



Feb 14, 2024

Categories:

[Publications](#)

[Technology Blog](#)

Authors:

[William S. Morriss](#)

Will AI Destroy the DMCA Copyright Compromise?

In the 1990s, as today, the internet ran on content, which meant that it both had, and was, a problem. It had a problem in that the possibility of being held liable for copyright infringement threatened to deter people who might otherwise provide the very services necessary for the internet to operate.^[1] It was also a problem in that the internet provided a hitherto unknown and, from a copyright perspective, potentially apocalyptic ability for members of the general public to make and distribute unlimited digital copies of creative works.^[2] For the internet to develop into the form we know today, driven in no small part by user-generated content, these problems needed to be resolved in a way that respected the rights of artists and authors while allowing for continued investment in, and expansion of, web technology.

Ultimately, the internet's content problem was resolved by the Digital Millennium Copyright Act (DMCA), which aligned the incentives of service providers and copyright holders through a safe harbor, notice and takedown framework. In effect, the DMCA established a mechanism under which companies that ran the technology that ran the internet couldn't be held liable for infringing user content if (a) they weren't aware that the content was infringing and if (b) they removed or disabled access to the infringing content upon being made aware of its existence.^[3] While there have been some hiccups along the way, this legal framework ultimately led to our current system, in which rightsholders can control their content by informing the appropriate service providers when infringement is detected, and service providers can have legal certainty and leeway to innovate so long as they remove infringing materials they are made aware of.

AI and the Limits of Existing Copyright Laws

Artificial intelligence today faces a content situation similar to that of the internet in the 1990s. Like the 90s internet, today's AI technology requires content to function. For example, ChatGPT was

trained using a corpus of about 300 billion words, and the MidJourney and DALL-E image generation apps were trained on a dataset of about six billion image-text pairs.^[4] Similarly, like the internet in the 1990s, today's AI technology is seen as posing an almost existential threat to artists and other content owners.^[5]

Similarities aside, there is a crucial difference from a legal perspective between today's AI models and the internet of the 1990s. The early internet stored creative works as identifiable pieces of content that could be removed if they were infringing. Today's AI technology does not. Instead of storing particular creative works, creative works are used to train an artificial intelligence model, with each individual work being reflected in infinitesimal changes spread across the values of billions of parameters (or trillions, in the case of OpenAI's GPT-4). Because of this, and given current technology, the only way to remove an infringing item from a generative AI like ChatGPT or MidJourney is to retrain the AI from scratch without the item(s) in question.^[6] Given that the cost for training such a model can run into the millions of dollars,^[7] this renders a DMCA-style notice and takedown style solution unworkable, especially considering the potential for different artists, authors, publishers and other rightsholders to submit multiple independent takedown notices. On the other hand, without DMCA-style liability protections, it's easy to see today's AI products, which appear to have been largely trained on datasets that include significant amounts of unlicensed third-party material, getting buried under, and potentially destroyed by, an avalanche of litigation.

Key Takeaways

The bottom line is that the legal regime that has insulated online service providers from copyright liability for the last 20 years does not seem to be applicable, even in principle, to online providers of artificial intelligence. Moreover, as more and more online service providers incorporate artificial intelligence into their products,^[8] potential infringement claims could undermine even the current, relatively peaceful coexistence between the technology and content industries. Perhaps this is as it should be. Rightsholders

have expressed concerns about the ability of small creators to police infringement using the notice and takedown portion of the DMCA, and even the U.S. Copyright Office itself has concluded that the intended balance between service providers and rightsholders has gone askew.^[9] On the other hand, it is difficult to see how to make continued AI development compatible with broad copyright liability, especially since no datasets currently exist that are both suitable to replace those which have led to infringement lawsuits and are fully licensed for AI training. As a result, it is not at all clear what kind of equilibrium can be reached between these competing interests, though it seems certain that both technical and legal innovations will be necessary to reach a solution to the AI/copyright conflict.

Frost Brown Todd has a dedicated team that advises clients on their own AI integrations, helping them institute governance programs, proactively address data security and privacy concerns, and navigate a host of other AI opportunities and risks unique to their industry. To learn more about how Frost Brown Todd can help empower your company's AI journey, contact the author of this article or any attorney with the firm's [Artificial Intelligence](#) team.

AI Legally Speaking

Explore the collection of articles below for more AI-related coverage and legal analysis.

- [***Managing Data Security and Privacy Risks in Enterprise AI***](#)
- [***New Expectations for AI: Governance is Key for AI Success in Health Care***](#)
- [***USPTO Guidance Suggests Two Strategies for AI Inventions***](#)
- [***FTC's New Rule on Consumer Reviews: Ensuring Compliance with Human- and AI-Generated Content***](#)
- [***Decoding Colorado's Artificial Intelligence Act***](#)
- [***Investment Management + AI = New Opportunities, Unique Risks***](#)
- [***AI Chatbots, Hallucinations, and Legal Risks***](#)
- [***AI and Energy: Network Innovation and Growth***](#)

- [**AI and Health Care Providers: Managing the Risks with Data**](#)
- [**Using Generative AI in Manufacturing: Three Key Considerations**](#)
- [**Three Steps to AI Success for CEOs**](#)

[Learn More](#)

[1] The early cases did indicate that bulletin board operators or similar service providers could be held liable for material uploaded by their users, though they were not consistent regarding what was necessary for such liability to be imposed. *Compare Playboy Enterprises, Inc. v. Frena*, 839 F.Supp. 1552 (M.D. Fla. 1993) (bulletin board operator liable for infringing photos uploaded by a user even though he was unaware of the infringement and removed the photos upon being made aware of the matter) with *Religious Technology Center v. Netcom On-Line Communication Services, Inc.*, 907 F.Supp. 1361 (N.D. Cal. 1995) (bulletin board operator not directly liable for infringing materials uploaded by a user, but may be liable for contributory infringement depending on factors such as whether the operator should have known of the infringement and whether the operator substantially participated in the infringement).

[2] *E.g.*, Section 512 of Title 17: A Report of the Register of Copyrights, May 2020, available at <https://www.copyright.gov/policy/section512/section-512-full-report.pdf>, (“Copyright Office Report”) page 1 (“Traditional content industries faced what many came to view as an existential threat, from the convergence of newly dominant and near-lossless digital media formats with a world-wide, interconnected network that facilitated the distribution of digital files of any type.”).

[3] 17 USC §§ 512(a)–(d).

[4] *Researchers Warn We Could Run Out of Data to Train AI by 2026. What then?*, The Conversation, November 7, 2023, available at <https://theconversation.com/researchers-warn-we-could-run-out-of-data-to-train-ai-by-2026-what-then-216741>.

[5] See, e.g., *Andersen et. al. v. Stability AI Ltd., et al.*, case no. 3:23-cv-00201, document 1 (complaint), available at

<https://stablediffusionlitigation.com/pdf/00201/1-1-stable-diffusion-complaint.pdf>, paragraph 5 (explaining that Midjourney and similar generative AI models allowed users to generate new works “in the style” of a particular artist, competing with the artist in the market and eliminating the need to go to the artist for commissioned works).

[6] Stephen Pastis, *A.I.’s un-learning problem: Researchers say it’s virtually impossible to make an A.I. model ‘forget’ the things it learns from private user data*, *Fortune* (August 30, 2023), available at <https://fortune.com/europe/2023/08/30/researchers-impossible-remove-private-user-data-delete-trained-ai-models/>.

[7] Chuan Li, *OpenAI’s GPT-3 Language Model: A Technical Overview*, *Lambda Labs* (June 3, 2020), available at <https://lambdalabs.com/blog/demystifying-gpt-3>.

[8] See, e.g., *Everyday AI in Microsoft 365*, available at <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE33c65> (explaining how Microsoft is applying artificial intelligence to make Office “easier to use, more collaborative and more secure.”).

[9] Copyright Office Report at 1: many OSPs spoke of section 512 as being a success, enabling them to grow exponentially and serve the public without facing debilitating lawsuits. Rightsholders reported a markedly different perspective, noting grave concerns with the ability of individual creators to meaningfully use the section 512 system to address copyright infringement and the “whack-a-mole” problem of infringing content reappearing after being taken down. Based upon its own analysis of the present effectiveness of section 512, the Office has concluded that Congress’ original intended balance has been tilted askew.

© 2026 FBT Gibbons LLP. May be considered attorney advertising in some jurisdictions.