Critics of Generative AI Are Worrying About the Wrong IP Issues

By Daniel Castro | March 20, 2023

Critics argue developers of generative AI systems such as ChatGPT and DALL-E have unfairly trained their models on copyrighted works. Those concerns are misguided. Moreover, restricting AI systems from training on legally accessed data would significantly curtail the development and adoption of generative AI across many sectors. Policymakers should focus on strengthening other IP rights to protect creators.

One of the most visible advancements in artificial intelligence (AI) is the development of generative AI—AI systems that can produce novel images, music, or text in response to user prompts. Users are still exploring potential applications of this technology in many fields, but early results are promising. Already people have used generative AI tools to draft news articles, press releases, and social media posts, create high-quality images, video, and music, and even write code. And many more applications in fields such as medicine, entertainment, and education are on the horizon.

However, some critics argue that generative AI poses a serious threat to content creators. For example, some visual artists have launched online protests denouncing AI and calling for online platforms to block AI-generated art.[1] One of their chief complaints is that when developers train generative AI systems on publicly accessible copyrighted content, they are unfairly exploiting the works of creators.[2] But these critics are wrong. Generative AI systems should not be exempt from complying with intellectual property (IP) laws, but neither should they be held to a higher standard than human creators.[3]

This report refutes five of the most common arguments made about how generative AI is unfair to creators:

1. Training generative AI systems on copyrighted content is theft.

2. Generative AI systems should not train on content without the copyright owner's explicit permission.
3. Generative AI systems should compensate copyright owners for training on their content.
4. Generative AI systems should not be allowed to produce content based on the style of an artist without their permission.
5. Generative AI systems use fragments of copyrighted content in their outputs.

The report also acknowledges that there are legitimate IP rights at stake. Specifically, it identifies five harmful activities:

1. Infringing on copyrights of AI-generated works
2. Distributing copyrighted content
3. Creating forgeries
4. Creating infringing content
5. Impersonating individuals

Finally, the report discusses the impact of generative AI on those harmful activities and recommends policymakers address concerns through robust enforcement of existing rights, offering guidance and clarity to those using these tools, new legislation to combat online piracy, and expanding prohibitions on the distribution of nonconsensual intimate images (sometimes referred to as "revenge porn") to include "deepfakes" (i.e., images and video created using generative AI).

CRITICS ARE WRONG THAT GENERATIVE AI IS UNFAIR TO CREATORS

The list of generative AI applications continues to grow (see table 1). As it does, AI-generated content has created a lot of praise and controversy.[4] Many welcome the advent of generative AI, seeing it as another powerful technology like software-based word processors and video editors that will empower creators to better express themselves. Others fear that AI will devalue artistic works, and artists themselves, by replacing vibrant human creativity with a cold, emotionless algorithm.[5]

The reality is more likely to be somewhere in the middle. While AI-generated content will likely serve as a useful substitute for certain purposes—simple marketing copy, stock images, and royalty-free music—it may hold less appeal for collectors of fine art, music connoisseurs, and literary aficionados. Indeed, art prices have historically operated differently than other goods. The price of fine art is not linked to production costs, but instead to abstract and subjective qualities such as the perceived quality of

the work, the reputation of the artist, and the opinions of gallery owners, collectors, and other authorities.[6] Likewise, some purists—whether they be writers or musicians— as a badge of honor, will likely refuse to use the technology to create content, while many others will adopt the technology because it makes their jobs easier. Indeed, if AI makes creators more productive, it will grow the economy.

Table 1: List of popular generative AI applications

Type	Examples
Image	DALL-E 2, Midjourney, Stable Diffusion
Text	ChatGPT, Bing, Bard
Music	Jukebox, MusicLM
Code	Copilot

But the biggest debate is about whether generative AI systems should be allowed to train their models on text, audio, images, and videos that are legally accessible to Internet users but are also protected by copyright. Some creators argue that it is unfair for developers to train their AI systems on content they have posted on the Internet without their consent, credit, or compensation. Their opposition is generally without merit, but their arguments are worth considering.

Before getting into these arguments, it is important to note that people do not have the right to use copyrighted content any way they want just because they can legally access it on the Internet. However, their not having the right to use it any way they want does not mean they cannot do anything with this content. Copyright law provides copyright owners certain exclusive rights, but these rights are subject to exceptions and limitations, including those under the fair use doctrine. For example, search engines can legally crawl websites without violating copyright laws. While it will ultimately be up to the courts to decide whether a particular use of generative AI infringes on copyright, there is precedent for them to find most uses to be lawful and not in violation of rightsholders' exclusive rights.[7]

1. Is Training Generative AI Systems on Copyrighted Content Theft?

Some argue that training AI systems on copyrighted content is theft plain and simple.[8] Indeed, stealing digital content is a serious problem—online piracy of movies, TV, music, games, and more costs creators billions in lost revenue annually.[9] But online piracy is clearly theft. There is little difference between someone watching a DVD they have shoplifted from a Walmart or Best Buy compared with someone watching a video they have

streamed illegally online—in both cases, they are watching a video they have not paid for.

But seeking inspiration and learning from others is not theft. It is not theft if someone watches a video legally, and that video inspires them to film their own unique creation. Indeed, TikTok and other social media platforms are filled with such videos inspired by related content. Similarly, writers, musicians, and other artists learn their craft by observing past creations. In fact, all creative works are shaped by past works, as creators do not exist in a vacuum. The inspections, impressions, and inspirations of the world around them are what give rise to new ideas. Calling this process theft is clearly inaccurate when applied to the way humans observe and learn, and it is equally inaccurate to describe training a generative AI system (discussed in more detail ahead).

2. Should Generative AI Systems Be Allowed to Train on Content Without the Copyright Owner's Explicit Permission?

Some argue that it is wrong to train AI systems on copyrighted content without first obtaining affirmative consent from the copyright holder. These critics say that even if training AI systems does not amount to theft, copyright owners should still have the right to determine how others use their works, since they created it. And indeed, the law does confer certain rights to copyright owners, such as the right to reproduce a work, the right to prepare derivative works, the right to perform a work publicly, and the right to display a work publicly.[10] However, this argument falls apart upon closer examination.

As noted, copyright owners have the right to decide whether to display or perform their works publicly. But if they choose to display their work in public, others can use their works in certain ways without their permission. For example, photographers can take pictures of sculptures or graffiti in public places even when those works are protected by copyright. Copyright prevents photographers from selling those images, but it does not require them to get permission from the copyright owner to take photos. Likewise, individuals do not need to get permission from a copyright owner to study a painting they see in a gallery or a song they hear on the radio. People are free to observe these works and use what they learn from them to create future content without the explicit permission of the copyright owners. There is no intrinsic rationale for why users of generative AI systems would need to obtain permission to train on copyrighted content they have legal access to. Musicians might practice a copyrighted song they heard on Spotify hundreds of times to learn to play an instrument or use their well-honed auditory memory to recall elements of pieces they have heard before. Learning from legally accessed works does not violate a copyright owner's exclusive reproduction and distribution rights. Unless human creators will be required to obtain permission before they can study another person's work, this requirement should not be applied to AI.

3. Should Generative AI Systems Compensate Copyright Owners for Training on Their Content?

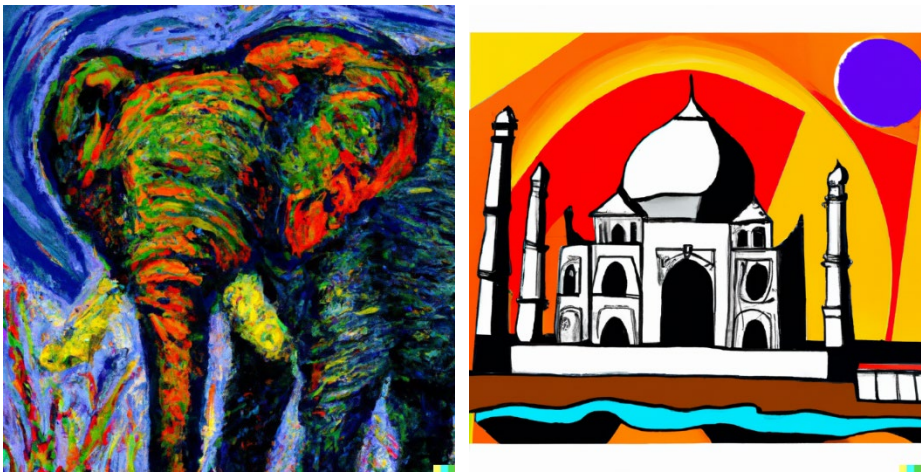
Some argue that people should pay copyright owners to train generative AI systems on their content because they are obtaining value from this process. But copyright owners do not have this same expectation when other human creators learn from their works. Budding musicians listen to hours of music, young writers closely study their favorite novels, and amateur painters spend hours looking at works in galleries. They do not have to pay copyright owners a separate fee to obtain the right to study the techniques, styles, and artistry of others. Indeed, when someone buys a ticket to an art museum or purchases a book, there is not one price for future creators and one for everyone else.

Critics of generative AI are also likely to overestimate individual contributions. Generative AI systems train on massive corpuses of data. For example, Stable Diffusion trained on a dataset of 600 million images.[11] Of those, out of a sample of 12 million of the most "aesthetically attractive images" (which presumably skew more toward works of art than other random images from the Internet), the most popular artist (Thomas Kinkade) appeared 9,268 times.[12] Put differently, the most popular artist in the dataset likely represented only 0.0015 percent of all images in the dataset. Or consider LaMDA, a large language model created by Google, that trained on 1.56 trillion words scraped from the Internet.[13] Given the size of these models, the contribution of any single person is miniscule.

4. Should Generative AI Systems Be Prohibited From Producing Content Based on the Style of an Artist Without Their Permission?

Generative AI systems allow users to request output that matches a specific style. For example, a user of DALL-E could generate an image using a prompt such as "Elephant in the style of Van Gogh" or "The Taj Mahal in the style of Picasso." (See results in figure 2.) Some have argued that generative AI systems should not be able to produce content that mimics a particular artist's distinctive visual style without their permission.[14] However, once again, such a demand would require holding AI systems to a different standard than humans. Artists can create an image in the style of another artist because copyright does not give someone exclusive rights to a style.[15] For example, numerous artists sell Pixar-style cartoon portraits of individuals.[16] And it is perfectly legal to commission someone to write an original poem in the style of Dr. Seuss or an original song in the style of Louis Armstrong. Users of generative AI systems should retain the same freedom.

Figure 1: Images generated by DALL-E in response to the prompts "Elephant in the style of Van Gogh" and "The Taj Mahal in the style of Picasso," respectively



5. Do Generative AI Systems Use Fragments of Copyrighted Content in Their Outputs?

Some argue that generative AI systems are "21st-century collage tool[s] that [remix] the copyrighted works of millions of artists whose work was used as training data."[17] Rather than producing unique output, these critics claim that generative AI systems are merely stitching together fragments of copyrighted content their algorithms have ingested. However, this argument reflects a poor understanding of how generative AI systems work.

Generative AI systems do not produce remixes of existing content. They are not, as some might mistakenly imagine, taking small samples of various works, altering them, and then recombining them in a new order. Instead, generative AI systems use massive amounts of training data to create incredibly complex prediction models that allow them to produce realistic content in response to specific prompts. For example, OpenAI's GPT-3 large language model trained on 45 terabytes of text with 175 billion parameters.[18] The DALL-E 2 image model trained on 250 million images with 3.5 billion parameters.[19]

When prompted to "write a story about a duck on the moon" or "create a picture of a duck on the moon," these models are not searching through existing data to find the closest match, but instead are generating new content that fits certain parameters based on the statistical patterns they have observed in the training data. For example, a "duck" consists of certain essential elements, such as a bird with a short neck, stout body, and webbed feet. Each of these elements has its own range of acceptable parameters: colors, proportions, etc. The AI model does not

understand any of these parameters (and indeed, the parameters are likely much more abstract than this example), but it is uses them to generate high-quality output.

GENERATIVE AI DOES NOT EXCUSE OTHER ILLEGAL ACTS

Many of the concerns about generative AI are misguided and reflect a tendency for fear to outpace understanding of emerging technologies.[20] Indeed, techno panics about AI are not new: Fears about AI taking jobs, destroying privacy, hiding bias, and subjugating humanity run rampant not only in dystopian science fiction novels but also among members of the press, policymakers, and professional pundits.[21] While critics are wrong to argue that generative AI presents a threat to copyright owners' legal rights, there are legitimate IP issues for policymakers to consider.

This section discusses some of the ways people can infringe on IP rights and the policy implications for generative AI.

1. Infringing on Copyrights of AI-Generated Works

Individuals who use AI to create content deserve copyright protection for their works. In the United States, the U.S. Copyright Office has developed initial guidance for registering works created by using AI tools.[22] The Copyright Office should not grant copyright requests to an AI system itself or for works in which there is no substantial human input.[23] Copyright protection for AI-generated content should function similarly to that of photographs wherein a machine (i.e., a camera) does much of the mechanical work in producing the initial image, but it is a variety of decisions by the human photographer (subject, composition, lighting, post-production edits, etc.) that shape the final result. Likewise, individuals who use AI tools to create content do more than just click a button, such as experimenting with different prompts, making multiple variations, and editing and combining final works. As generative AI becomes a mainstream tool used widely by content creators, policymakers should ensure copyright law fully protects their rights, both domestically and abroad, and offer regularly updated guidance and clarity for those using AI tools, especially as these technologies mature.

2. Distributing Copyrighted Content

Just as it is against the law to break into an art gallery or a music studio, it is also illegal to unlawfully gain access to private digital files. Many artists choose not to post their works publicly on the Internet. For example, many portrait and wedding photographers will only share their completed works with their clients. Other creators, including visual artists, musicians, and writers, use services such as Patreon, Ko-fi, or OnlyFans to limit who can access their works. If someone subverts access controls to gain entry to these systems, they are engaging in unlawful conduct—and are typically violations of the Computer Fraud and Abuse Act (CFAA). Likewise, if someone distributes copyrighted content without permission, such as by

posting it on the public Internet or sharing it on peer-to-peer file sharing services, they are committing a crime. Law enforcement should prosecute these crimes, including by working with foreign law enforcement agencies to prosecute crimes across borders.

Those training generative AI systems may unintentionally include pirated content that is available online in their datasets. The response to this problem should be to reduce the availability of infringing content online, not stop using generative AI. There are many steps policymakers should take to reduce online piracy. For example, law enforcement should work with Internet stakeholders, such as hosting sites, domain registrars, and Internet service providers (ISPs) to dedicate more resources for taking down infringing content. Congress should also pass legislation to more aggressively limit online piracy.[24] For example, it should make a law that provides rightsholders with the ability to obtain a legal injunction requiring ISPs in the United States to block access to websites whose primary purpose is to widely distribute copyright-protected content—a step many other countries have taken.[25] Congress should also pass legislation directing the U.S. Copyright Office to work with the private sector to designate standard technical measures online services should have to adhere to in order to qualify for the safe harbor provisions under the Digital Millennium Copyright Act.[26]

Once infringing content is removed from the public Internet, it will not be available to anyone to use, including those training generative AI systems. But there is no more reason to prohibit the use of generative AI systems that may have unintentionally included pirated content than there is a need to prohibit artists from creating works after seeing art found in museums later discovered to be stolen.[27]

3. Creating Forgeries

While generative AI allows users to create art similar to other artists, it does not allow anyone to misrepresent the creator or the provenance of the work. Just as it is illegal for artists, no matter how talented, to misrepresent their works as that of someone else, so too is it unlawful to use generative AI to misrepresent content as being created by another artist.[28] For example, someone might enjoy creating drawings of their own original characters in the style of Bill Watterson, the cartoonist behind the popular Calvin and Hobbes comic strip, but they cannot misrepresent those drawings as having been created by Watterson himself.[29] Addressing this type of problem is a long-standing issue in the art world. For example, forgeries of Thomas Kinkade paintings were mass produced in China and Thailand at the height of his popularity.[30] Law enforcement can and should prosecute individuals who create frauds and buyers should always conduct due diligence before purchasing.

4. Creating Infringing Content

Generative AI may allow creators to produce works that have similar styles to existing copyrighted works, but they do not allow creators to produce identical or nearly identical works. Copyright owners, including those of literary, musical, and artistic works, can claim infringement if someone produces a work that is substantially similar to their own because they have an exclusive right to produce derivative works. Courts have repeatedly intervened in these cases, including for sampling small portions of a song, such as when Queen and David Bowie successfully sued Vanilla Ice because the bass line in "Ice Ice Baby" came directly from "Under Pressure," and for replicating key elements of a song, such as when the estate of Marvin Gaye successfully sued Robin Thicke and Pharrell Williams for the similarities between "Blurred Lines" and "Got to Give It Up."[31] While the latest generative AI systems mostly produce novel content, it is possible for these systems to replicate content from training data.[32] Artists can and should continue to enforce their rights in court when someone produces nearly identical work that unlawfully infringes on their copyright, whether that work was created entirely by human hands or involved the use of generative AI.

5. Impersonating Individuals

The right of publicity is the IP right that protects individuals from the unauthorized commercial use of their identity. This right is especially important for celebrities, as it enables them to control how others use their likeness commercially, such as in advertisements or in film and TV. While generative AI—specifically deepfake technology—makes it easier to create content that impersonates someone else, the underlying problem itself is not new. Courts have repeatedly upheld this right, including for cases involving indirect uses of an individual's identity. In one notable case, game show hostess Vanna White won damages for an advertisement that depicted a robot meant to impersonate her.[33] In another, late-night television star Johnny Carson won a claim against a portable toilet company that used the phrase "Here's Johnny" without his permission.[34] Generative AI has not changed the fact that individuals should continue to enforce their publicity rights by bringing cases against those who violate their rights.

Generative AI also raises questions about who owns rights to certain character elements. For example, if a movie studio wants to create a sequel to a film, can it use generative AI to digitally recreate a character (including the voice and image) or does the actor own those rights? And does it matter how the film will depict the character, including whether the character might engage in activities or dialogue that could reflect negatively on the actor? These types of questions will likely be settled through contracts performers sign addressing who has rights to a performer's image, voice, and more.

Deepfake technology also makes it much easier to produce hyper-realistic fake nude and sexually explicit images and videos of individuals without their consent. While this problem is not entirely new, the scale of the problem is much greater than in the past. Legislation is still needed in many jurisdictions to address distribution of nonconsensual intimate images and videos (sometimes referred to as "image-based sexual abuse" or "revenge porn") including those created by deepfakes. While more jurisdictions have laws prohibiting distribution of this type of content, only a few of them address fake content.[35] Policymakers should update and expand these laws to better protect individuals.

CONCLUSION

Generative AI is an important technological advancement that offers a variety of useful applications many sectors of the economy should welcome. While there are many important considerations for how generative AI impacts IP rights and how policymakers can protect rightsholders, critics are wrong to claim that generative AI models should not be allowed to train on legally accessed copyrighted content. Moreover, imposing restrictions on training generative AI models on lawfully accessed content could unnecessarily limit its development. Instead, policymakers should offer guidance and clarity for those using these tools, focus on robust IP rights enforcement, create new legislation to combat online piracy, and expand laws to protect individuals from impersonation.

REFERENCES

1. Chloe Xiang, “Artists Are Revolting Against AI Art on ArtStation,” *Motherboard*, December 14, 2022, https://www.vice.com/en/article/ake9me/artists-are-revolt-against-ai-art-on-artstation.
2. Emma Roth, “Microsoft, GitHub, and OpenAI ask court to throw out AI copyright lawsuit,” *The Verge*, January 28, 2023, https://www.theverge.com/2023/1/28/23575919/microsoft-openai-github-dismiss-copilot-ai-copyright-lawsuit.
3. Daniel Castro, “Ten Principles for Regulation That Does Not Harm AI Innovation,” Center for Data Innovation, February 8, 2023, https://datainnovation.org/2023/02/ten-principles-for-regulation-that-does-not-harm-ai-innovation/.
4. Henry Kissinger, Eric Schmidt and Daniel Huttenlocher, “ChatGPT Heralds an Intellectual Revolution,” *Wall Street Journal*, February 24, 2023, https://www.wsj.com/articles/chatgpt-heralds-an-intellectual-revolution-enlightenment-artificial-intelligence-homo-technicus-technology-cognition-morality-philosophy-774331c6.
5. Yasmeen Kanaan, “A Future Dystopia: Will Human Artists Ever be Replaced by Artificial Intelligence?” *artmejo*, n.d., https://artmejo.com/a-future-dystopia-will-human-artists-ever-be-replaced-by-artificial-intelligence/.
6. Dominique Sagot-Duvaroux et al., “Factors Affecting Price on the Contemporary Art Market,” *Cultural Economics* (Berlin: Springer 1992), https://doi.org/10.1007/978-3-642-77328-0_11.
7. James Vincent, “Getty Images is suing the creators of AI art tool Stable Diffusion for scraping its content,” *The Verge*, January 17, 2023, https://www.theverge.com/2023/1/17/23558516/ai-art-copyright-stable-diffusion-getty-images-lawsuit; James Vincent, “AI art tools Stable Diffusion and Midjourney targeted with copyright lawsuit,” *The Verge*, January 16, 2023, https://www.theverge.com/2023/1/16/23557098/generative-ai-art-copyright-legal-lawsuit-stable-diffusion-midjourney-deviantart.
8. Nick Vincent and Hanlin Li, “ChatGPT Stole Your Work. So What Are You Going to Do?” *Wired*, January 20, 2023, https://www.wired.com/story/chatgpt-generative-artificial-intelligence-regulation/.
9. David Blackburn, Jeffrey A. Eisenach, and David Harrison Jr., “Impacts of Digital Video Piracy on the U.S. Economy,” NERA Economic Consulting and Global Innovation Policy Center, June 2019, https://www.theglobalipcenter.com/wp-content/uploads/2019/06/Digital-Video-Piracy.pdf.
10. “What is Copyright?” U.S. Copyright Office, n.d., https://www.copyright.gov/what-is-copyright/.
11. Andy Baio, “Exploring 12 Million of the 2.3 Billion Images Used to Train Stable Diffusion’s Image Generator,” Waxy.org, August 30, 2022, https://waxy.org/2022/08/exploring-12-million-of-the-images-used-to-train-stable-diffusions-image-generator/.
12. Ibid.
13. Romal Thoppilan et al., “LaMDA: Language Models for Dialog Applications,” Google, February 10, 2022, https://arxiv.org/pdf/2201.08239.pdf.

14. Will Knight, “Algorithms Can Now Mimic Any Artist. Some Artists Hate It,” *Wired*, August 19, 2022, https://www.wired.com/story/artists-rage-against-machines-that-mimic-their-work/.

15. Greg Kanaan, “You Can’t Copyright Style,” The Legal Artist, February 1, 2016, https://www.thelegalartist.com/blog/you-cant-copyright-style.

16. “Get The Best Disney Drawing Services,” Fiverr, n.d., https://www.fiverr.com/gigs/disney (accessed March 13, 2023); “Get the Best Pixar Services,” Fiverr, n.d., https://www.fiverr.com/gigs/pixar (accessed March 13, 2023).

17. Matthew Butterick, “We’ve filed a lawsuit challenging Stable Diffusion, a 21st-century collage tool that violates the rights of artists,” Stable Diffusion Litigation, January 13, 2023, https://stablediffusionlitigation.com/.

18. Tom B. Brown, et al., “Language Models are Few-Shot Learners,” *arXiv*, July 22, 2020, https://arxiv.org/pdf/2005.14165.pdf.

19. Aditya Ramesh et al., “Hierarchical Text-Conditional Image Generation with CLIP Latents,” *arXiv*, April 13, 2022, https://cdn.openai.com/papers/dall-e-2.pdf.

20. Daniel Castro and Alan McQuinn, “The Privacy Panic Cycle: A Guide to Public Fears About New Technologies,” Information Technology and Innovation Foundation, September 10, 2015, https://itif.org/publications/2015/09/10/privacy-panic-cycle-guide-public-fears-about-new-technologies/.

21. Robert D. Atkinson, “’It's Going to Kill Us!’ and Other Myths About the Future of Artificial Intelligence,” Information Technology and Innovation Foundation, June 6, 2016, https://itif.org/publications/2016/06/06/its-going-kill-us-and-other-myths-about-future-artificial-intelligence/.

22. “Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence,” *Federal Register*, Vol. 88, No. 51, March 16, 2023, https://www.govinfo.gov/content/pkg/FR-2023-03-16/pdf/2023-05321.pdf.

23. Jane Recker, “U.S. Copyright Office Rules A.I. Art Can’t Be Copyrighted,” *Smithsonian Magazine*, March 24, 2022, https://www.smithsonianmag.com/smart-news/us-copyright-office-rules-ai-art-cant-be-copyrighted-180979808/.

24. Nigel Cory and Daniel Castro, “Online Piracy Is Digital Looting, and It Needs to Stop,” *Inside Sources*, September 15, 2020, https://insidesources.com/online-piracy-is-digital-looting-and-it-needs-to-stop/.

25. Nigel Cory, “A Decade After SOPA/PIPA, It’s Time to Revisit Website Blocking,” Information Technology and Innovation Foundation, January 26, 2022, https://itif.org/publications/2022/01/26/decade-after-sopa-pipa-time-to-revisit-website-blocking/.

26. Jaci McDole, “Why the SMART Copyright Act Is a Smart Idea,” Information Technology and Innovation Foundation, October 7, 2022, https://itif.org/publications/2022/10/07/why-the-smart-copyright-act-is-a-smart-idea/.

27. Charly Wilder, “When a Visit to the Museum Becomes an Ethical Dilemma,” *The New York Times*, February 28, 2023, https://www.nytimes.com/2023/02/14/travel/museums-stolen-art.html.

28. Oscar Holland, “The husband-and-wife forgers who fooled the art market — and made millions,” *CNN*, February 7, 2023, https://www.cnn.com/style/article/wolfgang-helen-beltracchi-forgers/index.html.

29. Andrew Liptak, “Star Wars and Calvin & Hobbes is the perfect combination in these fantastic cartoons,” *The Verge*, March 5, 2017, https://www.theverge.com/2017/3/5/14731000/star-wars-calvin-hobbes-brian-kesinger-art-comics.

30. Malcolm Berko, “Thomas Kinkade paintings not worth much, if anything, because of oversaturation,” *Business Journal*, February 1, 2019, https://www.bizjournals.com/columbus/news/2019/02/01/berko-thomas-kinkade-paintings-not-worth-much-if.html.

31. “Recent Cases,” Musicians Institute Library, December 2, 2022, https://library.mi.edu/musiccopyright/currentcases.

32. Gowthami Somepalli et al., “Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models,” *arXiv*, December 12, 2022, https://arxiv.org/pdf/2212.03860.pdf.

33. *White v. Samsung Electronics America, Inc.*, 971 F.2d 1395 (9th Cir. 1992), https://law.justia.com/cases/federal/appellate-courts/F2/971/1395/71823/.

34. *John W. Carson v. Here’s Johnny Portable Toilets, Inc.*, 698 F.2d 831 (6th Cir. 1983), https://openjurist.org/698/f2d/831/carson-v-heres-johnny-portable-toilets-inc.

35. Karen Hao, “Deepfake porn is ruining women’s lives. Now the law may finally ban it.” *MIT Technology Review*, February 12, 2021, https://www.technologyreview.com/2021/02/12/1018222/deepfake-revenge-porn-coming-ban/.

ABOUT THE AUTHOR

Daniel Castro is the director of the Center for Data Innovation and vice president of the Information Technology and Innovation Foundation. He has a B.S. in foreign service from Georgetown University and an M.S. in information security technology and management from Carnegie Mellon University.

ABOUT THE CENTER FOR DATA INNOVATION

The Center for Data Innovation studies the intersection of data, technology, and public policy. With staff in Washington, London, and Brussels, the Center formulates and promotes pragmatic public policies designed to maximize the benefits of data-driven innovation in the public and private sectors. It educates policymakers and the public about the opportunities and challenges associated with data, as well as technology trends such as open data, artificial intelligence, and the Internet of Things. The Center is part of the Information Technology and Innovation Foundation (ITIF), a nonprofit, nonpartisan think tank.

Contact: info@datainnovation.org

datainnovation.org