Lessons Learned from 2024 and the Year Ahead in AI Litigation

January 13, 2025

Introduction

As generative AI platforms grow in sophistication, the initial era of text chatbots led by ChatGPT has evolved into a complex AI ecosystem of voice assistants and image and video creation platforms. Yet that is just the beginning; a world of autonomous AI agents is on the horizon. Generative AI has transformed how people around the world work; how they create; and what they see, hear, and watch online. But this new era of AI has not come without controversy, as authors and rights holders have launched waves of litigation against the companies that trained and released generative AI models, as well as their investors and affiliates, alleging violations of intellectual property rights.

Despite the busy 2024 litigation year against companies offering AI platforms in 2024, significant intellectual property questions remain unanswered as the calendar turns to 2025. The past year was marked by many more filed cases than decisions, and those decisions that were issued largely demonstrated how well-known pitfalls will also hamper this new wave of AI lawsuits.

A look back at 2024 shows that plaintiffs are learning from the initial wave of cases filed against AI developers, and that plaintiffs are now adapting their claims—as well as the defendants they are suing—to meet the shifting legal terrain. Although plaintiffs have been dealt some setbacks, they have responded by bringing new legal claims against a broader class of defendants, making clear that there is no end in sight to these disputes. Though much early media attention focused on copyright infringement claims, more recent cases have advanced a variety of claims, including trademark dilution, false advertising, right of publicity, and unfair competition claims that pose a new set of challenges for AI developers and companies that use generative AI outputs in their advertising and elsewhere.

2025 may bring *some* clarity to the legal status of AI, including through highly anticipated guidance from the U.S. Copyright Office on fair use, as well as judicial decisions tackling some of the thorniest unresolved legal questions. But courts may take

divergent paths on those issues, especially given the fact-specific nature of many of the plaintiffs' challenges, which depend not only on their specific claimed rights but also on the way each AI company has trained their model and how those models function. There are also likely to be fundamental disagreements among judges as to the strength of core defenses like fair use, which in the past have split appellate courts and even the Supreme Court. Earlier cases will reach new phases, bringing discovery disputes and class certification challenges. And the AI industry will not rest on its laurels while these legal issues are hashed out in court: new innovations may lead to new types of plaintiffs or claims, especially as more businesses adopt AI as part of their operations and marketing. With new cases being filed nearly every week in jurisdictions around the country and the world, what is most certain is that there will be a long road ahead.

Looking Back at 2024: Facing Setbacks, Plaintiffs Adapt and Evolve

Companies and individuals suing AI developers for IP infringement faced several stumbling blocks in 2024, as judges across the country began pruning plaintiffs' weaker arguments, largely finding fault with inadequately pled copyright claims:

- **Unregistered Works.** To state a federal Copyright Act claim, a plaintiff must have a registered copyright in the claimed work at the time the lawsuit is filed, a step many authors of copyrightable work have not taken. In *Millette v. OpenAI, Inc.*, two YouTube video creators alleged that OpenAI transcribed their videos without permission and fed the transcripts into training datasets used to develop its generative AI models. However, only one of the YouTube creators had registered a copyright for her video content with the United States Copyright Office. Although this plaintiff was able to bring a direct copyright infringement claim against OpenAI, her co-plaintiff was forced to limit his claims to unfair competition and unjust enrichment.[1]

- **Proving Defendants' Use of Training Data Inputs.** To support claims their copyrights were violated when used as an input for AI training data, plaintiffs must allege that their specific copyrighted works were ingested by defendants' AI model. This obstacle to "training data" claims has proven most problematic for plaintiffs filing suits against AI developers who have not publicized the datasets they used to train their models, leading plaintiffs to rely on shaky evidence like search results

[1] *Millette v. OpenAI, Inc.*, No. 24 CV 4710, First Amended Complaint, ECF No. 47 (N.D. Cal. Oct. 18, 2024). These plaintiffs likely learned from the 2023 decision in *Andersen v. Stability AI Ltd.*, where two artists and named plaintiffs were dismissed from a case because their artwork was unregistered. 700 F. Supp. 3d 853 (N.D. Cal. Oct. 30, 2023).

from the website "haveibeentrained.com" to support their claims.[2] But many AI training datasets are not publicized, and AI developers have argued that the mechanisms and data they use to train their models are trade secrets. Defendants have fought to file documents under seal, reducing the chance that discovery could be leveraged by subsequent plaintiffs.

- **Showing Substantial Similarity of Generative AI Outputs.** Courts have widely rejected claims that AI models themselves are unlawful derivatives of the copyrighted works they were trained on.[3] This has left plaintiffs to argue that AI models' outputs infringed their copyrighted works, but many plaintiffs have struggled to produce evidence of verbatim reproduction of their copyrighted works – a phenomenon known among AI developers as "regurgitation." Most AI models' outputs are not exact or even near replicas of their training data, especially now that AI developers have caught on to the significance of the regurgitation problem, adjusted their controls, and added guardrails to prevent models from generating certain content. This has made it difficult for plaintiffs bringing "infringing output" claims to show examples of outputs that wholly copy from their copyrighted works, forcing plaintiffs to argue only that AI platforms could hypothetically produce substantially similar outputs.

- **Meeting the DMCA's Statutory Requirements.** Digital Millennium Copyright Act ("DMCA") claims typically argue that AI developers remove or omit copyright management information ("CMI") like author names or copyright notices in their datasets and in generative AI outputs that utilize the copyrighted works in question. Plaintiffs have struggled to bring a successful DMCA claim against AI developers, however, because they need to allege that the developers knew or had reason to know that their actions would "induce, enable, facilitate, or conceal" copyright infringement. Courts have long interpreted this standard to require plaintiffs to establish that defendants *first*, knew they removed or altered CMI, and *second*, knew that their CMI removal or alternation would likely result in copyright infringement.[4] In addition, plaintiffs must also show that defendants wholly reproduced the copyrighted works at issue without any changes outside of the CMI removal or alteration.[5] This has so far been a difficult standard for plaintiffs to meet.

  In one (atypical) case filed in 2024, *Raw Story Media, Inc. v. OpenAI Inc.*, the plaintiff alleged *only* a violation of the DMCA. Unlike the federal Copyright Act, the DMCA

---

2 *See id.*

3 *Kadrey v. Meta Platforms, Inc.*, No. 23 CV 3417, 2023 WL 8039640 (N.D. Cal. Nov. 20, 2023).

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

5 *Kirk Kara Corp. v. Western Stone and Metal Corp.*, No. CV 2-20-1931, 2020 WL 5991503 (C.D. Cal. Aug. 14, 2020) (no DMCA violation because the work with the removed CMI, while "substantially similar," was not an "identical" copy of the plaintiff's work).

does not require the works at issue to be registered, which may have been the reason the *Raw Story Media* plaintiffs pursued only that basis for relief. Raw Story Media has collectively published over 400,000 news articles and features, and likely has not registered all of them with the U.S. Copyright Office. In its complaint, Raw Story Media accused an AI developer of removing CMI (including author names and copyright notices) when using Raw Story Media articles as AI training data, arguing this facilitated unauthorized use and concealed infringement. Defendants successfully moved to dismiss, however, arguing that Raw Story Media could not show that it had suffered any harm and therefore lacked standing to pursue its claims.[6]

Despite these challenges, plaintiffs are not slowing down, and new lawsuits were filed at a steady clip over the course of 2024. More corporate plaintiffs filed complaints in 2024 than ever before. Plaintiffs of all kinds pursued claims against an increasingly diverse pool of defendants, including newer AI-developers and corporate investors in AI. Although many of those complaints focused on the core copyright and trademark infringement theories present in the earliest AI lawsuits, plaintiffs have begun diversifying their claims to include false advertising and unfair competition, along with other state-law causes of action. The most recent complaints also demonstrate that plaintiffs are learning from the struggles of those who filed suits before them and are modifying their strategies to avoid the same pitfalls.

More Plaintiffs: Media Corporations Enter the Fray, Alongside New Individual Claimants

Although the earliest suits against AI developers were largely class actions filed by individual creators, corporate plaintiffs with large IP assets increasingly entered the legal fray in 2024, suing AI developers in a bid to protect their keystone properties and building on the legal theories first advanced in early class action cases.

The New York Times opened the floodgates for corporate IP plaintiffs in the final week of 2023, filing a complaint against Microsoft and Open AI alleging they had trained their generative AI models using millions of The New York Times' copyrighted works without permission.[7] Since then, other corporations like The Daily News, The Chicago Tribune, Universal Music Group, Warner, Sony, Atlantic Records, The Intercept Media, and most recently Dow Jones have followed suit and filed similar claims.

Although these corporate plaintiffs dominated the news, individual plaintiffs continued to file new cases as well. Voice actors, YouTubers, authors, and other artists initiated

---

[6] *Raw Story Media, Inc. v. OpenAI, Inc.*, No. 24 CV 1514, 2024 WL 4711729 (S.D.N.Y. Nov. 7, 2024).

[7] *The New York Times Co. v. Microsoft Corp.*, No. 23 CV 11195, Complaint, ECF No. 1 (S.D.N.Y. Dec. 27, 2023).

new lawsuits in 2024, hoping their claims might generate relief from new or existing defendants.

More Defendants: New Developers Become Targets, and Investors and Partners Face Contributory Claims

As a second generation of AI developers began releasing their products to the market, they were met with the same suits early AI companies like OpenAI, Anthropic, and Meta faced back in 2023. Plaintiffs filed suits in 2024 against AI developers including MosaicML, Suno, Uncharted Labs, Perplexity AI, and Lovo, largely relying on the claims and strategies that had been pioneered in earlier suits against the first generation of AI developers.

Although these new suits signaled the persistence of rights owners in bringing claims against AI developers large and small, plaintiffs in 2024 also began trying new tactics to target a wider class of defendants. The early wave of cases involving AI developers generally targeted those developers directly. More recently, plaintiffs have pursued defendants beyond the developers themselves, asserting theories of contributory liability to bring new plaintiffs into the mix. Complaints against parent or partner companies, like Databricks and Microsoft—who are large investors in AI technology but not directly responsible for the underlying AI model—have highlighted the financial and technical support given to AI developers and the material benefits these partner defendants received from AI developers' alleged infringement in arguing they have vicarious and contributory liability for facilitating the AI developers' work.[8]

As more companies begin to fund and provide critical infrastructure for AI projects going forward, they may become targets of plaintiffs pursuing this broader approach to naming defendants. Courts have yet to clarify how indirect infringement theories will apply in these cases, or just how broad plaintiffs' claims might reach.

More Claims: Complaints Reach Beyond Copyright for Novel Theories of Liability

The most common claims brought in the first wave of AI litigation were grounded in copyright law. Plaintiffs frequently argued that using their copyrighted works as AI training data violated their copyrights, and that AI models were outputting infringing reproductions or derivative works. As courts began to clarify the viability of those arguments—most notably, casting doubt on plaintiffs' derivative work theories—plaintiffs in 2024 reached beyond copyright law to assert claims based on trademark

---

[8] *Makkai v. Databricks, Inc.*, 24 CV 2653, Complaint, ECF No. 1 (N.D. Cal. May 2, 2024); *Daily News LP v. Microsoft Corp.*, No. 24 CV 3285, Complaint, ECF No. 1 (S.D.N.Y. Apr. 30, 2024).

infringement and dilution, false advertising, and state-law right of publicity and unfair competition theories.

These "kitchen sink" complaints have routinely faced motions to dismiss, though these motions have rarely succeeded in ending litigation entirely. Some claims, especially trademark claims and copyright claims based on training data, are surviving into discovery. Courts, likely cognizant of the unsettled terrain, have also been permissive in granting leave to replead even where motions to dismiss succeed against an initial (or even an amended) complaint. For example, in *Andersen v. Stability AI Ltd.*, No. 23 CV 201 (N.D. Cal.), the court allowed the plaintiffs to file multiple amended complaints to try to remedy the deficiencies that had been identified after rounds of motions to dismiss. This has given plaintiffs multiple bites at the apple in formulating their claims – and delayed appellate review of emerging trends in district court opinions.

Some of the earliest AI cases included trademark claims alongside copyright claims, but plaintiffs in 2024 began reaching beyond classic infringement theories to argue dilution—a type of trademark claim that requires a plaintiff prove the fame of their mark, but which then permits the plaintiff to pursue defendants without having to prove consumer confusion. For example, a group of media companies claimed that AI "hallucinations" (incorrect or misleading information generated by an AI model that was attributed to the plaintiffs' publications) damaged their brands and reputations.[9] Plaintiffs have also asserted Lanham Act false advertising and deceptive practices claims against AI developers, alleging misrepresentation of their data sources and permissions.[10]

In addition to these federal causes of action, plaintiffs have sought to use state law to avoid some of the pleading requirements that caused difficulty in earlier cases. For example, the authors of works that were never registered (which is required to bring a federal copyright claim)—like YouTube creators—have used state unfair competition and unjust enrichment claims to pursue AI companies, arguing they profited unfairly when using the plaintiffs' content as training data without permission. These claims have generally been met with skepticism by courts where they are duplicative of claims under the Copyright Act given that the Act preempts state laws that grant rights equivalent to federal copyright rights.[11] Voice actors have also sued AI companies for

---

9 *Daily News v. Microsoft Corp.*, No. 24 CV 3285, Complaint, ECF No. 1 (S.D.N.Y. Apr. 30, 2024).

10 *Lehrman v. Lovo, Inc.*, No. 24 CV 3770, Complaint, ECF No. 1 (S.D.N.Y. May 16, 2024).

11 *Millette v. OpenAI*, No. 24 CV 4710, Complaint, ECF No. 1 (N.D. Cal. Aug. 2, 2024). In December 2024, OpenAI filed a motion to dismiss against the YouTube video creators in *Millette* arguing that the creators' complaint is "a carbon copy of the complaint in *Tremblay, et al. v. OpenAI*," and arguing that the court in *Tremblay* and every other case addressing state-law theories like unfair competition and unjust enrichment have "reached the same conclusion: that the use of copyrighted material to train AI models is governed exclusively by federal copyright law." *Millette*, No. 24 CV 4710, Motion to Dismiss, ECF No. 55 (N.D. Cal. Dec. 16, 2024).

state-law right of publicity violations, alleging their voices were cloned without consent to develop generative AI text-to-speech platforms.[12]

Looking Ahead to 2025: What's Next for AI Litigation

The surge of complaints filed in 2024 means that, in 2025, we expect defendants—and courts—to steal the spotlight from the plaintiffs that drew headlines over the past year with their filings. Initial rulings are expected on AI developers' core fair use defense, which could either open litigation floodgates or slow the stream of new cases and increase the pace at which existing cases are resolved.

In the cases that survived early motion practices, judges will confront tricky discovery challenges, including determining whether developers' training data and procedures can remain confidential.

Meanwhile, we expect plaintiffs will continue filing claims against AI developers and continue their efforts from 2024 to expand the playing field—raising disputes over licensing agreements, seeking expansive discovery, moving for class certification, and fighting across jurisdictions.

Clarity on the Strength of AI Developers' Fair Use Defenses

As plaintiffs continue to file new lawsuits, they await clarity on the strength of AI developers' fair use defenses with trepidation. The doctrine of fair use allows use of copyrighted material without permission when the user's purpose is transformative or educational rather than commercial. AI developers have consistently argued that the inclusion of copyrighted works in training datasets constitutes fair use because the AI developers first transform the works into small tokens of data that are used to teach AI models to generate new content. AI developers also point out that they use vast datasets for training, and the individual contribution of any single copyrighted work to a models' functionality is minimal. Publishers and creators contend that AI models are fundamentally commercial and compete with their copyrighted works in the marketplace.

Courts have avoided conclusively ruling on the viability of the fair use defense in generative AI IP cases so far, but a ruling is likely to arrive sooner rather than later as parties continue filing pretrial motions. The U.S. Copyright Office is also poised to

[12] *Lehrman*, Complaint.

release a comprehensive report on how the doctrine of fair use should apply to generative AI early in 2025.

The first signs from the judiciary might come from a District of Delaware judge who is poised to rule on a renewed motion for summary judgment in *Thompson Reuters v. Ross Intelligence* (Case No. 20 CV 613), or a Northern District of California judge who is poised to rule on a preliminary injunction in *Concord Music Group v. Anthropic* (Case No. 23 CV 1092). These cases are factually quite distinct—the *Thompson Reuters* case involves a more unusual set of facts involving training a legal AI platform using headnotes from the well-known Westlaw platform, while the *CMG* case poses a more common type of challenge to the use of copyrighted works in AI training data. The outcome of these motions could shed light on the viability of fair use defenses going forward—although the distinct sets of facts could result in a split outcome that would only deepen uncertainty surrounding fair use in the AI context.

A key issue judges and regulators weighing fair use questions will consider is how much weight to place on transformativeness and commercialism in their decisions—and ultimately any jury instructions—regarding fair use. In 2023, the Supreme Court elevated the importance of commercialism in fair use analysis in *Andy Warhol Foundation v. Goldsmith*,[13] reaffirming that courts should consider whether the defendants' ultimate use of the allegedly transformed work is commercial when evaluating the purpose and character of the use at issue. However, Judge Stephanos Bibas, who is overseeing *Thompson Reuters v. Ross Intelligence*, cautioned against "overread[ing]" this decision in an early opinion, suggesting that the Supreme Court's prior focus on transformativeness in *Google v. Oracle*,[14] may be more pertinent than the *Warhol* decision in cases related to emerging technologies like AI.[15]

Although Judge Bibas ultimately did not resolve the application of the fair use doctrine in the *Thomson Reuters* case, holding that factual questions required a trial, those confronting the fair use defense in 2025 will have to address the difficult question of whether the transformativeness of AI models outweighs their commercial nature.

Discovery Disputes over AI Training Data and Trade Secrets

Although courts have allowed plaintiff's infringement claims to survive a motion to dismiss based on evidence that content creators' names have appeared on websites like "haveibeentrained.com," copyright holders will likely need more direct proof that AI developers trained models using their copyrighted works as the litigations progress. As

---

[13] 143 S. Ct. 1258 (2023).

[14] 141 S. Ct. 1183 (2021).

[15] *Thompson Reuters Enterprise Centre GmbH v. Ross Intelligence*, 694 F. Supp. 3d 467, 482 (D. Del. 2023).

cases proceed into discovery, courts will begin to rule on AI developers' arguments that their algorithms should be protected under seal.

While plaintiffs struggle to prove that AI developers trained their AI models using the plaintiffs' registered works, AI developers argue that their training datasets should not be publicized. According to the developers, the mechanisms and data they use to train their models constitute trade secrets that must remain proprietary for the developers to maintain advantages over competitors. In recent cases, however, judges have been skeptical of AI developers' attempts to keep this information under seal, which has provided prospective plaintiffs additional information about where their works may have been used.

Although some AI developers have previously published their training data, this practice may come to a halt if the trade secret arguments are successful—and if publicizing training data is used to strengthen lawsuits against the developers.

The Next Wave of AI Claims: Looking Beyond Copyright

As decisions clarify the relative strengths and weaknesses of plaintiffs' claims and AI developers' defenses, we expect plaintiffs to continue to pursue an ever-larger pool of defendants, including with a broader array of claims:

- One area we expect to see increase in importance are **trademark claims** arguing that chatbots fueled by generative AI infringe brand owners' trademark rights when they hallucinate or generate outputs that use or alter existing trademarks. The initial round of these claims flew largely under the radar, as they were generally not targeted by defendants' initial motions to dismiss. These claims may be challenging to prove given the bespoke nature of generative AI outputs, but also could pose difficulties for AI developers whose models cannot be easily changed to avoid use of trademarks and logos.

- We also expect to see more **right of publicity claims** in 2025, both against AI developers as well as companies who use AI outputs that mimic celebrities in their advertising. In one well-publicized instance in 2024, Scarlett Johansson sent a cease-and-desist letter to OpenAI after the company released an AI voice assistant called "Sky" that Johansson alleged was modeled on her voice even after she refused to work with OpenAI. Although OpenAI claimed the model was trained using a different voice actress and the dispute was quickly resolved, the increasing capacity of generative AI will make it easier for the technology to knockoff the voices and appearances of public figures, which may lead to more public figures needing to resort to lawsuits to protect their rights to their likenesses.

- As AI-generated content becomes widely used in online advertising, companies may also face **false advertising claims** for distributing AI-generated advertisements through chatbots and other marketing platforms. Ensuring proper controls around the creation and use of AI-generated advertising materials, and vetting such materials to ensure that the advertising claims can be substantiated, has been a key compliance focus that will only grow in importance in the coming years.

- As more and more companies enter the AI space, we also expect to see an increase in **trade secrets litigation** involving AI models. In March 2024, Meta settled a trade secrets misappropriation claim with Neural Magic, after Meta hired one of Neural Magic's employees and then published a new compiling program that allegedly contained Neural Magic's proprietary technology. Although the case settled confidentially—not long after the court refused to exclude a plaintiff's expert who calculated potential damages at $766 million—it is likely a bellwether for future trade secrets claims among AI companies.

The Next Battlegrounds for AI Licensing and Litigation

Licensing

As litigation begins to clarify the contours of creators' rights—and the costs of litigating those disputes—AI developers and owners of copyrighted works may begin to negotiate widespread licensing schemes with increased fervor, as they seek to minimize the risks of an unfavorable legal ruling. Although licensing content from specific data repositories for use in training data is already a common practice, broad licensing schemes may offer massive payouts for would-be corporate plaintiffs, along with the news publications and music publishers that have already begun litigation. Establishing licensing schemes with AI developers could enable a more stable market where content creation and generative AI coexist with less animosity; however, AI developers, especially smaller companies, may be unwilling or unable to pay robust licensing fees.

Class Actions

The early wave of cases against AI developers included many class actions, which will soon reach the class certification stage. We expect to see hard-fought battles over whether class treatment is appropriate, and AI developers who are unsuccessful in motions to dismiss may attempt to escape these suits by asking courts to resolve class certification issues early in the litigation process. The hurdles posed by class certification, especially for output-based claims that may depend on the specifics of how individual plaintiffs' works appeared in generated content, may halt many lawsuits before they are able to reach the merits of the claims.

International Disputes

Although early lawsuits against AI developers were largely filed in the U.S., given the nationality of the major early players in the AI space, we expect to see plaintiffs' fights

with AI developers spill onto other jurisdictional battlegrounds in 2025. In November 2024, an India-based multimedia news company called Asian News International filed an "infringing input" suit against OpenAI in the High Court of Delhi. A few plaintiffs have also filed claims against AI developers in Canada, Germany, and the United Kingdom. We expect to see this trend continue – especially if cases in the U.S. begin to result in victories for AI developers.

* * *

As more and more companies enter the generative AI space, both as developers and as users, they will be new fodder for lawsuits from a plaintiffs' bar that is growing more sophisticated at drafting complaints that can survive motions to dismiss. This is especially true for well-resourced corporate plaintiffs, who will likely play an even bigger role going forward. In the meantime, individual creators may struggle to overcome the hurdles discussed above—or come up short of the resources needed to conduct expensive discovery against AI developers that have had little difficulty raising funds.

Absent an extraordinary intervention by legislatures in the U.S. and around the world, the viability of intellectual property claims against AI developers will not be resolved in 2025. Cases are proceeding across too many different jurisdictions, many raising highly fact-dependent arguments. Even decisive rulings in prominent cases will not prove to be the final word, given the near certainty that the losing side will quickly appeal.

Companies that are already actively working in the AI space—and those waiting on the sidelines to see how the legal risks shake out—will need to continue to watch closely as these hard-fought battles play out around the world.

* * *

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

Samuel J. Allaman
Associate, New York
+ 1 212 909 6026
sjallama@debevoise.com

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

Joshua Plastrik
Associate, New York
+1 212 909 6953
jhplastrik@debevoise.com

Carl Lasker
Law Clerk, New York
+ 1 212 909 6185
calasker@debevoise.com

Chloé Sweeney-Rowan
Law Clerk, New York
+ 1 212 909 6240
cpsweeneyrowan@debevoise.com