds-latest-careerbuilder-study-300474228.html> [<https://perma.cc/PM42-EM9J>] [Hereinafter *Number of Employers*].

52. *Id.*

53. *Id.*

54. *Id.*

55. See Danielle Keats Citron, *Sexual Privacy*, 128 YALE L.J. 1870, 1927-28 (2019); see generally Coralie Kraft, *Trolls Used Her Face to Make Fake Porn. There Was Nothing She Could Do.*, N.Y. TIMES MAG. (July 31, 2024), <https://www.nytimes.com/2024/07/31/magazine/sabrina-javellana-florida-politics-ai-porn.html> (After explicit deepfake images of her were shared online, Sabrina Javellana decided against taking the state teaching-certification exam knowing the risk of damaging a future school's reputation).

56. See *id.*

57. See Victoria Rousav, *Sexual Deepfakes and Image-Based Sexual Abuse: Victim-Survivor Experiences and Embodied Harm* 107 (May 2023) (M.A. thesis, Harvard University), https://www.researchgate.net/publication/370324689_Sexual_Deepfakes_and_Image-Based_Sexual_Abuse_Victim-Survivor_Experiences_and_Embodied_Harms [<https://perma.cc/69R4-6GM5>].

58. See Citron, *supra* note 55, at 1927.

59. See Caramela, *supra* note 33.

60. See Sophie Compton & Reuben Hamlyn, *Opinion: The rise of deepfake pornography is devastating for women*, CNN (Oct. 29, 2023, at 12:07 EDT), <https://www.cnn.com/2023/10/29/opinions/deepfake-pornography-thriving-business-compton-hamlyn/index.html> [<https://perma.cc/J8V4-DE4B>].

61. *Id.*

circulating amongst her peers at school.⁶² This image exacerbated the bullying she was already experiencing at school until she tragically took her life; she was only 15 years old.⁶³ Although a victim's sexual privacy is not technically violated per se, explicit deepfake content intrudes upon and violates a person's sexual identity.⁶⁴ Explicit deepfake images exercise power over a person's sexuality, create a sexual identity not of that person's curation, and exhibit this false identity without that person's consent.⁶⁵

III. CURRENT LEGAL LANDSCAPE

Currently, there is no federal legislation enacted that prohibits or regulates deepfake content.⁶⁶ Traditional torts and efforts by states to create legislation may provide avenues for victims to hold the content creator liable.⁶⁷ However they do not provide a way to remove the content online or to hold platforms liable for refusing to remove the content. Federal legislation concerning deepfakes has been discussed, but multiple attempts at bills have so far failed.⁶⁸ One of the first attempts was the DEEP FAKES Accountability Act of 2019 which would require watermarks on all deepfake content and impose criminal penalties for creation of content that was sexual, intended to incite or cause physical harm, or used in the criminal conduct related to fraud.⁶⁹ This bill did not receive a vote and was never enacted.⁷⁰ In the summer of 2024, the Disrupt Explicit Forged Images and Non-Consensual Edits Act of 2024 (DEFIANCE Act) passed the Senate.⁷¹ This pending legislation would provide a civil remedy for victims against offenders who create or possess with intent to distribute or offenders who distribute if they knew or disregarded the lack of consent from the victim.⁷² In the absence of federal legislation, victims may turn to traditional torts or state legislation to pursue legal remedies.

62. See Isabella Ross, *Social Media Summit unpacks impact of explicit deepfake image abuse on young girls*, ABC NEWS (Oct. 12, 2024, at 17:13 EDT), <https://www.abc.net.au/news/2024-10-13/social-media-deepfake-abuse-impact-on-kids/104457222> [<https://perma.cc/6KSQ-XT4R>].

63. *Id.*

64. See Citron, *supra* note 55, at 1921.

65. See *id.*

66. See Michelle M. Graham, *Deepfakes: Federal and state regulation aims to curb a growing threat*, THOMSON REUTERS (June 26, 2024), <https://www.thomsonreuters.com/en-us/posts/government/deepfakes-federal-state-regulation/> [<https://perma.cc/RGU8-AHJQ>].

67. See *infra* note 102 and Section III.A.

68. See Andrew R. Chow, *Congress May Finally Take on AI in 2025. Here's What to Expect*, TIME (Dec. 23, 2024, at 07:00 EST), <https://time.com/7203040/congress-ai-preview-2025/> [<https://perma.cc/Z6KT-RWCM>].

69. See DEEP FAKES Accountability Act, H.R. 3230, 116th Cong. (2019).

70. See *id.*

71. See Kat Tenbarge, *The Defiance Act passes in the Senate, potentially allowing deepfake victims to sue over nonconsensual images*, NBC NEWS (July 24, 2024, at 15:28 EDT), <https://www.nbcnews.com/tech/tech-news/defiance-act-passes-senate-allow-deepfake-victims-sue-rcna163464> [<https://perma.cc/9653-PN7N>].

72. See *id.*

A. Traditional Torts

Victims of explicit deepfake content may turn to traditional torts to pursue remedies against their offenders. Through the tort of Intentional Infliction of Emotional Distress (IIED) a victim can hold an offender liable for emotional distress caused by their extreme and outrageous conduct.⁷³ The Restatement poses a closely analogous illustration where someone is held liable for the shame, embarrassment, and humiliation that a woman feels after they caused her to appear nude in public amongst people she was not familiar with.⁷⁴ In a similar manner a creator of explicit deepfake content that posts online causes a person to be “naked” amongst strangers and may cause them similar humiliation and embarrassment. One issue a victim may face in bringing up this claim is whether posting such content will constitute extreme and outrageous behavior.⁷⁵ IIED has been successfully used in the context of non-consensual pornography.⁷⁶ One potential gap for victims filing such a claim is whether their state considers explicit deepfake content as outrageous and proving their emotional distress in connection to the content.

Victims may also seek recourse through defamation. A communication is considered defamatory if it detrimentally harms their reputation or causes others to avoid association.⁷⁷ Due to the sensitive nature that a person’s sexual reputation or perception affects how other people or society view them, explicit deepfake content would meet this classification.⁷⁸ A victim may also seek legal remedies through the tort of false light.⁷⁹ An offender can be held liable for giving publicity that places another in a false light if (1) the false light would be “highly offensive to a reasonable person” and (2) the offender knew or “acted in reckless disregard” about the falsity of the publicity and false light that the other person would be placed in.⁸⁰ The first requirement would likely be satisfied because serious offense is reasonably expected by a reasonable person when there is a “major misrepresentation of [their] character,... activities or beliefs.”⁸¹ Explicit deepfake content can create a false representation of a person’s personal beliefs concerning sexual activity and portray victims engaging in activities they did not actually partake in.⁸² The second requirement would also be satisfied because an offender would have knowledge that they were purposely creating a false sexual image of the

73. RESTATEMENT (SECOND) OF TORTS § 46 (AM. L. INST. 1965).

74. RESTATEMENT (SECOND) OF TORTS § 46, illus. 3 (AM. L. INST. 1965).

75. RESTATEMENT (SECOND) OF TORTS § 46, cmt. d (AM. L. INST. 1965).

76. See *K.I. v. Tyagi*, No. 1:23-2383-JRR, 2024 U.S. Dist. LEXIS 203666, at *30–31 (D. Md. Nov. 8, 2024) (explaining that New York courts have found public disclosures or threats of publicly disclosing sexually explicit content of a person without their consent constitutes extreme and outrageous conduct). See *Taylor v. Franko*, No. 09-00002 JMS/RLP, 2011 U.S. Dist. LEXIS 75128, at *13–14 (D. Haw. July 11, 2011) (awarding plaintiff \$425,000 for her claims of IIED, NIED, defamation, and public disclosure of private facts against defendant who posted nude photos of her online).

77. RESTATEMENT (SECOND) OF TORTS § 559 (AM. L. INST. 1977).

78. See Rousay, *supra* note 57.

79. See RESTATEMENT (SECOND) OF TORTS § 652E (AM. L. INST. 1977).

80. See RESTATEMENT (SECOND) OF TORTS § 652E(a)–(b) (AM. L. INST. 1977).

81. See *id.* at cmt. c.

82. See Citron, *supra* note 55, at 1921–22.

victim by incorporating nude body parts that are not of the victim.⁸³ A potential issue and barrier for filing a claim is proving the element of publicity.⁸⁴ The Restatement defines publicity as a matter that is communicated to “the public at large, or to so many persons that the matter must be regarded as substantially certain to become one of public knowledge.”⁸⁵ As a result, it is not an invasion of a victim’s privacy if explicit deepfake content is communicated to a small group of people.⁸⁶ A victim may then have to prove that enough people had viewed this explicit deepfake content to overcome this barrier. However, a “publication in a newspaper or a magazine, even of small circulation,” is sufficient to meet this publicity requirement.⁸⁷ If a court determines that posting explicit deepfake content on a website, even with limited viewership, is a fitting analogue, a victim may not have to prove actual viewership.

B. State Legislation

Current legislative proposals concerning deepfake content are all at the state level and some of them address deepfake content used to spread disinformation about a candidate during an election.⁸⁸ Several states have also taken initiative to enact legislation to grant victims a remedy against offenders who generate explicit deepfake content using their image.⁸⁹ California’s explicit deepfake content statute allows a victim to hold a person liable if they create or intentionally distribute explicit deepfake content of them without consent.⁹⁰ This statute also increases the maximum level of damages awarded if the victim can prove the creation or distribution was conducted with malice.⁹¹ Virginia updated its non-consensual pornography statute to prosecute persons who disseminate explicit deepfake content of another person.⁹² This amendment specifies “another person” includes “a person whose image was used in creating, adapting or modifying a videographic or still image with the intent to depict an actual person and who is recognizable as an actual person by the person’s face, likeness or other distinguishing characteristic.”⁹³

83. See Caramela, *supra* note 33. See generally Ajder, *supra* note 28.

84. See RESTATEMENT (SECOND) OF TORTS § 652E(a)–(b) (AM. L. INST. 1977).

85. RESTATEMENT (SECOND) OF TORTS § 652D, cmt. a (AM. L. INST. 1977).

86. See RESTATEMENT (SECOND) OF TORTS § 652D, cmt. a (AM. L. INST. 1977).

87. RESTATEMENT (SECOND) OF TORTS § 652D, cmt. a (AM. L. INST. 1977).

88. See *Tracker: State Legislation on Deepfakes in Elections*, PUBLIC CITIZEN, <https://www.citizen.org/article/tracker-legislation-on-deepfakes-in-elections/> [<https://perma.cc/4E3R-S5CL>] (As of January 22, 2025, 23 states have enacted legislation regulating the use of deepfake content in elections).

89. Vittoria Elliot, *The US Needs Deepfake Porn Laws. These States Are Leading the Way*, WIRED (Sep. 5, 2024, at 06:00 EDT), <https://www.wired.com/story/deepfake-ai-porn-laws/> [<https://perma.cc/YW5Q-FWBX>].

90. See CAL. CIV. CODE § 1708.86 (2025).

91. See *id.*

92. See Michael Grothaus, *Virginia Updates Its Revenge Porn Laws to Include Deepfakes*, FAST CO. (July 2, 2019), <https://www.fastcompany.com/90372079/virginia-updates-its-revenge-porn-laws-to-include-deepfakes>. See VA CODE § 18.2-386.2 (2019).

93. VA CODE § 18.2-386.2 (2019).

New York also amended its non-consensual pornography statute to prosecute producers of explicit deepfake content.⁹⁴ However, their statute still requires showing an intent to harm the victim emotionally, financially, or physically.⁹⁵ A defendant could possibly argue they never intended for the victim to view it and therefore there was no intent to cause that type of harm.

Kentucky incorporated deepfake provisions into a statute concerning sexual crimes against minors, making it a felony to produce explicit deepfake content of a minor.⁹⁶

Although many states have enacted some type of legislation concerning explicit deepfake content, there is great variety in what each statute requires for a claim and allows for remedies.⁹⁷ Some only apply to minors,⁹⁸ some allow a civil remedy,⁹⁹ and others only address prosecution.¹⁰⁰ Depending on where a victim lives, they may not have access to a civil remedy in state laws.¹⁰¹

C. Shortcomings of Available Legal Remedies

All these current remedies address holding persons who create the images accountable.¹⁰² This comes with the difficulties of having to identify and track down perpetrators online, which can be difficult, expensive, and potentially dangerous for victims.¹⁰³ Although victims with money and resources may find success in holding their offenders responsible,¹⁰⁴ the average victim does not and may be left without redress.¹⁰⁵

Current legislation also does not address how a victim may force the takedown of images from the sites. Even if a victim is able to hold the offender accountable for creating the content, the continual existence of the content

94. See S. B. S1042, 2023 Leg. Sess. (NY. 2023).

95. See N.Y. PENAL LAW §245.15 (2023).

96. See H.B. 207 2024 Gen. Assemb., Reg. Sess. (KY. 2024).

97. See e.g., N.Y. PENAL LAW §245.15 (2023); VA CODE § 18.2-386.2 (2019); H.B. 207 2024 Gen. Assemb., Reg. Sess. (KY. 2024); CAL. CIV. CODE § 1708.86.

98. See H.B. 207 2024 Gen. Assemb., Reg. Sess. (KY. 2024).

99. See CAL. CIV. CODE § 1708.86.

100. See e.g., N.Y. PENAL LAW §245.15 (2023); VA CODE § 18.2-386.2 (2019).

101. See Julia Scammahorn, *Missouri lawmakers consider 'Taylor Swift Act' to criminalize deepfake images of minors*, KCTV (Feb. 24, 2026, at 14:26 EST), <https://www.kctv5.com/2026/02/24/missouri-lawmakers-advance-taylor-swift-act-criminalize-deepfake-images-minors/> [<https://perma.cc/QG2F-A88T>] (Missouri is one of three states that do not have any legislation concerning the abuse of artificial intelligence such as deepfakes).

102. See e.g., N.Y. PENAL LAW §245.15 (2023); VA CODE § 18.2-386.2 (2019); H.B. 207 2024 Gen. Assemb., Reg. Sess. (KY. 2024); CAL. CIV. CODE § 1708.86.

103. See Laurie Segall, *Opinion: The Taylor Swift AI photos offer a terrifying warning*, CNN (Jan. 31, 2024, at 20:06 EST), <https://www.cnn.com/2024/01/31/opinions/taylor-swift-deepfakes-ai-segall/index.html> [<https://perma.cc/M45U-X57E>].

104. See *Florida senator extorted*, *supra* note 43.

105. See Segall, *supra* note 103.

online may cause the victim to repeatedly suffer mental harm and anguish.¹⁰⁶ As long as it exists online, anyone has the ability to see it.¹⁰⁷ One young woman attempted to have an explicit deepfake image removed from a website by contacting the website master and was instead blackmailed to send nude photos in exchange for the removal.¹⁰⁸

In short, a victim is at the mercy of a platform's goodwill to remove this content.¹⁰⁹ Some platforms have taken initiative to help fight against this.¹¹⁰ In 2023, Facebook and Instagram became founding members of the platform Take it Down, in collaboration with the National Center for Missing and Exploited Children.¹¹¹ This platform allows people under 18, or people whose images were taken when they were under 18, to "submit a case that will proactively search for their intimate images" on participating platforms.¹¹² This great initiative is limited in scope as only participating companies such as TikTok, YouTube, and Pornhub are moderating this content.¹¹³ Additionally, this tool is unavailable for persons whose image was taken at 18 or older.¹¹⁴ Attempts to moderate content on the Internet can be difficult and met with resistance.¹¹⁵

IV. ISSUES WITH CONTENT MODERATION

A. *Communications Decency Act*

Content moderation has been a hot topic since the growing popularity of the Internet in the 1990s.¹¹⁶ Early attempts to broadly regulate inappropriate content were struck down,¹¹⁷ but one safe harbor section withstood legal review and has protected platform hosts from liability for third-party

106. See Becky Freeth, *Jennifer Lawrence says 'trauma' of having her nude photos leaked will last 'forever'*, COSMOPOLITAN (Nov. 24, 2021), <https://www.cosmopolitan.com/uk/body/a38328843/jennifer-lawrence-trauma-nude-photo-leak/> [<https://perma.cc/P2XY-JHK2>] (Actress Jennifer Lawrence describes how she will experience trauma for the rest of her life because anyone can look up her nude photos at any time).

107. See *id.*

108. See Sturmer & Bradley, *supra* note 6.

109. See *id.*

110. See Antigone Davis, *New Updates to Help Prevent the Spread of Young People's Intimate Images Online*, META (Feb. 27, 2023), <https://about.fb.com/news/2023/02/helping-prevent-the-spread-of-young-peoples-intimate-images-online/> [<https://perma.cc/6MUU-H4FE>].

111. See *id.*

112. See *id.*

113. See *Participating Online Platforms*, TAKEITDOWN, <https://takeitdown.ncmec.org/participants/> [<https://perma.cc/KQE9-E2JT>].

114. See Davis, *supra* note 110.

115. See generally *Ashcroft v. ACLU*, 542 U.S. 656, 661, 670 (2004) (Congress attempted to prohibit knowingly posting content that is harmful to minors on the Internet, but the Supreme Court struck the legislation down because it was not narrowly tailored and the government failed to prove alternatives were less effective).

116. Romain Badouard & Anne Bellon, *Introduction to the special issue on content moderation on digital platforms*, INTERNET POL'Y REV., Mar. 23, 2025, at 2, 4.

117. See *Reno v. ACLU*, 521 U.S. 844, 877–879 (1997).

content.¹¹⁸ However, this broad interpretation of immunity from liability has left victims of harmful content like non-consensual pornography unable to force content removal or to hold platforms liable.¹¹⁹

In response to the growing access and use of the Internet, Congress enacted the Communications Decency Act (CDA) as part of the Telecommunications Act of 1996.¹²⁰ Modernization and easy access to the Internet created an issue of people abusing the Internet and making it less safe for other users, especially minors, to use.¹²¹ The Senate Conference report noted that there was an increase in reports that the Internet was being used to send explicit content, engage minors in inappropriate contact with adults, and steal personal information.¹²² In part, provisions of the CDA were intended to “modernize the existing protections against obscene, lewd, indecent, or harassing uses of a telephone.”¹²³ Decency provisions were supposed to increase penalties for abusive uses of the Internet, protect privacy, and protect families from unwanted programming on television that they deem inappropriate for their children.¹²⁴ This bill would subject a person to a fine or prison if they used the Internet to display sexual content in a manner available to persons under the age of 18.¹²⁵

The Supreme Court struck down the portions of this act regulating indecency as unconstitutional due to its overbreadth.¹²⁶ The Court reasoned that the indecency provisions would prohibit nonpornographic material with serious educational value such as information about birth control.¹²⁷ The Court acknowledged the government interest in protecting minors from harmful content, but this interest would not justify suppression of a broad content of speech addressed to adults.¹²⁸ Ironically, the original intent of the CDA to attempt regulation of pornography and indecent material on the Internet did not come into fruition.¹²⁹

118. See Chris Cox, *Policing the Internet: A Bad Idea in 1996 -- and Today*, REALCLEARPOLITICS (June 25, 2020), https://www.realclearpolitics.com/articles/2020/06/25/policing_the_internet_a_bad_idea_in_1996_--_and_today.html.

119. See *GoDaddy.com, LLC v. Hollie Toups*, 429 S.W.3d 752, 762 (Tex. App. 2014).

120. See Telecommunications Act of 1996, Pub. L. No. 104-104 § 501 (1996).

121. See S. REP. NO. 104-23, at 59 (1995).

122. See *id.*

123. *Id.*

124. See *id.*

125. See Telecommunications Act of 1996, Pub. L. No. 104-104, § 502, 110 STAT. 134 (1996).

126. See *Reno v. ACLU*, 521 U.S. 844, 877–879 (1997).

127. See *id.* at 877–878.

128. See *id.* at 875.

129. See S. REP. NO. 104-23, at 59 (1995); see Telecommunications Act of 1996 at §§ 501–509.

B. Safe Harbor Provision

Following the Supreme Court's decision to strike down provisions of the CDA, Section 230 became the sole part of the Act to survive.¹³⁰ The safe harbor provision ensures that a platform host will not be "treated as the publisher or speaker of any information provided by" a third party.¹³¹ This section of the CDA originated in a bill called the Internet Freedom and Family Empowerment Act sponsored by two representatives; Republican Christopher Cox and Democrat Ron Wyden.¹³² Part of the driving force behind this bill was to promote development of the Internet and allow the free market to flourish without excessive limitations from government regulation.¹³³ An additional reason why this bill was created was in response to two New York court cases that discussed platform liability for third-party content.¹³⁴ In 1991 CompuServe, an electronic library service that hosts forums, was sued for libel by Cubby Inc. for allegedly defamatory statements posted on its journalism forum.¹³⁵ The district court compared CompuServe to a newspaper vendor and ruled that because they did not know or have reason to know of the statements, they could not be held liable; they were simply a distributor.¹³⁶

In contrast, a few years later another New York court held that Prodigy, another Internet service provider, was liable for the content posted by third parties.¹³⁷ The court distinguished *Cubby, Inc. v. Compuserve, Inc* from *Stratton Oakmont v. Prodigy Servs. Co.* because they had implemented moderation controls and held itself out to the public as controlling the content on its bulletin boards.¹³⁸ Prodigy was considered a publisher rather than a distributor.¹³⁹

Following these court cases, platforms now began to face a new dilemma: they could either attempt to moderate content but risk liability for remaining content posted by third parties, or they could choose not to moderate and risk offensive, damaging, or obscene content proliferating

130. See Ambika Kumar & Tom Wyrwich, *The Test of Time: Section 230 of the Communications Decency Act Turns 20*, DAVIS WRIGHT TREMAINE LLP, (Sep. 2016), <https://www.dwt.com/blogs/media-law-monitor/2016/08/the-test-of-time-section-230-of-the-communications> [<https://perma.cc/SAS8-3B99>].

131. 47 U.S.C. § 230(c)(1).

132. See Chris Cox, *Policing the Internet: A Bad Idea in 1996 -- and Today*, REALCLEARPOLITICS, (June 25, 2020), https://www.realclearpolitics.com/articles/2020/06/25/policing_the_internet_a_bad_idea_in_1996__and_today.html.

133. See 47 U.S.C. § 230(b)(1) – (2).

134. See Cox, *supra* note 132.

135. See *Cubby, Inc. v. CompuServe, Inc.*, 776 F. Supp. 135, 137–38 (S.D.N.Y. 1991).

136. See *id.* at 140–41.

137. See *Stratton Oakmont v. Prodigy Servs. Co.*, No. 31063/94, 1995 N.Y. Misc. LEXIS 229, at *12–13 (Sup. Ct. May 24, 1995).

138. See *id.* at *10.

139. See *id.* at *1.

instead.¹⁴⁰ As described by then-Representative Cox, the jurisprudence of *Prodigy* and *CompuServe* would have “created a powerful and perverse incentive for platforms to abandon any attempt to maintain civility on their sites.”¹⁴¹ Representative Cox argued that platforms should instead be encouraged to filter and moderate content as much as they can.¹⁴² To encourage self-moderation, this bill would allow platforms to act as “Good Samaritans” and moderate the content that they host if they want to in the manner they choose.¹⁴³

The “Good Samaritan” provision of the bill would protect computer providers from liability on the basis that they will make a good faith effort to restrict access to material that is obscene, excessively violent, or otherwise objectionable.¹⁴⁴ A year following enactment of Section 230, the Fourth Circuit became the first federal court of appeals to interpret the law and its scope.¹⁴⁵ Mr. Zeran sued AOL Online for its delay in removing defamatory messages and refusal to post retractions and screen for similar subsequent postings.¹⁴⁶ In *Zeran v. Am. Online, Inc.*, the Fourth Circuit held that Section 230 immunized AOL Online from liability.¹⁴⁷ The court explained that even if a user provides notice to the platform of damaging or offensive content, the platform is not liable for its decision to moderate or retain the content because it would go against the purpose of enacting the statute.¹⁴⁸

C. Broad Interpretation of Section 230

Following this decision, many courts followed suit in broadly interpreting Section 230(c)(1) and solidifying a very broad definition of immunity for platforms under Section 230.¹⁴⁹

140. See Department of Justice’s Review of Section 230 of The Communications Decency Act Of 1996, DEP’T OF JUST., <https://www.justice.gov/archives/ag/departments-justice-s-review-section-230-communications-decency-act-1996> [https://perma.cc/M24L-SW64] [hereinafter *Review of Section 230*].

141. See Cox, *supra* note 132.

142. See 141 CONG. REC. 22044, 22045 (1995) (statement of Rep. Christopher Cox).

143. See Cox, *supra* note 132; see generally 47 U.S.C. § 230(c)(2) (this bill later becomes codified).

144. See Cox, *supra* note 132.

145. See Valerie Brannon & Eric Holmes, CONG. RSCH. SERV., R46751, SECTION 230: AN OVERVIEW 9-10 (2024).

146. See *Zeran v. Am. Online, Inc.*, 129 F.3d 327, 330 (4th Cir. 1997).

147. See *id.* at 332.

148. See *id.* at 333.

149. See Valerie Brannon & Eric Holmes, CONG. RSCH. SERV., R46751, SECTION 230: AN OVERVIEW 10 (2024); see also *Universal Commc’n Sys. v. Lycos, Inc.*, 478 F.3d 413, 418-19 (1st Cir. 2007) (joining the Fourth, Ninth, and Tenth Circuits in their broad interpretation of Section 230 immunity).

As a result, many courts will find preemption under Section 230 for lawsuits that would impose liability based on third-party content.¹⁵⁰ When courts have not found preemption, it has primarily been because the claim at issue was dismissed on other grounds or the court found the defendant was the actual source or responsible for the content at issue.¹⁵¹ The currently broad interpretation of Section 230 means that platforms have little incentive or reason to closely monitor their content or to even remove illicit content once identified.¹⁵²

In more recent years Section 230 has come under fire for the broad immunity grant for platforms.¹⁵³ Some argue that Section 230 has been taken advantage of and often used as a ‘get out of jail free card’ and its current use to escape liability was not aligned with the original intent of passing the Safe Harbor provision.¹⁵⁴ Victims of non-consensual pornography have found no relief in holding platforms liable for hosting such illicit images on their websites.¹⁵⁵ Due to the broad definition of immunity under Section 230, courts have immunized platforms who knew or should have known that their platforms were being used to host violations of privacy, civil rights, and defamation.¹⁵⁶ Even worse, some individuals have used Section 230 as a shield to assert that their conduct, which would horrify some people, is protected by the CDA.¹⁵⁷

Although the original purpose of the CDA was to help moderate content on the Internet, it is now weaponized to help facilitate harmful and negative

150. See, e.g., David S. Ardia, *Free Speech Savior or Shield for Scoundrels: An Empirical Study of Intermediary Immunity under Section 230 of the Communications Decency Act*, 43 LOY. L.A. L. REV. 373, 435, 438–40 (2010) (reporting courts found preemption under section 230 for at least one claim in almost 60% of decisions. Additionally, nearly 70% of unreversed federal court decisions found preemption when addressing the application of section 230 in summary judgement motions).

151. See *id.* at 442.

152. See Danielle Keats Citron, *Section 230's Challenge to Civil Rights and Civil Liberties*, KNIGHT FIRST AMEND. INST. (Apr. 6, 2018), <https://knightcolumbia.org/content/section-230s-challenge-civil-rights-and-civil-liberties> [<https://perma.cc/V54Q-D3Z9>] [hereinafter Citron, *Section 230's Challenge*].

153. See Rebecca Kern, *White House renews call to ‘remove’ Section 230 liability shield*, POLITICO, (Sep. 9, 2022, at 12:39 EDT), <https://www.politico.com/news/2022/09/08/white-house-renews-call-to-remove-section-230-liability-shield-00055771> [<https://perma.cc/4VWQ-YVYB>].

154. See Citron, *Section 230's Challenge supra* note 152.

155. See GoDaddy.com, LLC v. Hollie Toups, 429 S.W.3d 752, 762 (Tex. App. 2014) (reversing lower court’s refusal to grant summary judgement because plaintiffs were seeking to hold GoDaddy liable as publisher for the nude content on their website, which is barred by Section 230).

156. See DANIELLE KEATS CITRON, HATE CRIMES IN CYBERSPACE 171 (2014) [hereinafter CITRON, HATE CRIMES].

157. See Kashmir Hill, *Hunter Moore Will Post Your Nude Photos but Will Only Include Your Home Addresses If He Thinks You’re a Horrible Person*, FORBES, (Dec. 5, 2012, at 17:16 EST), <https://www.forbes.com/sites/kashmirhill/2012/12/05/hunter-moore-is-going-to-start-posting-your-nude-photos-again-but-will-only-post-your-home-address-if-he-thinks-youre-a-horrible-person/> [<https://perma.cc/9TZH-D8CF>].

actions on the Internet.¹⁵⁸ Calls for amending Section 230 have increased over the years but to no avail.¹⁵⁹ The CDA has only been amended twice since its enactment in 1996, with the most recent amendment in 2018 creating an exception for certain cases that involve sex trafficking.¹⁶⁰ As explicit deepfake content proliferates on the Internet, there will be more victims who are unable to force the takedown of that content or hold platforms liable; it is solely the platform's decision.¹⁶¹ In the case of celebrities, there may not be a desire to remove explicit deepfake content due to their popularity and interest.¹⁶²

It's clear that Section 230 is both a saving grace to platforms and a weapon for bad actors to avoid liability.¹⁶³ To some, it feels like a nearly impossible task to get a website to remove this explicit, harmful content.¹⁶⁴ However, there are a few exceptions to Section 230 immunity.¹⁶⁵ One such exception is the immunity from liability does not apply to intellectual property laws.¹⁶⁶ Through this exception there is a potential avenue to ensure that platforms either remove explicit deepfake content or face liability for their actions.

V. ANALYSIS

A. *Right of Publicity Law*

The Right of Publicity is commonly known as a property right over the use of one's name, image, or likeness.¹⁶⁷ This right exists only in state common law and statutes as there is no federal right of publicity.¹⁶⁸ This right may provide a person injunctive relief or damages if their name, image, or

158. See Alina Selyukh, *Section 230: A Key Legal Shield For Facebook, Google Is About To Change*, NPR, (Mar. 21, 2018, at 05:11 EST), <https://www.npr.org/sections/alltechconsidered/2018/03/21/591622450/section-230-a-key-legal-shield-for-facebook-google-is-about-to-change> [<https://perma.cc/SN52-QAAH>], (quoting Cox). See generally CITRON, HATE CRIMES, *supra* note 156, at 168 (discussing how a revenge porn site operator told a woman that she could pay him \$75 to take down her photo and that he is untouchable for liability because of federal law).

159. See *Review of Section 230*, *supra* note 140; see also CITRON, HATE CRIMES, *supra* note 156, at 177 (proposing amending Section 230's Safe Harbor provision to exclude bad actors like platforms that exclusively host cyber stalking or non-consensual pornography).

160. See Allow States and Victims to Fight Online Sex Trafficking Act of 2017 (FOSTA), Pub. L. No. 115-164, § 4, 132 Stat. 1253 (2018); see also 47 U.S.C. §§ 230(b)(5), (e)(5).

161. See Sturmer & Bradley, *supra* note 6.

162. See Conger & Yoon, *supra* note 2.

163. See generally CITRON, HATE CRIMES, *supra* note 156, at 168.

164. See Sturmer & Bradley, *supra* note 6.

165. See 47 U.S.C. § 230(e).

166. See *id.* at § 230(e)(2).

167. See Jennifer E. Rothman, *The Inalienable Right of Publicity*, 101 GEO. L.J. 185, 187 (2012).

168. See *id.*; see also Mark Roesler & Joey Roesler, *Patchwork Protections: The Growing Need for a Federal Right of Publicity Law*, ABA (July 10, 2024), https://www.americanbar.org/groups/intellectual_property_law/resources/landslide/2024-summer/patchwork-protections-growing-need-federal-right-publicity-law/ [<https://perma.cc/UD2D-MHWD>] (describing need for a Federal Right of Publicity to harmonize and provide consistency for the current patchwork landscape of differing state laws).

likeness is used without their permission.¹⁶⁹ Because the right of publicity varies based on the state, some laws only encompass commercial use, while others only require that the use was made for any type of advantage.¹⁷⁰ With the rise of AI creating more realistic images and videos, recent federal legislation has attempted to create rights of publicity.¹⁷¹ The NO FAKES Act of 2024 would create a civil action for a person whose image, voice, or likeness was used to create computer-generated content or the publishing of such content without their consent.¹⁷² This legislation would require actual knowledge and would exclude common copyright fair uses such as scholarship, satire, and parody.¹⁷³ The No AI Fraud Act would create a civil action for conduct that makes available to the public a digital voice replica without someone's consent or a personalized cloning service.¹⁷⁴

1. Circuit Split on CDA Section 230

There is currently a circuit split concerning whether right of publicity laws fall within the intellectual property exception of CDA Section 230's safe harbor provision.¹⁷⁵ The Ninth Circuit rejected the right of publicity as falling within intellectual property under Section 230.¹⁷⁶ This court held that including the right of publicity would undermine the broad grant of immunity provided by Section 230.¹⁷⁷ As a result, the Ninth Circuit construed intellectual property in Section 230 to mean federal intellectual property statutes which primarily encompass copyright or trademark.¹⁷⁸ The Ninth Circuit further narrowly construed the meaning of intellectual property to mean those that are specifically established such as trademark, patent, and copyright.¹⁷⁹ One argument against including the right of publicity under intellectual property is that because the goal of Section 230 is to help "preserve the vibrant and competitive free market that presently exists for the Internet and other interactive computer services, unfettered by Federal or

169. *See id.*

170. *See id.* at 211.

171. *See* NO FAKES Act of 2024, H.R. 9551, 118th Cong. (2024); *see* No AI FRAUD Act, H.R. 6943, 118th Cong. § 3(b)(1) – (2) (2024).

172. *See* H.R. 9551 §§ 2(a)(1), (c).

173. *See id.* at §2(c).

174. *See* H.R. 6943 § 2(c).

175. *See* Ned T. Himmelrich, *Right of Publicity May (or May Not) Be Intellectual Property Under Section 230*, GORDON FEINBLATT LLC (Feb. 22, 2024), <https://www.gflaw.com/what-we-do/insights/right-publicity-may-or-may-not-be-intellectual-property-under-section-230> [<https://perma.cc/778M-F2T9>].

176. *See* Perfect 10, Inc. v. CCBill LLC, 488 F.3d 1102, 1118-19 (9th Cir. 2007).

177. *See id.* at 1119 n.5.

178. *See id.* at 1119. *See generally* 15 U.S.C. §§ 1051-1127 (federal trademark statutes); 17 U.S.C. §§ 101-122 (federal copyright statutes).

179. *See generally* Enigma Software Grp. USA, LLC v. Malwarebytes, Inc., 946 F.3d 1040, 1053 (9th Cir. 2019) (establishing that the intellectual property "exception does not apply to false advertising claims brought under § 1125(a) of the Lanham Act, unless the claim itself involves intellectual property").

State regulation” inclusion of this right that varies state to state would stifle the market.¹⁸⁰

Although the First Circuit did not directly discuss if the right of publicity would fall within intellectual property, this circuit found that a state trademark dilution claim was not barred by Section 230 because it was a claim based on intellectual property.¹⁸¹ The Third Circuit found that the right of publicity fell within intellectual property exception of Section 230(e)(2).¹⁸² Within its analysis, the court highlighted that two legal dictionaries, including Black’s, define the right of publicity as an intellectual property and that it has been described as analogous to copyright, trademark, and patent.¹⁸³

2. Right of Publicity is an Intellectual Property

This circuit split must be resolved in favor of the right of publicity falling out of Section 230 immunity under the intellectual property exception. First, the plain text of statute reads to include state intellectual property statutes as opposed to just federal, which includes right of publicity.¹⁸⁴ Section 230 specifically states that the immunity will not have any effect on federal criminal statutes, intellectual property law, or state law.¹⁸⁵ If Congress intended that only intellectual property included in federal statutes applied, they would have said so.¹⁸⁶ Because the meaning of this statute is plain, there is no further need to ascertain other potential interpretations.¹⁸⁷

Another reason is the property and economic interests in one’s personal image and reputation. In *Hepp*, the court highlights that the plaintiff had “dedicated considerable time, effort and money into building her brand.”¹⁸⁸ Additionally some state courts have recognized that individuals have property interests in their personas.¹⁸⁹ This interest need not be exclusive to public figures because of significant interests that private persons also have in their personas.¹⁹⁰ As mentioned earlier, employers may make or alter hiring decisions based on perceptions of a person’s image and reputation.¹⁹¹ Even if the effects of a tarnished reputation are not immediate, they may affect a

180. See 47 U.S.C. § 230(b)(2); see also *Hepp v. Facebook*, 14 F.4th 204, 211 (3d Cir. 2021).

181. See *Universal Commc’n Systems, Inc.*, 478 F.3d at 422–23 (citing 47 U.S.C. § 230(e)(2)).

182. See *Hepp*, 14 F.4th at 214.

183. See *id.* at 213.

184. See 47 U.S.C. § 230(e)(2).

185. See 47 U.S.C. § 230(e)(1) – (3).

186. See *Hepp*, 14 F.4th at 211 (explaining that when Congress wanted to limit the interpretation about state law within CDA §230, it knew how to do so and did so explicitly).

187. See *Bostock v. Clayton Cty.*, 590 U.S. 644, 673–74 (2020) (explaining when the meaning of the text is plain, the Court’s job of interpretation is done).

188. See *Hepp*, 14 F.4th at 212.

189. See *id.*

190. See *Number of Employers*, *supra* note 51.

191. See *id.*

person in the future.¹⁹² Additionally, a private person today may become a public figure in the future.¹⁹³ One such example was Vanessa Williams who was the first Black woman to be crowned Miss America.¹⁹⁴ Only two months before the end of her reign, she was forced to relinquish her title because the magazine *Penthouse* announced they would publish scandalous photos she had posed for two years prior.¹⁹⁵ These photos were taken when she was still a private figure; she was working as an assistant to a photographer.¹⁹⁶ She had also been assured at the time the photos were taken that they were merely silhouettes, that she would not be identifiable, and that the photos would not leave the studio.¹⁹⁷ Despite this, these photos that were supposed to be private became public and burned Vanessa Williams' reputation.¹⁹⁸ In this era of social media and the world wide web, anything posted on the Internet can exist into perpetuity.¹⁹⁹

Lastly, there have been past and present references to the right of publicity as an intellectual property right within the federal government.²⁰⁰ In *Zacchini v. Scripps-Howard Broad*, the Supreme Court considered the right of publicity as "closely analogous to... patent and copyright" because it focuses "on the right of the individual to reap the reward of his endeavors."²⁰¹ In the introduced bill NO FAKES Act of 2024, the bill's purpose is stated as "[t]o protect intellectual property rights in the voice and visual likeness of individuals, and for other purposes" describing the right of publicity as a type of intellectual property.²⁰² The right of publicity should accordingly be considered an intellectual property that falls out of Section 230 immunity.

B. A (Narrow) Federal Right of Publicity

The enactment of a narrow federal right of publicity through an amendment to the Digital Millennium Copyright Act would clearly answer

192. See generally Jennifer Latson, *The Scandal That Cost a Miss America Her Crown*, TIME, (July 23, 2015, at 10:30 EDT), <https://time.com/3961120/miss-america-scandal-vanessa-williams/> [<https://perma.cc/KS7T-68YT>].

193. See *id.*

194. See *id.*

195. See *id.*; see Lester Fabian Brathwaite, *How Vanessa Williams lost the Miss America crown but won over Hollywood*, ENT. WKLY. (July 23, 2024, at 23:57 EDT), <https://ew.com/how-vanessa-williams-lost-miss-america-crown-but-won-over-hollywood-8682640>.

196. See Latson, *supra* note 192.

197. See *id.*

198. See Janine Rubenstein, *How Vanessa Williams Went from Being 'Canceled' amid Miss America Scandal to Proving Her Haters Wrong (Exclusive)*, PEOPLE, (July 27, 2024, at 09:20 EDT), <https://people.com/how-vanessa-williams-proved-haters-wrong-after-miss-america-scandal-exclusive-8684148> [<https://perma.cc/5E9C-Q4SW>].

199. See generally Freeth, *supra* note 106.

200. See *Zacchini v. Scripps-Howard Broad. Co.*, 433 U.S. 562, 573 (1977); see NO FAKES Act of 2024, H.R. 9551, 118th Cong. (2024).

201. *Zacchini*, 433 U.S. at 573.

202. NO FAKES Act of 2024, H.R. 9551, 118th Cong. (2024); see also No AI FRAUD Act, H.R. 6943, 118th Cong. § 3(b)(1) – (2) (2024) (asserting that every person has a property right in their own likeness and that it constitutes an intellectual property right).

the circuit split question on whether a right of publicity is an intellectual property right.²⁰³ The right of publicity would be automatically granted upon one's birth and exist until 10 years past one's death. This right would allow a person to control the use of their image or likeness and would only apply to explicit deepfake content. Because only several states recognize a right of publicity and they differ from state to state, some individuals either do not have a right of publicity or have a limited right that would not allow them to bring a claim for non-commercial use.²⁰⁴ Providing a narrow federal right of publicity will ensure that every individual, regardless of jurisdiction, is able to prevent their image from being used to create explicit deepfake content and to hold others liable for this offense. People may still create deepfake content to create parodies or other related content in the absence of further legislation.²⁰⁵ However, explicit deepfake content using someone's image or likeness cannot be created without the permission of the right-holder.

This federal right of publicity would create a "floor" of minimum protections.²⁰⁶ A very narrow federal right of publicity will also ensure that existing state laws concerning right of publicity are not affected.²⁰⁷ States would be free to increase their individual right of publicity privileges as they see fit. This proposal will not establish that all state right of publicity laws are exempt from Section 230 immunity due to the potential issues with individual First Amendment rights.²⁰⁸ However, this proposal will exempt such laws from Section 230 immunity as applied to explicit deepfake content.

Because this right of publicity places a significant restriction on the creation of explicit deepfake content, it may face a First Amendment challenge because content-based speech restrictions are subject to strict scrutiny.²⁰⁹ This restriction would successfully overcome such a challenge because (1) there is a compelling government interest and (2) this right is narrowly tailored to achieve this interest.²¹⁰ Here the compelling government interest is protecting citizens from the harms associated with non-consensual explicit deepfake content. One's sexual reputation or perception can have

203. *Compare* Hepp v. Facebook, 14 F.4th 204, 214 (3d Cir. 2021), *with* Perfect 10, Inc. v. CCBill LLC, 488 F.3d 1102, 1119 (9th Cir. 2007).

204. *See* Roesler & Roesler, *supra* note 168 (explaining that "37 states recognize the right of publicity, with 25 of these states recognizing the right via statute and 12 by common law" and some states only recognize a right of publicity for public figures); *see* LA. STAT. ANN. §470.3 (2022) (Louisiana statute provides that "every individual has a property right in connection with the use of that individual's identity for commercial purposes"); *see also* *Right of Publicity Statutes & Interactive Map*, RIGHT OF PUBLICITY, <https://rightofpublicity.com/statutes> [<https://perma.cc/YBG8-B7DK>].

205. *See generally* Graham, *supra* note 66 (explaining there is no federal legislation regulating the creation of deepfake content).

206. *See* AUTHORS, ATTRIBUTION, AND INTEGRITY: EXAMINING MORAL RIGHTS IN THE UNITED STATES, UNITED STATES COPYRIGHT OFFICE 118 (Apr. 2019).

207. *See id.*

208. *See generally* Reed v. Town of Gilbert, 576 U.S. 155, 163–164 (2015); *see generally* Rothman, *supra* note 167, at 2223 (explaining a common defense to a right of publicity is the First Amendment).

209. *See* Reed, 576 U.S. at 163–164.

210. *See id.* at 155 (asserting that content-based speech restrictions can only be justified if the law is narrowly tailored to achieve a compelling government interest).

particularly detrimental consequences.²¹¹ As highlighted earlier, sexual content of a person can ruin their reputation,²¹² sabotage employment opportunities,²¹³ and harm their mental health to the point of death.²¹⁴ This government interest significantly outweighs a person's individual interest in creating explicit deepfake content using another person's image and likeness.

This parallels with non-consensual pornography, where several state supreme courts have found preventing intrusions on individual privacy and protecting individuals from the harms of non-consensual pornography to be compelling government interests.²¹⁵ In these cases, the courts found that the statutes prohibiting the dissemination of non-consensual pornography survived strict scrutiny because they were narrowly tailored to achieve the compelling interest.²¹⁶ This right of publicity is similarly narrowly tailored to only affect the use of one's image and likeness in the context of explicit deepfake content; this would achieve the government's interest without broadly affecting other uses such as creating parody or satire content.²¹⁷ This right of publicity would not completely bar an individual from creating the explicit deepfake content outright; they would still be able to use their own image and likeness and that of another person who gives them consent. Establishing a narrow federal right of publicity would be the first step towards content moderation of explicit deepfake content; the next and more critical step would be to incorporate it into a process to ensure identified images are removed upon identification.

C. *Digital Millennium Copyright Act*

The Digital Millennium Copyright Act (DMCA) is a copyright law that implemented two 1996 treaties of the World Intellectual Property Organization.²¹⁸ Passed in 1998, the DMCA amended Title 17 of the United States Code and created a limitation on the liability of platforms for copyright infringement committed by their users.²¹⁹ This limited liability, however, is contingent on a platform responding expeditiously to remove or disable access to material that has been identified by a copyright owner as infringement.²²⁰ A platform must take down content that constitutes copyright

211. See Rubenstein, *supra* note 198.

212. See Moreau & Rourke, *supra* note 45.

213. See *Number of Employers*, *supra* note 51.

214. See Ross, *supra* note 62.

215. *Accord* People v. Austin, 155 N.E.3d 439, 462 (Ill. 2019); *State v. Katz*, 179 N.E.3d 431, 455–56 (Ind. 2022);

State v. Casillas, 952 N.W.2d 629, 642–43 (Minn. 2020); *State v. VanBuren*, 214 A.3d 791, 811 (Vt. 2018).

216. *Accord* Austin, 155 N.E.3d at 466; *Katz*, 179 N.E.3d at 460; *Casillas*, 952 N.W.2d at 644; *VanBuren*, 214 A.3d at 814.

217. See *generally* Austin, 155 N.E.3d at 466, 474.

218. Digital Millennium Copyright Act, Pub. L. No. 105-304 (1998).

219. See *Digital Millennium Copyright Act*, U.S. GEN. SERVICES AGENCY, <https://digital.gov/resources/digital-millennium-copyright-act> [https://perma.cc/K8DH-7NNN].

220. Digital Millennium Copyright Act Pub. L. No. 105-304 § 202 (1998); see 17 U.S.C. § 512(b)(2)(E).

infringement once notice has been given or else can be held liable for the copyright infringement.²²¹ A person who physically takes a photo is the copyright owner.²²² As the owner, only they can sue for copyright violations, not the person(s) featured in the photo.²²³ This could prove especially difficult when the person who has taken the photo was a stranger. As currently written, the DMCA would not provide a person the ability to hold a platform to the notice and takedown measures of the DMCA for explicit deepfake content of themselves on the platform.²²⁴ But amending the DMCA to include a federal right of publicity would provide a victim with this powerful tool to get this content removed or to hold the platform responsible for noncompliance.²²⁵

1. Incorporating the Right of Publicity into DMCA Through Amendment

The DMCA must be amended to include a narrow federal right of publicity for explicit deepfake content. Rather than create new legislation, Congress can amend this statute that already contains a notice and takedown mechanism; this would provide victims with an avenue to remove explicit deepfake content from platforms.²²⁶ This amendment would allow a person whose image or likeness is used to create explicit deepfake content to contact a platform and request the removal of the content.²²⁷ This would also bypass the issue of needing an actual copyright owner to file the notification.²²⁸ The ability to utilize the notice and takedown mechanism in the DMCA would be limited to intellectual property that is specifically identified in this Act: copyright and (the proposed) right of publicity. This would ensure that this form of content moderation would only occur where Congress has specifically spoken and alleviate concerns that the various state regulations would stifle content in the free market.²²⁹ There would also be a requirement to utilize notice and takedown and (1) wait 30 days or (2) receive a denial from the platform host (whichever is earlier) before pursuing legal liability against the platform host for explicit deepfake content.

This framework would solve legal problems that existing remedies are unable to address. Current remedies only provide victims with a way to hold the actual content creator liable for making the explicit deepfake content.²³⁰ They do not provide a way to remove the content, throwing victims at the

221. See 17 U.S.C. § 512(c)(1)(C).

222. See CITRON, HATE CRIMES, *supra* note 156, at 121–122.

223. See *id.*

224. See generally 17 U.S.C. § 512(c)(1)(C) (only identified copyright infringement is eligible for the notice and takedown process).

225. See *id.* at § 512(c)(1)(C).

226. See *id.*

227. See *id.*

228. See CITRON, HATE CRIMES, *supra* note 156, at 121–122.

229. See generally *Hepp v. Facebook*, 14 F.4th 204, 211 (3d Cir. 2021).

230. See e.g., N.Y. PENAL LAW §245.15 (2023); VA CODE § 18.2-386.2 (2019); H.R. 207 2024 Gen. Assemb., Reg. Sess. (KY. 2024); CAL. CIV. CODE § 1708.86; see discussion *supra* Sections III.A, III.C.

mercy of the platform to remove it of their goodwill.²³¹ Section 230 stands as a nearly impenetrable barrier against forcing the removal of explicit deepfake content because of its safe harbor provision, allowing platforms to escape liability.²³² However, as an intellectual property right, the right of publicity falls out of the Section 230 safe harbor provisions.²³³ Incorporating this right into the DMCA gives victims a mechanism to force the takedown of explicit deepfake content and prevent further harm.²³⁴ In the scenario where a platform refuses to comply with this request, victims can seek legal recourse against the platform without Section 230's usual bar against civil liability.²³⁵ The ability to hold platforms liable will prevent bad faith actors from using Section 230 as a shield to further harm victims.²³⁶

2. Challenges to Right of Publicity Framework

A possible challenge against this federal right of publicity framework within the DMCA is that an offender may assert their First Amendment right as a defense. A platform may then argue that since the content creator may escape liability from existing state right of publicity remedies through this defense, they too should escape liability from this framework. Courts have applied several tests to analyze a first amendment defense against a right of publicity claim.²³⁷ In the “relatedness test”, the use of a person’s name, image, or likeness is not barred unless it is “wholly unrelated” to the underlying work.²³⁸ In *Rogers v. Grimaldi*, the court held that the use of Ginger Rogers’ name in a movie title was permitted because it was clearly related to the content of the movie.²³⁹ Within the context of explicit deepfake content, this claim should fail because the nature of such content is creating a false sexual identity for the actual person, unrelated to their actual identity.²⁴⁰ A defendant may argue that the use of a person’s face is enough of a hook to be related to the person’s image or likeness. Even if a court were to determine that the use of one’s face is sufficient to meet the threshold of relatedness, there is also a strong argument that this test would not even be appropriate in this context.²⁴¹ The Third Circuit rejected the use of the relatedness test, considering it better suited for trademark-like claims.²⁴² The Ninth Circuit agreed with this view, asserting that the relatedness test is used to protect the consumer from

231. See Sturmer & Bradley, *supra* note 6.

232. See 47 U.S.C. § 230(c)(2); see also Zeran, 129 F.3d at 332.

233. See e.g., 47 U.S.C. § 230(e)(2); Hepp v. Facebook, 14 F.4th 204, 214 (3d Cir. 2021).

234. See 17 U.S.C. § 512(b)(2)(E).

235. See 47 U.S.C. § 230(c)(2).

236. See Hill, *supra* note 157.

237. See Robert C. Post & Jennifer E. Rothman, *The First Amendment and the Right(s) of Publicity*, 130 YALE L.J. 86, 129–130 (2020).

238. See *id.* at 130.

239. See *Rogers v. Grimaldi*, 875 F.2d 994, 1005 (2d Cir. 1989).

240. See Citron, *supra* note 55, at 1921.

241. See *Hart v. Elec. Arts, Inc.*, 717 F.3d 141, 157 (3d Cir. 2013).

242. See *id.*

confusion not the right of publicity holder.²⁴³ A defense that attempts to utilize this test should fail.

In the “transformative test”, the court uses analysis similar to the transformative factor of the fair-use doctrine test in copyright law.²⁴⁴ The court considers whether the work adds significant creative elements so that it is transformed beyond an imitation or likeness.²⁴⁵ Within this test, explicit deepfake content should also fail. A defendant may argue that the images they use to create the content transform the victim into having a body that they do not have in reality, therefore meeting this threshold. However, the goal of the underlying network that creates deepfakes is to create content that is indistinguishable from reality.²⁴⁶ As a result, the content created will look very realistic. The Third Circuit asserted that when the use is “entertainment that is merely a copy or imitation, even if skillfully and accurately carried out” there is no creative component and therefore no transformational purpose.²⁴⁷ In a similar manner, creating explicit deepfake content is attempting to create an indistinguishable image of a person with sexual content. With this consideration, explicit deepfake content should be considered closer to an imitation or likeness and fail this test.

VI. CONCLUSION

Explicit deepfake content is unfortunately an issue that will continue to grow with more advances in technology. While there have been recent movements to create state legislation and closer attempts to create federal legislation, there is still significant work to be done to protect victims and minimize harm. Legal remedies to either prosecute the offender or create a civil action to seek damages may face barriers if the offender is unable to be identified or found. More significantly, the continual presence of this content online may be a reminder of this deeply invasive violation. Some attempts to have websites remove analogous non-consensual pornography have been thwarted due to the broad definition of immunity under Section 230 of the CDA.²⁴⁸ Although broad, this immunity is not absolute.

Creating a narrow federal right of publicity helps to utilize the intellectual property exception that exists under Section 230. Many legal sources have defined the right of publicity as intellectual property or an analogue to trademark, patent, and copyright. By narrowing the scope of this right to the context of explicit deepfake content, we can ensure that the compelling government interest in protecting citizens from the harm of non-consensual explicit deepfake content is achieved in the least restrictive manner. Incorporating this right of publicity into the DMCA will provide

243. See *Keller v. Elec. Arts Inc. (In re NCAA Student-Athlete Name & Likeness Licensing Litig.)*, 724 F.3d 1268, 1280 (9th Cir. 2013).

244. See Post & Rothman, *supra* note 237, at 129. See 17 U.S.C. § 107.

245. See Post & Rothman, *supra* note 237, at 129.

246. See discussion *supra* Section II.A.

247. *Hart v. Elec. Arts, Inc.*, 717 F.3d 141, 164 (3d Cir. 2013).

248. See CITRON, HATE CRIMES, *supra* note 156, at 168. See generally 47 U.S.C. § 230(c)(2).

victims with an avenue to force the takedown of this content and the ability to hold platforms liable if they refuse. Platforms will not be able to hide behind Section 230 or use it as a weapon against victims.

