Article

Ayelet Gordon-Tapiero*

Unreal and Unjust: An Enrichment-Based Approach to the Deepfake Dilemma

https://doi.org/10.1515/jtl-2025-0031
Received September 24, 2025; accepted September 24, 2025; published online October 27, 2025

Abstract: Generative AI technology is taking the world by storm. The development of technology enabling creation and manipulation of content has give facilitated a substantial rise in the proliferation of deepfakes. Whereas in the past content creation and manipulation required a certain level of expertise, today deepfake technology is easily accessible and enables the quick and seamless creation of highly believable content. This technology has many positive applications: for example, in healthcare, education, cultural preservation and the entertainment industry. But deepfake technology is also used to deceive and cause harm. Deepfake technology is used to create sexual deepfakes that humiliate and harm primarily women and girls, to engage in fraudulent activities, and to generate disinformation undermining trust in democratic processes and institutions. This Article suggests analyzing the deepfake dilemma through the lens of the doctrine of unjust enrichment. Under this doctrine a party which has become enriched at the expense of another must make restitution of benefits it received. An enrichment-based approach may offer several advantages, particularly when compared to harm-based remedies. First, it may be easier to identify the defendant in an unjust enrichment case, as the defendant is the company developing the underlying technology and there is no need to identify the individual who created the deepfake. Second, a lawsuit under unjust enrichment may be filed by various plaintiffs, thus, obivating the need to identify the individual

The author thanks Vivian Eichler and Yotam Kaplan for helpful comments. This work was supported by the European Union under ERC grant 101125913. Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

***Corresponding author: Ayelet Gordon-Tapiero**, Fellow, HUJI Benin School of Computer Science and Engineering and The Federmann Cyber Security Center, Jerusalem, Israel,
E-mail: Ayelet.Gordon@mail.huji.ac.il

harmed by a deepfake. Finally, it may be easier to identify and quantify the profits which are monetary, current and concentrated within a small number of companies. We argue that under certain circumstances viewing the deepfake challenge through the lens of unjust enrichment can allow for a realignment of the financial incentives of the companies developing deepfakes technology with broader social goals and values, encouraging them to develop technology that is less harmful and more responsible.

Keywords: deepfake; unjust enrichment; restitution; generative AI; disinformation; law and technology

1 Introduction

The debate about whether generative AI is a 'normal technology' or whether it represents a fundamental shift is currently the subject of intense debate.[1] What is clear, even at this point, is that similar to other technological developments, generative-AI driven technologies "are not in themselves good or bad; it is the way they are used that determines their value."[2] Generative AI has taken a giant leap in recent years resulting in a reality where anyone with a computer and internet connection can access tools that enable the creation of deepfakes easily, quickly and freely. While the ability to generate text, images, video and audio has advantageous applications, these tools can also be used nefariously.[3] Malicious actors use generative-AI driven deepfake generators to create sexual deepfakes of girls, to defraud companies and individuals, to spread disinformation and to undermine trust in democratic institutions and processes.[4]

This harmful reality is not inevitable.[5] Some of the harms stemming from deepfake technology can be mitigated by integrating safeguards into the

1 Arvind Narayanan & Sayash Kapoor, *AI As Normal Technology,* Knight First Amen. Inst. at Colum. U. (2025) (arguing that AI represents a normal technological development).

2 This quote is widely attributed to David Sarnoff, *see e.g.* as cited in Marshall McLuhan, Understanding Media: The Extensions of Man 10 (1964).

3 Orly Lobel, *The Law of AI for Good*, 75 Fla. L. Rev. 1073 (2023) (describing positive applications of generative AI).

4 A word on terminology – in this Article we use the term *deepfake generators* to mean technology driven by generative AI that enables the creation of deepfakes, whether they are in the form of text, images, videos or audio.

5 Ayelet Gordon-Tapiero, Yotam Kaplan & Gideon Parchomovsky, Deepfake Liability, N.C. L. Rev. (2025, Forthcoming) suggesting analyzing the liablity of companies developing deepfake technology through the doctrine of products liability.

underlying technology. For example, including a watermark indicating that content was generated by AI may limit some of the harms – particularly where such outputs are used to defraud individuals and companies. In other cases, integrating tools that enable tracking the provenance of a file (such as watermarking or requiring user registration) may increase the ability to prosecute users who create harmful deepfakes.[6] Unfortunately, many of the deepfake generators accessible online today make minimal efforts to increase the safety of their tools, leaving the door open for malicious actors to harm individuals and society.

Victims of deepfakes have limited recourse. They are often unable to identify the creator of the deepfake,[7] and platforms hosting deepfakes are largely immune from liability under Section 230 of the Communications Decency Act as they have been created by (often anonymous) third parties.[8] This Article draws attention to the role of the companies developing and deploying deepfake technology. The leading generative-AI companies have become immensely valuable, generating large profits at the expense of the victims who fall prey to their technology. It suggests that in addition to traditional harm-based remedies, it is important to note that enrichment-based remedies may offer an effective way forward. Under the doctrine of unjust enrichment, courts can instruct that profits derived unjustly at the expense of another be disgorged. We argue that disgorging profits from companies developing and deploying generative-AI technology has the potential to incentivize them to change their behavior by integrating safeguards into their technology, making it safer and minimizing the ensuing harms.

The Article proceeds as follows. Section 2 provides a description of deepfake technology, its promises and perils. It describes the leading companies developing and deploying deepfake technology and the main fields in which deepfakes have already created substantial harm. The Section highlights that harms created by deepfakes impact both specific individuals as well as society at large. Section 3 offers a discussion of the application of the doctrine of unjust enrichment as a basis for deepfake liability. It describes the elements comprising the doctrine and suggests how it may be applied in the context of deepfake technology. Section 4 analyzes the

6 Ayelet Gordon-Tapiero, Yotam Kaplan & Gideon Parchomovsky, *Deepfake Liability*, 104 N.C. L. Rev. 1, 39 (2025) (describing tools whose integration would increase the safety of deepfake technology).

7 Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 Cal. L. Rev. 1753, 1792 (2019) ("Civil liability cannot ameliorate harms caused by deep fakes if plaintiffs cannot tie them to their creators.").

8 47 U.S.C. § 230(c) (1) (2012).

advantages offered by enrichment-based remedies compared to harm-based remedies such as tort law. A short conclusion follows.

2 Deepfake Technology

In an era increasingly shaped by AI, the line between fact and fiction is becoming perilously blurred. The phenomenon of manipulated and falsified content is not new, though the advent of generative AI has accelerated the scale, speed and sophistication with which individuals can generate highly reliable yet completely false content. In the past, manipulation of media required expertise, skills and access to special tools.[9] At its early stages the technology used to generate deepfakes was "too cumbersome for the average computer user … the process require[d] above average computer literacy, including understanding torrenting, path configuration, file structures, and application versioning."[10] The development of generative-AI technology in recent years, coupled with its increased accessibility, have democratized the ability to quickly and simply generate highly reliable deepfakes.[11] It is so simple and cheap, in fact, that even young children with minimal digital literacy can do so.[12]

---

9 Melissa Heikkilä, *An AI Startup Made a Hyperrealistic Deepfake of Me That's So Good It's Scary*, MIT Tech. Rev. (Apr. 25, 2024), https://www.technologyreview.com/2024/04/25/1091772/new-generative-ai-avatar-deepfake-synthesia/ ("Until now, all AI-generated videos of people have tended to have some stiffness, glitchiness, or other unnatural elements that make them pretty easy to differentiate from reality.").

10 Jessica Ice, *Defamatory Political Deepfakes and the First Amendment*, Case W. Res. L. Rev. 417, 425 (2019) (Noting that in 2019 the technology to generate deepfakes was "too cumbersome for the average computer user … the process requires above average computer literacy, including understanding torrenting, path configuration, file structures, and application versioning."). *See* Regina Rini & Leah Cohen, *Deepfakes, Deep Harms*, 22 J. Ethics & Soc. Phil. 143 (2022).

11 Matt Burgess, *Deepfake Creators are Revictimizing GirlsDoPorn Sex Trafficking Survivors*, Wired (Jun. 25, 2024), https://www.wired.com/story/girlsdoporn-deepfake-victim-videos/ ("As deepfake technology has become increasingly capable of creating realistic imagery and easier to use, hundreds of websites and apps designed to create or host deepfake sexual abuse have appeared.").

12 John Werner, *Kids Can Use AI, Too – Look What They're Coming Up With…* Forbes (Nov. 22, 2023), https://www.forbes.com/sites/johnwerner/2023/11/22/kids-can-use-ai-toolook-what-theyre-coming-up-with/ ("The simplicity of the interface … is going to provide more access to kids."); Kevin Kelly, *Picture Limitless Creativity at Your Fingertips*, Wired (Nov. 17, 2022), https://www.wired.com/story/picture-limitless-creativity-ai-image-generators/ ("Not everyone can write, direct, and edit an Oscar winner like *Toy Story 3* or *Coco*, but everyone *can* launch an AI image generator and type in an idea."); Heikkilä, *supra* note 8 ("almost anyone will now be able to make a digital double…"); Ice, *supra* note 9, at 423 ("using deepfake technology is relatively simple.").

Generative AI has many promising applications.[13] It has been used in the entertainment industry, reducing costs, accelerating production, enabling previously impossible and prohibitively expensive visual effects, and developing new modes of story creation and storytelling.[14] Generative AI is playing an increasingly large role in educational settings. It enables teachers to structure learning processes, personalize teaching, create tailor-made content for their lesson plans and educational goals, bring historical figures to life and animate past events, as well as to design adaptive assessments and feedback mechanisms.[15] But generative AI also has a dark side. It is used in ways that are harmful to both individuals and society. It has been used to shame, embarrass and extort young girls through the creation of sexual deepfakes, to commit fraud and identity theft, to generate harmful disinformation, manipulating elections and undermining trust in democracy.[16]

The term *deepfake* stems from a combination of the terms *deep learning* and *fake content,* and refers to the outputs of the digital creation of false content or manipulation of real content.[17] Chesney and Citron define deepfakes as "hyper-realistic

13 Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 Cal. L. Rev. 1753, 1768 (2019); Orly Lobel, The Equality Machine 140 (2022) (describing the benefits that AI brings to the medical field); Lobel, *supra* note 3, at 1073 (describing benefits of generative AI); Kim Martineau, *Generative AI Could Offer a Faster Way to Test Theories of how the Universe works*, IBM (Mar. 14, 2024), https://research.ibm.com/blog/time-series-AI-transformers.

14 Stewart Townsend, *Exploring the Impact of AI on Film Production in 2024*, Medium (Mar. 2, 2024), https://medium.com/@channelasaservice/exploring-the-impact-of-ai-on-film-production-in-2024-f02da745af00; Rick Spair, *The Rise of AI in Hollywood: How Technology is Changing the Movie Industry #innovation #technology #management #data*, Medium (May 22, 2024), https://medium.com/@rickspair/the-rise-of-ai-in-hollywood-how-technology-is-changing-the-movie-industry-innovation-technology-0677faa67886#:~:text=Critics%20argue%20that%20relying%20solely%20on%20AI%20algorithms%20may%20result,job%20losses%20in%20the%20industry.

15 Dan Patterson, *Deepfakes for Good? How Synthetic Media is Transforming Business*, Tech Informed (Oct. 5, 2023) https://techinformed.com/deepfakes-for-good-how-synthetic-media-is-transforming-business/ ("Deepfake algorithms can animate historical photos and footage, allowing influential figures to give speeches and presentations as if they were in the classroom.").

16 Benjamin L.W. Sobel, *A Real Account of Deepfakes*, Mich. L. Rev. 1, 22 (2025); Chesney & Citron, *supra* note 11, at 1771–1785; Danielle Keats Citron, *Sexual Privacy*, 128 Yale L.J. 1870, 1922 (2019); Danielle Keats Citron, The Fight for Privacy: Protecting Dignity, Identity and Love in the Digital Age 38 (2022).

17 Lucas Whittaker, Kate Letheren & Rory Mulcahy, *The Rise of Deepfakes: A Conceptual Framework and Research Agenda for Marketing*, 29 Austl. Mktg J. 204 (2021). See also Richard L. Hasen, *Deep Fakes, Bots and Siloed Justices: American Election Law in a "Post-Truth" World*, 64 St. Louis U. L.J. 535 (2020); Robert Chesney & Danielle Citron, D*eepfakes: A Looming Crisis for National Security, Democracy and Privacy?* Lawfare (Feb. 21, 2018), https://www.lawfaremedia.org/article/deepfakes-looming-crisis-national-security-democracy-and-privacy; Morgan Meaker, *Deepfake Audio is a Political Nightmare*, Wired (Oct. 9, 2023), https://www.wired.com/story/deepfake-audio-keir-starmer/; Rini & Cohen, *supra* note 9.

digital falsification of images, video, and audio."[18] Creators of deepfakes "can make a politician, celebrity, or anyone else appear to say or do anything the manipulator wants."[19] They enable their creators to construct representations of actions, events, scenarios and conversations that never took place.[20] Some people believe that they are able to identify when an image, video or audio has been fabricated.[21] In reality, research has shown that deepfakes can be extremely difficult for laypeople to detect.[22] Even AI experts and AI driven tools have a hard time identifying high-quality deepfakes.[23] This is not an inevitable reality. There are various technologies that could allow marking of AI-generated content as such, signaling to viewers that they are encountering fabricated content.[24] Companies developing and deploying deepfake technology largely choose not to integrate such capabilities into their technology. The result of this design choice is that it is challenging to detect that a file has been algorithmically generated and is often virtually impossible to identify who has created the deepfake, making it almost impossible to hold the creator accountable for the harms they have generated.

Deepfakes have become extremely popular and widespread. One of the areas in which deepfakes are generating immense harms is non-consensual pornography, also referred to as sexual deepfakes. In April 2024, the New York Times noted that the prevalence of such deepfakes was so widespread, it identified it as an epidemic.[25] The

18 Chesney and Citron, *supra* note 12, at 1757.

19 Hasen, *supra* note 16, at 542.

20 *See* Rini & Cohen, *supra* note 9; Ice, *supra* note 9.

21 Nils C. Köbis, Barbora Doležalová & Ivan Soraperra, *Fooled Twice: People Cannot Detect Deepfakes but Think They Can*, 24 iScience 1, 5 (2021) ("Results reveal that participants have exaggerated beliefs in their detection abilities when such beliefs are elicited in an unincentivized way.").

22 Kimberly T. Mai et al., *Warning: Humans Cannot Reliably Detect Speech Deepfakes*, PLOSONE (Aug. 2, 2023), https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0285333.

23 KEATS CITRON, *supra* note 15, at 38, (citing Hany Farid: "In January 2019, deep fakes were buggy and flickery. Nine months later, I've never seen anything like how was they're going."); Momina Masood, Marriam Nawaz, Khalid Mahmood Malik, Ali Javed & Aun Irtaza, *Deepfakes Generation and Detection: State-of-the-Art, Open Challenges, Countermeasures, and Way Forward*, 53 APPLIED INTELL. 3974 (2023); KEATS CITRON, *supra* note 15, at 38 ("Deepfakes are so sophisticated that even experts struggle to distinguish them.").

24 Hany Farid, *Creating, Using, Misusing, and Detecting Deep Fakes*, 1 J. ONLINE TRUST & SAFETY 1, 11 (2022).

25 Natasha Singer, *Teen Girls Confront an Epidemic of Deepfake Nudes in Schools*, N.Y. TIMES (Apr. 8, 2024), https://www.nytimes.com/2024/04/08/technology/deepfake-ai-nudes-westfield-high-school.html. Victims of sexual deepfakes suffer an array of harms. *See* Welsh v. Martinez, 114 A.3d 1231, 1242 (Conn. App. Ct. 2015) (describing that victims of sexual deepfakes often experience fear); Danielle Keats Citron & Mary Anne Franks, *The Internet as a Speech Machine and Other Myths Confounding Section 230 Reform*, 2020 U. CHI. LEGAL F. 45 (2020); Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PENN. L. REV. 477, 548 (2006).

overwhelming majority of sexual deepfakes are of women and girls, making it very much a gender-based epidemic.[26] In several highly publicized cases female politicians and artists were the target of sexual deepfakes.[27]

Another field in which deepfakes are prevalent is in the creation of political disinformation.[28] Such deepfakes target politicians, showcasing them in an embarrassing light and discrediting them. Political deepfakes are especially pervasive in the period leading up to an election to sway public opinion or confuse voters. In 2024, days before the Democratic primaries in New Hampshire, President Biden called voters over the phone, advising them to refrain from voting in the primaries. These calls were, of course, audio deepfakes of the President's voice.[29] Deepfakes can also be used to influence the outcomes of wars. In 2022 a video of Ukrainian President Zelenskyy appeared in which he allegedly conceded defeat to the Russians and called on troops to surrender.[30] The video was part of an effort to harm Ukraine.[31]

Finally, deepfakes are often used in fraud and identity theft, presenting new challenges for individuals, institutions and law enforcement.[32] In one frightening case, the CEO of a UK based company was instructed by his boss to transfer $243,000 to one of the company's suppliers.[33] It later turned out that the phone call was an

---

26 Danielle Keats Citron, The Fight for Privacy: Protecting Dignity, Identity and Love in the Digital Age 39 (2022) ("for intimate privacy violations, women and minors are more likely to be the victims.").

27 Anastasia Powell, Adiran J. Scott, Asher Flynn & Asia A. Eaton, *Whether of Politicians, Pop Stars or Teenage Girls, Sexualised Deepfakes Are on the Rise. They Hold a Mirror to Our Sexist World*, The Conversation (Feb. 7, 2024), https://theconversation.com/whether-of-politicians-pop-stars-or-teenage-girls-sexualised-deepfakes-are-on-the-rise-they-hold-a-mirror-to-our-sexist-world-222491.

28 Nilesh Christopher & Varsha Bansal, *Indian Voters Are Being Bombarded with Millions of Deepfakes. Political Candidates Approve.* Wired (May 20, 2024), https://www.wired.com/story/indian-elections-ai-deepfakes/; Cristina Criddle, *Political Deepfakes Top List of Malicious AI Use, DeepMind Finds*, Financial Times (Jun. 25, 2024), https://www.ft.com/content/8d5bc867-c69d-44df-839f-d43c92785435.

29 Ali Swenson & Will Weissert, *New Hampshire Investigating Fake Biden Robocall Meant to Discourage Voters Ahead of Primary*, AP (Jan. 23, 2024), https://apnews.com/article/new-hampshire-primary-biden-ai-deepfake-robocall-f3469ceb6dd613079092287994663db5.

30 Farid, *supra* note 23, at 11 (describing the incident).

31 *Id.*

32 FS-ISAC Artificial Intelligence Risk Working Group, *Deepfakes in the Financial Sector; Understanding the Threats, Managing the Risks* (2024).

33 Catherine Stupp, *Fraudsters Used AI to Mimic CEO's Voice in an Unusual Cybercrime Case*, Wall St. J. (Aug. 30, 2019), https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402. *See* also Heather Chen & Kathleen Magramo, *Finance Worker Pay Out $25 Million After Video Call Wit Deepfake 'Chief Financial Officer,'* CNN World (Feb. 4, 2024), https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk (reporting on another case of fraud by deepfake).

audio deepfake generated by a fraudster. A retired grandmother in New Zealand fell for a crypto scam, losing $224,000 after being convinced to invest in the coin by a deepfake of the country's prime minister.[34]

Finally, people also use deepfake technology for less nefarious purposes such as for bringing back a dead relative,[35] or enabling the victim of a murder to provide testimony in court.[36]

Regardless of the settings in which deepfakes appear, they are having a profound impact on our perception of reality and truth.[37] The fact that one can no longer believe what they see or hear, causes people to question all content, prompting them to be suspicious and to question the existence of an epistemic value to perception or meaning to truth.[38] From a collective perspective, such distrust causes societies to become increasingly fragmented, with individuals retreating into their echo chambers, limiting their interactions to people that believe content that aligns with their views as real, and content that contradicts their beliefs as fake.[39]

Generative-AI technology is an extremely lucrative business. The global market for generative-AI driven technology was valued at 7.58 billion USD in 2024 and is expected to reach 38.45 billion USD by 2032.[40] The companies leading the generative-AI market, including OpenAI, Google, Stability AI and Nvidia, are some of the most

---

34 Lane Nichols, *Pensioner Loses $224k After Being Tricked by AI Deepfake Chirstopher Luxon Cryptocurrency Investment Scam*, New Zealand Herald (Oct. 19, 2024), https://www.nzherald.co.nz/nz/pensioner-loses-224k-after-being-tricked-by-ai-deepfake-christopher-luxon-cryptocurrency-investment-scam/YLG3EQMOAZATVARBL5ITDRL2DA/.

35 Christopher Intagliata, Avery Keatley & Scott Detrow, *People Are Creating Deepfakes of their Dead Relatives,* NPR (May 12, 2024), https://www.npr.org/2024/05/12/1250835619/deepfakes-for-the-dead.

36 Matthew Gault & Jason Koebler, *'I Loved that AI:' Judge Moved by AI-Generated Avatar of Man Killed in Road Rage Incident*, 404 (May 7, 2025), https://www.404media.co/i-loved-that-ai-judge-moved-by-ai-generated-avatar-of-man-killed-in-road-rage-incident/.

37 Jules Roscoe, *Deepfake Scams Are Distorting Reality Itself*, Wired (Jun. 4, 2025), https://www.wired.com/story/youre-not-ready-for-ai-powered-scams/?utm_source=chatgpt.com.

38 Spencer McKay & Chris Tenove, *Disinformation as a Threat to Deliberative Democracy*, 74 Pol. Rsch. Q. 703, 708 (2020) (the authors suggest calling this effect 'epistemic cynicism'); Heikkilä, *supra* note 8.

39 Tomer Shadmy, *Content Traffic Regulation: A Democratic Framework for Addressing Misinformation*, Jurimetrics 1, 11 (2022); Shreeharsh Kelkar, *Post-Truth and the Search for Objectivity: Political Polarization and the Remaking of Knowledge Production*, 5 Engaging Sci. Tech. & Soc'y 86 (2019) (discussing how misinformation hinders the belief in a shared truth).

40 *Deepfake Technology Market Size, Share & Industry Analysis, By Component (Software and Services), By End User (Government Organizations, Media and Entertainment, Retail and E-commerce, Legal, and Others), and Regional Forecast, 2025–2032*, Fortune Business Insights (Jun. 9, 2025), https://www.fortunebusinessinsights.com/deepfake-technology-market-109936.

valuable in the world.[41] Some deepfake generators are offered for free, often with a paid subscription-based option. Premium options offer higher quality outputs, generation of longer video/audio generation, and even removal of watermarks.[42] The full extent of monetization models of AI still remains elusive.[43] What is clear is that generative-AI companies collect and store the content input by users, and can use this data to further train and improve their models. In an age where generative-AI companies are scrambling to get more data[44] and also to exclude their competitors from accessing this data, collecting a large amount of data from user prompts can offer companies a competitive advantage.[45]

The enormous financial benefits stemming from the development and deployment of deepfake technology are expected to increase and grow the market in future years. At the same time, the harms stemming from use of these technologies are having a profound impact on the lives of individuals and societies. This reality in which financial interests drive substantial harms makes it essential for policymakers to grapple with the challenges arising from the market for deepfake technology. In the next Section we suggest analyzing the profits generated by companies developing deepfake technology through the doctrine of unjust enrichment. We argue that applying enrichment-based remedies to the challenges stemming from deepfake technology offers a promising way forward.

---

41 Thomas Babychan, Top 10 AI Companies with the Highest Net Worth, (Jul. 2, 2025), https://techstory.in/top-10-ai-companies-with-the-highest-net-worth/; Sabrina Ortiz, *The Best AI Image Generators Are Getting Scary Good at Things they Used to be Terrible At*, ZDNET (May 9, 2025), https://www.zdnet.com/article/best-ai-image-generator/; Ema Lukan, *The 12 Best AI Video Generators (Freed & Paid) of 2025*, Synthesia (Jun. 11, 2025), https://www.synthesia.io/post/best-ai-video-generators. Smaller companies also offer deepfake technology – *see* Caroline Haskins, *A Deepfake Nude Generator Reveals a Chilling Look at Its Victims*, Wired (Mar. 25, 2024), https://www.wired.com/story/deepfake-nude-generator-chilling-look-at-its-victims/; *New Freedoms of Imagination*, LumaLabs, https://lumalabs.ai/dream-machine; Midjourney, https://www.midjourney.com/home.

42 *See infra* Section 3.2.1.

43 We may be vieweing the beginning of such monetaization efforts. In October, 2025, Meta announced that it would be using users chats with AI to personalize ads for them. Clare Duffy, Meta Will Soon Use Your Conversations with its AI Chatbot to Seel You Stuff, CNN Business (Oct. 2, 2025), https://edition.cnn.com/2025/10/01/tech/meta-ai-chatbot-targeted-ads#:~:text=Meta%20will%20soon%20use%20what,with%20even%20more%20personalized%20ads.

44 Ayelet Gordon-Tapiero, Katrina Ligett & Kobbi Nissim, On The Rival Nature of Data: Tech and Policy Implications, CSLAW '25 Proc. of the 2025 Symp. on Comp. Sci. and L. 17 (2025).

45 Bilbao European Encounters, *AI Is Setting Off a Great Scramble for Data*, The Economist (Aug. 13, 2023), https://www.economist.com/business/2023/08/13/ai-is-setting-off-a-great-scramble-for-data; Steven Melendez, *In the AI Era, Data is Gold. And These Companies are Striking it Rich*, Fast Company (Jul. 1, 2024), https://www.fastcompany.com/91148997/data-is-gold-in-ai-era.

3 Unjust Enrichment

The financial incentives driving the market for deepfake technology are huge. At the same, the harms stemming from this technology cannot be ignored. In this part we explore the option of addressing the challenges stemming from deepfake technology through the lens of the doctrine of unjust enrichment. We argue that given the current legal landscape and the incentives driving the companies develoiping deepfake technology, the profits derived from technology that harms others should be considered unjust and should therefore be disgorged. This Section opens with an introduction to the doctrine of unjust enrichment and concludes by applying it to the case of deepfake technology.

3.1 The Law of Unjust Enrichment

The law of unjust enrichment provides that a party unjustly enriched at the expense of another must make restitution of any benefits generated unjustly.[46] The need to ensure that one does not enjoy benefits it received unjustly is a basic moral principle and an important cornerstone of many legal systems. Since the law of unjust enrichment functions as a flexible, residual legal category, it is often used in contexts where other legal remedies do not offer effective remedies and is therefore especially appropriate for implementation in the context of developing technologies.[47,48]

A claim in unjust enrichment is typically comprised of three elements: (1) enrichment by the defendant; (2) the enrichment was generated at the expense of the plaintiff; (3) the enrichment was unjust. This doctrine allows for the analysis of various factual scenarios. For example, if a person mistakenly receives a payment they were not entitled to, they are considered to have been unjustly enriched at the expense of the person who made the payment. Thus, they are not allowed to enjoy the

46 Restatement (Third) of Restitutions and Unjust Enrichment § 1 (2010) ("a person who is unjustly enriched at the expense of another is subject to liability in restitution"); Ward Farnsworth, Restitution: Civil Liability for Unjust Enrichment 1 (2014).

47 *See* Emily Sherwin, *Restitution and Equity: An Analysis of the Principle of Unjust Enrichment*, 79 Tex. L. Rev. 2083, 2107 (2001) ("what makes unjust enrichment both powerful and dangerous … is it open-endedness."); Charlie Webb, Reason and Restitution: A Theory of unjust Enrichment 39 (2016). *See e.g.* proposals to view harmful algorithmic personalization and generative AI training as unjust enrichment: Ayelet Gordon & Yotam Kaplan, *Generative AI Training as Unjust Enrichment*, 86 Ohio St. L. J. (2025); Ayelet Gordon-Tapiero & Yotam Kaplan, *Unjust Enrichment by Algorithm*, 92 Geo. Wash. L. Rev. 305 (2024).

48 Ayelet Gordon-Tapiero & Yotam Kaplan, Generative AI Training as Unjust Enrichment, 86 Ohio St. L. J. 285 290 (2025).

financial windfall but must return it to the payor.[49] In another example, imagine that a person is walking down the street and suffers a heart attack. Luckily, a doctor nearby performs life-saving medical treatment. The person saved has been enriched at the expense of the doctor's service and must pay fair market value for the services rendered despite the fact that the two had no formal contract before the incident.[50] Note that both these cases demonstrate that for enrichment to be unjust it is not necessary that the benefitting party is a wrongdoer: it is simply that they received a benefit they should not be allowed to retain. The unjustness refers to the mechanism that caused their enrichment, and not to their behavior.

3.2 Unjust Enrichment and Deepfake Technology

In this section we apply the basic tenets of the law of unjust enrichment to the companies developing and deploying deepfake technology. We note that there are several mechanisms through which companies developing and deploying deepfake technology become enriched and identify the ways in which this enrichment is generated at the expense of plaintiffs. Finally, we offer guidelines as to when and why this enrichment should be viewed as unjust.

3.2.1 Enrichment by the Defendant

There are (at least) four mechanisms through which companies developing and deploying deepfake technology become enriched. First, companies which develop and deploy high-quality generative-AI technologies, especially those who are first to do so, add immense value to the company and its shareholders.[51] As of March 2025, OpenAI, the company leading the generative AI market, was valued at $300 billion.[52] OpenAI developed Sora, a video generator and Dall-E, an image generator. Google's Deepmind, that developed Imagen, an image generator and Veo 3, a video generator,

49 This provision is subject to some exceptions defense rules – such as if the enriched party changed their position based on the payment. In such a case, if the receiver acted in good faith, they may be exempt from returning (some of) the sum.

50 K.A.L. v. Southern Med. Bus. Servs., 854 So. 2d 106 (Ala. Civ. App. 2003); Bingham Mem. Hosp. v. Boyd (in Re Estate of Boyd), 8 P.3d 664 (Ida. App. 2000) (both cases debate events in which individuals received medical care where no contract had been concluded prior to the treatment and the courts ruled that the patients became enriched at the expense of the hospitals that treated them; Peter Birks, Unjust Enrichment 40–41 (2d ed. 2005).

51 *See* Section 2, *supra*.

52 Cade Metz, *OpenAI Completes Deal That Values Company at $300 Billion*, N.Y. Times (Mar. 31, 2025), https://www.nytimes.com/2025/03/31/technology/openai-valuation-300-billion.html.

was valued at $67.05 billion.[53] Smaller companies, such as Stability AI and Adobe offer high-quality image and video generators and are also extremely valuable.[54]

The fact that many companies developing deepfake technology largely fail to invest sufficient resources to develop and integrate tools that would make their products safer, and less susceptible to manipulation in the form of harmful deepfakes can be viewed as a second enrichment mechanism.[55] Any resources that are 'saved' by not investing in deveoping such tools and integrating them into the companmy's technology can be viewed as having been unjustly generated at the expesne of the people harmed by deepfakes. Watermarking the output of deepfake generators would at least enable sophisticated users and law enforcement to determine that a file has been algorithmically generated or manipulated. Moreover, watermarking can attest not only to the fact that a file has been generated by AI but it can also "be integrated into devices that people use to make digital contents to create immutable metadata for storing originality details such as time and location of multimedia contents as well as their untampered attestment."[56] Knowing that a user's details are embedded into the deepfakes they create and they could be held accountable for their generation, may lower the motivation to generate harmful

53 *DeepMind Platform Co., Ltd*. Yahoo Finance (Jun. 26, 2025), https://nz.finance.yahoo.com/quote/223310.KQ/key-statistics/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAALyLlNhSBmRPgFrdGFDN2gWOW1ZiSy3slXWHwL8bVqo8eE9Ur1iG3tltO5VNyJpWwBxQdiyapb4xbYVt3vVQs6frAJoHH8zJ0KT0PecJUiwDpvuMrufCoaABEmu3gJrmgeN3_m2vzm_AJ25aRcF0h2U5mKOjZye1OYMV41OBF_u.

54 Adobe is valued at over $150 billion, *Adobe Inc*. Yahoo Finance (Jul. 31, 2025), https://finance.yahoo.com/quote/ADBE/key-statistics/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAALyLlNhSBmRPgFrdGFDN2gWOW1ZiSy3slXWHwL8bVqo8eE9Ur1iG3tltO5VNyJpWwBxQdiyapb4xbYVt3vVQs6frAJoHH8zJ0KT0PecJUiwDpvuMrufCoaABEmu3gJrmgeN3_m2vzm_AJ25aRcF0h2U5mKOjZye1OYMV41OBF_u; the last publicly available valuation of StabilityAI valued the company at $1 billion in 2022, during the early days of the democratization of generative AI, *see* Akash Sriram & Krystal Hu, *Cash-Strapped Stability AI Raises $80 mln with New CEO and Board*, Reuters (Jun. 26, 2024), https://www.reuters.com/technology/artificial-intelligence/cash-strapped-stability-ai-raises-80-mln-with-new-ceo-board-2024-06-25/?utm_source=chatgpt.com.

55 Some companies have begun integrating watermarks into some of their outputs – *see* Xuandong Zhao et al., *SoK: Watermarking for AI-Generated Content*, IEEE Symp. Security and Privacy (2025); Bram Rijsbosch, Gijs van Dijck & Konrad Kolling, *Adoption of Watermarking Measures for AI-Generated Content and Implications Under the EU AI Act*, arXiv: 2503.18156v2 (2025).

56 Thanh Thi Nguyen et al., *Deep Learning for Deepfakes Creation and Detection: A Survey*, 223 Comput. Vision & Image Understanding, 103525 (2022); Robert Chesney & Danielle Citron, *Disinformation on Steroids*, Council Foreign Relations (Oct. 2018), https://www.cfr.org/report/deep-fake-disinformation-steroids; Tiffany Hsu, *Google Joins Effort to Help Spot Content Made With A.I.*, N.Y. Times (Feb. 8, 2024), https://www.nytimes.com/2024/02/08/business/media/google-ai.html; *see* also Jae Young Hwang & SangHoon Oh, *A Brief Survey of Watermarks in Generative AI*, 14th Int'l Conf. Info. & Comm. Tech. Convergence 1157 (2023).

deepfakes to begin with. Some companies include restrictions regarding the outputs their models may generate.[57] These are, however, often circumventable.[58]

The third enrichment mechanism stems from the collection of data input the models collect from user prompts. Companies developing generative AI models are continuously looking for new sources of data on which they can train their models, increasing demand for new sources of data.[59] Companies that have exclusive access to certain data and are able to exclude other companies from using the data to train their models (for example through exclusive licensing deals) will find themselves with a competitive advantage. Companies which are able to attract a large number of users to use their deepfake technology can use the prompts input by the users (whether text, image, video or audio) as training material for their newer models. Companies that have better (i.e. a larger amount, higher quality and more diverse) training material can use it to develop better models compared to other actors in the market and attract more users. This in turn will raise the valuation of the company and generate more training material for it through the collection of data from a larger number of user prompts.

Finally, a fourth source of enrichment for companies developing generative-AI technology is subscription fees. Most companies developing and deploying deepfake technology offer users a free version of their technology. For a subscription fee users can get an upgraded experience including – faster response times (OpenAI),[60] priority access during high demand (OpenAI, Midjourney),[61] access to more

57 Will Knight, *This Uncensored AI Art Tool Can Generate Fantasies – and Nightmares*, Wired (Sep. 21, 2022), https://www.wired.com/story/the-joy-and-dread-of-ai-image-generators-without-limits/ ("But most of these image generators are designed to restrict what users can depict, banning pornography, violence, and pictures showing the faces of real people."); Benjamin Sobel, *Elements of Style: Copyright, Similarity, and Generative AI*, 38 Harv. J. Law & Tech. 49, 60 (2024) ("All of the major image-generating services … place contractual and technological limits on functionality. Their terms of service typically set certain categories of imagery – such as violence and pornography – off-limits. [Some] employ filtering technology that automatically blocks certain prompts from causing output to be generated.").

58 These restrictions, despite being circumventable, serve as a form of friction, and require more effort and expertise to remove. On the advantages of friction *see* Ayelet Gordon-Tapiero, Paul Oh & Ashwin Ramaswami, *Fact and Friction: A Case Study in the Fight Against False News*, 57 U.C. Davis L. Rev. 171, 183 (2023) (describing the purposeful of integration into technology).

59 Ayelet Gordon-Tapiero, Katrina Ligett & Kobbi Nissim, *On the Rival Nature of Data: Tech and Policy Implications*, ACM Symp. on Comput. Sci. & Law 17, 24 (2025).

60 *Pricing*, OpenAI, https://openai.com/chatgpt/pricing/.

61 *Id*; *Comparing Midjourney Plans*, Midjourney, https://docs.midjourney.com/hc/en-us/articles/27870484040333-Comparing-Midjourney-Plans.

sophisticated models (Gemini, Midjourney, Runway, Adobe Firefly),[62] removal of watermarks (Kapwing),[63] the ability to use generative AI to conduct research (Microsoft Copilot, Gemini),[64] generation of podcasts (Microsoft Copilot),[65] custom voice cloning (ElvenLabs, HeyGen, Kapwing),[66] generation of longer videos (Deepfakes Web) and more.[67]

3.2.2 At the Expense of the Plaintiff

There are two main categories of plaintiffs harmed by deepfake technology. The first are individuals whose likeliness has been used and manipulated. The victims of sexual deepfakes, often girls and women, suffer immense harms following the publication and distribution of deepfakes.[68] Politicians are also often the subject of deepfakes – with their likelihood being manipulated to humiliate them, to spread misinformation or manipulate elections or other political processes.[69] Celebrities are also often the subject of deepfakes – whether with the goal of humiliating them or of using their likelihood for financial goals.[70] A fourth group of individuals who are often the subject of deepfakes are people whose likelihood is manipulated in fraud or who fall prey to such scams, believing that the voice they are hearing, or the video

---

62 Midjourney, *Id*; *Get More Out of Gemini*, Gemini, https://gemini.google/subscriptions/; *Choose the Best Plan for You*, Runway, https://runwayml.com/pricing; *Compare Firefly Plans*, Adobe, https://www.adobe.com/il_en/products/firefly/plans.html.

63 *Pricing*, Kapwing, https://www.kapwing.com/pricing.

64 *Microsoft Copilot Pro*, Microsoft, https://www.microsoft.com/en-us/store/b/copilotpro; Gemini, *supra* note 58.

65 *Id.*

66 *Pricing*, IIElevenLabs, https://elevenlabs.io/pricing?utm_source=google&utm_medium=cpc&utm_campaign=t1_brandsearch_brand_english&utm_id=21809606381&utm_term=eleven%20voice%20ai&utm_content=brand_-_semi-brand&gad_source=1&gad_campaignid=21809606381&gbraid=0AAAAAp9ksTGjPlmkzYqCT47brhGtC6PFa&gclid=Cj0KCQjwyvfDBhDYARIsAItzbZHo_49UiSrlJwbsh4gBOGkP7ZaitgokX8JqWEfM9w08gCENWxYZXbkaAmW1EALw_wcB; *Flexible Pricing Plans*, HeyGen, https://www.heygen.com/pricing?utm_source=GPM&utm_campaign=22783027563&gad_source=1&gad_campaignid=22788915971&gbraid=0AAAAABiJW6ba4hVhQJxBySD2V1_LD6z6L&gclid=Cj0KCQjwyvfDBhDYARIsAItzbZEOLgiw7Yn13pQeHl5xi5GDkBHBm1xL1AGarr68TuW8E0vu9bjoPlUaAsR5EALw_wcB; Kapwing, *supra* note 59.

67 *Deepfake AI Generator*, DeepfakesWeb, https://deepfakesweb.com/#pricing.

68 Citron, *supra* note 6, at 1870.

69 Shannon Bond, *How AI Deepfakes Polluted Elections in 2024*, NPR (Dec. 21, 2024), https://www.npr.org/2024/12/21/nx-s1-5220301/deepfakes-memes-artificial-intelligence-elections.

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

they are seeing is of a loved one or of an authority figure who requires the subject to conduct a financial transaction or to reveal private information.[71]

Another type of victim of deepfake technology is society at large. Societies can be harmed by deepfakes when these are aimed at undermining important values guiding them – in particular those underlying democratic societies.[72] This happens when political deepfakes influence public opinion and democratic processes. Deepfakes that promote disinformation undermine the ability of individuals and societies to determine truth from lies.[73] Such deepfakes promote a climate of truth decay – a reality whereby individuals increasingly doubt the existence of an objective truth or the ability to reach one or agree what is true.[74]

The enrichment of AI developers is unjust in the sense that these companies profit though the loss of others. Part of what makes deepfake technology appealing for some users (and thus profitable for the companies) is the ability to harm others, for example through the creation of deepfakes. To counter this effect, it is important to ensure that companies do not profit by making their models more harmful.

3.2.3 The Enrichment was Unjust

Companies developing generative-AI technologies that enable easy generation of deepfakes are becoming enriched at the expense of the victims of their technology. This outcome is the result of dangerous design choices made by companies developing and deploying deepfake technology. A reality where a small number of companies generate immense profits at the expense of individual victims and society at large is unjust and unsustainable. Legal claims brought under unjust enrichment can realign the financial incentives that drive companies to prefer their immense financial profits over the basic wellbeing of individuals and the functioning of societies. By imposing enrichment-based liability, unjust enrichment can change companies' incentives and establish the necessary levels of deterrence against

---

71 *See* examples in Section 2, *supra*.

72 Michael Waldman, *The Danger of Deepfakes to Democracy*, Brennan Inst. Just. (Mar. 26, 2024), https://www.brennancenter.org/our-work/analysis-opinion/danger-deepfakes-democracy.

73 Spencer McKay & Chris Tenove, *Disinformation as a Threat to Deliberative Democracy*, 74 Pol. Rsch. Q. 703, 708 (2020) (suggesting this effect be referred to as 'epistemic cynicism'); Heikkilä, *supra* note 8.

74 Shadmy, *supra* note 38, at 86 (discussing how misinformation hinders the belief in a shared truth); Ahmed Maati et al., *Information, Doubt, and Democracy: How Digitization Spurs Democratic Decay*, 31 Democratization 922 (2025); Simon Chesterman, *Truth Decay: Evolving Legislative Response to Address Online Misinformation, Disinformation and "Fake News,"* Nat'l U. Singapore Working Paper No. 2025/003 (2024).

deployment of such technology in its current state, as well as guidelines for companies on how to adapt their technology in the future.[75]

Under the principles outlined in this section, courts can award restitution designed to ensure that companies developing generative-AI technologies used to create harmful deepfakes do not profit unjustly at the expense of individuals and society. The doctrine of unjust enrichment leaves broad discretion to courts to identify categories of unjust behavior on a case-by-case basis. In the next Section we describe the advantages of filing a claim against developers of deepfake technology through unjust enrichment compared to doing so based on other legal doctrines.

4 Comparative Advantages of Unjust Enrichment

In this Section we describe the comparative advantages of pursuing enrichment-based remedies through unjust enrichment claims against companies developing and deploying deepfake technology compared to harm-based remedies. We begin by describing the nature of the expected defendants under an unjust enrichment claim and then proceed to discuss the identity of potential plaintiffs in such a lawsuit. We note the ways in which unjust enrichment is particularly suited for the nature of the defendants and plaintiffs in the context of the development and deployment of deepfake technology. Finally, we analyze the advantages of a gain-based remedy compared to a harm-based one. We note that the harms stemming from deepfake technology are often non-monetary, are only expected to crystalize sometime in the future and are often difficult to quantify. The gains for the companies, on the other hand, materialize in the present, are monetary and are therefore much easier to quantify. These reasons all give rise to comparative benefits of unjust enrichment-based claims compared harm-based ones.

We do not suggest that unjust enrichment is a panacea. There are certainly cases where harm-based claims may be more appropriate and give rise to a more fitting remedy to a particular plaintiff. In such cases, harm-based claims should be pursued. What we do argue is that it is important to offer potential plaintiffs the benefits that an unjust enrichment claim offers. As an equity-based doctrine it is particularly well suited to deal with "novel legal claims."[76]

---

75 Vanessa Casado Pérez et al., *Climate Lies & Unjust Profits*, Emory L.J. 1, 46 (Forthcoming) (highlighting the role of unjust enrichment in deterring companies from behaving unjustly); Daniel Friedmann, *Restitution of Benefits Obtained Through the Appropriation of Property or the Commission of a Wrong*, 80 Colum. L. Rev. 504, 55 (1980) ("Principles of deterrence are accordingly relevant to a claim for restitution.").

76 Emily Sherwin, *Reparations and Unjust Enrichment*, 84 B.U. L. Rev. 1443, 1448 (2004).

4.1 Identifying Defendants

When a person files a tort claim for a harm they suffered, they need to identify the wrongdoer responsible for it. In a setting where there is a single wrongdoer, it may be relatively straightforward to prove that the actions of the wrongdoer directly and exclusively caused harm to the injured party.[77] In the case of a chain of potential wrongdoers, questions of comparative fault and joint liability must be diligently analyzed and proven, making the case much more complex.[78] As concisely described by Kyar: "The [harmful] outcome may be the result of not one action or series of actions by a single actor, but rather a confluence of multiple actions by multiple actors, given the ability of complex, adaptive systems to combine and magnify causal impacts."[79] An individual claiming to have been harmed by a deepfake would face serious challenges when facing the need to prove the elements of a tort claim: they would need to identify the contribution of each potentially liable defendant to the resulting harm: the company developing the harmful technology, the creator of the deepfake, the people who shared it, the platforms who hosted the harmful content and perhaps more, depending on the facts of each case. Determining the contribution of each of these parties to the resulting harm and therefore assigning liability to each would be extremely challenging under accepted tests of causation.[80]

Approaching this scenario from an unjust enrichment perspective sidesteps many of these challenges, since unjust enrichment focuses on the enrichment generated by the defendant and not on the harm suffered by the plaintiff.[81] There is, therefore, no requirement to demonstrate the precise harm caused to the plaintiff, nor to attribute a precise part of it to specific actions by the defendant. Identifying the defendant in an unjust enrichment claim involves the plaintiff discovering who has become unjustly enriched at their expense. In the context of deepfake technology, it seems typically easier to identify a party that has become enriched than a party who

77 Jane Stapleton, *The Two Explosive Proof-of-Causation Doctrines Central to Asbestos Claims*, 74 Brook. L. Rev. 1011, 1012 (2009) ("Under orthodox common law rules concerning causation, a tortfeasor is liable for an indivisible injury that would not have happened absent that party's breach.").

78 Neal C. Stout & Peter A. Valberg, *Bayes' Law, Sequential Uncertainties, and Evidence of Causation in Toxic Tort Cases*, 38 U. Mich. J. Reform 781, 889–90 (2005) ("Generally, the more alternative possible causes there are for the injury, or the more likely an alternative possible cause explains the injury, the more explanation (specificity) courts should require from the causation expert as to why the subject agent is the probable cause."); Troyen A. Brennan, *Causal Chains and Statistical Links: The Role of Scientific Uncertainty in Hazardous-Substance Litigation*, 73 Cornell L. Rev. 469 (1988) (noting the complexity of determining causation).

79 Douglas A. Kysar, *What Climate Change Can Do About Tort Law*, 41 Env't. L. 1, 62 (2011).

80 *Id.*

81 Douglas Laycock, *The Scope and Significance of Restitution,* 67 Tex. L. Rev. 1277 (1989) (noting that restitution, unlike compensatory damages, focuses on the gains by the defendant).

is liable under traditional requirements of tort law. The defendant in such a claim would be the company that developed and deployed the technology used to generate the harmful deepfake. A company that has created a product that generates immense risks, allows creators of deepfakes to hide undetected behind the technology, and does not include sufficient tools that would allow viewers to identify that the deepfake has been algorithmically generated, has behaved in an unjust way. The enrichment by such companies is strongly connected to the mechanism causing the harm, and such profits should, therefore, be subject to disgorgement.[82]

4.2 Identifying Plaintiffs

Under tort law, an individual that suffers harm can bring a claim against the alleged wrongdoer. In the context of harms caused by deepfakes, there is sometimes a particular individual harmed by the technology. In other cases, however, especially when thinking about political deepfakes that undermine trust in democracy and its institutions there is no identifiable individual directly harmed. In a tort case, this lack of identifiable victim would create an insurmountable challenge, that could prevent courts from assigning liability to the companies developing and deploying deepfake technology.

We suggest that filing an unjust enrichment claim against companies developing and deploying deepfake technology allows for a much broader list of potential plaintiffs. In particular, we imagine two groups of plaintiffs: private plaintiffs and public plaintiffs. The first category includes individuals at whose expense the defendant became enriched. A single individual whose likelihood was used in the creation of a deepfake could sue on her own, or as part of an aggregate lawsuit including many people at whose expense the company became enriched in a lawsuit that is similar to a class action.[83] Since the enrichment could be extremely high, this creates an incentive for individuals to not only join such a lawsuit, but to lead it.[84]

82 Pérez et al., *supra* note 71, at 50.

83 Alon Harel & Alex Stein, *Auctioning for Loyalty: Selection and Monitoring of Class Counsel*, 22 Yale L. & Pol'y Rev. 69, 81 (2004) (describing the class action agency problem and ways to overcome it).

84 Rodriguez v. W. Publ'g Corp., 563 F.3d 948, 958 (9th Cir. 2009) ("Incentive awards are fairly typical in class action cases. Such awards are discretionary and are intended to compensate class representatives for work done on behalf of the class, to make up for financial or reputational risk undertaken in bringing the action, and, sometimes, to recognize their willingness to act as a private attorney general.").

Public plaintiffs could also file an unjust enrichment lawsuit. In cases where it is difficult to identify individuals at whose expense the companies became enriched, or if such plaintiffs do not wish to file a lawsuit,[85] or are unable to do so,[86] a public plaintiff could step in. Attorneys general have been known to bring unjust enrichment claims, for example in tobacco litigation cases and in cases alleging that social media sites knowingly harmed children.[87] Gilboa, Kaplan and Sarel have suggested that under a similar mechanism, attorneys general could bring claims in the context of climate litigation.[88] Another party that could serve as a public plaintiff are organizations representing the type of person at whose expense the companies became enriched.

In a combination of private action with public outcomes – courts could grant *cy pres* relief in a class action.[89] Under the *cy pres* doctrine, a court can instruct the plaintiff in a class action lawsuit to donate some of the settlement stemming from the disgorgement of unjust profits to a charitable organization connected to the lawsuit's matter.[90] This could include organizations advocating on behalf of sexual assault survivors, organizations seeking to increase digital literacy etc.[91]

---

85 A plaintiff who was the subject of a sexual deepfake, for example, may refrain from becoming involved in such a lawsuit for several reasons: e.g. it may force them to relive the trauma and humiliation caused by the creation and circulation of the deepfake, they may fear that filing such a lawsuit will draw increased attention to the deepfake, in what has become known as the *Streisand Effect*. The term was coined in Mike Masnick, *Since When Is It Illegal to Just Mention a Trademark Online?* Techdirt (Jan. 5, 2005), https://www.techdirt.com/2005/01/05/since-when-is-it-illegal-to-just-mention-a-trademark-online/.

86 Keith N. Hylton, *Litigation Costs and the Economic Theory of Tort Law*, 46 U. Miami L. Rev. 111, 113 (1991) (explaining how the costliness of litigation can prevent plaintiffs from suing).

87 Jessica Hill, *Nevada Attorney General Files Lawsuit Against YouTube,* Las Vegas Rev. J. (Jun. 16, 2025), https://www.reviewjournal.com/news/politics-and-government/nevada/nevada-attorney-general-files-lawsuit-against-youtube-3385945/ (reporting that the case included an unjust enrichment claim against YouTube for designing a highly addictive platform and become enriched at the expense of the children users of the platform).

88 Maytal Gilboa, Yotam Kaplan & Roee Sarel, *Climate Change as Unjust Enrichment*, 112 Geo. L. J. 1039 (2024).

89 Stewart R. Shepherd, Comment, *Damage Distribution in Class Actions: The Cy Pres Remedy*, 39 U. Chi. L. Rev. 448, 452 (1972) (describing the *cy pres* remedy).

90 Martin H. Redish, Peter Julian & Samantha Zyontz, *Cy Pres Relief and the Pathologies of the Modern Class Action: A Normative and Empirical Analysis*, 62 Fla. L. Rev. 617, 634 (2010) ("funds [from the settlement] would be used to create a charitable trust, and that trust would be used either to create a charitable foundation or donate to a pre-existing charitable organization related in some way (however loosely) to the subject of the class action suit".).

91 Thomas E. Kadri & Sonja West, *Deepfake Torts: Emerging Torts Framework in U.S. Deepfake Regulation*, J. Tort L. 1, 18 (2025) (noting that a Minnesota bill seeking to address the harms of deepfakes offers this option as a remedy).

4.3 Calculating Harms v. Gains

The law of unjust enrichment focuses on gains, not on harms.[92] Under unjust enrichment doctrine, proof of harm is not a central requirement; the core element of such a claim is the benefit to the defendant.[93] Unjust enrichment is therefore advantageous when harms are difficult to measure and gains are more easily identifiable and quantifiable. This is especially true in the context of the harms stemming from deepfake technology. They are often non-monetary, spread over a large (and sometimes unidentified) number of people and are often anticipated future harms whose value can only be estimated at the present time.[94]

In cases where a concrete sum has been lost due to a deepfake, such as in fraud cases, the calculation of the harm caused may be straightforward. In other cases, the type of harms caused by deepfakes do not lend themselves to precise calculations as easily.[95] The emotional harm suffered by survivors of sexual deepfakes is not easily quantifiable; the fear of being the subject of a sexual deepfake is perhaps even more challenging to calculate; the loss of dignity resulting from a deepfake is immeasurable,[96] and the loss of trust in democratic institutions due to the harms caused by deepfakes is perhaps impossible to quantify.[97] This does not mean that some of these harms cannot be quantified or that courts do not have the tools to overcome such challenges. Since tort remedies are monetary, however, a legal framework that requires quantifying non-monetary harms puts plaintiffs at a disadvantage compared to a legal framework that is based on quantifying concrete enrichment.[98]

The challenges involved with quantifying non-monetary harms that may materialize in the future, is starkly in contrast with the relative ease at which concrete, current, monetary gains can be counted. The enrichment generated from the development and deployment of deepfake technology is concentrated within a small number of companies that are easier to identify and is current and concrete.

---

92 *See* Laycock, *supra* note 77, at 1283.

93 Ernest J. Weinrib, Corrective Justice 117–118 (2012).

94 Kadri & West, supra note 87, at 7 ("Quantifying the harms of sexual deepfakes can be a challenge.").

95 Michael Goodyear, *Dignity and Deepfakes,* Arizona St. L. J. 1 (2025).

96 *Id.*

97 *Id*, at 8 ("While deepfakes can inflict economic and societal harms, [they can] also harm individual victims at an emotional and sociological level, including stripping them of control over their own self-presentation, damaging their reputations, and causing them to internalize shame and even ostracize themselves from society.").

98 *See* Yotam Kaplan, *The Other View of* The Cathedral, 82 Md. L. Rev. 479, 508–18 (2023) (noting that the task to determining compensation in cases of harm is subject to judicial bias).

All these reasons make enrichment-based remedies more practical compared to harm-based ones in the context of the challenges created by deepfakes.

5 Conclusions

Generative AI driven technology has taken a giant leap forward in recent years. The developments driving the technology have facilitated faster, more accessible technology able to generate higher quality, realistic-looking outputs. While this technology has many advantages and holds promises for future innovations, it also has a frightening dark side. Generative AI allows malicious actors to use the technology in harmful ways creating deepfakes that promote political disinformation, facilitate manipulation and fraud and enable the creation of humiliating sexual deepfakes. In this Article we suggested that while harm-based tort claims may be helpful and effective in certain cases, the harms of deepfake technology may be more effectively addressed through the application of the law of unjust enrichment. Enrichment-based remedies offer several benefits compared to harm-based ones. First, they make it easier to identify the defendant. Second, they allow for a multiplicity of plaintiffs, and finally it is easier to quantify concrete current gains concentrated within a small number of actors than to assess future, anticipated harms spread over a large number of victims. Addressing the enrichment generated by companies developing and deploying deepfake technology has the potential to realign the companies' incentivizes driving them to develop safer tools that will limit the harms stemming from their technology while allowing individuals and society to enjoy the benefits the technology has created.