

Updates

April 20, 2023

First Lawsuits Arrive Addressing Generative AI



This is the second of a three-part series on the hot legal topics surrounding generative artificial intelligence (AI) (see Part 1: [The Latest Chapter in Copyrightability of AI-Generated Works](#)).

As the quality of generative AI tools has soared, copyright and other intellectual property (IP) issues around generative AI tools have attracted increased attention. Some artists, creators, and performers have raised concerns about the use of their content or identity in connection with these technologies and fear that these technologies could in some sense replace them. Developers and users of these tools, however, point to their benefits, and the value of innovation, and highlight the need for access to broad data resources to facilitate that innovation. Now that the initial lawsuits involving these technologies have been filed, these issues may be addressed by the courts for the first time.

The Cases Begin

GitHub Copilot Lawsuit ([Complaint](#))

The first case to be filed involving generative AI is a class-action lawsuit filed in November against GitHub, Microsoft, and OpenAI, involving GitHub's [Copilot](#) tool. Copilot is an AI-powered tool that suggests new lines of code in real time based on what a programmer has already written. This case does not raise any copyright infringement claims, but instead focuses mostly on breach of contract and privacy-related claims. The plaintiffs allege that Copilot copies code from publicly available software repositories on GitHub without meeting the requirements of the open-source licenses applicable to such code (e.g., by failing to provide attribution, copyright notices, and a copy of the license terms, and by not making Copilot itself open source). The Complaint includes other claims, such as violation of 17 U.S.C. § 1202 for the alleged removal of copyright management

information (CMI), claims relating to GitHub's handling of "personal data" and "personal information," a claim of wrongful interference with the plaintiff's business interests and expectations, and claims of fraud, false designation of origin, unjust enrichment, and unfair competition.

***Andersen et al. v. Stability AI Ltd. et al* ([Complaint](#))**

In January, the same legal team that filed the Copilot lawsuit filed another class-action lawsuit relating to three image generation tools: [Stable Diffusion](#) (developed by Stability AI), [Midjourney](#) (developed by Midjourney), and DreamUp (developed by DeviantArt). These AI-powered tools produce images in response to text inputs from users. The plaintiffs claim that the models powering these tools were trained using copyrighted images (including those owned by the plaintiffs), which they allege were scraped from the internet. They describe this technology as "merely a complex collage tool," and claim the training images are actually stored in the tools as compressed copies. A number of technologists have [criticized](#) this characterization as containing technical inaccuracies and misrepresenting how generative AI tools function. These critics argue that it is incorrect to say that the models store any portion of the training data, and that the new images are created based on patterns and other information that the models have learned from training data sets.

Unlike the Copilot lawsuit, this lawsuit raises copyright infringement claims, including claims that the defendants directly infringed the plaintiffs' copyrights by using the plaintiffs' works to train the models and by creating unauthorized derivative works and reproductions of the plaintiffs' work in connection with the images generated using these tools. The plaintiffs also raise claims of vicarious copyright infringement (based on the activities of users of the tools), violation of 17 U.S.C. § 1202 (for the alleged removal of copyright management information), unfair competition, violation of the plaintiffs' statutory and common law right of publicity (based in part on the alleged ability to use the tool to generate work "in the style of" specific artists), breach of contract claims against defendant Deviant Art (for allegedly violating its own terms of service and privacy statement), and unfair competition.

Getty Images Lawsuits ([Complaint](#))

A few days after the *Andersen* class-action lawsuit was filed, Getty Images filed a separate lawsuit against Stability AI in the High Court of Justice in London. According to a [press release](#), Getty Images claims that Stability AI had "unlawfully copied and processed millions of images protected by copyright and the associated metadata owned or represented by Getty." The Complaint is not yet publicly available as of the publication of this Update. However, in early February, Getty Images filed another lawsuit against Stability AI in the U.S. District Court for the District of Delaware, which raised a variety of copyright and trademark-related claims, including claims of direct copyright infringement, violation of 17 U.S.C. § 1202 (for both allegedly removing copyright management information and providing false copyright management information), trademark infringement, unfair competition, trademark dilution, and deceptive trade practices. These claims stem from both the training of the models used in Stable Diffusion (which it alleges was trained on "at least 12 million" copyrighted images and accompanying text and metadata from the Getty Images website) as well as images generated by Stable Diffusion users. The suit alleges that some of the generated images contained the Getty Images watermark and claims that in some cases, the watermark was modified and applied to "bizarre or grotesque synthetic imagery that tarnished Getty Images' hard earned reputation... ."

What Are the Key Questions These Cases Are Likely To Address?

These cases raise multiple issues of first impression regarding the use of AI tools, which could have critical implications for the future of generative AI technology.

Here are some of the key questions at issue in these cases:

1. Does training a model on copyrighted material require a license?

Generative AI relies on models that are trained on vast amounts of data, and such training generally entails making interim copies of the training materials as part of the training process. This training enables the algorithms to "learn" patterns and the statistical relationships between elements (e.g., for images, things like size, shape, proportion, color, and relative position), which enables AI models to gain an understanding, for example, of what makes a cat "catlike." This information can then be used to create a new picture of a cat. One of the questions likely to be addressed in these cases is whether such interim copying requires a license.

Proponents of these tools often argue that such interim copies constitute fair use because they are made for the purpose of extracting and gaining an understanding of unprotected elements of the training materials (e.g., factual and statistical information), rather than to copy protected expression. Fair use is an analysis that requires examining four factors that are set forth in the fair use section of the Copyright Act and applying them to the specific facts involved. Although there is not yet any applicable case law applying fair use to the process of training machine learning models, some point to cases in other areas (such as reverse engineering video games) that have held that interim copying of a work to gain understanding of unprotected elements is fair use.^[1] Others, however, argue fair use should not apply to generative AI tools. They argue that these tools are used to generate works that are of a similar nature to the works used to train the model and that therefore, this type of use is not sufficiently "transformative" (a key fair use factor).

2. Does the output generated by a generative AI tool infringe the copyright in the materials on which the model was trained?

These cases also raise infringement claims in connection with the generation of images or other output that results from the use of generative AI tools. Questions that may arise here include whether such output constitutes a derivative work of, or infringes the reproduction right in, the training data. Courts may consider factors such as whether or not any alleged similarities between the output and the training data are merely due to similarities in unprotected elements, or if there are substantial similarities in protected expression, and whether or not the use of a specific content to train a model could be considered a de minimis use. If AI output is found to be infringing, there may also be questions as to who is liable for such infringement.

3. Does generative AI violate restrictions on removing, altering, or falsifying copyright management information?

[Section 1202](#) of the Digital Millennium Copyright Act (DMCA) provides certain restrictions regarding the alteration or removal of copyright management information (CMI) and regarding the provision, distribution, or importation of "false" CMI. CMI is defined in Section 1202(c) and includes, among other things, the copyright notice, title, and other information identifying a work, the name and other identifying information about the creators and copyright owners of the work, and information regarding the terms for use of the work.

Section 1202(a) of the DMCA prohibits providing, distributing, or importing for distribution false CMI if it is done "knowingly" and with "the intent to induce, enable, facilitate, or conceal infringement." Getty Images' Complaint alleges that Stability AI provides false CMI, based on examples it provides showing output from the Stable Diffusion tool containing modified versions of the Getty Images watermark. This raises questions such as (1) whether generated output that includes someone else's watermark constitutes false CMI under Section 1202, (2) whether Stability AI is "providing" the false CMI or whether it is an unintended result of an automated process initiated by the user, and (3) what is required to provide the requisite knowledge and intent.

Section [1202\(b\)](#) of the DMCA prohibits (1) intentionally removing or altering any CMI, (2) distributing CMI that one knows to have been altered or removed, or (3) distributing or publicly performing copies of works knowing that the CMI has been removed or altered, provided that in each case, a defendant must also be shown to have known, or had reason to know, that its actions would "induce, enable, facilitate, or conceal an infringement." Both the *Copilot* and *Getty Images* lawsuits raise claims for violation of Section 1202(b). The *Getty Images* suit alleges the defendants intentionally removed or altered CMI in the form of watermarks and metadata associated with images Stability AI allegedly copied from the Getty Images website. One issue these cases may address is the level of proof necessary to establish that the removal or alteration was an "intentional" act and that the defendants knew or had reason to know that their actions would induce, enable, facilitate, or conceal an infringement.

4. Does generating work in the "style" of a particular artist violate that artist's right of publicity?

Right of publicity law varies considerably from state to state, but generally speaking, it prohibits the commercial use of an individual's name, image, voice, signature, or likeness (and in certain states this extends to broader aspects of "identity" or "persona"). Some states have specific statutes addressing right of publicity and other states rely on common law rights of publicity (and in some states, like California, there may be both). In all states, rights of publicity must be balanced against First Amendment-protected speech, especially where the use is in connection with an expressive work. Right of publicity statutes often have specific carveouts for certain types of expressive works, and courts have developed various tests to balance these competing interests.

The *Anderson* class-action suit raises both statutory and common law right of publicity claims under California law. First, the Complaint alleges that the defendants "used Plaintiffs' names and advertised their AI tool's ability to copy or generate work in the artistic style that the plaintiff's popularized in order to sell Defendant's products and services." Based on this initial pleading, this appears to be a traditional right of publicity claim based on use of the artists' names in advertising the defendants' products and services. The Complaint also focuses on the user's ability to use a text prompt to request that the generated images be "in the style of" a specific artist, and this claim appears to be based, at least in part, on the alleged use of artistic "style" (which is not mentioned expressly in the California statute). The common law claims raised in the suit appear to be making the argument that the plaintiffs' artistic "identities" extend to their body of work and their specific artistic styles and that the plaintiffs' identity is used every time art is generated that reflects their "style." Although California common law has recognized a somewhat broad definition of "identity," (including impersonations of a professional singer's distinctive voice^[2]), we do not yet have case law on whether the California common law right of publicity protects an artist's "style" based solely on the use of the artist's artwork itself.

5. Does the incorporation of a trademark in generated output constitute trademark infringement or give rise to a dilution claim?

The Getty Images Complaint alleges that Stability AI has infringed several of Getty Images' registered and unregistered trademarks in its generation of images and that such use is likely to cause confusion that Getty Images has granted Stability AI the right to use its marks or that Getty Images sponsored, endorsed, or is otherwise associated, affiliated, or connected with Stability AI and its synthetic images. The Complaint also brings a claim for federal trademark dilution under [15 U.S. Code § 1125\(c\)](#) and alleges that Stability AI included a "Getty" watermark on generated images that lack the quality of images that a customer would find on the Getty Images website. It alleges that in some cases, the watermark appeared in connection with low quality, bizarre, and grotesque images. The Complaint argues that these uses cause both dilution by blurring (by lessening the capacity of the plaintiff's mark to identify and distinguish goods or services) and by tarnishment (by harming the reputation of the mark by association with another mark).

6. How do open-source or creative commons license terms apply in connection with use for training AI models and distributing resulting output?

In the *Copilot* case, the plaintiffs claim that the defendants violated open-source license terms by (1) using materials governed by open-source licenses to train Copilot and republishing such materials without providing attribution, copyright notices, and a copy of the license terms; and (2) not making Copilot itself open source. This is a question of first impression.

Takeaways

- These lawsuits will be important to watch for companies creating or using AI tools and technologies as they address questions of first impression regarding a number of issues surrounding the intersection of AI and intellectual property rights.
- Companies involved with generative AI tools (developers or users) should understand how the tools work and are being used and evaluate whether they would benefit from additional risk mitigation strategies.

Endnotes

[1] See, e.g., *Sega Enterprises Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1518 (9th Cir. 1992), as amended (Jan. 6, 1993); *Sony Computer Ent., Inc. v. Connectix Corp.*, 203 F.3d 596, 608 (9th Cir. 2000).

[2] *Midler v. Ford Motor Co.*, 849 F.2d 460, 463 (9th Cir.1988) (holding common law but not statutory cause of action applicable to appropriation of singer's voice by voice impersonator).

© 2023 Perkins Coie LLP

Authors

Explore more in

[Intellectual Property Law](#) [Technology Transactions & Privacy Law](#) [Trademark, Copyright, Internet & Advertising](#) [Artificial Intelligence & Machine Learning](#) [Digital Media & Entertainment, Gaming & Sports Interactive Entertainment](#)

Related insights

Update

[Two Tools for Trump To Dismantle Biden-Era Rules: the Regulatory Freeze and the Congressional Review Act](#)

Update

The FY 2025 National Defense Authorization Act: What's New for Defense Contractors