

AI Intellectual Property Disputes: The Year in Review

December 2, 2025

As we approach the end of 2025, the flurry of developments in artificial intelligence (“AI”) and intellectual property (“IP”) litigation that [we tracked throughout 2024](#) is showing no signs of slowing down. Courts are finally beginning to confront the substantive merits of plaintiffs’ infringement claims and defendants’ fair use defenses. The decisions that have been released so far this year, with many more expected to come, are shaping both the generative AI industry, as well as businesses of all stripes that are increasingly embedding generative AI technology into their processes. As new and improved generative AI features emerge seemingly every week—from improved generative video models to even more powerful deep reasoning models—the facts of many of these cases seem almost quaint, yet the impact of plaintiffs’ legal theories could still be substantial.

Debevoise is tracking more than 50 lawsuits between IP holders (from individual artists to large corporate publishing houses and national newspapers) and AI developers that are pending in federal courts around the country. Regardless of recent rulings on summary judgment motions, settlements and precedent-setting licensing negotiations, IP rights holders continue to file new claims, amend existing pleadings to include new facts brought to light in discovery, and cure deficiencies that led to previous motions to dismiss.

This article reviews the most significant developments in AI litigation so far this year, including:

- The first major fair use rulings involving AI technology arrived, with courts finding generative AI training generally transformative and protected, but key questions remain unresolved, including the legality of training on pirated works, whether outputs actually threaten the market for original works, and the extent to which AI-powered tools directly compete with rights holders’ commercial offerings.
- A coalition of corporate media plaintiffs filed the largest coordinated action to date, and other high-profile plaintiffs like Disney and Universal have followed suit,

signaling that litigation strategy is evolving to include a wider class of plaintiffs and a broader portfolio of IP assets at issue.

- Recent rulings signal a new openness to output-based infringement claims.
- A possible court split is emerging regarding the pleading requirements for a viable Digital Millennium Copyright Act (the “DMCA”) claim in the context of generative AI.
- Courts across the country continue to grant plaintiffs wide latitude to amend their pleadings—often using new facts from discovery to revive and strengthen previously dismissed claims. However, a recent summary judgment ruling shows that these procedural successes do not guarantee ultimate victory.
- In discovery, while plaintiffs have gained access to internal datasets and documents, AI developers have largely succeeded in shielding those materials from public view through protective orders and privilege claims.
- Many state-law theories continue to be consistently rejected by courts, narrowing plaintiffs’ common practice of “throwing the kitchen sink.”
- Class certification has quickly become a high-stakes battleground: the Bartz plaintiffs obtained class certification and quickly thereafter a \$1.5 billion settlement agreement, underscoring the immense pressure certification can place on AI developers seeking to break cases into individual actions to gain settlement leverage.
- Trademark claims continue as a parallel front in AI litigation, with plaintiffs using Lanham Act theories to challenge AI-generated misattribution and branding, and early decisions indicating that allegations of consumer confusion and reputational harm can proceed past the pleading stage.

The rest of the year promises even more action for IP holders and AI developers alike, as courts continue to grapple with the most contested substantive and procedural issues at the heart of these disputes. We discuss these updates below, as well as our thoughts on what the industry can expect in the months ahead as these hard-fought cases continue to wind through the courts.

First Fair Use Rulings Arrive

As [we predicted](#) at the end of last year, 2025 brought the [first major fair use rulings](#) involving generative AI—and they signal important early trends for litigation in this space. In two closely watched decisions from the U.S. District Court for the Northern District of California, courts directly addressed whether training generative AI models constitutes fair use, issuing narrow rulings holding that training can be fair use of copyrighted works. Earlier in the year, the U.S. District Court for the District of Delaware [also addressed fair use](#) in the context of non-generative AI, coming out the other way in a case between direct competitors.

Generative AI Training: Two Early Wins for Defendants, but Questions Remain

The first major fair use rulings in generative AI training cases arrived within days of each other in June 2025, offering early wins for defendants while revealing fundamental disagreements about how courts should analyze AI training. In *Bartz v. Anthropic*, using copyrighted works to train large language models (“LLMs”) was held to be fair use, but was distinguished from storing pirated datasets. In *Kadrey v. Meta*, the court similarly found training to be fair use even when using pirated sources, though this holding was pointedly narrow and fact-specific.

In *Bartz v. Anthropic*, on June 23, 2025, Judge Alsup held that AI developer Anthropic’s use of copyrighted works to train its LLM was fair use.¹ Judge Alsup reasoned that training AI using copyrighted works was highly transformative, analogizing the process to human learning and memory, and emphasizing the difference in purpose and character between Anthropic’s AI system and plaintiffs’ original works.

Importantly, Judge Alsup distinguished Anthropic’s **training uses** from its **storage** of pirated works, concluding that merely maintaining a database of pirated materials was not transformative and not protected by fair use. Judge Alsup separately found Anthropic’s scanning and digital storage of legally purchased books was a transformative format change, noting this eased storage and enabled searchability without infringing copyright holders’ exclusive rights. But he reserved the question of liability for maintaining a central digital library of those scanned copies, citing unresolved factual issues that required a trial.

Just two days later, in *Kadrey v. Meta Platforms*, Judge Chhabria similarly concluded that Meta’s use of copyrighted works—including those obtained from pirated sources such as shadow libraries—for training its LLM was fair use, but his holding was pointedly

¹ *Bartz v. Anthropic PBC*, No. 24 CV 05417-WHA, Order on Fair Use, ECF No. 231 (N.D. Cal. June 23, 2025).

narrow.² Judge Chhabria underscored that the outcome hinged heavily on the plaintiffs' failure to present empirical evidence that the outputs from Meta's LLM diluted or substituted the market for their original works, rather than on a broadly applicable legal principle. Notably, Judge Chhabria suggested that when considering whether AI training on copyrighted materials is illegal, "in most cases the answer will likely be yes."³

The courts split on critical analytical approaches: *Bartz* analogized AI training to human learning and dismissed concerns about indirect market competition, while *Kadrey* rejected this analogy and emphasized that generative AI's unique potential to "flood the market with competing works" makes market dilution "highly relevant" to the fair use analysis. The courts also diverged on whether data acquisition and training should be analyzed together or separately, with implications for how defendants' data collection methods will be evaluated. Significantly, *Bartz* found that storing pirated works was not fair use as a matter of law, while the *Kadrey* court clarified that its finding of fair use hinged on plaintiffs' failure to present empirical evidence of market harm, providing a detailed roadmap for future evidence development that could result in a different outcome. The two judges' areas of disagreement, even in opinions that reached the same result, signal that deeper fissures may emerge as litigants learn from these early rulings.

A Different Approach to Fair Use: Direct Competitors Training Non-Generative AI

Earlier in the year, Judge Bibas (sitting by designation in the District of Delaware) [addressed fair use](#) in the context of non-generative AI in the *Thomson Reuters v. ROSS Intelligence* case. Thomson Reuters' legal search engine, Westlaw, provides copies of cases and statutes along with original headnotes to support legal research. ROSS, a legal research startup, sought to build its own AI-driven search platform for lawyers and trained its tool using Westlaw's copyrighted headnotes.

Judge Bibas held that ROSS's use of Westlaw's copyrighted headnotes to train its AI-powered legal research tool did not qualify as fair use.⁴ Relying heavily on the Supreme Court's 2023 decision in *Andy Warhol Foundation v. Goldsmith*, Judge Bibas found that ROSS's use was neither sufficiently transformative nor distinct in purpose, and he emphasized that the commercial nature of ROSS's product, combined with its potential to compete directly with Westlaw, also weighed against fair use. Despite recognizing some factors in ROSS's favor—such as the non-creative nature of Westlaw's

² *Kadrey v. Meta Platforms, Inc.*, No. 23 CV 03417-VC, Order Denying the Plaintiffs' Motion for Partial Summary Judgment and Granting Meta's Cross-Motion for Partial Summary Judgment, ECF No. 598 (N.D. Cal. June 25, 2025).

³ *Kadrey v. Meta Platforms, Inc.*, 788 F. Supp. 3d 1026, 1034 (N.D. Cal. 2025).

⁴ *Thomson Reuters Enterprise Centre GmbH & West Publishing Corp. v. ROSS Intelligence Inc.*, No. 120 CV 613-SB, Memorandum Opinion, ECF No. 770 (D. Del Feb. 11, 2025).

headnotes—Judge Bibas ultimately concluded the balance favored protecting the IP holder’s rights. (Notably, Judge Alsup distinguished this holding in the *Bartz* opinion, emphasizing that the lack of direct market competition between Anthropic’s AI system and plaintiffs’ original works presented a very different situation than the direct commercial competition between ROSS and Thomson Reuters.)

Judge Bibas’ decision that the fair use question was ripe for a ruling at summary judgment was somewhat surprising given the court’s earlier signals that fair use would likely be a question for a jury. Combined with the *Bartz* and *Kadrey* cases, these holdings signal that courts may be willing to evaluate the merits of transformative use defenses if the record is sufficiently clear.

Unresolved Questions and Future Rulings Ahead

These decisions underscore that the fair use analysis for AI will remain highly fact-intensive and context-specific. Although *Bartz*, *Kadrey*, and *Thomson Reuters* are significant early rulings, they each illustrate a different approach to the fair use analysis, and none fully resolves the broader legal questions surrounding fair use in AI. Critically, the analyses could shift if plaintiffs present stronger evidence of market harm stemming from AI-generated outputs or direct competition with original works. Judge Chhabria explicitly provided a roadmap for plaintiffs to develop evidence that could overcome the fourth fair use factor, which other plaintiffs are already beginning to leverage in ensuing actions.

The contrasting approaches taken by these district courts set the stage for differing conclusions reached by near-certain appellate review, to include an interlocutory appeal of Judge Bibas’ fair use ruling already pending before the Third Circuit.

Additional fair use rulings are on the horizon. Based on current scheduling, upcoming fair use rulings are likely to include decisions in *UMG Recordings v. Suno* (District of Massachusetts), *In re Google Generative AI Litigation* (Northern District of California), and *Concord Music v. Anthropic* (Northern District of California). The *UMG v. Suno* case concerns whether AI music-generation tools infringed copyrighted sound recordings, whereas *In re Google* involves multiple copyright owners challenging Google’s use of protected text and images to train Gemini and related AI systems. Finally, *Concord Music v. Anthropic* focuses on whether Anthropic’s model infringes musical lyrics. These rulings are expected in mid- to late-2026, with *UMG v. Suno* likely first, followed by *Google* and *Anthropic* later in the year.

Corporate Plaintiffs Join Forces in Filing New Cases

At the end of 2024, [we wrote about](#) how IP holders spent the year filing “kitchen sink” complaints against AI developers, reaching beyond copyright law for novel theories of liability. We also noted that large corporate plaintiffs with substantial IP portfolios were increasingly entering the fray. In 2025, IP holders doubled down on these tactics, and we do not expect these trends to slow down.

In February, a coalition of 14 major media companies, including Condé Nast, The Atlantic, Forbes and The Guardian, sued AI developer Cohere, Inc. for direct and secondary copyright infringement, trademark infringement, and false designation of origin.⁵ The media companies’ complaint alleges that Cohere unlawfully scraped their copyrighted articles, used them to train its AI models, and generated outputs that reproduce or closely summarize paywalled news content. The media companies also assert that Cohere’s AI “hallucinations” falsely attribute fabricated content to legitimate news organizations, damaging their brands.

While this case mirrors the allegations in the headline-grabbing *New York Times* case, it is nonetheless significant as it represents the largest collective action by media organizations against an AI developer to date and reinforces the trend of well-resourced corporate plaintiffs pursuing multi-pronged legal challenges to protect their IP assets.

The trend of corporate plaintiffs collaborating against AI developers gained further momentum in June, when major studios Disney and Universal City Studios filed a highly visible lawsuit against Midjourney in the U.S. District Court for the Central District of California.⁶ Unlike recent lawsuits focusing primarily on the inputs used to train generative AI models, this suit prominently features alleged infringements in Midjourney’s outputs, highlighting the studios’ ability to leverage their extensive portfolios of iconic visual IP. The plaintiffs’ complaint meticulously documents Midjourney-generated images that reproduce famous copyrighted characters—including Darth Vader, Elsa, Spider-Man, Shrek and others—at times closely mimicking specific movie stills.

Disney and Universal’s decision to team up underscores the strategic advantage gained from pooling resources to confront well-funded AI defendants and highlights their distinctive position as plaintiffs capable of substantiating claims based on recognizable visual assets—something individual artists or smaller copyright holders may struggle to achieve. Moreover, the studios explicitly allege that Midjourney failed to implement

⁵ *Advance Local Media LLC v. Cohere Inc.*, No. 25 CV 01305, Complaint, ECF No. 1 (S.D.N.Y. Feb. 13, 2025).

⁶ *Disney Enterprises Inc. v. Midjourney Inc.*, No. 25 CV 05275, Complaint, ECF No. 1 (C.D. Cal. June 11, 2025).

adequate guardrails despite clear cease-and-desist demands, potentially bolstering their claim of willful infringement and setting the stage for heightened liability.

The cases against Cohere and Midjourney both signal important developments about the direction of AI litigation more broadly.

First, they reflect that corporations with extensive IP portfolios are increasingly viewing generative AI as a potential threat to the value of those assets. Rather than waiting on the sidelines to see how other cases unfold, these corporations appear prepared to pursue litigation themselves to preserve the value of their proprietary content.

Second, the decision by multiple corporate plaintiffs to join forces in single actions demonstrates a strategic recognition that pooling resources may help manage the high cost and complexity of discovery in AI litigation—especially against very well-resourced defendants. Cases involving large-scale model training often require the production and review of massive datasets, along with internal communications regarding model development and data selection. Discovery may also implicate multiple model iterations, each with its own set of training and fine-tuning data. These requests frequently raise trade secret concerns and require the negotiation of detailed security protocols, all of which can drive up litigation costs. For corporations seeking to establish that their works were used to train AI models, joint litigation may reduce the burden on any individual plaintiff by offering a more efficient path forward.

Third, consolidated complaints of this kind may strengthen the credibility of the plaintiffs' allegations. When numerous companies operating in the same industry allege similar harm from the same conduct, it can sharpen arguments around market impact and reduce uncertainty around standing or individualized injuries. These dynamics may also serve as a roadmap for IP holders in other industries—such as software developers, video game publishers, film studios, streaming platforms, and educational content providers—that rely heavily on proprietary works and are closely monitoring how courts respond to claims of unauthorized scraping and reproduction of protected content.

As courts begin to test the limits of AI developers' liability, companies with substantial IP interests are unlikely to remain passive. If early collective suits gain traction, similar joint actions may emerge across industries, further accelerating the pace and scope of AI litigation.

On the other hand, not all corporate plaintiffs are opting to continue litigating. For example, Warner Music Group ("WMG") resolved two separate actions it had been pursuing against AI-music startups Udio and Suno in the past several days. WMG reached a settlement with Udio on November 19, 2025 and with Suno on November 25.

Immediately following these resolutions, WMG publicly announced new partnerships with both companies to develop and commercialize AI-generated music leveraging WMG's catalog and artist relationships. These coordinated settlements and subsequent collaborations suggest a strategic shift: rather than seeking solely to enjoin AI developers, WMG appears to be positioning itself to participate in and shape emerging AI-driven markets.

These developments may foreshadow the direction in which disputes involving large corporate IP holders are heading. While individual creators and smaller rights holders may continue to pursue their claims in court, major industry stakeholders increasingly seem to be evaluating not only the risks posed by generative AI, but also the potential benefits of integrating AI technologies into existing business models. For these entities, leveraging their substantial IP portfolios through licensing, partnerships or co-development arrangements may ultimately prove more advantageous than protracted litigation aimed at halting or limiting AI innovation.

Renewed Interest in Output-Based Infringement Claims

In a recent shift from a trend we wrote about [last year](#), IP holders and courts are beginning to give renewed attention to copyright claims based on AI-generated outputs. Until this fall, IP holders had struggled to advance such theories, and courts had consistently dismissed them for failing to allege substantial similarity or identify specific infringing examples.

This pattern changed in October, when Judge Sidney Stein in the *In re OpenAI Copyright Litigation* multidistrict proceeding sustained IP holders' output-based infringement claims at the pleading stage. Denying OpenAI's motion to dismiss, Judge Stein held that the class plaintiffs had plausibly alleged both actual copying and substantial similarity between ChatGPT-generated text and their copyrighted works.⁷ The opinion reproduced detailed ChatGPT-generated summaries and derivative outlines, including a multi-page synopsis of *A Game of Thrones*, to illustrate how a reasonable jury could find overlap in protectable elements such as plot, characters and tone. Applying the "more discerning observer" test, the court concluded that at least some AI outputs plausibly incorporated expressive elements of the underlying works.

In November, Judge McMahon joined Judge Stein in sustaining output-based infringement claims at the pleading stage, denying Cohere's partial motion to dismiss

⁷ *In re OpenAI Copyright Litigation*, No. 23 CV 08292-SHS-OTW, Order on Motion to Dismiss, ECF No. 716 (S.D.N.Y. Oct. 27, 2025).

Advance Local Media’s “substitutive summaries” theory of infringement.⁸ Cohere argued that summaries generated by its AI model do not copy any protectable expression, but Judge McMahon rejected this argument, relying on evidence presented by the publishers showing verbatim regurgitation and close paraphrasing of their news articles. In one example cited by the court, Cohere’s AI allegedly produced an output which “directly copied eight of ten paragraphs from a *New Yorker* article with very minor alterations.”⁹

For AI developers, these rulings underscore the growing risk that the outputs themselves, not just the training process, could form the basis for costly discovery and fact-intensive litigation. For rights holders, they represent judicial recognition that output-based copying allegations may be legally sufficient, providing a roadmap for future pleadings and a potential foothold for expanding the scope of viable infringement theories in AI litigation.

Amended Pleadings Show Early Dismissals Are Not Deterring Plaintiffs

Plaintiffs are continuing to amend their pleadings to cure deficiencies identified in initial dismissal orders and account for newly discovered facts. At the end of 2024, [we noted](#) that IP holders’ DMCA claims against AI developers were facing significant hurdles as courts across the country have wrestled with how the decades-old DMCA should be applied in cases involving generative AI.

In 2025, however, courts have begun reaching different conclusions on DMCA claims, with several plaintiffs having learned the lessons of those earlier cases and filing complaints that are surviving motions to dismiss. These recent decisions reveal a pattern that offers renewed hope for plaintiffs, highlighting the importance of precise pleading strategies: DMCA claims may survive dismissal when plaintiffs provide specific, detailed allegations about intentional CMI removal during training, but will likely fail when they rely on more generalized allegations or focus solely on the distribution of fragmentary outputs.

The Legal Standard for DMCA Claims

As background, when copyright owners include copyright management information (“CMI”) alongside their copyrighted work, the DMCA prohibits altering or removing that CMI under § 1202(b)(1). It also prohibits distributing copies of the copyrighted

⁸ *Advance Local Media LLC v. Cohere Inc.*, No. 25 CV 01305, Decision and Order Denying Defendant’s Partial Motion to Dismiss, ECF No. 59 (S.D.N.Y. Nov. 13, 2025).

⁹ *Advance Local Media LLC v. Cohere Inc.*, No. 25 CV 01305, Decision and Order Denying Defendant’s Partial Motion to Dismiss, ECF No. 59 at 8 (S.D.N.Y. Nov. 13, 2025).

work where the CMI has been altered or removed under § 1202(b)(3). To adequately state a DMCA claim based on CMI removal, a copyright holder must establish that the defendant knew, or had reasonable grounds to know, that their actions would “induce, enable, facilitate, or conceal” copyright infringement.

Recent rulings have involved claims based on two different alleged actions by defendants: (1) the alleged *removal* of CMI during model training and (2) the alleged *distribution* of CMI-less outputs. Courts have generally been more receptive to claims concerning CMI removal—especially where plaintiffs identify specific tools or processes used during AI training—while distribution-based claims have faced more scrutiny, particularly when based on fragmentary or non-identical AI outputs.

The legal standards are not entirely settled, however. Some (but not all) circuit courts impose a “double scienter” requirement, meaning that plaintiffs must establish that defendants *first*, knew they removed or altered CMI; and *second*, knew that their CMI removal or alteration would likely result in copyright infringement.¹⁰ In addition to these statutory requirements, plaintiffs in every circuit must also meet the constitutional Article III standing requirement of establishing that they suffered an injury in fact.

Plaintiffs’ More Detailed Pleadings Begin to Reverse the Trend of DMCA Claim Dismissals

In late 2024, in *Raw Story Media v. OpenAI*, the U.S. District Court for the Southern District of New York [dismissed a group of journalists’ DMCA claim against an AI developer for inadequately alleging harm](#), holding that the IP holders did not have standing because they had failed to show that defendants’ removal of CMI likely resulted in actual copyright infringement.¹¹ According to the court, the key flaw in journalists’ pleading was the absence of any evidence showing that the developer’s generative AI model had in fact disseminated (or was even likely to disseminate) plaintiffs’ copyrighted work without their CMI present. The court held that merely alleging that an AI developer removed CMI from training data was insufficient to sustain a claim.

Despite that precedent, in early 2025, IP holders successfully defeated motions to dismiss their DMCA claims in both the Northern District of California and the Southern District of New York. These more successful plaintiffs provided specific factual allegations that the plaintiffs in *Raw Story* lacked.

¹⁰ *Mango v. BuzzFeed*, 970 F.3d 167, 171 (2nd Cir. 2020); *Stevens v. Corelogic, Inc.*, 899 F.3d 666, 675 (9th Cir. 2018).

¹¹ *Raw Story Media, Inc. v. OpenAI Inc.*, No. 24 CV 01514-CM, Order on Motion to Dismiss, ECF No. 117 (S.D.N.Y. Nov. 7, 2024).

In the *Kadrey* case (discussed above on the issue of fair use), a class of authors successfully moved to file a third amended complaint, reviving their previously dismissed DMCA claim by using specific facts about Meta’s data processing methods learned through discovery into their other claims.¹² This case illustrates a broader trend of plaintiffs using discovery to strengthen pleadings, including by reviving claims that were previously dismissed.

The *Kadrey* authors originally alleged that Meta removed CMI from the outputs its LLMs produced. Like in *Raw Story Media*, however, the court dismissed this claim because the authors failed to demonstrate that Meta’s alleged removal of CMI resulted in distribution of the authors’ works without attribution.¹³ In response, the *Kadrey* authors used evidence they obtained during the discovery process (which proceeded while the partial dismissals were being litigated in parallel) as a hook for overcoming dismissal of their DMCA claim, shifting their argument to focus on the removal of CMI from their books during Meta’s model training process. In their third amended complaint, the authors used this new evidence to allege more concretely that Meta had deliberately stripped metadata from their books before using them as inputs, obscuring the origins of any outputs Meta’s model might ultimately generate.

In denying Meta’s motion to dismiss the re-plead DMCA claim, the *Kadrey* court ruled that plaintiffs’ new allegations established a sufficient injury because the removal of CMI was an adequately concrete injury for the purposes of standing. The *Kadrey* court cited to *Raw Story Media*, but declined to follow its lead and appeared to endorse a more permissive interpretation of concrete harm under Article III. In the *Kadrey* court’s opinion, “whether the DMCA seeks to prevent the exact injuries the plaintiffs allege is a separate question from whether the plaintiffs have been concretely injured by Meta’s CMI removal. Article III standing requires that the injury to the plaintiff—not their cause of action or the statute under which it arises—have a close relationship to a traditional harm.”¹⁴ Ultimately, however, this victory for the *Kadrey* plaintiffs proved merely preliminary, as the *Kadrey* court’s holding on fair use (discussed above) meant Meta was entitled to summary judgment on the DMCA claim, as well as the copyright claims.

Plaintiffs also defeated a motion to dismiss a DMCA claim in *The Intercept Media v. OpenAI*, where the U.S. District Court for the Southern District of New York held that The Intercept had sufficiently alleged that OpenAI intentionally removed CMI during

¹² *Kadrey v. Meta Platforms, Inc.*, No. 23 CV 03417-VC, Third Amended Complaint, ECF No. 407 (N.D. Cal. Jan. 21, 2025).

¹³ *Kadrey v. Meta Platforms, Inc.*, No. 23 CV 03417-VC, Order on Motion to Dismiss, ECF No. 56 (N.D. Cal. Nov. 20, 2023).

¹⁴ *Kadrey v. Meta Platforms, Inc.*, No. 23 CV 03417-VC, Order on Motion to Dismiss, ECF No. 56, p. 2 (N.D. Cal. Nov. 20, 2023).

the model training process and knew that this removal could facilitate future infringement, meeting the DMCA’s “double scienter” requirement and overcoming the threshold for pleading concrete harm.¹⁵ Like the authors in *Kadrey* and unlike the plaintiffs in *Raw Story Media*, The Intercept provided the detailed allegations about defendants’ data processing methods.

Both the *Kadrey* and *Intercept* opinions were narrow, and far from complete victories for plaintiffs. The courts dismissed distribution-based DMCA claims, and rejected arguments that the developers removed CMI to facilitate third-party infringement. But they also illustrate that at least some DMCA claims can survive dismissal for lack of standing when plaintiffs provide sufficiently detailed allegations about CMI removal during training, suggesting that such claims are not entirely foreclosed for IP holders who continue to “throw the kitchen sink.”

A subsequent decision in the consolidated *New York Times v. Microsoft* litigation illustrates how pleading specificity determines success in DMCA claims against AI developers. While the Southern District of New York held that all plaintiffs had standing to pursue both training and distribution-based DMCA claims, the court’s analysis of the substantive allegations revealed the critical importance of detailed factual pleading. The court dismissed *The New York Times*’ DMCA claims in their entirety, finding that the complaint failed to satisfy the double scienter requirement because it did not allege with sufficient specificity how CMI was intentionally removed during training.

By contrast, the same court allowed DMCA claims brought by co-plaintiffs *Daily News* and *The Center of Investigative Reporting* to proceed because their pleadings met the double scienter requirements. These plaintiffs provided detailed allegations about OpenAI’s use of web-scraping tools allegedly designed to strip CMI during data ingestion. The court found those concrete allegations sufficient to support plausible inferences that defendants knew they were removing CMI and knew such removal could facilitate downstream infringement.¹⁶

These decisions reveal a growing consensus at the District Court level: courts are receptive to training-based DMCA claims when supported by specific allegations about intentional CMI removal, but distribution-based claims face significant hurdles. The key distinction lies in pleading specificity—concrete facts about defendants’ data processing methods during model training can satisfy both standing and double scienter

¹⁵ *The Intercept Media, Inc. v. OpenAI Inc.*, No. 24 CV 01515-JSR, Opinion and Order, ECF No. 127 (S.D.N.Y. Feb. 20, 2025).

¹⁶ *The New York Times Company v. Microsoft Corporation*, No. 23 CV 11195-SHS, Opinion re: Motion to Dismiss the Amended Complaint, ECF No. 514 (S.D.N.Y. May 4, 2025).

requirements, while generalized allegations of CMI removal or claims based on fragmentary outputs typically cannot.

This emphasis on factual specificity creates both challenges and opportunities for DMCA plaintiffs in the AI context. When outputs are fragmented, rephrased or incomplete, distribution-based theories are difficult to sustain. Success instead depends on plaintiffs' ability to obtain detailed discovery about defendants' data processing methods and to craft precise allegations about CMI removal during training. This dynamic favors well-resourced plaintiffs capable of conducting extensive fact development and underscores that surviving motions to dismiss requires meticulous factual pleading rather than novel legal theories.

Permissive Leave to Amend Likely to Continue

The authors' amended complaint in *Kadrey* is emblematic of a broader trend in AI litigation: courts are increasingly granting IP holders multiple opportunities to introduce new claims and revive previously dismissed ones, even after motions to dismiss succeeded against an initial or even an amended complaint.¹⁷ This drawn out process, often while discovery is proceeding in the background, has allowed IP holders to adapt their claims based on defendants' internal documents and courts' early rulings on pleading deficiencies, ensuring that new and previously dismissed theories remain in play. In *Concord Music Group v. Anthropic*, for example, the court initially dismissed the publishers' secondary-liability allegations but later allowed them to proceed after plaintiffs amended their complaint to add detailed factual allegations about Anthropic's dataset curation and internal guardrails.¹⁸

With the benefit of parallel fact discovery, these early partial dismissals have served as a roadmap for plaintiffs to shape discovery requests to help them clarify and reassert their theories. For defendants, early wins on motions to dismiss have offered only limited protection. As a result, AI developers are frequently unable to escape litigation prior to summary judgment, unless they are opposing amendments that are unsupported by new factual detail.

Of course, even significant momentum gained through eventual pleading-stage victories does not guarantee plaintiffs ultimate success at the summary judgment stage. The *Kadrey* court ultimately granted summary judgment to Meta when considering the merits of plaintiffs' DMCA claim. Judge Chhabria held that because Meta's copying

¹⁷ See, e.g., *Dow Jones & Company, Inc. v. Perplexity AI, Inc.*, No. 24 CV 07984-KPF, Second Amended Complaint, ECF No. 46 (S.D.N.Y. Jan. 28, 2025); *Andersen v. Stability AI Ltd.*, No. 3:23-cv-00201, Second Amended Complaint, ECF No. 238 (N.D. Cal. Oct. 31, 2024).

¹⁸ *Concord Music Group, Inc. v. Anthropic PBC*, No. 24 CV 03811-EKL, Order Denying Motion to Dismiss First Amended Complaint, ECF No. 461 (N.D. Cal. Oct. 6, 2025).

constituted fair use, it could not satisfy the DMCA's requirement that the defendant knew the removal of CMI would induce, enable, facilitate or conceal infringement.¹⁹ Judge Chhabria emphasized that liability under § 1202(b)(1) cannot attach to an act performed in furtherance of fair use, reinforcing the critical interplay between fair use findings and the viability of DMCA claims in generative AI contexts.

Discovery Battles Shape AI Litigation Strategy

As generative AI cases move beyond the pleadings stage, discovery is becoming the central battleground—one that poses distinct challenges for both sides. IP holder plaintiffs have notched early wins in 2025, compelling AI developers to disclose certain details about how their models were trained. In *Tremblay v. OpenAI*, a class of authors [won a discovery dispute](#) requiring OpenAI to disclose subsets of its training datasets and internal communications regarding model development.²⁰ Meanwhile, in *Kadrey v. Meta*, authors obtained unredacted statements that they claim show Meta knowingly collected training data from LibGen, an online shadow library known to house pirated books and other copyrighted materials. The authors in *Kadrey* also succeeded in compelling responses to requests for admission concerning Meta's use of their books as training data for its LLMs.²¹ And in November, Magistrate Judge Wang issued an opinion in the consolidated *In re OpenAI ChatGPT Litigation* holding that OpenAI waived its attorney-client privilege as to communications related to its deletion of training datasets, and ordering both the production of such communications, as well as depositions of the attorneys involved.²²

These rulings are significant because they provide plaintiffs with access to information that can strengthen or sharpen their core claims, as discussed above. Obtaining access to pre-training datasets and AI developers' internal discussions has allowed some IP holders to trace whether their copyrighted works were used without permission and to assess what AI developers knew about potential infringement risks. Defendants, in turn, have resisted these efforts by arguing that broad disclosure imposes an undue burden and implicates trade secret protections. Courts have shown some receptivity to defendants' arguments, particularly where plaintiffs seek expansive datasets or technical records that may include substantial amounts of irrelevant or proprietary information. Courts have also allowed AI developers like Meta to shield some internal discussions

¹⁹ *Kadrey v. Meta Platforms, Inc.*, No. 23 CV 03417-VC, Order Granting Meta's Motion for Partial Summary Judgment as to the Plaintiffs' DMCA Claim, ECF No. 601 (N.D. Cal. June 27, 2025).

²⁰ *See Tremblay v. OpenAI, Inc.*, No. 23 CV 03223-AMO, Joint Discovery Letter Brief, ECF No. 254 (N.D. Cal. Jan. 17, 2025).

²¹ *Kadrey v. Meta Platforms, Inc.*, No. 23 CV 03417-VC, Public Version of Discovery Order, ECF No. 374 (N.D. Cal. Jan. 17, 2025).

²² *In re OpenAI, Copyright Litigation*, 2025 WL 3270133 (S.D.N.Y. Nov. 24, 2025).

based on attorney-client privilege, especially when communications relate to the risks of training AI models with web-scraped content.

A separate but related set of disputes has emerged regarding whether discovery materials should be made public. Even when courts compel disclosure, they are frequently allowing defendants to file documents under seal or to limit access to secure environments, recognizing that broad disclosure of AI developers' confidential information could expose proprietary algorithms to competitors.

Courts have adopted a range of approaches to balance IP holders' discovery needs with AI developers' confidentiality concerns—from requiring training data and source code to be made available in siloed data rooms with strict security controls limiting public access in *Bartz v. Anthropic* to allowing AI developers to seal certain proprietary details in their required disclosures in *Concord v. Anthropic*.²³ Similar rulings have been issued in *UMG Recordings v. Uncharted Labs* and *Authors Guild v. OpenAI*, where protective orders limited public access to model source code and internal communications containing proprietary business information.²⁴

The common thread across these cases is that plaintiffs have largely succeeded in gaining access to internal data, but defendants have largely succeeded in shielding that data from public view. This distinction matters. While the compelled production of evidence may increase litigation risk in the short term, protective orders and sealing rulings may blunt the downstream impact by limiting how that information can be used or weaponized outside the immediate litigation. For AI developers, discovery disputes are becoming not just about whether IP holders will gain access to the information they request, but also about how far that access will extend beyond the four corners of the case.

Courts appear to be navigating toward a middle ground: granting plaintiffs the information needed to pursue their claims while honoring defendants' claims of confidentiality. For IP holders, discovery victories may provide the factual support needed to advance toward summary judgment, class certification, or a challenge to fair use. For AI developers, success in controlling the use and visibility of disclosed materials may mitigate broader reputational or commercial risks.

²³ *Bartz v. Anthropic PBC*, No. 24 CV 05417-WHA, Joint Stipulation and [Proposed] Order Regarding Training Data and Source Code Inspection Protocol, ECF No. 24 CV 05417-WHA, ECF 83 (N.D. Cal. Feb. 11, 2025); *Concord Music Group, Inc. v. Anthropic PBC*, No. 24 CV 03811-EKL, Protocol for Producing Documents and Electronically Stored Information as Modified, ECF No. 298 (N.D. Cal. Feb. 13, 2025).

²⁴ *UMG Recordings, Inc. v. Uncharted Labs, Inc.*, No. 24 CV 04777-AKH, Stipulated Source Code Protocol, ECF No. 68 (S.D.N.Y. Jan. 8, 2025); *Authors Guild v. OpenAI Inc.*, No. 23 CV 8292-SHS-OTW, Order Granting Letter Motions to Seal, ECF No. 305 (S.D.N.Y. Jan. 10, 2025).

AI Developers Notch Victories Against Kitchen-Sink Complaints, but Challenges Persist

State-Law Claims Prove Difficult to Plead

As IP holders continue to file new claims and invest in discovery, the challenges [we wrote about](#) in 2024 still persist. AI developers continue to secure legal victories at the motion to dismiss stage, narrowing the scope of litigation—at least on paper. Their continued success, particularly in knocking out state-law theories, signals that the “kitchen sink” style complaints brought by IP holders may not be effective at establishing new ways of thinking about state laws. However, the breadth of different claims IP holders have asserted has successfully prevented most cases from being dismissed in their entirety at the pleading stage.

In *Kadrey v. Meta*, the plaintiffs alleged that Meta unlawfully accessed their books in violation of the California Comprehensive Computer Data Access and Fraud Act the CDAFA, a California state law that provides for a private right of action and that criminalizes unauthorized access to data without permission and with the intent to cause harm. The court was not receptive to the *Kadrey* plaintiffs’ claim and ultimately dismissed it on preemption grounds, concluding that there was no way to distinguish the alleged violation from conduct already governed by the Copyright Act. Because the claim was premised entirely on Meta’s access to plaintiffs’ digitized, copyrighted books, the court found it could not proceed as a standalone theory of liability.²⁵

In *Millette v. OpenAI*, OpenAI successfully moved to dismiss all of the YouTube video creators’ state-law claims, including unfair competition and unjust enrichment, narrowing their case down to a single federal copyright claim.²⁶ In its motion to dismiss, OpenAI cited previous rulings rejecting state-law theories in *Tremblay v. OpenAI*, *Doe v. GitHub*, *Andersen v. Stability AI*, and *Kadrey v. Meta*.²⁷ The plaintiffs did not oppose the dismissal, likely recognizing that the claims were unlikely to survive in light of the prevailing case law. As a result of the state-law theories being dismissed, one plaintiff who did not hold a registered copyright for her video content was foreclosed from obtaining relief.

²⁵ *Kadrey v. Meta Platforms, Inc.*, No. 23 CV 03417-VC, Order Granting in Part and Denying in Part Motion to Dismiss, ECF No. 471 (N.D. Cal. Mar. 7, 2025).

²⁶ *Millette v. OpenAI, Inc.*, No. 24 CV 04710-EJD, Plaintiffs’ Statement of Non-Opposition to Defendants’ Motion to Dismiss, ECF No. 60 (N.D. Cal. Feb. 7, 2025).

²⁷ *Millette v. OpenAI, Inc.*, No. 24 CV 04710-EJD, Defendants’ Motion to Dismiss, ECF No. 55 (N.D. Cal. Dec. 12, 2024).

AI developers sued more recently are following the same playbook, moving early to challenge novel state-law theories.²⁸ The growing line of rulings rejecting these theories has reinforced the view that many state-law causes of action—particularly those related to unfair competition, attribution or unjust enrichment—are either preempted by federal copyright law or too loosely connected to the alleged misconduct to support relief.

While these dismissals are favoring defendants, their practical benefit to defendants has been limited. Dismissing state-law claims rarely reduces discovery burdens or meaningfully alters the trajectory of a case where core copyright claims remain intact, and does little to reduce the cost or complexity of litigation. In many instances, these state-law claims appear to be included as part of the “kitchen sink” approach and are dismissed without significant resistance.

Nonetheless, the pattern of dismissals may carry longer-term consequences for the composition of future lawsuits by narrowing the available claims and the class of viable plaintiffs. Individuals or entities without registered copyrights, such as content aggregators, platform users, or those asserting broader reputational harms, may find themselves without a clear path to relief.

A New Battleground for Defendants—Class Certification

Looking ahead, AI developers have also begun to target class certification as key to their broader defense strategy. In the consolidated *In re Google Generative AI Copyright Litigation*, Google moved to strike class allegations, claiming that plaintiffs’ proposed class is insufficient because its definition depends on the ultimate legal determination of whether each IP holder was harmed, though plaintiffs have responded that the purported dependence of the class definition on individualized harm does not preclude certification at this stage, emphasizing that the existence of common questions—such as whether Google’s model training practices constitute infringement—predominates over individual variations in damages or harm.²⁹

AI developers are expected to argue that copyright claims are too individualized—turning on ownership, licensing and harm—to support class treatment. But it remains far from clear that courts will agree. Given the shared factual questions surrounding how AI systems collect and use creative works, plaintiffs may be able to demonstrate

²⁸ See, e.g., *Lehrman v. Lovo, Inc.*, No. 24 CV 03770-JPO, Reply Memorandum in Further Support of Defendant’s Motion to Dismiss the Amended Class Action Complaint, ECF No. 39 (S.D.N.Y. Jan. 31, 2025).

²⁹ *In re Google Generative AI Copyright Litigation*, No. 23 CV 03440-EKL, Defendants’ Reply in Support of Motion to Strike Class Allegations, ECF No. 111 (N.D. Cal. Feb. 21, 2025).

that these common issues predominate, making it difficult for defendants to prevent class certification at this early stage.

The first major ruling on class certification in AI litigation came in July, when Judge Alsup certified a “LibGen & PiLiMi Pirated Books Class” in *Bartz v. Anthropic*. The class is limited to actual or beneficial owners of timely registered, ISBN/ASIN-bearing works that Anthropic downloaded from pirate libraries. The court held that common metadata and hashing methods could reliably establish class-wide infringement, rejecting arguments that ownership disputes or file errors would defeat predominance. At the same time, Judge Alsup denied certification for Books3 and Scanned Books classes, finding their metadata too unreliable and their claims too fragmented.

The *Bartz* ruling underscores that class certification is available for plaintiffs that can marshal cohesive evidence of common proof.³⁰ Just weeks after the rulings on class certification and fair use, the parties in *Bartz* reached a preliminary settlement agreement, which was approved by the court on September 25. Plaintiffs and commentators have described the risks associated with class certification as a “death knell” scenario for defendants, highlighting why certification may quickly become one of the most consequential battlegrounds in AI litigation.

Trademark Claims Maintain Momentum as a Parallel Front

In addition to the steady cadence of copyright-based challenges, trademark claims have continued as a meaningful—and increasingly frequent—front in AI litigation. Though copyright theories have dominated headlines, plaintiffs are signaling that Lanham Act claims offer a powerful complementary tool in disputes over AI model training and outputs. Several high-profile actions in 2025 have included trademark infringement, false designation of origin, and dilution allegations alongside more traditional copyright claims. For example, when Encyclopaedia Britannica and Merriam-Webster sued Perplexity in September, they alleged that the company’s models routinely misrepresented their works by falsely attributing AI-generated content to their trademarked publications or omitting key material in ways that misleadingly suggested an association or endorsement.³¹ The complaint asserted that Perplexity knew its outputs were being attributed to plaintiffs’ marks, underscoring how false attribution

³⁰ *Bartz v. Anthropic PBC*, No. 24 CV 05417-WHA, Order on Class Certification, ECF No. 244 (N.D. Cal. July 17, 2025).

³¹ *Encyclopaedia Britannica, Inc. v. Perplexity AI, Inc.*, No. 25 CV 07546, Complaint, ECF No. 1 (S.D.N.Y. Sep. 10, 2025).

and “hallucinated” references to well-known brands are becoming fertile ground for litigation.³²

Recent rulings also confirm that Lanham Act theories can survive early procedural challenges. In *Advance Local Media v. Cohere*, for instance, a coalition of news organizations successfully defeated a motion to dismiss their trademark infringement and false designation of origin claims, which stemmed from AI-generated text that allegedly reproduced trademarks and misattributed content to the publishers.³³ Cohere argued both that no likelihood of confusion existed and that any trademark use was protected by nominative fair use. The court rejected these arguments at the pleading stage, holding that plaintiffs sufficiently alleged consumer confusion arising from AI outputs and that the nominative fair use doctrine did not apply on the facts presented. This ruling suggests that, where plaintiffs can plausibly allege confusion or reputational harm tied to AI-generated misattribution, Lanham Act claims still present a viable path forward—particularly as courts grapple with the ways in which large-scale model outputs can mimic brand identifiers or create misleading associations.

Finally, AI developers are learning that trademark exposure is not limited to the conduct of their models: traditional trademark clearance obligations continue to apply to product names, app features and branding decisions. Beyond lawsuits asserting trademark infringement through AI outputs, several recent actions involve more conventional disputes over brand identity, demonstrating that even industry leaders must conduct routine trademark due diligence. In November, OpenAI was sued by OverDrive, the maker of the Sora library app, which alleged that OpenAI’s Sora video-generation tool infringes its existing SORA trademark.³⁴ Days later, a federal court issued a temporary restraining order preventing OpenAI from using the mark CAMEO for one of its features, after the celebrity-video platform Cameo asserted the name infringed on its own mark.³⁵ These suits underscore that trademark issues in the AI landscape are not confined to cutting-edge questions about model outputs, but also encompass the everyday branding choices associated with rapidly evolving AI products—choices that can trigger immediate injunctive relief where established marks are implicated.

We expect a number of decisions will be issued in the coming months that will be pivotal for IP holders and AI developers alike, with courts expected to weigh in on more discovery disputes, renewed motions to dismiss, class certification challenges, and fair

³² *Id.* at 52.

³³ *Advance Local Media LLC v. Cohere Inc.*, No. 25 CV 01305-CM, Decision and Order Denying Defendant’s Partial Motion to Dismiss, ECF No. 59 (S.D.N.Y. Nov. 13, 2025).

³⁴ *OverDrive Inc v. OpenAI OpCo LLC*, No. 25 CV 02520, Complaint, ECF No. 1 (N.D. Ohio Nov. 19, 2025).

³⁵ *Baron App, Inc. v. OpenAI, Inc.*, No. 25 CV 09268-EKL, Order Granting Application for Temporary Restraining Order, ECF No. 42 (N.D. Cal. Nov. 21, 2025).

use defenses that could shape the course of future AI litigation. As more high-profile IP holders join the fray and continue to advance novel claims, AI developers will likely face growing pressure to negotiate licensing deals or mount aggressive legal defenses.

We will also be watching closely as IP holders press new theories that could further reshape the landscape, from claims that pirating and storing shadow library copies is distinct from training, to arguments over whether “leeching” and “seeding” works used in AI training infringes distribution rights, to carefully developed evidence of market dilution that courts have signaled could prove decisive in future fair use battles.

* * *

Please do not hesitate to contact us with any questions.



Megan K. Bannigan
Partner, New York
+1 212 909 6127
mkbannigan@debevoise.com



Christopher S. Ford
Counsel, New York
+1 212 909 6881
csford@debevoise.com



Carl Lasker
Associate, New York
+1 212 909 6185
clasker@debevoise.com



Abigail Liles
Associate, New York
+1 212 909 6387
aeliles@debevoise.com