

AI Insights

If you have any questions regarding the matters discussed in this memorandum, please contact the following attorneys or call your regular Skadden contact.

Stuart D. Levi

Partner / New York
212.735.2750
stuart.levi@skadden.com

Pramode Chiruvolu

Counsel / Palo Alto
650.470.4569
pramode.chiruvolu@skadden.com

This memorandum is provided by Skadden, Arps, Slate, Meagher & Flom LLP and its affiliates for educational and informational purposes only and is not intended and should not be construed as legal advice. This memorandum is considered advertising under applicable state laws.

One Manhattan West
New York, NY 10001
212.735.3000

525 University Avenue
Palo Alto, CA 94301
650.470.4500

Court Largely Grants Defendants' Motion To Dismiss in AI Training Data Case

On Monday, a district court largely granted the defendants' motions to dismiss in *Andersen et al. v. Stability AI et al.*, one of a series of putative class action cases that content creators have recently filed alleging that their content was used without permission to train artificial intelligence (AI) models. While the court (Judge William Orrick of the U.S. District Court for the Northern District of California) gave the plaintiffs leave to amend their complaint, a number of statements by the court may make it challenging for these plaintiffs to articulate a viable claim on the main issues. The decisions will likely be cited by other current and future defendants in similar cases, but the facts of *Andersen*, including that the initial pleadings were so vague, may limit the precedential value of this decision.

Background

The plaintiffs filed suit against the defendants for, among other claims, direct infringement, vicarious infringement, violation of the Digital Millennium Copyright Act (DMCA) and violations of the right of publicity under California law, based on the following alleged activity.

- The plaintiffs alleged that in August 2022, defendant Stability AI (Stability) created and released a "general-purpose" AI software program called Stable Diffusion that was trained on billions of third-party images scraped from the internet without permission, including images created by Sarah Andersen and the other plaintiffs, a group of illustrators and artists.
- The plaintiffs further alleged that Stable Diffusion was a software library providing "image-generating services" to products such as that produced and distributed by Midjourney, another defendant in the case. These products allowed users to enter a text prompt, including prompts that sought images "in the style of" a particular artist, and generate an image.
- The plaintiffs conceded that none of the generated output images was likely a close match to any specific image in the original training data set, but asserted that these generated images were nonetheless infringing derivative works of their original works, given how the defendants used the training data to extract information through a series of mathematical processes to build their model.
- The plaintiffs also alleged that some of their images were scraped from defendant DeviantArt, an online community where some of their work had been posted, and that DeviantArt was itself offering DreamUp, a for-pay image-generation product that relies

Court Largely Grants Defendants' Motion To Dismiss in AI Training Data Case

on Stable Diffusion. The plaintiffs alleged that this activity violated DeviantArt's terms of service against using content for "commercial" purposes without consent, as well as its privacy policy.

Each of the defendants moved to dismiss.

The Court's Decision

Scope of Works at Issue

As an initial matter, the court dismissed all claims concerning works that were not registered for copyright, a prerequisite to filing a federal copyright claim. This limited the case to 16 collections of works by Andersen.

The court next rejected the defendants' assertion that the plaintiffs had failed to identify with any specificity which of the plaintiffs' works had been used as training data. Andersen had relied on searching her name on the "ihavebeentrained.com" website to support the plausibility and reasonableness of her assertion that her works were used in the training data set at issue. The court found this sufficient evidence at the pleading stage, and noted that the defendants can try to establish during fact discovery that the plaintiffs' works were not used.

Direct Infringement

The plaintiffs' theory of direct infringement was that Stability scraped the plaintiffs' images from the internet and used them to train Stable Diffusion. The court denied Stability's motion to dismiss this claim, holding that at this stage of the case, the plaintiffs had properly claimed their works were directly copied for use as training data.

The more complicated and interesting issue was whether the plaintiffs had failed to properly allege direct infringement against DeviantArt. As noted, DeviantArt was an alleged source of the training data as opposed to a party engaged in copying any images. The plaintiffs had instead posited three theories as to how DeviantArt had engaged in direct infringement: (1) DeviantArt had distributed Stable Diffusion, which contains compressed copies of the plaintiffs' training images; (2) DeviantArt had created and distributed the DreamUp image-generation product, which is itself an infringing derivative work; and (3) DeviantArt had generated and distributed AI-generated images, each of which were themselves infringing derivative works.

DeviantArt asserted that Andersen's "compressed copy" theory was implausible given that there were five billion images in the training data set, and that five billion images could not possibly be compressed into an active program. DeviantArt also relied heavily on the plaintiffs' concession that none of the generated output images was likely a close match to any specific image in the original training data set.

The court held that the plaintiffs' complaint was unclear as to how the diffusion process used by Stable Diffusion was tantamount to storing compressed images of training data. As the court explained, the plaintiff needed further clarity and plausible facts to advance a theory that Stable Diffusion "contains mathematical or statistical methods that can be carried out through algorithms or instructions in order to reconstruct the [t]raining [i]mages" to create output images. The court also directed the plaintiffs to better allege why the mere offering of the Stable Diffusion library by DeviantArt constituted direct infringement through the acts of copying or distribution.

In reaching its decision, the court also offered a general observation on the relative merits of a direct infringement claim in this context. According to the court, it is unclear whether entities such as DeviantArt or Midjourney could be liable for direct infringement if the Stable Diffusion "algorithms and instructions" generate images that include only a few elements of the plaintiffs' images. However, the court noted that if the plaintiffs could plead that the defendants' products allow users to create new works by expressly referencing the plaintiffs' works by name, then "the inferences about how and how much of Andersen's protected content remains in Stable Diffusion or is used by the AI end-products might be stronger."

With respect to its allegation that all generated outputs were, by definition, infringing derivatives of the original training data, the defendants asserted that the plaintiffs had failed to allege substantial similarity between their works and the works generated by the defendants' products; indeed, as noted, the plaintiffs went so far as to concede this likely would never be the case. The plaintiff relied instead on a line of cases that substantial similarity is not required where there is evidence of direct copying, as is allegedly the case here. The court rejected this argument, noting that in all of the cases cited by the plaintiff, the allegedly infringing work still contained some elements of, or similarity to, the original work — a fact that the plaintiffs here were unable to show. The court also noted that it was implausible, as the plaintiffs alleged, that every output image relied on copyrighted training data and that all outputs were therefore, by definition, derivative works. The court granted the plaintiffs leave to amend, rejecting the defendants' argument that no plausible allegation could be made given the plaintiffs' concession that the AI-generated images were unlikely to resemble the training data. Instead, the court noted that the complaint had alleged that the generated images might be so similar to the plaintiffs' artistic identities or style to be misconstrued as fakes, and that this might be sufficient to state a claim in an amended complaint clarifying the plaintiffs' theories and factual allegations.

The plaintiffs' claims of direct infringement against Midjourney largely tracked those against DeviantArt, except the plaintiffs also alleged that Midjourney had itself scraped the plaintiffs' images. However, the court found that the complaint was not clear as to

Court Largely Grants Defendants' Motion To Dismiss in AI Training Data Case

whether Midjourney had scraped the images itself or Midjourney was simply alleged to be using Stable Diffusion's scraped images.

The court therefore granted DeviantArt's and Midjourney's motions to dismiss on the direct infringement claim, with leave to amend.

Vicarious Infringement

In order to establish vicarious infringement, a plaintiff must allege that the defendant (1) has the right and ability to supervise the infringing conduct and (2) has a direct financial interest in the infringing activity.

The court dismissed, with leave to amend, the plaintiffs' vicarious infringement claim against DeviantArt and Midjourney because it had not alleged direct infringement claims against them, and against Stability because of the plaintiffs' failure to properly plead direct infringement as discussed above.

DMCA

The plaintiffs alleged that the defendants had deleted the copyright management information (CMI) from their works in violation of the DMCA. The court dismissed this claim, with leave to amend, because the plaintiffs had failed to allege with any specificity in their complaint that their images included CMI, that the defendants had removed the CMI, or that the defendants both knew about the removal and that it would "induce, enable, facilitate, or conceal infringement" as required by the statute.

Right of Publicity

The plaintiffs alleged that their rights of publicity had been violated under California law because the defendants knowingly use their names by allowing users to request art in "the style of" their names. At the hearing on the motion to dismiss, however, the plaintiffs recast their argument to allege that their names had been used to advertise or promote the defendants' products. The court noted that nowhere in the complaint was this "advertising" allegation made. With respect to the actual allegation in the complaint, the court was not persuaded that users would believe that using the plaintiffs' names in a text prompt would generate images that were actually created by the plaintiffs, especially since the plaintiffs had conceded that none of the AI-generated images would likely be a close match for any of their own images.

DeviantArt separately moved to dismiss the right of publicity claim on First Amendment grounds, arguing that the plaintiffs' publicity rights needed to be balanced against the right of free expression given the transformative use of the outputs. Here, the court sided with the plaintiffs that this was a fact-specific issue not appropriate for a motion to dismiss, but noted that the plaintiffs first needed to better articulate their right of publicity claim.

The court dismissed this claim with leave to amend.

Unfair Competition Claim

The court dismissed, with leave to amend, the plaintiffs' unfair competition claims on a variety of grounds, including that the plaintiffs failed to allege plausible facts that a user would think a generated image originated from or was endorsed by the plaintiffs and failed to allege any fraudulent misappropriation by the defendants.

Breach of Contract

The plaintiffs alleged breach of contract by DeviantArt based on a theory that the DeviantArt terms of service prevented the use of content on the site for any commercial purpose. The court dismissed this claim with leave to amend, noting that these terms did not restrict DeviantArt's own use of content. The court noted that the plaintiff would need to allege specific provisions of the terms that were breached and that it was the intended third-party beneficiary of those provisions.

Key Points

While the *Andersen* decision touches on some of the most important issues at the intersection of copyright law and the development and use of AI, the facts of the case did not provide an opportunity for the court to delve into these issues in any meaningful way, at least at the motion to dismiss stage. Below are the main takeaways.

- With respect to almost every claim, the court found that the plaintiffs had simply failed to allege sufficient plausible facts to support their claims.
- The court was also clearly frustrated by the plaintiffs' repeated pivoting in their opposition to the motion to dismiss to legal theories different than the ones actually stated in their complaint.
- Finally, the plaintiffs were boxed in by their concession that none of the images generated by the defendants' services and products were likely to replicate the plaintiffs' copyrighted images.

Other pending AI cases may present different fact patterns. For example, in *Getty Images v. Stability AI*, Getty proffered evidence in its complaint of AI-generated images that included distorted but distinct indicia of the Getty watermark. That case is currently in jurisdictional discovery. Similarly, in *Authors Guild v. OpenAI Inc.*, which was filed in mid-September, the complaint included numerous examples of AI-generated works that were arguably derivatives of the plaintiffs' books (e.g., book summaries or next installments of a series that used the same characters as the original). Decisions, and the ultimate verdicts, in these other cases may provide greater clarity on how courts view the allegedly unauthorized use of copyrighted content as training data.