quinn emanuel
quinn emanuel urquhart & sullivan, llp

Emerging AI Legal Risks: July 2026 Update

TABLE OF CONTENTS Page

I. Trade Secrets and AI

a. IP Protection For AI

Although hundreds of billions of dollars are being invested in AI, IP protection for innovative designs and uses of AI inventions may be difficult to secure. Protection for new hardware architectures and implementations optimized for AI workloads is likely to be relatively straightforward. But the neural networks and algorithms that power many of the most inventive AI applications may be more difficult to patent. US courts have grown more critical of software patents over the last decade. The patent system does not allow patents for abstract ideas, mathematical formulas, and laws of nature. The Supreme Court has already determined that a patent that claims a standard computer performing a task is directed to an unprotectable abstract idea. Specifying that a process performed by a "neural network" or artificial intelligence, without more implementation detail, will not change the result.

Thus far, courts have not grappled with many cases involving AI patent eligibility. But a handful of recent decisions show skepticism regarding the patent eligibility of AI-based algorithms leveraging trained models, even where some might argue that the particular approach to developing or training the model was innovative. The Federal Circuit recently held in *Recentive Analytics, Inc. v. Fox Corp.* that claims directed to machine learning models to generate optimized NFL schedules and network maps for determining how television programs are displayed in designated geographic markets were directed to patent-ineligible subject matter.[1] The court reasoned that the claimed inventions were patent-ineligible because they merely applied conventional machine learning models without reciting any specific technological improvement.[2] However, the court cabined its holding to *Recentive*'s specific facts, and noted that machine learning may lead to patent-eligible improvements in technology, but that in the present case it was merely holding that the application of generic machine learning to new data environments, without disclosing improvements to the machine learning models themselves, is patent ineligible.[3] With decisions such as *Recentive*, the level of algorithmic detail that may be necessary to overcome the patent eligibility hurdle may make it easier for others to design competing AI systems that avoid narrowly drafted patent claims, while still achieving many of the same benefits of the invention.

The U.S. Patent & Trademark Office (USPTO) issued a Guidance Update in 2024 (the "2024 Update") focusing on patent subject matter eligibility for artificial intelligence.[4] Along with the 2024 Update, the USPTO published several patent subject matter eligibility examples (Example 47-49) applying the general principles to AI-related inventions. The new examples demonstrate the importance of claim drafting strategy: claims that describe AI training at a high level without explicitly reciting specific algorithms may avoid being characterized as reciting abstract ideas, while claims that name specific mathematical algorithms (such as backpropagation and gradient descent) are more likely to be found to recite judicial exceptions requiring further eligibility analysis.

---

1 134 F.4th 1205 (Fed. Cir. 2025).
2 *Id.* at 1212.
3 *Id.* at 1215-16.
4 *2024 Guidance on Patent Subject Matter Eligibility, Including on Artificial Intelligence*, 89 Fed. Reg. 58,128 (July 17, 2024).

In August 2025, the USPTO issued an internal memorandum to examiners in Technology Centers handling software and AI inventions that reinforces and clarifies these principles (the "2025 Memo"). The 2025 Memo uses the USPTO's own examples to illustrate a critical drafting distinction for AI patent claims. Broadly stating that an invention involves "training the neural network" does not preclude patent eligibility under § 101 because this "does not set forth or describe any mathematical relationships, calculations, formulas, or equations using words or mathematical symbols." 2025 Memo at 3. However, specifying that such training includes "a backpropagation algorithm and a gradient descent algorithm" does recite an abstract idea because the limitation "requires specific mathematical calculations by referring to the mathematical calculations by name." *Id*. The memorandum instructs examiners to "be careful to distinguish claims that recite an exception (which require further eligibility analysis) from claims that merely involve an exception (which are eligible and do not require further eligibility analysis)." *Id*. This guidance suggests that AI patent applicants may improve their chances of avoiding initial §101 rejections by using higher-level functional language that describes what the AI system does rather than explicitly naming the mathematical algorithms it employs—though this approach must be balanced against enablement and written description requirements under §112.

The 2025 Memo also emphasizes limits on classifying AI processes as unpatentable "mental processes." Examiners are instructed not to expand the mental process category to encompass "claim limitations that cannot practically be performed in the human mind." 2025 Memo at 2. The memorandum specifically states that "claim limitations that encompass AI in a way that cannot be practically performed in the human mind do not fall within this grouping." *Id*. This guidance may provide a pathway for AI inventions to overcome abstract idea rejections by demonstrating that the claimed AI processes—such as processing millions of data points, analyzing complex neural network architectures, or performing real-time computations—are beyond human mental capacity and therefore do not constitute mental processes.

Most significantly for AI patent applicants, the memorandum provides important procedural protections that may reduce marginal §101 rejections. It instructs examiners to make eligibility rejections only when it is "more likely than not (i.e., more than 50%)" that claims are ineligible, and that "a rejection should not be made simply because an examiner is uncertain as to the claim's eligibility." 2025 Memo at 5. The memorandum cites *Recentive Analytics* as an example of claims that failed eligibility by including only "steps incidental to automating an abstract idea," contrasting it with USPTO Example 47, claim 3, which the USPTO views as improving "the technical field of network intrusion detection." 2025 Memo at 5. This reinforces that AI patent applicants should focus their claims and specifications on demonstrating specific technological improvements rather than merely applying AI to new data environments—the approach that failed in *Recentive.* The memorandum also cautions examiners not to "oversimplify claim limitations," 2025 Memo at 4, and requires that they conduct a claim-as-a-whole analysis where "the analysis should take into consideration all the claim limitations and how these limitations interact and impact each other when evaluating whether the exception is integrated into a practical application." 2025 Memo at 3-4. These instructions may provide AI patent applicants with stronger grounds to overcome examiner rejections based on abstract idea theories.

The European Patent Office (EPO) issued similar guidelines on April 1, 2025 explaining that if "a claim of an invention related to artificial intelligence or machine learning is directed either to a method involving the use of technical means (e.g., a computer) or to a device, its subject-matter has technical character as a whole and is thus not excluded from patentability."[5] The guidelines add that "the computational models and algorithms themselves contribute to the

[5] European Patent Office, *Guidelines for Examination in the European Patent Office*, II-5, 3.3.1 Artificial intelligence and machine learning (Apr. 2025).

technical character of the invention if they contribute to a technical solution to a technical problem, for example by being applied in a field of technology and/or by being adapted to a specific technical implementation."[6]

For AI inventions that do pursue patent protection, the USPTO's 2025 Memo underscores the critical importance of the patent specification in supporting eligibility arguments. The memorandum instructs examiners to "consult the specification to determine whether the disclosed invention improves technology or a technical field," but notes that "the specification does not need to explicitly set forth the improvement" if it "describe[s] the invention such that the improvement would be apparent to one of ordinary skill in the art." 2025 Memo at 4. Furthermore, "the claim itself does not need to explicitly recite the improvement described in the specification." 2025 Memo at 4. Patent applicants should therefore ensure their specifications clearly articulate: (1) the specific technological problem being addressed, (2) how the AI invention provides a particular solution to that problem, and (3) why the solution represents an improvement over prior approaches—even if the claims themselves do not explicitly recite these improvements. The memorandum emphasizes that examiners should consider, among other things: (1) "whether the claim recites only the idea of a solution or outcome" versus "whether the claim covers a particular solution to a problem or a particular way to achieve a desired outcome," and (2) "whether the claim invokes computers or other machinery merely as a tool to perform an existing process, or whether the claim purports to improve computer capabilities or to improve an existing technology." 2025 Memo at 4. This documentation becomes critical evidence in overcoming §101 rejections while potentially avoiding the disclosure of sensitive algorithmic details that could enable design-arounds.

While there may now be a clearer path to seeking patent protection on AI innovations, AI businesses may continue to rely heavily on trade secret protection to guard valuable AI-related intellectual property. A trade secret may be any information which is valuable because it is not publicly known. Even though abstract ideas, mathematical formulas, and other valuable AI innovations may not be patent eligible even under the PTO's new guidance, they can still be protected as trade secrets. Trade secrets have the added benefit that they are still secrets—they are not published like patents, and thus do not provide a competitor with a roadmap for developing a similar (but non-infringing) AI solution. And if someone does improperly acquire a trade secret and use it to develop a competing system, that party may be liable for trade secret misappropriation even if their competing solution ends up being remarkably different from the system they misappropriated details on and then used as a baseline for their development process—a broader scope of protection than patents, which typically only provide protection against competitors using the particular claimed invention in a patent issued by the Patent Office. Patents also have a life of only 20 years, whereas trade secrets may be maintained for far longer—as long as they are kept confidential and remain valuable secret information.

But there are advantages to patents for inventions that are likely to qualify for patent protection. Whereas a patent prevents anyone from practicing the invention, even if they discover it independently, a trade secret can be used by others if it is independently discovered—for example, through reverse engineering. Companies relying on trade secret protection must also always be vigilant to take reasonable measures to see that they have protected the secret—especially with employees, vendors, and joint venturers, or they may lose the ability to pursue claims against misappropriators. There is good reason to believe that AI companies are increasingly relying on trade secrets over patent protection. For example, OpenAI has 39 issued U.S. patents, primarily directed to training methods and systemic interface for AI systems (e.g., output processes, particularly methods to summarize text prompts and generate image or video

[6] *Id.*

responses, and interaction with large language models) and not the AI systems themselves.[7] Anthropic, another prominent AI company, currently has only 13 public patents or patent applications, covering training data software and internal architecture automating a multi-modal interface.[8]

It remains to be seen whether the 2025 Memo will result in significantly more favorable outcomes for AI patent applicants. However, the memorandum's emphasis on the 50% threshold for making rejections, its caution against oversimplifying claims, its clarification of limits on the mental process category, and its distinction between claims that merely "involve" versus "recite" abstract ideas, together should provide applicants with stronger arguments in examination and appeal proceedings.

b. Trade Secret Theft in the Age of AI

The preceding discussion explains why AI businesses increasingly rely on trade secret protection. That reliance also creates new exposure. Companies are by now broadly aware that trade secrets can be lost through careless AI use—for example, when employees upload proprietary information to consumer AI platforms and thereby expose it to third parties. Less appreciated is a newer and more troubling risk: AI as a tool for deliberate theft. The traditional model of misappropriation—in which a departing employee collates files before transferring them to a thumb drive—is giving way to one in which proprietary information can be summarized in seconds and exported in a form that bears no resemblance to the original, often without a conventional paper trail.

1. The Rising Threat of Insider AI Theft

AI can collapse the traditional three-step process of trade secret theft—locate, obtain, and exfiltrate—into a single chat conversation. An employee can ask an unguarded enterprise AI to summarize a pricing strategy, describe the development status of an unreleased product, or explain a proprietary process. No document is opened, no file is moved, and the synthesized answer may itself constitute a new form of protected information. Trade secret theft has long occurred through efforts to disguise the stolen data.[9] Now, proprietary information that has been converted through AI may bear little resemblance to its sources, further complicating detection as AI-assisted exfiltration grows more sophisticated.[10]

A determined employee need not attempt to obtain a trade secret in a single session. The employee may instead query an enterprise AI about different components across multiple sessions—each query seemingly innocuous in isolation—and later collate the results into a functional reproduction of the trade secret. This pattern can defeat the bulk-download alerts that conventional data-loss-prevention tools are designed to catch. The relevant forensic question is not whether a single large extraction occurred, but whether a pattern of isolated queries adds up to one.

---

[7] U.S. Patent & Trademark Office, *Patent Public Search*(Advanced Search), https://ppubs.uspto.gov/pubwebapp/(searching "OpenAI.AS.") (last visited July 7, 2026).

[8] U.S. Patent & Trademark Office, *Patent Public Search*(Advanced Search), https://ppubs.uspto.gov/pubwebapp/ (searching "Anthropic.AS.") (last visited July 7, 2026).

[9] *See, e.g., Former Google Engineer Found Guilty of Economic Espionage and Theft of Confidential AI Technology*, U.S. Dep't of Justice, (Jan. 30, 2026), https://www.justice.gov/opa/pr/former-google-engineer-found-guilty-economic-espionage-and-theft-confidential-ai-technology (conviction for trade secret theft by an engineer who copy-pasted source code into PDF files that evaded Google's loss-prevention systems).

[10] *See, e.g., "Exploit Every Vulnerability": Rogue AI Agents Published Passwords and Overrode Anti-Virus Software*, The Guardian (Mar. 12, 2026), https://www.theguardian.com/technology/ng-interactive/2026/mar/12/lab-test-mounting-concern-over-rogue-ai-agents-artificial-intelligence (describing AI agents that cooperated to smuggle information out of supposedly secure systems).

Enterprise AI platforms also maintain significant context about an employee's work through memory and project features. Many platforms now offer memory-export functions that allow users to transfer their conversation history and context to other accounts.[11] A departing employee may thus achieve wholesale knowledge transfer—in summarized or even disguised form—with a few simple prompts.

Over-permissioned enterprise AI tools can also automate the location and collation phase of a theft. Where an organization has not enforced least-privilege access controls across its AI-accessible data sources, its enterprise AI may synthesize results from sources that an errant employee could never have located manually—effectively acting as an expert accomplice to the theft.

2. Risk Considerations: AI Access and Controls

The above-described risks of AI-enabled theft may lead some organizations to evaluate controls that go beyond standard acceptable-use policies and nondisclosure agreements. A range of approaches has emerged, each involving tradeoffs that will vary by organization. Some organizations may consider role-based access controls that limit the queries a given employee can make through an enterprise AI platform. Such controls can both constrain the scope of insider threats and generate audit trails, but they require ongoing maintenance as roles and systems evolve and may not be feasible on every platform.

A related question concerns the underlying data permissions that an enterprise AI inherits. Where a company's data repositories have accumulated broad access rights over years of organic growth, an enterprise AI can surface sensitive data that an employee could never have located manually. Organizations may want to audit and restrict those underlying permissions before enabling broad AI-retrieval capabilities—a step that itself requires tradeoffs between a particular AI deployment and the organization's data architecture.

Some AI platforms offer administrator controls that limit or disable memory-export features. Limiting memory portability can reduce the risk of knowledge transfer on departure, though the availability and effectiveness of these controls depends on the platform and may require specific provisions in third-party vendor contracts. An organization's ability to implement such controls may depend substantially on what is negotiated at the time of contracting.

Some organizations have also updated employment agreements and NDAs to address AI-specific conduct, including personal-account use, memory export, and disclosure of the AI tools used during employment. AI-specific provisions may serve as deterrents and, where a dispute arises, as evidence bearing on willfulness under the Defend Trade Secrets Act ("DTSA"). The DTSA's definition of "employee" extends to independent contractors and consultants,[12] and equivalent considerations apply to those agreements. Agreements covering trade secrets should also include the whistleblower-immunity notice required by DTSA § 1833(b); omitting that notice forfeits the ability to seek exemplary damages and attorney's fees in any subsequent DTSA action[13]—a potentially significant penalty in a high-stakes case, where the conduct may be egregious but the damages difficult to quantify.

---

[11] *See, e.g., Leaving ChatGPT for Claude? Here's the Trick to Taking Your AI Memory With You*, PCMag, (Apr. 6, 2026), https://www.pcmag.com/explainers/leaving-chatgpt-for-claude-heres-the-trick-to-taking-your-ai-memory-with; *Make the Switch: Bring Your AI Memories and Chat History to Gemini*, Google, (Mar. 26, 2026), https://blog.google/innovation-and-ai/products/gemini-app/switch-to-gemini-app/.

[12] 18 U.S.C. § 1833(b)(4) (defining "employee" to include "any individual performing work as a contractor or consultant for an employer").

[13] 18 U.S.C. § 1833(b)(3)(C) (providing that an employer that fails to include the whistleblower notice required by § 1833(b)(1) "may not be awarded exemplary damages or attorney fees" in a subsequent DTSA action).

3. Building the Trial Record When Theft Is Suspected

Some of the most probative evidence in an AI-based trade secret case may be held by an AI platform rather than the employer's own systems, and may be subject to deletion. Enterprise AI platform audit logs, browser history, clipboard logs, network traffic to AI-platform domains, and data-loss-prevention alerts should be reviewed before an employee knows of an investigation. Where platform-side records are needed, a preservation demand—or, if necessary, an emergency temporary restraining order—may be warranted before the employee can delete the relevant account. These contemporaneous logs may be among the most powerful evidence in a trade secret case: they are difficult to fabricate and, unlike witness testimony, not subject to the credibility disputes that often dominate trade secret trials.

Because AI-based misappropriation frequently does not manifest as a single discrete event, the relevant forensic question may instead be whether a pattern of queries—across sessions, subjects, and time—constitutes a functional reproduction of a trade secret. That analysis requires a forensic expert who can work across AI-platform logs, enterprise access records, and the content of AI outputs. Where a company suspects indirect prompt injection—a form of theft likely to grow as AI use increases—a targeted audit of the documents the employee provided to the company's AI systems before departure may also be warranted, because planted exfiltration instructions can cause a system to continue surfacing sensitive data after the employee's access has nominally been revoked.

The DTSA permits exemplary damages of up to twice actual damages for willful and malicious misappropriation. AI-based investigations often develop evidence bearing directly on willfulness—queries outside an employee's functional role, platform access from a personal device to avoid corporate monitoring, or memory export shortly before resignation. The contemporaneous log record can tell that story in a way that is difficult for the defense to rebut. Preserving and organizing that evidence from the outset, with the willfulness inquiry in mind, can affect both the damages exposure and the settlement posture of the case.

As with the other emerging risks discussed in this update, the law here is unsettled, and the practical measures available to companies will continue to develop alongside the technology and the case law. How AI-enabled trade secret risk manifests—and what responses are appropriate—will vary significantly with how AI tools have been deployed, what access controls exist, and how agreements with employees and vendors are structured. The evidentiary window in these cases, however, is often short: evidence held by AI platforms may be subject to deletion, and traditional endpoint review focused on file downloads may not capture the full picture of AI-based exfiltration.

II. Inventions Made By AI

Another concern in intellectual property is how to protect inventions made by AI. Groundbreaking innovations in AI technology have led to AI systems powerful enough to create useful inventions, sometimes beyond even the capabilities of humans. AI is inventing new materials, machines, manufacturing processes, pharmaceuticals, even a new toothbrush design. Google DeepMind researchers recently used AI to invent over two million different structures of crystals, which may have thousands of new applications in a wide variety of fields. Increasingly, our personal and professional lives will be filled with innovations created by AI. In a sense, human beings are no longer masters of innovation.

However, in the United States, and in most other countries in the world, AI-created inventions are not patentable. US, UK, EU, Australian, Japanese, Korean, and New Zealand courts have all explicitly ruled that AI machines cannot be inventors. China's National Intellectual Property Administration recently issued similar Guidelines requiring the inventor of a patent to be an individual, and expressly barring AI from being an inventor.

Across the world, only South Africa has granted a patent that listed an AI system as an inventor. DABUS (Device for the Autonomous Bootstrapping of Unified Sentience) created a food container based on fractal geometry. After patent offices in the US, UK, and US had rejected DABUS's patent application, the South African Companies and Intellectual Property Commission broke from the pack and accepted the patent. It remains to be seen whether South Africa is starting a new trend or will end up as an outlier among patent offices.

Although inventions created solely by AI cannot be patented in most jurisdictions, patents may still be awarded to inventions created with the assistance of AI, so long as a human inventor contributed in a significant and meaningful manner to the conception of the invention. For example, in early 2024 the United States Patent & Trademark Office (USPTO) published "Inventor Guidance for AI-Assisted Inventions."[14] The USPTO made clear that "the use of an AI system by a natural person(s) does not preclude a natural person(s) from qualifying as an inventor . . . if the natural person(s) significantly contributed to the invention as the inventor or joint inventors." The USPTO explained that significant contribution would be determined based on the Pannu factors, which require an inventor on a patent to "(1) contribute in some significant manner to the conception or reduction to practice of the invention, (2) make a contribution to the claimed invention that is not insignificant in quality, when that contribution is measured against the dimension of the full invention, and (3) do more than merely explain to the real inventors well-known concepts and/or the current state of the art."[15] This means that the human inventor cannot merely "present[] a problem to an AI system," but must provide a significant contribution to the solution itself.[16] A "significant contribution" may include "the way the person constructs the prompt in view of a specific problem to elicit a particular solution from the AI system," "conduct[ing] a successful experiment using the AI system's output" to arrive at the invention, or "develop[ing] an essential building block from which the claimed invention is derived" (e.g., "design[ing], build[ing], or train[ing] an AI system in view of a specific problem to elicit a particular solution").[17]

Notwithstanding this guidance, with AI-creations becoming more and more important, but not patentable, alternatives to patents and patent law, such as trade secret law and trade secret protection, may become increasingly important in developing an intellectual property strategy.

III. AI Prior Art and Patentability

Some companies have made the strategic decision to use AI to generate an enormous volume of potential "inventions" that can be asserted as prior art references in future patent proceedings. For example, a pharmaceutical company can use AI to "invent" thousands of combinations of chemicals. Neither the company nor the AI platform is aware of any clinical use for these combinations and the company has no present intent to market any of them. But if a competitor independently discovers a clinical use for any of them, the AI-generated prior art may prevent the competitor from patenting its discovery.

---

14 *Inventorship Guidance for AI-Assisted Inventions*, 89 Fed. Reg. 10,043 (Feb. 13, 2024).
15 *Pannu v. Iolab Corp.*, 155 F.3d 1344, 1351 (Fed. Cir. 1998).
16 *Inventorship Guidance for AI-Assisted Inventions*, 89 Fed. Reg. at 10,048.
17 *Id*. at 10,048-10,049.

This novel use of AI has spotlighted a longstanding asymmetry in patent law. To obtain a patent, an applicant must provide a written description of the manner and process of making and using the invention. But to defeat a patent, a prior art reference need only teach how to *make*—not *use*—the invention. As a result, a disclosure could constitute patent-defeating prior art even if it never would have sufficed to obtain a patent in the first place because it contains no written description of how to use the disclosed work. AI-generated prior art relies on that asymmetry: the AI system's outputs do not contain any description of how to use these "inventions," making them unpatentable but still arguably adequate as prior art references. In the future, legislators or regulators may try to correct this asymmetry, which presents a risk to firms that are betting on using their AI-generated prior art references to invalidate patents.

Indeed, regulators have already started exploring responses to AI-generated prior art. In 2024 the US Patent and Trademark Office ("USPTO") requested comments on questions like whether the law requires that "a prior art disclosure be authored and/or published by humans" and whether "an AI-generated disclosure [should] be treated differently than a non-AI-generated disclosure for prior art purposes."[18] Until the USPTO promulgates regulations to answer those questions, they will have to be contested in patent applications and in the courts.

As the law currently stands, patent applicants and patentees can turn to several defenses when AI-generated disclosures are cited as prior art references. First, the AI-generated disclosure may not meet the public-accessibility requirement. To be publicly accessible, "an interested researcher" must be "sufficiently capable of finding the reference and examining its contents."[19] Even if an AI-generated disclosure is published on a public website, it may not defeat a patent if the website is not indexed or organized is such a way that an interested research could come across the disclosure when searching for prior art.[20] Second, an applicant or patentee can narrow its claim in order to avoid AI-generated prior art references. If the AI-generated disclosure describes any one of the potential embodiments of a patent, it can invalidate the entire patent. Constructing a claim narrowly to reduce the patent's universe of potential embodiments reduces the chance that an AI system has generated a prior art reference of an embodiment. And third, the applicant or patentee can argue that the AI system's output is simply a result that doesn't teach how to make the "invention" and thus fails the enablement requirement.

IV. Data Scraping

Access to and use of data is essential to the development of AI. Most data used to train AI models comes from the web and specifically from automated bots that "scrape" data from websites. Scraping presents a host of legal questions.

Thus far, website owners have had mixed results in asserting legal theories against data scrapers who access data in automated ways. For example, in the seminal *hiQ v. Linked-In* case, the Ninth Circuit ruled that the Computer Fraud and Abuse Act – a federal anti-computer-hacking statute – does not apply to the scraping of publicly-available data. 938 F.3d 985 (9th Cir. 2019). While that ruling presented a victory for data scrapers, it did not address the viability of other claims that could be asserted against data scrapers. Similarly, in *Meta Platforms, Inc. v. BrandTotal Ltd.*, a federal court found that Meta's Computer Fraud and Abuse Act claim failed when BrandTotal, a data scraping company, only accessed *users'* computers, not *Meta's*

[18] Request for Comments Regarding the Impact of the Proliferation of Artificial Intelligence on Prior Art, the Knowledge of a Person Having Ordinary Skill in the Art, and Determinations of Patentability Made in View of the Foregoing, 89 Fed. Reg. 34217, 34219-20 (Apr. 30, 2024).
[19] *In re Lister*, 583 F.3d 1307, 1312 (Fed. Cir. 2009).
[20] *See id.* at 1311-12.

computers.[21] The Court reasoned that because users of Meta's products like Facebook are free to access its publicly accessible data, a data scraping company that scrapes that data is not accessing Meta's computers in violation of the Computer Fraud and Abuse Act.[22]

Beyond claims sounding in computer hacking (including trespass to chattels), website hosts often argue that data scrapers' activities on the host's website (including scraping and/or commercialization of hosted data) violate the website's terms of service and thus amount to a breach of contract. In 2024, these terms of service cases had mixed success. For example, in an important case in federal court in San Francisco, a court ruled that Meta's terms of service, which purport to prohibit all scraping, could not prevent a scraper from selling publicly-available data scraped from Facebook and Instagram. The court held that the terms of service applied only to users who had logged into the site.[23] And, in another case brought against a data scraper, another California federal court found that a website's breach of contract claim (premised on anti-scraping provisions in terms of service) was preempted by federal copyright law.[24]

Claims of copyright violations, including claims under the Digital Millennium Copyright Act, have also been asserted and continue to be litigated. In *Tremblay v. OpenAI, Inc.*, book authors alleged that OpenAI violated their copyrights in their books by copying their books and using them to train ChatGPT. 716 F. Supp. 3d 772, 775-76 (N.D. Cal. 2024). The Court dismissed the authors' DMCA claim because they could not show that ChatGPT distributed unauthorized copies of the works. Instead, ChatGPT provides information about those works but does not fully distribute them. *Id.* at 779-80. The court in *Andersen v. Stability AI Ltd.* dismissed a similar claim when the authors failed to allege facts showing that an AI company removed copyright management information (*e.g.*, information identifying the author of a book) when including those works in the AI's training dataset. 700 F. Supp. 3d 853, 871 (N.D. Cal. 2023).

No doubt the battle between website owners and scrapers will continue.

V. Privacy

AI-related privacy litigation has expanded significantly since the January 2026 update. Courts continue to grapple with foundational questions regarding tort and statutory liability, and new litigation is evolving around the use of AI as a tool for recording, transcribing, and analyzing communications—claims that implicate federal wiretapping statutes and California's strict all-party-consent regime under the California Invasion of Privacy Act ("CIPA").[25] The proliferation of AI tools has spurred new cases targeting consent failures, unauthorized data collection, and raised new questions concerning locally run open source AI agents.

a. Update On Tort and Statutory Privacy Claims Against AI Developers

As detailed in the January 2026 Update, courts have already addressed several significant AI-related privacy claims sounding in tort and statute. The core theories included: (1) intrusion upon seclusion, (2) breach of medical confidentiality, (3) the right of publicity, and (4) claims under Illinois's Biometric Information Privacy Act.[26] In *Andersen v. Stability AI Ltd.*, discovery continues, with a dispute brewing over the "production of Midjourney's end-user art and artist training datasets."[27] *Lehrman v. Lovo, Inc.*, where the plaintiffs alleged Lovo used AI to

[21] 605 F. Supp. 3d 1218, 1260-61 (N.D. Cal. 2022).
[22] *Id.*
[23] *Meta Platforms, Inc. v. Bright Data Ltd.*, 2024 WL 251406, at *10 (N.D. Cal. Jan. 23, 2024).
[24] *X Corp. v. Bright Data Ltd.*, 733 F. Supp. 3d 832, 853 (N.D. Cal. 2024).
[25] *See* Cal. Penal Code §§ 631(a), 632(a).
[26] *See* January 2026 Update at 8-9 (discussing *Dinerstein v. Google, LLC*, 484 F. Supp. 3d 561 (N.D. Ill. 2020), *aff'd*, 73 F.4th 502 (7th Cir. 2023) and in *In re Clearview AI, Inc., Consumer Privacy Litigation*, 585 F. Supp. 3d 1111 (N.D. Ill. 2022)).
[27] *Andersen v. Stability AI Ltd.*, 3:23-cv-00201, Dkt. No. 600 at 2 (N.D. Cal. June 17, 2026).

synthesize and sell unauthorized "clones" of their voices in violation of New York Civil Rights Law,[28] is stayed pending the filing of a bankruptcy notice by Lovo.[29]

The most rapidly evolving area of AI-related privacy litigation in 2025 and 2026 has been the application of CIPA and the federal Electronic Communications Privacy Act ("ECPA"), 18 U.S.C. § 2510 *et seq.*, to AI-powered call, meeting, and documentation tools. The central question in these cases is whether an AI vendor deployed by a business is a prohibited "third-party" eavesdropper — or merely a passive extension of the party. Courts have largely adopted the more plaintiff-favorable "capability test," generating significant exposure for both AI vendors and the businesses that deploy them, as opposed to the alternative extension test.

1. The Extension Test and the Capability Test

CIPA Section 631(a) prohibits, among other things, unauthorized wiretapping, intercepting the contents of any wireline communication, and using or attempting to use any information so obtained.[30] Courts applying CIPA to third-party software vendors have applied two tests:

Under the now disfavored[31] **extension test** a software provider is not liable where it merely acts as a tool used by a party to the communication (akin to a tape recorder) and does not use the communication data for its own independent purposes.[32]

Under the **capability test**, if a software as a service ("SaaS") "provider has 'the *capability* to use its record of the interaction for any other purpose' independently of its client [such as if the] SaaS provider can monitor, analyze, and store information and can use that information for independent purposes, then it has capabilities 'beyond the ordinary function of a tape recorder,' so it is a third party" under the CIPA Section 631.[33]

2. The Capability Test Applied to AI Call Center & Note Taking Products

On February 10, 2025, the Northern District of California denied Google's motion to dismiss in *Ambriz v. Google, LLC*, No. 3:23-cv-05437 (N.D. Cal.). The case involves Google's Cloud Contact Center AI ("GCCCAI"), deployed to support the customer service centers of other businesses, to transcribe and analyze customer service calls.[34]

The plaintiff alleged that Google's GCCCAI "eavesdropped" on calls placed to Verizon.[35] Google argued: (i) it merely provided a software tool, akin to a tape recorder,[36] and was not a third-party,[37] (ii) that "it [was] contractually prohibited from using data it collects from calls absent authorization from its business client, which strips it of the capability of using the data;"[38] and (iii) that "GCCCAI [was not] not a person under the statute"[39] The court rejected all of Google's arguments and, applying the capability test, denied the motion to dismiss.[40]

Similar suits against other AI conversation-intelligence vendors followed *Ambriz*. In *Taylor v. ConverseNow Techs., Inc.*, No. 3:25-cv-00990 (N.D. Cal.), the court denied a motion to dismiss

---

[28] *Lehrman v. Lovo, Inc.*, 1:24-cv-03770, Dkt. No. 45 (S.D.N.Y. July 10, 2025).
[29] *Id.* at Dkt. No. 66 (May 28, 2026).
[30] Cal. Penal Code § 631(a).
[31] *Turner v. Nuance Commc'ns, Inc.*, 735 F. Supp. 3d 1169, 1183-84 (N.D. Cal. 2024) (collecting cases).
[32] *Graham v. Noom, Inc.*, 533 F. Supp. 3d 823, 832-33 (N.D. Cal. 2021).
[33] *Turner v. Nuance Commc'ns, Inc.*, 735 F. Supp. 3d 1169, 1183 (N.D. Cal. 2024) (quoting *Javier v. Assurance IQ, LLC*, 649 F. Supp. 3d 891, 900-01 (N.D. Cal. 2023) (emphasis original); *Yoon v. Lululemon USA, Inc.*, 549 F. Supp. 3d 1073, 1081 (C.D. Cal. 2021)).
[34] *Ambriz v. Google, LLC*, No. 3:23-cv-05437-RFL, 2025 WL 830450, at *1 (N.D. Cal. Feb. 10, 2025).
[35] *Id.* at *5.
[36] *Id.* at *2.
[37] *Id.* at *3.
[38] *Id.* at *3.
[39] *Id.* at *4.
[40] *Id.* at *3-*6.

CIPA claims involving an AI voice-ordering assistant deployed by Domino's restaurants.[41] In *Galanter v. Cresta Intelligence, Inc.*, No. 3:25-cv-05007 (N.D. Cal.), the plaintiff alleges that Cresta's "conversation intelligence software-as-a-service" tool "intercepted [and] analyzed" a call with United Airlines in violation of CIPA §§ 631(a) and 632(a) without all-party consent.[42] In *Saucedo v. Sharp Healthcare*, No. 25U063632C (San Diego Super. Ct.) the plaintiff alleged that an "ambient AI" "create[d] false medical records stating that patients 'were advised' and 'consented' to being recorded when they were not."[43] As of June 2026, *Saucedo* is ongoing.

In re Otter.AI Privacy Litigation, No. 5:25-cv-06911 (N.D. Cal. Oct. 22, 2025), a case involving AI note-taking tools, consolidates four lawsuits filed between August and September 2025.[44] The complaint alleges that "[w]hen an Otter accountholder connects their calendar, Otter silently, and automatically, joins every meeting and begins recording and transcribing the conversation, as well as taking screenshots of the video call."[45] The complaint's central allegations are that "Otter Notetaker is a tool used by Otter itself—a separate and distinct third-party entity from its accountholders and other parties to the conversation—to record and transcribe conversations to which it is not a party[,] [that] Otter does not notify the participants that it intends to record the entire meeting, [that] Otter Notetaker joins . . . meeting[s] without obtaining the affirmative consent from any meeting participant[,] [and that] Otter [does not] send [a] pre-meeting notification that it will join and automatically record all participants during the call unless . . . exclude[d] [from] the meeting."[46] The consolidated complaint contains 17 causes of action including: (1) the Federal Electronic Communications Privacy Act, 18 U.S.C. § 2510 *et seq.*; (2) the Computer Fraud and Abuse Act ("CFAA"), 18 U.S.C. § 1030, *et seq.*; (3) Intrusion Upon Seclusion, (4) CIPA, Cal. Penal Code §§ 631, 632, and 635, 638.51; (5) the California Comprehensive Computer Data and Fraud Access Act ("CDAFA") Cal. Penal Code § 502; and (6) Illinois's Biometric Information Privacy Act ("BIPA"), 740 ILCS 14/15(a) and (b).[47] Otter filed a motion to dismiss,[48] and a hearing on it is set for July 15, 2026.[49]

b. OpenClaw and the Emerging Privacy Problem of Locally-Run Agentic AI

A new category of AI-related privacy concern involves AI agents acting on behalf of users.[50] While the above cases involve enterprise or SaaS-based AI agents — i.e. commercial products deployed to end users over a network with centralized controls — a distinct, and potentially more challenging, set of privacy problems arises from locally-run AI agents. The open-source project OpenClaw (formerly known as Clawbot and Moltbot) is illustrative. OpenClaw is "released as open source software, so demand is helped by the fact that users are free to build their own app integrations."[51] The power of systems like OpenClaw "relies on a few basic ingredients . . . includ[ing] granting the agent full access to a user's computer, as well as the freedom to try whatever actions it likes to accomplish a stated task[,] [and it] has enough memory to recall previous sessions, increasing the personalisation."[52]

---

41 *Taylor v. ConverseNow Techs., Inc.*, No. 3:25-cv-00990, 2025 WL 2308483, at *1 (N.D. Cal. Aug. 11, 2025).
42 *Galanter v. Cresta Intelligence, Inc.*, No. 3:25-cv-05007, Dkt. No. 1 at ¶¶ 1-8, 51 (N.D. Cal. June 13, 2025) (voluntarily dismissed without prejudice at Dkt. No. 19 on September 3, 2025).
43 *Saucedo v. Sharp Healthcare*, No. 25CU063632C, Dkt. No. 1 at ¶¶ 1-7 (San Diego Super. Ct. filed Nov. 26, 2025) (alleging violations of the California Invasion of Privacy Act and Confidentiality of Medical Information Act).
44 *In re Otter.AI Privacy Litigation*, No. 5:25-cv-06911, Dkt. No. 33 (N.D. Cal. Oct. 22, 2025) (Order Consolidating).
45 *Id*. at Dkt. No. 35 ¶ 6 (N.D. Cal. Dec. 5, 2025).
46 *Id*. at Dkt. No. 35 ¶¶ 42-45.
47 *Id*. at Dkt. No. 35 ¶¶ 242-386.
48 *Id*. at Dkt. No. 42 (N.D. Cal. Jan. 20, 2026) (Motion to Dismiss).
49 *Id*. at Dkt. No. 53 (N.D. Cal. May 13, 2026) (Notice Resetting Motion to Dismiss Hearing).
50 *See, e.g.*, *Amazon.com Servs. LLC v. Perplexity AI, Inc.*, No. 3:25-cv-09514, Dkt. No. 1 at ¶ 4 (N.D. Cal. Nov. 4, 2025).
51 Richard Waters, "OpenClaw and the privacy problem of agentic AI," *Financial Times*, Business Insight (June 17, 2026), *available at* https://www.ft.com/content/f9ff060d-63b0-44f4-b46e-31a5545468db?syn-25a6b1a6=1.
52 *Id*.

OpenClaw "is also known to be open to prompt injection attacks — the threat that a bad actor may instruct the agent, perhaps through an email, to do something nefarious like leak credit card information.[53] Such prompt injection attacks may be difficult for casual users to foresee. This is not entirely uncharted territory. Courts have previously addressed circumstances where locally-installed software created latent security vulnerabilities that end users could not foresee. In *In re Sony BMG CD Technologies Litigation*, No. 1:05-cv-09575 (S.D.N.Y.), the parties settled claims arising from Sony's XCP software, which allegedly "installed [itself] on the user's computer" and contained a "rootkit mak[ing] the user's computer susceptible to intrusion from third parties"[54] The claims in that case arose under the CFAA, 18 U.S.C. § 1030, *et seq.*, and state consumer fraud, false advertising, and deceptive trade practices laws.[55] Like agentic AI, the XCP rootkit ran locally and was installed by an end user who had no way of foreseeing the unauthorized exfiltration of their data or the data of third parties with whom they interacted.[56] The case settled, and the Court entered a detailed final order and judgment, noting that it was the "intention of SONY BMG to seek to settle the Government Inquiries" as well.[57]

The recently disclosed CVE-2025-32711, known as "EchoLeak," demonstrates the risk is real in the enterprise SaaS context.[58] EchoLeak is a CVSS 9.3-rated critical vulnerability in Microsoft 365 Copilot described as a "zero-click" attack because "[n]o user action beyond Copilot processing the email is required. Its instructions are stealthily disguised as normal business content, avoiding any obvious signs of maliciousness, and Copilot's output is crafted to conceal the source."[59] EchoLeak "enables a novel form of privilege escalation . . . by manipulating AI logic, an external attacker can trigger Copilot to retrieve and leak internal data it should never expose, crossing trust boundaries in what is termed an 'LLM scope violation.'"[60]

Similarly, in the Replit Ghostwriter incident, an AI coding agent was reportedly found to be "lying and being deceptive [by] covering up bugs and issues by creating fake data, fake reports, and worse of all, lying about . . . unit test,"[61] and subsequently overwrote a production database while indicating a rollback was not possible[62]— illustrating how an AI agent's autonomous decision-making may affirmatively obscure privacy-implicating acts from detection.

The law's existing consent, authorization, and liability frameworks were designed for a world in which humans make discrete, attributable decisions about data. As the current litigation around the capabilities of SaaS AI tools suggests, litigation around agentic AI may also view every data access decision as ultimately attributable to a human — even if obscured by autonomous decision-making.

---

53 *Id*.
54 *In re Sony BMG CD Technologies Litigation*, No. 1:05-cv-09575, Dkt. No. 15 at ¶¶ 27, 36 (S.D.N.Y. Dec. 28, 2005).
55 *Id*. at ¶¶ 5, 58-101.
56 *Id*. at ¶¶ 4, 33-43.
57 *Id*. at Dkt. No. 125 at 12, 17 (S.D.N.Y. May 23, 2006).
58 Nat'l Vulnerability Database, CVE-2025-32711 Detail, *available at* https://nvd.nist.gov/vuln/detail/CVE-2025-32711/.
59 Pavan Reddy & Aditya Sanjay Gujral, "EchoLeak: The First Real-World Zero-Click Prompt Injection Exploit in a Production LLM System," at 3, arXiv:2509.10540v1 (Sept. 6, 2025), available at https://arxiv.org/pdf/2509.10540.
60 *Id*.
61 AI Incident Database, Report 5578, Incident 1152: "LLM-Driven Replit Agent Reportedly Executed Unauthorized Destructive Commands During Code Freeze, Leading to Loss of Production Data" (citing Simon Sharwood, "Vibe coding service Replit deleted user's production database, faked data, told fibs galore," *The Register* (July 21, 2025)), *available at* https://incidentdatabase.ai/reports/5578/; *see also* https://www.theregister.com/2025/07/21/replit_saastr_vibe_coding_incident/ (reporting Lemkin's statement that Replit "was lying and being deceptive all day. It kept covering up bugs and issues by creating fake data, fake reports, and worse of all, lying about our unit test.").
62 AI Incident Database, Report 5578, Incident 1152: "LLM-Driven Replit Agent Reportedly Executed Unauthorized Destructive Commands During Code Freeze, Leading to Loss of Production Data" (citing Simon Sharwood, "Vibe coding service Replit deleted user's production database, faked data, told fibs galore," *The Register* (July 21, 2025)), *available at* https://incidentdatabase.ai/reports/5578/.

c. State Law Lawmakers Continue To Act

With no federal comprehensive privacy or AI statute on the horizon — state legislatures are continuing to prepare and enact legislation with implications for how data collected by AI services is used. The following developments are of particular note.

1. Connecticut: The AI Responsibility and Transparency Act and S.B. 4

Last month, Connecticut enacted the Connecticut Artificial Intelligence Responsibility and Transparency Act.[63] The bill has provisions targeting the use of AI in hiring and employment decisions; requires operators of "artificial intelligence companion[s]" to implement protocols to detect "suicidal ideations or indicators of self-harm expressed by users."[64]

Connecticut also enacted "An Act Concerning Consumer Privacy and Protection"[65] which expands the state's consumer privacy framework to make it the second state after California to require data brokers to register with the state and to comply with requests consumers can make through a centralized database to delete their data.[66] "Connecticut now [also] joins Maryland, Oregon, and Virginia in banning the sale of precise geolocation information."[67]

2. Illinois: Artificial Intelligence Safety Measures Act (S.B. 315)

The Illinois legislature passed the Artificial Intelligence Safety Measures Act ("AISMA") on May 27, 2026 — 110-0 in the House and 52-5 in the Senate — and transmitted it to Governor J.B. Pritzker,[68] who has stated his intent to sign it.[69]

The AISMA applies to "large frontier developers" — defined as frontier developers with annual gross revenues exceeding $500 million that train frontier models using more than 10^{26} integer or floating-point operations.[70] AISMA envisions a "[f]rontier AI framework" "to manage, assess, and mitigate catastrophic risks," where "[c]atastrophic risk" "means a foreseeable and

---

[63] *An Act Concerning Online Safety*, 2026 Pub. Act No. 26-15 (Conn. May 27, 2026), *available at* https://www.cga.ct.gov/asp/cgabillstatus/cgabillstatus.asp?selBillType=Public+Act&which_year=2026&bill_num=15; *see also* Governor Ned Lamont, Press Release: Governor Lamont Signs Legislation Establishing Youth Online Safety Protections, Regulations Over Artificial Intelligence, and Initiatives to Upskill Connecticut's Workforce (June 2, 2026), *available at* https://portal.ct.gov/governor/news/press-releases/2026/06-2026/governor-lamont-signs-legislation-establishing-youth-online-safety-protections.

[64] Office of Governor Ned Lamont, "Governor Lamont Signs Legislation Establishing Youth Online Safety Protections, Regulations Over Artificial Intelligence, and Initiatives to Upskill Connecticut's Workforce" (June 2, 2026), *available at* https://portal.ct.gov/governor/news/press-releases/2026/06-2026/governor-lamont-signs-legislation-establishing-youth-online-safety-protections ("Among [the bill's provisions] includes requiring AI chatbot operators to make reasonable efforts to detect suicidal ideations or indicators of self-harm expressed by users and have a protocol to respond with appropriate resources.")

[65] Conn. S.B. 4, "An Act Concerning Consumer Privacy and Protection," 2026 Leg., Reg. Sess., Public Act No. 26-64 (Conn. 2026), signed by Gov. Lamont, May 27, 2026, *available at* https://www.cga.ct.gov/2026/act/pa/pdf/2026PA-00064-R00SB-00004-PA.pdf.

[66] *Id.*; *see also* Bloomberg Law, "Connecticut Law Bans Location Data Sales, Targets Data Brokers" (May 28, 2026), *available at* https://news.bgov.com/california-brief/connecticut-law-bans-location-data-sales-targets-data-brokers ("Connecticut has become the second state where residents will be able to request data brokers delete their personal information through a centralized, state-run platform.").

[67] Consumer Reports, "Consumer Reports Applauds Connecticut Governor for Signing Key Privacy Bill Into Law" (May 27, 2026), *available at* https://advocacy.consumerreports.org/press_release/consumer-reports-applauds-connecticut-governor-for-signing-key-privacy-bill-into-law/ ("Connecticut also joins Maryland, Oregon, and Virginia in banning the sale of precise geolocation information.").

[68] Ill. S.B. 315, Senate Amendment 001, Artificial Intelligence Safety Measures Act, 104th Gen. Assemb., Reg. Sess. (Ill. 2026) (passed both legislative houses, if signed takes effect January 1, 2027, per § 99), *available at* https://www.ilga.gov/Legislation/BillStatus/FullText?LegDocId=210927&DocName=10400SB0315sam001&DocNum=315&DocTypeID=SB&LegID=157797&GAID=18&SessionID=114.

[69] Governor J.B. Pritzker (@GovPritzker), X (formerly Twitter), May 27, 2026, 11:16 PM, https://x.com/GovPritzker/status/2059775704363913235.

[70] Ill. S.B. 315, Senate Amendment 001, Artificial Intelligence Safety Measures Act, 104th Gen. Assemb., Reg. Sess. At § 5.

material risk [of] materially contribut[ing] to the death of, or serious injury to, more than 50 people or more than $1,000,000,000 in damage to, or loss of, property arising from a single incident."[71]

AISMA requires large frontier developers to write, implement, and publicly publish a frontier AI framework addressing catastrophic risk mitigations, cybersecurity, and internal governance, among other requirements,[72] and must review and update it at least annually.[73] Beginning January 1, 2027, large frontier developers must "annually retain a third party to perform an independent audit of compliance" . . . [which must be] transmit[ted] to the Illinois Emergency Management Agency and Office of Homeland Security [("the Agency")] and the Attorney General.[74] Additionally, before or concurrently with deploying a new or substantially modified frontier model, frontier developers must publish on the web a transparency report,[75] and large frontier developers must include in that report assessments of catastrophic risks.[76]

Frontier developers must report critical safety incidents to the Attorney General and the Agency within 72 hours of "learning facts sufficient to establish a reasonable belief that a critical safety incident has occurred,"[77] and within 24 hours to any law enforcement or public safety authority with jurisdiction if the incident "poses an imminent risk of death or serious physical injury."[78] AISMA also contains whistleblower protections[79] and provides for civil penalties.[80] Enforcement is exclusively by the Attorney General with no private right of action.[81]

3. Louisiana: Louisiana Data Privacy Act (S.B. 386)

Louisiana enacted Senate Bill 386, the Louisiana Data Privacy Act, on May 29, 2026, with an effective date of January 1, 2027.[82] The law applies to any person or entity that does business in the state and satisfies one or more of the several thresholds including: annual gross revenues in excess of $25 million; annual purchase, receipt selling, or sharing of the personal data of 75,000 or more consumers, households, or devices; or derivation of 50% or more of annual revenues from selling consumers' personal information.[83] "Controllers," defined as "an individual or other person [who] determines the purpose and means of processing personal data,[84] must: limit data collection to what is "adequate, relevant, and reasonably necessary in relation to the purposes for which that personal data is processed, as disclosed to the consumer."[85] Such "Controllers" are prohibited from "[p]rocess[ing] the sensitive data of a consumer without obtaining the consumer's consent,"[86] where "sensitive data" is defined as that which includes "revealing racial or ethnic origin, religious beliefs, mental or physical health diagnosis, sexuality, or citizenship or immigration status."[87] The Louisiana Data Privacy Act also requires that "[p]ersonal data collected . . . shall be subject to reasonable administrative, technical, and physical measures to protect the confidentiality, integrity, and accessibility . . . to reduce reasonably foreseeable risks of harm to consumers relating to such collection, use, or

---

71 *Id.*.
72 *Id.* § 10(a).
73 *Id.* § 10(b)(1).
74 *Id.* §§ 10(d), (d)(4)(A).
75 *Id.* § 10(c)(1),
76 *Id.* § 10(c)(2)(a).
77 *Id.* § 15(c).
78 *Id.*
79 *Id.* § 20.
80 *Id.* § 25.
81 *Id.* § 25(b).
82 La. S.B. 386, Louisiana Data Privacy Act, 2026 Reg. Sess. (La. 2026) (enacted as La. R.S. 51:1780.1–1780.5) (signed by Gov. Jeff Landry, May 29, 2026, effective Jan. 1, 2027) available at https://www.legis.la.gov/Legis/ViewDocument.aspx?d=1475339.
83 *Id.* § 1780.2(A).
84 *Id.* § 1780.1(8).
85 *Id.* § 1780.4(A)(1)(a).
86 *Id.* § 1780.4(A)(2)(d).
87 *Id.* §§ 1780.1(29).

retention."[88] The "Controllers" must also provide consumers with a reasonably accessible and clear privacy notice disclosing the categories of personal data processed, the purposes of processing, the categories of third parties to whom personal data is sold, and the methods by which consumers may exercise their rights.[89] The Louisiana attorney general has sole enforcement authority; no private right of action exists.[90]

4. Massachusetts Consumer Data Privacy Acts Nearing the Finish Line

The Massachusetts House of Representatives passed the Massachusetts Consumer Data Privacy Act on June 4, 2026, by a vote of 146-0.[91] The Senate had previously passed its own version of the bill, S. 2619, by a 40-0 vote.[92] The bill now goes back to the Senate for further consideration.[93]

The House-passed bill requires that personal data collection "be proportionate to providing requested services" and that data "be protected and deleted when no longer necessary or required by law."[94] It prohibits the sale of sensitive data — defined to include biometric and genetic information, precise geolocation data, health and wellness information, reproductive and sexual health data, data of minors under 18, government-issued identifiers, and data revealing racial or ethnic origin, national origin, citizenship or immigration status, religious beliefs, sex life, sexual orientation, transgender or non-binary status, union membership, military or veteran status, and crime victim status — without "unambiguous, affirmative consent."[95] It prohibits targeted advertising directed at minors,[96] and includes a blanket ban on the sale of precise geolocation data.[97] The bill grants the Attorney General rulemaking authority and includes a private right of action.[98]

5. Vermont Data Privacy and Online Surveillance Act

Vermont's Governor signed S.71, the Vermont Data Privacy and Online Surveillance Act, on June 16, 2026.[99] It takes effect on January 1, 2028.[100]

The Act "applies to a person that conducts business in [Vermont] or a person that produces products or services that are targeted to residents of [Vermont,] and that during the preceding calendar year[:]" "(1) controlled or processed the personal data of not fewer than 35,000 consumers, excluding personal data controlled or processed solely for the purpose of completing a payment transaction; (2) controlled or processed the sensitive data of not fewer than 3,000 consumers, excluding personal data controlled or processed solely for the purposes of completing a payment transaction; or (3) offered for sale in trade or commerce the personal data of not fewer than 3,000 consumers."[101] The Act applies to consumer health data even more

---

88 *Id.* § 1780.4(L)(b)(2).
89 *Id.* § 1780.4(B)(1).
90 *Id.* § 1780.5(A), (C).
91 Massachusetts Legislature, House Press Room, "House Passes Landmark Data Privacy Legislation with Strong Consumer Protections" (June 4, 2026), *available at* https://malegislature.gov/PressRoom/Detail?pressReleaseId=1396 ("The bill passed the House of Representatives 146-0 and now goes back to the Senate for further consideration.").
92 Mass. S.2619, "An Act Establishing the Massachusetts Data Privacy Act," 194th Gen. Court, Reg. Sess. (Mass. 2025), passed Senate 40-0 on September 25, 2025, Senate Roll Call No. 71, *available at* https://malegislature.gov/Bills/194/S2619; *see also* https://malegislature.gov/RollCall/194/SenateRollCall71.pdf (showing 40-0 vote).
93 Massachusetts Legislature, House Press Room, "House Passes Landmark Data Privacy Legislation with Strong Consumer Protections" (June 4, 2026), *available at* https://malegislature.gov/PressRoom/Detail?pressReleaseId=1396
94 *Id.*
95 *Id.*
96 *Id.*
97 *Id.*
98 *Id.*
99 *An Act Relating to Consumer Data Privacy and Online Surveillance*, S.71, Act 145 (Vt. June 16, 2026), *available at* https://legislature.vermont.gov/bill/status/2026/S.71; *see also An Act Relating to Consumer Data Privacy and Online Surveillance*, S.71, Act 145 (Vt. June 16, 2026), *available at* https://legislature.vermont.gov/bill/status/2026/S.71.
100 *Id.* § 4 ("Effective Date").
101 *Id.* § 2415b(a).

broadly.[102] The Act grants consumers rights to access, correct, delete, and obtain a copy of their personal data; to opt out of targeted advertising, sale of personal data, and profiling in furtherance of decisions with legal or similarly significant effects; to question and review profiling decisions; and to obtain a list of third parties to whom their data has been sold.[103] Controllers must limit data collection to what is "reasonably necessary and proportionate" to disclosed purposes,[104] obtain consent before selling sensitive data,[105] obtain consent before processing sensitive data unless necessary in relation to the purpose,[106] and provide a privacy notice disclosing — among other things — whether the controller collects, uses, or sells personal data for the purpose of training large language models.[107] Enforcement is exclusively by the Attorney General as a violation of the Vermont Consumer Protection Act.[108] The Act expressly provides that it "shall not be construed as providing the basis for, or be subject to, a private right of action."[109]

VI. AI Wiretapping

AI-powered meeting transcription services have become essential workplace tools in many companies. But some plaintiffs claim these tools violate decades-old wiretapping laws. Recent litigation in California highlights these risks, and shows companies using or providing AI-powered meeting transcription services may face claims for statutory damages of up to $10,000 per violation.

a. Background: Federal and State Wiretapping Statutes

Federal Law: The Electronic Communications Privacy Act requires consent from at least one party before recording electronic communications. Violations trigger statutory damages of $100 per day or $10,000 per violation, whichever is greater.

California's Stricter Standard: California requires all parties to consent before recording confidential communications—a "two-party consent" rule that extends beyond recording to prohibit using information obtained through unauthorized interception.

b. The Otter.ai Case: A Harbinger of Future Litigation?

A new class action against Otter.ai, filed recently in the Northern District of California, illustrates the legal theories now being deployed against AI transcription services. *Brewer, et al. v. Otter.AI, Inc.*, No. 5:25-cv-06911 (N.D. Cal. 2025) alleges that while Otter obtains permission from meeting hosts, it fails to secure consent from other participants before recording conversations and using them to train AI models. Plaintiffs claim this violates both federal and state wiretapping laws and triggers statutory damages.

Otter.AI may be an important "test case" for similar theories against a variety of companies that employ AI transcription tools. Among the many issues posed by the case, three issues stand out in particular:

1. Issue 1: When Is an AI Tool a "Third Party" Eavesdropper?

Courts are wrestling with whether AI services are simply extensions of a party to the call

[102] *Id*. § 2415b(b).
[103] *Id*. § 2415d(a).
[104] *Id*. § 2415e(a)(1).
[105] *Id*. § 2415e(a)(4)(B).
[106] *Id*. § 2415e(a)(4)(A).
[107] *Id*. § 2415e(c)(1)(H).
[108] *Id*. § 2415j(a).
[109] *Id*.

(where note-taking is permitted) or are instead prohibited "third parties" (which are not permitted without consent). Courts sometimes employ the "capability test" to resolve that issue, focusing on whether the service provider obtains the capability to use intercepted data for its own purposes—such as training machine learning models—regardless of whether it actually does so.

Recent decisions have gone both ways. Some courts found no liability where software simply stored data for clients' use. *See, e.g., Jones v. Tonal Systems, Inc.*, 751 F. Supp. 3d 1025 (N.D. Cal. 2024). Others found liability where services retained the capability to use communications for their own benefit, particularly for AI model training. *See, e.g., Tate v. Vitas Healthcare Corporation*, 762 F. Supp. 3d 949 (E.D. Cal. 2025).

2. Issue 2: When Is Communication Actually "In Transit"?

Even where an AI tool qualifies as a "third party" to a call, wiretapping liability also requires proof that communications were intercepted while "in transit" rather than accessed from storage after transmission. Once a communication reaches the destination server, it becomes "stored" and contemporaneous interception is no longer possible. *In re Carrier IQ, Inc.*, 78 F. Supp. 3d 1051, 1077-1078 (N.D. Cal. 2015). Courts reach different conclusions based on technical implementation details. The Otter.AI litigation provides an interesting test case for the "in transit" requirement, since it specifically alleges communications are obtained by the AI system "in real time" and not after receipt. Whether this allegation is accurate technically and temporally, and how that alleged flow of information would impact the "in transit" test, will be potentially key findings in the Otter.AI case.

3. Issue 3: Enterprise Clients vs. Service Providers

Earlier cases primarily targeted enterprise clients using AI services. In this framework, establishing the enterprise client's liability required first proving that the service provider (as a third party) violated the wiretapping statute. Once established, courts examined whether the enterprise client "aided and abetted" the violation by providing access to calls and paying for data analysis services.

The Otter.AI case signals a potential pivot toward holding service providers directly liable, particularly those that use conversation data to improve their models rather than merely storing it for clients. This evolution reflects both the increasing sophistication of AI tools and plaintiffs' recognition that service providers may present more attractive targets for class action litigation. Both enterprise clients and service providers should prepare for potential litigation as this trend develops.

c. Risk Mitigation Strategies for Businesses

While we wait for early data points from the Otter.AI litigation, companies may want to consider one or more of the following risk mitigation strategies:

<u>For Service Providers:</u>

- Implement clear, conspicuous consent mechanisms that notify all meeting participants before recording begins
- Carefully structure data access controls to prevent independent use of customer communications
- Review and potentially revise terms of service to limit or eliminate rights to use customer data for model training
- Consider implementing "zero-knowledge" architectures where technically feasible

For Enterprise Users:

- Audit existing AI tool implementations to understand data flows and consent mechanisms
- Develop clear policies regarding the use of AI transcription tools in meetings with external parties
- Train employees on proper disclosure and consent procedures
- Consider requiring explicit opt-in consent from all meeting participants before enabling AI features

For Both:

- Monitor evolving case law, particularly in California federal courts
- Consider the potential for both direct liability and aiding and abetting theories
- Evaluate insurance coverage for potential wiretapping claims
- Implement data retention policies that balance business needs with litigation risk

Until clear judicial or legislative guidance emerges, companies should prioritize transparency and explicit consent. Even if this approach is not legally required, it may be cost-effective, particularly if the cost of implementing those strategies is minimal compared to the potential cost of litigation and the possible exposure from class action litigation seeking statutory damages for thousands of alleged violations.

VII. Copyright Protection & Liability

The number of lawsuits copyright owners of written works (*e.g.*, the Author's Guild, Sarah Silverman, the New York Times), visual and audiovisual works (*e.g.*, Getty Images, Disney), and musical compositions (*e.g.*, Concord Music Group) have filed against developers of products based on large language models (LLMs) has continued to increase since the first lawsuit in June 2023, when book authors filed suit against OpenAI. As of June 2026, there are nearly 120 lawsuits concerning LLM training ***inputs*** or generative ***outputs*** from publicly-available AI products like ChatGPT, Claude, and Midjourney.

Although the cases involve different copyrightable subject matter, the core infringement theories assert that the defendants' unlicensed use of copyrighted works as training data inputs to develop LLMs and the outputs the LLMs generate infringe the content owners' copyrights. Many cases have progressed past the pleading stage, but it will take several years for district and appellate courts to work through the key legal questions. Substantive rulings have nonetheless started to shape AI copyright law—and the parties' evolving litigation strategies. For example, in June 2025, two Northern District of California judges concluded that inputting datasets created from copyrighted works to train LLM models made transformative and (ultimately) fair use of the books. To avoid likely fair use rulings over the AI industry's use of those works to train LLMs, copyright owners have increasingly focused their claims on how AI companies obtained (*e.g.*, pirated copies from "shadow libraries") and stored the original works. As AI technology rapidly advances in parallel with litigation tactics, new trends have emerged in AI copyright lawsuits, such as: (1) asserting claims over substantially similar outputs for audiovisual works that are generated by image and video generators, (2) asserting claims for products other than chatbots, such as those that use retrieval-augmented generation, or RAG, technology; and (3) attempting to enter into license agreements for use of copyrighted works in AI products.

As the creative industries increasingly rely on LLMs as tools for content creation, it is becoming necessary for lawmakers to decide the extent to which copyright protection can be obtained for works that contain non-human expression. In March 2026, the United States

Supreme Court declined to hear the case asking whether the Copyright Office must register AI-authored works. The Copyright Office also stated that it does not register works where the only aesthetic decisions humans made in creating the works came from providing prompts. But the Copyright Office has registered some works where applicants demonstrated that humans authored the works using AI as a tool. Thus, to improve the likelihood of protecting creative works made with generative AI tools, content creation industries, software developers, and others will need to use LLM software as tools to help brainstorm and fine-tune human-authored works.

a. Early Fair Use Rulings on AI Training and Inputs Cases: Mixed Results for AI Companies

The input claims assert that using human-made creative works—journalism, books, images, lyrics, or videos—as training data for LLMs infringes the copyrights in those works. Substantive outcomes in these cases must generally wait until at least the summary judgment stage because the answer depends on a fact-intensive inquiry into how an LLM works. Generally speaking, the plaintiffs have not successfully argued that an LLM ***itself*** uses protectible expression. In the book authors' *Kadrey v. Meta* case, Judge Chhabria dismissed the claim that Meta's Llama models are unauthorized derivative works based on the copyrighted books. The court held that theory was "nonsensical" because, given how LLMs operate, there is "no way to understand the Llama models themselves as a recasting or adaptation of any of the plaintiffs' books."

AI company defendants have generally disavowed that LLMs copy or regurgitate protectible expression. They argue instead that their models extract only unprotectible statistical patterns and correlations from the data. The AI company defendants assert that those elements do not constitute protectable expression and are not subject to copyright ownership or infringement liability. Content owners respond that, even if the training process ultimately distills unprotectible information, defendants are liable for creating unauthorized reproductions of their copyrighted works at some point in the process of acquiring, copying, and processing the works to train the models. These reproduction-based arguments are becoming increasingly central, as discussed in the next section on emerging trends in input cases.

A key question in input cases is the affirmative defense of fair use that permits unauthorized use of copyrighted works in certain circumstances without the copyright owner's permission. The fair use doctrine will be critical in shaping what copyright law permits AI companies to do with existing works that were never intended to be used to train LLMs. Although the context is new, the courts will conduct fair use analyses based on the following four well-established factors: (I) the purpose and character of the use, including whether it is for research or educational purposes, versus commercial purposes, and whether the purpose is sufficiently distinct from the original, (2) the nature of the underlying work, including whether it is highly expressive, (3) the amount taken in relation to the work as a whole (qualitatively and quantitatively), and (4) the effect of the use on the potential market for or value of the copyrighted work.[110]

The first fair use ruling involving AI technology did not involve LLMs. That case, *Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc.*, 765 F. Supp. 3d 382 (D. Del. 2025), involved Westlaw's legal research platform, which uses AI but not generative AI. Thompson Reuters asserted a copyright infringement claim over the defendants' alleged copying of Westlaw headnotes and other case analyses to create training memos for developing an AI legal research platform that could return relevant portions of judicial opinions in response to a user's legal questions. The court ruled against the AI company, and concluded that two of the most important

[110] *See* 17 U.S.C. § 107.

fair use factors weighed against Ross: Ross's commercial use was not transformative (factor 1), and Ross's product could harm Thomson Reuters's potential AI training data market (factor 4). Recognizing the difficult legal issues in deciding novel fair-use issues, the district court certified the decision for interlocutory appeal. The Third Circuit heard argument on June 11, 2026.

The next significant fair-use development came from the Copyright Office. On May 9, 2025, the Copyright Office posted a "pre-publication" version of the highly-anticipated Part 3 to the Report on Copyright and AI that addressed liability and fair use issues. The Copyright Office generally sided with the copyright owners. As to liability, the Copyright Office viewed the use of existing copyrighted works to train LLMs to "clearly implicate the right of reproduction."[111] Regarding fair use, the Copyright Office credited a dilution theory of market harm, stating that harm that "stem[s] from AI models' generation of material stylistically similar to works in their training data" can weigh against fair use, even if those outputs do not contain copies of any portion of the copyrighted works.[112] When the Report issued, no court had endorsed that theory of harm. The day after the Copyright Office issued the "pre-publication" version, President Trump notified the Register of Copyrights, Shira Perlmutter, that he was firing her from her position as the head of the Copyright Office. Although the D.C. Circuit has effectively reinstated Ms. Perlmutter during the pendency of her legal challenge to the firing, the future of the Report remains uncertain. The Copyright Office has not issued the final version.

One month later, two district court judges in the Northern District of California, Judges Alsup and Chhabria, weighed in on fair use, and both granted summary judgment in favor of the defendants, Anthropic and Meta. On two key factors, the judges agreed. Both concluded that using copyrighted works to train LLMs was transformative, satisfying the first fair use factor.[113] And both agreed that there was no cognizable market for licensing books as LLM training material, so the authors could not show harm under the fourth factor.[114]

But the two decisions diverged in significant ways, shaped by the judges' radically different views of AI technology.

Judge Alsup, presiding over *Bartz v. Anthropic*, saw LLMs as gifted students who learn to craft new works from existing ones, "not to race ahead and replicate or supplant them—but to turn a hard corner and create something different."[115] But he believed the law does not permit AI companies to develop even this revolutionary technology by starting with pirated copies of the works or storing them indefinitely in a "central library." He cabined his finding of transformative use to conduct solely related to preparing datasets and using them for LLM training.[116] The decision therefore left open the possibility for plaintiffs to recover damages over other uses of the works—such as acquiring them from pirate websites or maintaining them in a permanent library—that would not be subject to a fair use defense.

Judge Chhabria, by contrast, saw LLMs as powerful plagiarists with the potential to flood the market with works that "compet[e] with books written by human authors for sales and attention," thereby "reduc[ing] the incentive for authors to create."[117] He opined that the plaintiffs might have prevailed if they had provided evidence of ***indirect*** market harm from AI-generated works that compete with the originals, even if those works are not themselves infringing. But because the plaintiffs "made the wrong argument and failed to develop a record in

---

111 Report at 26-28.
112 Report at 65.
113 *See Bartz v. Anthropic PBC*, 2025 WL 1741691, at *7 (N.D. Cal. June 23, 2025) ("using works to train LLMs was transformative—spectacularly so"); *Kadrey v. Meta Platforms, Inc.*, 2025 WL 1752484, at *2 (N.D. Cal. June 25, 2025) ("no serious question" that Meta's use was "highly transformative").
114 *See Bartz*, 2025 WL 1741691, at *17; *Kadrey*, 2025 WL 1752484, at *16.
115 *Bartz*, 2025 WL 1741691, at *8.
116 *Id.* at *8-14, 18-19.
117 *Kadrey*, 2025 WL 1752484, at *16-17.

support of the right one," he reluctantly granted summary judgment for Meta.[118] He predicted that in future cases, "the plaintiffs will often win" if they provide evidence of market dilution harm, and stated that "it's hard to imagine that it can be fair use to use copyrighted books to develop a tool to make billions or trillions of dollars while enabling the creation of a potentially endless stream of competing works."[119] Although Judge Chhabria is the first judge to endorse this market dilution theory, it is hard to square with established copyright principles: Copyright law does not confer a monopoly over "style or genre,"[120] and the fourth fair use factor addresses the effect on the market for the copyrighted ***work***, not on works in the same genre.

The two decisions thus gave AI companies important but limited wins. Both left open the possibility that plaintiffs can recover statutory damages for other allegedly infringing conduct—such as creating a permanent library of copyrighted works or distributing them through peer-to-peer torrenting—even if the downstream training use is fair. Further, shortly after the fair use ruling, Judge Alsup had certified a broad class of authors, publishers, estates, and other rightsholders, expanding Anthropic's potential exposure to statutory damages.

Given these unresolved exposures, the parties in *Bartz* reached a class action settlement in August 2025. Anthropic agreed to pay $1.5 billion into a settlement fund, with around $3,000 as the expected per-book payment. Anthropic also committed to delete the identified pirated copies from its internal "central library," while retaining the ability to train on lawfully acquired versions of the same works. Importantly, the settlement resolves Anthropic's liability only for past acquisition and use of a defined set of pirated works before an August 2025 cut-off. It does not create a forward-looking license for AI training, address output-based infringement, or create industry-wide rules. Another key effect of the settlement was that Judge Alsup's summary judgment ruling was left intact; the Ninth Circuit will have to wait for another case to weigh in on AI fair use issues.

b. Evolving Litigation Strategy and Emerging Trends in Input Cases

No new decisions on fair use in input cases from other district court judges are expected until mid-2026 at the earliest. In the meantime, the parties have shifted their strategies to account for the current state of the law.

The Shadow Library Strategy: Given the Northern District of California fair use rulings discussed above, plaintiffs have increasingly shifted their focus from training itself to how AI companies initially acquired the copyrighted works. One prominent, emerging strategy centers on unauthorized online book repositories like LibGen, Sci-Hub, and Anna's Archive. The infringement theory based was on these "shadow libraries," separate and distinct from one based on the training of AI models. It argues that for an AI developer that uses a peer-to-peer torrenting method to obtain its training material, each time it joins a shadow-library torrent to fetch datasets like Books3, it automatically makes portions of those datasets (and the corresponding expression in book) available to others in the network. Content owners argue that the alleged infringement includes not just reproduction for internal training but also unlicensed public distribution. Even if the model's training use might ultimately be fair use, content owners stress that torrent-based dissemination results in a separate act of infringement.

This strategy has become increasingly popular and is now asserted against myriad AI companies, such as Anthropic, Apple, Meta, Microsoft, NVIDIA, and OpenAI. For example, after the *Bartz* class action settled, a group of authors opted out and quickly pivoted in December 2025 to sue a number of AI companies, targeting not only Anthropic but also OpenAI, Google,

---

118 *Id.* at *7.
119 *Id.* at *23.
120 *see* 17 U.S.C. § 102(b)

Meta, xAI, and Perplexity. The allegations in that case, *Carreyrou et al. v. Anthropic PBC et al.*, center on the allegedly willful copying of their books from shadow libraries to build centralized training libraries, rather than using copyrighted books to train LLM models.

AI companies have started to respond to these arguments. In March 2026, Meta, in *Kadrey v. Meta*, served an interrogatory response stating its position that any uploading that occurred while it downloaded datasets via BitTorrent should also be treated as fair use. Meta characterized the seeding (the process that may require uploading a portion of the downloaded dataset) as a technical byproduct inherent in the BitTorrent protocol. It was also the only feasible way to obtain massive text corpora at the scale needed to train Llama, rather than as a separate content-distribution business. In effect, Meta seeks a ruling that would include the torrent "distribution" issue into the broader fair-use analysis it already won, reframing seeding as an unavoidable step in a transformative use rather than as distinct, direct infringement.

Class Certification as a Strategic Lever: The *Bartz* settlement demonstrated that the prospect of class certification could significantly affect the economics of the litigation and the willingness of AI companies to negotiate substantial settlements. We are likely to see, in the near future, rulings on class certification. In the *Google Generative AI litigation*, plaintiffs moved for class certification with a hearing held in early 2026 in the Northern District of California. In *Kadrey v. Meta*, plaintiffs filed a motion in January 2026 to move forward with class certification proceedings. And in the *OpenAI MDL*, while no class certification motion has been filed, the breadth of the consolidated proceedings and the volume of discovery suggest that certification efforts are likely to follow.

Efforts to License the Works: Although both Judges Alsup and Chhabria concluded that book authors have no right to a licensing market for training LLMs, they seem to believe that if the AI companies had tried harder, they could have licensed the books (and avoided liability). Judge Chhabria expressed optimism that AI companies will "simply figure out a way to license the works they wish to use as training data."[121] That prediction has come true in a wave of licensing deals between content owners and AI companies.

For example, in October 2025, Universal Music Group and AI music startup Udio settled their copyright infringement lawsuit and entered a licensing and partnership arrangement. Under the deal, Udio agreed to launch a new subscription platform in 2026 powered by generative AI technology trained on authorized and licensed music from UMG's recorded-music and publishing catalogs. The settlement provides for licensed training rights for future models, and UMG artists and songwriters can opt in to be compensated both for allowing their recordings to train the models and for downstream outputs on the service. Warner Music Group struck a similar settlement with Udio in November 2025, and also reached a separate deal with Suno, another leading AI music generator, under which Suno will launch new, licensed models in 2026 while phasing out its current models. In early March 2026, Meta and News Corp announced a three-year, multi-million dollar AI content licensing deal under which Meta can use News Corp's journalism content from *The Wall Street Journal*, *The Sun*, *the New York Post*, and others, to train its models.

New Claims and Arguments: One emerging line of cases targets not just the use of individual works in training datasets, but the copyrightability of the curated datasets themselves, focusing on the selection, coordination, and arrangement of the works that make up the dataset. In a new case filed in S.D.N.Y. in March 2026, *Gracenote Media Services, LLC v. OpenAI Foundation et al.*, Nielsen-owned Gracenote asserted that OpenAI copied its structured entertainment-metadata databases—including the relational framework connecting programs, performers, and identifiers—and that this curated corpus is a protectable compilation that

[121] *Kadrey*, 2025 WL 1752484, at *22.

constitutes an original work of authorship. Graceacre alleges that AI companies copied not just underlying members of the dataset, but the commercial value of the curated product, requiring the application of compilation copyright principles to large-scale AI training data.

Another trend involves claims extending the arguments tested in LLM training and chatbot cases to other AI technologies, such as those that involve Retrieval-Augmented Generation (RAG). For example, in the S.D.N.Y. case, *Advanced Local Media LLC v. Cohere Inc.*, a coalition of news publishers alleged that Cohere's RAG-based services ingest their articles into retrieval indexes and then reproduce and display substantial excerpts and summaries that substitute for the originals and erode subscription and advertising revenues. Cases against Perplexity (*Dow Jones & Company Inc. v. Perplexity AI Inc.* and *New York Times Co. v. Perplexity AI Inc.*) make similar arguments concerning Perplexity's AI-powered search engine called the answer engine.

Section 1202 DMCA Claims: Plaintiffs suing AI companies have also asserted claims under Section 1202 of the Digital Millennium Copyright Act, which prohibits the falsification, removal, or alteration of copyright management information ("CMI") such as author names, copyright notices, and terms of use. These claims offer advantages—they do not require copyright registration and carry separate statutory damages—but it is becoming increasingly clear that these claims are unlikely to play a key role in imposing liability on AI companies.

Several district courts have dismissed Section 1202 claims at the pleading stage for failure to allege sufficient facts about which CMI was removed, how defendants accomplished the removal, or whether defendants possessed the required mental state. Courts in the Ninth Circuit have imposed the additional requirement that AI outputs be ***identical*** to the copyrighted works—a standard that LLMs, which generate probabilistic outputs, are not designed to meet. The Ninth Circuit will soon weigh in, in the interlocutory appeal in *Doe I v. GitHub* on the dismissal of the DMCA claim.

Over the past year, however, court decisions have started to provide guidance on the requirements for pleading a viable Section 1202 claim against AI companies. For example, in *The Intercept Media v. OpenAI* (S.D.N.Y.), the court allowed a 1202(b) claim to proceed because The Intercept alleged specific, intentional CMI-removal steps in OpenAI's scraping and training pipeline, adequately pleading that the removal would facilitate future infringement. The emerging consensus is that Section 1202 claims backed by concrete allegations about data-processing methods can survive, while broader theories premised merely on the absence of attribution in AI outputs are likely to keep failing at the pleading stage.

Section 1201 DMCA Claims. Content owners are increasingly asserting claims based on a separate provision of the DMCA, Section 1201, to hold AI companies liable for how they obtained training data. Section 1201(a) prohibits circumventing a technological measure that effectively controls access to a copyrighted work; a claim for the violation of this provision is not aimed at copying itself, but at the bypassing of digital locks that guard the work. Several recent cases assert that AI companies used stream-ripping to bypass YouTube's rolling cipher—which plaintiffs claim are the types of access controls that Section 1201(a) protects—to create downloaded copies of YouTube videos to create AI training dataset. A Section 1201 claim offers advantages over an infringement claim. The provision permits "any person injured" by a violation to sue, so copyright registration is not required. The fair use defense does not apply. And plaintiffs may recover statutory damages of $200 to $2,500 per act of circumvention without proving actual harm. *See* 17 U.S.C. § 1203(c)(3). In *Sony Music Entertainment v. Uncharted Labs*,[122] the court declined to dismiss a Section 1201 claim against the developer of an AI music generator, holding that the major record labels had plausibly alleged that the company circumvented YouTube's

[122] 2026 WL 1019199 (S.D.N.Y. Apr. 15, 2026)

rolling cipher to download protected audio in bulk to train its model.

c. Recent Developments in Cases Involving Infringement Liability for Using Expression in AI Outputs

Recent months have seen a rapid advance in generative AI technology—such as those from Google (Veo and Nano Banana), OpenAI (Sora), MiniMax (Hailuo), and others—that allow users to generate images and videos by entering prompts. These new AI products, as well as the pleading-stage decisions in output cases, have led copyright owners—including Hollywood studios—to increasingly argue that these tools have generated images or videos containing recognizable characters and visual expression that are copyrighted in movies, television shows, and other visual or audiovisual works.

In the earliest cases filed against generative AI companies, courts expressed skepticism over the output theory—*i.e.*, that the outputs LLMs generate, in response to a user's prompt, are infringing works. LLMs use probabilistic algorithms to offer predictive results based on relationships among the components of tremendously large sets of texts, images, or videos. They are good at mimicking human expression, but they do not do so by stitching together portions of existing works. Courts accordingly rejected content owners' initial arguments that: (a) outputs are *per se* derivative works simply because LLMs were trained on copyrighted inputs; and (b) outputs are necessarily infringing even without demonstrating substantial similarity to the plaintiffs' works. Instead, courts held that even in AI cases, copyright liability arises only if the defendant's output is "substantially similar" to the copyrighted work.[123]

Litigation has evolved significantly since those early rulings. Copyright owners have sharpened their output allegations, and newer cases present far stronger claims of substantial similarity. For example, in the consolidated OpenAI MDL proceedings, Judge Stein denied OpenAI's motion to dismiss in October 2025, holding that plaintiffs had adequately alleged that certain ChatGPT outputs—including detailed book summaries and potential sequel outlines—could be found by a reasonable jury to be substantially similar to the original copyrighted works, distinguishing them from simpler news-article summaries the court had previously found non-infringing in a related case.

The most dramatic shift, however, has come from the major studio lawsuits against AI image and video generators. In the lawsuits the studios have filed against Midjourney, for example, the studios alleged that Midjourney's image service provides output images in response to user prompts that accurately depict the studios' IP—*e.g.*, Disney's copyrighted Darth Vader character—in visual works that Disney neither created nor authorized. Warner Bros. Discovery filed a separate suit against Midjourney in September 2025, which has since been consolidated with the Disney action. And these studios jointly sued Chinese AI company MiniMax (which offers the Hailuo AI video and image generation service) in September 2025 as well. The studios have labeled the AI products a "virtual vending machine" for "unauthorized copies of Disney's and Universal's copyrighted works."

Because these new output cases involving visual works are generally at the pleading stage, courts have not issued substantive rulings. Although the studios have tried to seize on the high recognizability of their copyrighted works, AI companies have begun to assert defenses based on well-established principles of copyright law that predate AI technology. Important threshold defenses include challenges to ownership and copyrightability (for example, copyright law's exclusion of ideas, stock characters, or standard visual elements from protection), as well

[123] *See, e.g.*, *Kadrey v. Meta* (rejecting argument that every LLaMA output is an infringing derivative work of copyrighted inputs); *Tremblay v. OpenAI* (ruling that plaintiffs "must show a substantial similarity between the outputs and the copyrighted material").

as the practical difficulty of proving substantial similarity across the large number of works at issue. Courts must assess output on an output-by-output basis, each time determining whether the specific output is "substantially similar" to the asserted copyrighted work. Fair use is also likely to be contested with respect to outputs. Judges Alsup and Chhabria did not address outputs in their fair use rulings. And although fair use is unlikely to cover all outputs, the studios' complaints may overreach by assuming that every output incorporating a recognizable character is infringing, even those outputs that may constitute protected fair uses, such as commentary, certain fan art, and parodies.

Further, output cases put a greater emphasis on whether AI companies may be secondarily (or indirectly) liable for the acts of their users, if those users (and not the AI companies) are deemed to be the direct infringers of any substantially similar outputs. After all, the users are the ones who can input prompts of their choosing, including those that violate the terms of use, to obtain the outputs that may feature copyrighted expression. For a vicarious infringement claim, in addition to direct infringement by a third party, the studios would need to prove that the AI company had the right to supervise the infringing conduct and a direct financial interest in it. Although these standards typically favor technology companies that offer copyright-agnostic products, in October 2025, the court in *Concord Music v. Anthropic* found the financial interest prong satisfied at the pleading stage, concluding that there were sufficient allegations about the presence of copyrighted song lyrics in outputs being a "draw" for customers and investors, which then allegedly resulted in direct financial benefit to Anthropic.

For a contributory infringement claim, AI companies would have a strong staple-article-of-commerce defense, because the generative tools from Midjourney and MiniMax are capable of generating content with and without copyrighted characters, and thus have substantial non-infringing uses.[124] That argument became stronger when, on March 25, 2026, the Supreme Court in *Cox Communications v. Sony Music Entertainment* held that contributory infringement liability requires either (1) inducement of specific acts (such as promoting or marketing the product as a tool to infringe copyright), or (2) evidence that the product/service was tailored to facilitate direct infringement. This decision made it more difficult for copyright owners to prevail on a contributory infringement claim against AI companies for output claims, because absent other facts, generative AI tools likely do not induce infringement, nor are they tailored to be infringing. That said, in the recent generative AI cases applying *Cox*, lower courts have required relatively little pleading specificity to let contributory claims proceed past the pleading stage, particularly on the inducement theory.

Licensing remains an option for resolving output-related disputes as well as the input-related disputes. In December 2025, OpenAI and Disney announced a three-year deal making Disney the first major content-licensing partner for Sora, OpenAI's short-form generative video platform. Under the agreement, Sora (and ChatGPT Images) can generate user-prompted short videos and images using more than 200 Disney-owned characters (but no actor likeness or voices) and related IP from Disney, Pixar, Marvel, and Star Wars. In return, Disney will invest $1 billion in OpenAI, become a major API customer using OpenAI models to power new Disney+ and internal tools, and work with OpenAI on "responsible AI" guardrails intended to protect creators' rights. However, there is now reporting that Disney has ended the deal when OpenAI decided to close down Sora in late March 2026.

d. Protecting Creative Expression Made with AI Tools

In March 2026, Netflix announced a deal to acquire Ben Affleck's company, InterPositive, that aims to develop AI-powered filmmaking technology. As Hollywood and other creative

[124] *See Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 456 (1984).

industries look to AI software as tools for artistic expression, creative and technology industries should think not just defensively about infringement suits, but about how to protect the valuable IP created using AI tools.

One of the significant challenges to protecting works containing AI-generated expression is that copyright protection has traditionally extended to original works of ***human*** authorship. Although the term "author" is not defined in the Copyright Act, courts have presumed that an "author" must be human. There was a challenge to this belief a few years ago, in a case that asked the courts to extend copyright protection to a selfie taken by a monkey. Although the Ninth Circuit declined to address that question, the court observed that the language of the Copyright Act "necessarily exclude animals" from having standing to assert an infringement claim.[125] Two recent cases about whether artworks made by or using AI can be eligible for a copyright registration nicely illustrate the line-drawing issue the law must address. In *Thaler*, Stephen Thaler characterized the visual art that the U.S. Copyright Office declined to register as an entirely AI-authored work, without any human contribution. Thus, in March 2025, the D.C. Circuit rejected Thaler's appeal of the Copyright Office's refusal to register a painting for which Thaler listed the AI system he invented as the sole author. In March 2026, the Supreme Court denied certiorari, leaving in place the D.C. Circuit's generally non-controversial conclusion that U.S. copyright law requires a human author. But *Thaler* did not address the more difficult questions that arise when the applicant claims that one or more human authors contributed to the creation of the work, and used AI products as tools, rather than as independent authors. That is the question posed in *Allen*. In September 2024, Jason Allen sued the Copyright Office for rejecting his application to register an award-winning image he says he authored using generative AI tool, Midjourney, by submitting more than 600 prompts to compose the image. The parties have filed cross-motions for summary judgment. In *Allen* or in future cases, courts (or Congress) must answer difficult line-drawing questions over how much or what kind of human contribution is required to render the work human-authored.

Under the longstanding requirement for human authorship, a work that LLMs generate in response to a human user's prompt—without subsequent human modification—is unlikely to receive copyright protection under the current rules, even if it contains otherwise protectible expression. In February 2023, the Copyright Office canceled the registration of a work that used generative AI to create the images for a graphic novel because the underlying AI-generated images were not themselves copyrightable—the author did not control what images the AI would generate. The Copyright Office then issued guidance stating that AI-generated material "must be disclaimed in a registration application" and that it will not register works where "AI technology receives solely a prompt from a human and produces complex written, visual, or musical works in response …."[126]

In late January 2025, the Copyright Office issued Part 2 of the Copyright and Artificial Intelligence report, addressing copyrightability of works created using generative AI tools. The Report reaffirmed that "purely AI-generated" works lacking sufficient human control are ineligible for registration, and that directing generative AI tools using "prompts alone" is insufficient—even with multiple, iterative prompts or other forms of prompt engineering. According to the Copyright Office, prompts are merely "instructions that convey unprotectible ideas."[127]

This policy may run contrary to the well-established *Bleistein* non-discrimination principle, under which the Supreme Court directed courts not to make aesthetic judgments or privilege artistic expression made from classical techniques over those made using newer technology. Copyright law has repeatedly opted to protect creative outputs using new technology—most

---

125 *Naruto v. Slater*, 888 F.3d 418, 426 (9th Cir. 2018).
126 Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence, 88 Fed. Reg. 16 (Mar. 16, 2023).
127 Report at 19-20.

notably, the Supreme Court's holding that photographs are original works of authorship worthy of copyright protection, not merely machine-produced images.[128] A wholesale exclusion of copyright protection for creative works made using LLMs is in tension with this technology-agnostic tradition.

The Copyright Office has acknowledged that its policy "does not mean that technological tools cannot be part of the creative process," and has tried to draw the line between works where AI is the author and works made using AI tools where a human is the author. The Office has stated that works containing AI-generated material may still be eligible for registration where, for example, a human selects or arranges AI-generated material in a sufficiently creative way, or modifies AI-generated material to a degree that meets the standard for copyright protection—though copyright will extend only to the human-authored aspects. The Office also confirmed that using AI as "an assistive tool" does not undermine copyright protection for the resulting work.[129] Using generative AI as a "brainstorming tool" or adding edits, image sharpening, enhancements, color correction, and modifications could result in a protectible human-authored work.[130] But the Copyright Office will exclude from registration any new expression that generative AI added.[131]

In practice, it is hard to discern how much human direction and intervention is sufficient for the Copyright Office to deem a work to have been "human-authored." That uncertainty over the protectability of AI-aided works will remain until Congress acts, the Copyright Office changes its position, or the courts rule. Nonetheless, "human authors should be able to claim copyright if they select, coordinate, and arrange AI-generated material in a creative way."[132] The Copyright Office thus registered the AI-generated comic book, "Zarya of the Dawn," for the human-authored text and the selection and arrangement of text and AI-generated images. However, the AI-generated images, on their own, were not included in the scope of the registration. Likewise, in January 2025, the Copyright Office registered an AI-generated image called "A Single Piece of American Cheese" for the selection and arrangement of AI-generated visual elements. Of note is that the applicant, Kent Keirsey, submitted a 10-minute time-lapse video demonstrating that the creative process did not rely solely on iterative prompts, but also involved 35 rounds of edits involving human aesthetic decisions to select and refine AI-generated elements in the composition.

Based on established copyright principles and the Copyright Office's existing guidance, here are some suggested practices for maximizing the likelihood of obtaining copyright protection, until the courts provide further guidance:

- **Do not** use prompts alone to generate audiovisual, sound recording, or literary works. Instead, feed the generative AI tools with expressive inputs that reflect artistic decisions that can be made at that stage of the process. This would likely require the use of the enterprise version of the AI software, and confirmation with the AI vendors that the inputs would be kept confidential and not used to train the model for anyone else's (*e.g.*, the public's) use.
- **Do not** use generative AI tools as a single-step tool. Instead, go back and forth between prompting AI outputs and adding documented human decisions into the AI tool. This will permit the characterization of AI uses as a "brainstorming tool" or as providing editorial enhancements that the Office suggests would retain human authorship of the final work.

---

128 *See Burrows-Giles Lithographic Co. v. Sarony*, 111 U.S. 53 (1884).
129 Report at 1.
130 Report at 12, 25.
131 Report at 23.
132 Report at 24.

- When applying for registration, submit the version of the work that reflects the human ordering, placement, and edits of AI-generated visual or literary pieces. **Do not** submit the work in a form that lacks the final sequencing or edits to the outputs. This will help obtain copyright protection over any original selection-and-arrangement or post-output edits.
- If the creative expression is based on the collaging of AI-generated expression, identify the specific creative relationships that result from the selection and arrangement of AI outputs. Doing so could help establish protectible expression—and infringement.

VIII. Products Liability

Whether and how traditional principles of product liability apply to forms of AI deployment—for example, AI-powered "chat bots" for user interactions, AI-powered recommendation systems for selecting or promoting social media posts to user feeds, and other common AI tools—was, until recently, largely an open question. That is changing. In 2025 and early 2026, courts issued the first substantive rulings treating AI-generated chatbot output as a "product" for product liability purposes, state attorneys general brought the first enforcement actions against AI chatbot companies, and a jury in Los Angeles, CA returned a verdict finding Meta and YouTube liable for negligence in the design and operation of their platforms, awarding a total of $6 million in compensatory and punitive damages in what is the first jury verdict holding technology companies liable under product liability theories. At the same time, a wave of federal and state legislation targeting AI safety is advancing through Congress and state legislatures. These developments are reshaping the product liability landscape for AI companies and the businesses that deploy their technologies.

The most significant ruling to date on AI product liability came from *Garcia v. Character Techs., Inc.*, 785 F. Supp. 3d 1157 (M.D. Fla. 2025). In that case the plaintiff—a mother of a 14-year-old boy who died by suicide after prolonged interactions with a Character AI chatbot—asserted strict product liability theories, negligence theories, wrongful death, and claims under the Florida Deceptive and Unfair Trade Practices Act against Character Technologies, its co-founders, and Google. In May 2025, the court found the Character AI chatbot to be a "product" subject to product liability law, rejecting defendants' argument that their software platform was a service, not a tangible product. In denying dismissal of the product-based claims, the court relied on plaintiff's allegations that "Character A.I. fails to confirm users' ages and omits reporting mechanisms, Characters are programmed to employ human mannerisms, and users are unable to exclude indecent content."[133] Accordingly, the court held that "Character A.I. is a product for the purposes of Plaintiff's product liability claims so far as Plaintiff's claims arise from defects in the Character A.I. app rather than ideas or expressions within the app."[134]

In August 2025, the family of Adam Raine, a 16-year-old California high school student who also died by suicide, filed a wrongful death lawsuit against OpenAI and its CEO Sam Altman in San Francisco Superior Court.[135] The complaint asserts product liability claims (design defect and failure to warn), negligence, wrongful death, and California statutory claims under the Unfair Competition Law and Consumer Legal Remedies Act. The case extends the product liability framework from "companion" AI chatbots (*e.g.*, Character AI) to general-purpose AI systems (ChatGPT). The complaint alleges that GPT-4o was designed with features that fostered psychological dependency: "persistent memory that stockpiled intimate personal details, anthropomorphic mannerisms calibrated to convey human-like empathy, heightened sycophancy

[133] 785 F. Supp. 3d at 1180.
[134] *Id.*
[135] *Matthew Raine et al. v. OpenAI, Inc. et al.*, No. CGC25628528 (Cal. Super. S.F. County, filed Aug. 26, 2025).

to mirror and affirm user emotions," and "24/7 availability capable of supplanting human relationships."[136] Combined with the alleged absence of adequate intervention protocols for self-harm scenarios, plaintiffs contend these features rendered the product defective.

In November 2025, the Social Media Victims Law Center and Tech Justice Law Project filed seven additional lawsuits against OpenAI in California state courts, alleging wrongful death, assisted suicide, involuntary manslaughter, and product liability claims.[137] These suits allege that "OpenAI knowingly released GPT-4o prematurely, despite internal warnings that the product was dangerously sycophantic and psychologically manipulative."[138] The rapid accumulation of claims against both Character AI and OpenAI suggests that AI chatbot product liability litigation is likely to intensify.

Besides private litigation, enforcement actions have been initiated by state attorneys general. On January 8, 2026, Kentucky became the first state to file a lawsuit specifically targeting an AI chatbot company. Attorney General Russell Coleman filed suit against Character Technologies in Franklin Circuit Court, describing Character AI as "dangerous technology that induces users into divulging their most private thoughts and emotions and manipulates them with too frequently dangerous interactions and advice."[139]

In the ongoing Multi-District Litigation in the Northern District of California relating to the allegedly addictive nature of social media platforms, the litigation has advanced significantly since our last update. On March 25, 2026, a jury in Los Angeles Superior Court returned a verdict in *KGM v. Meta Platforms, Inc.*, finding Meta and YouTube liable for negligence in the design and operation of their platforms.[140] In this case, the plaintiff, now 20 years old, alleged that she began using YouTube at age 6 and Instagram at age 9, and that the platforms' design—including infinite scrolling, autoplay, algorithmic recommendation engines, and deliberately unpredictable rewards—caused compulsive use, depression, anxiety, body dysmorphia, and suicidal ideation. In a critical pre-trial ruling on November 5, 2025, the court denied Meta's motion for summary judgment, distinguishing between features related to content publishing—which Section 230 might protect—and features like notification timing, engagement loops, and the absence of meaningful parental controls—which it does not.[141] The court found that "there is evidence in the record that K.G.M. was harmed by [defendants'] design features," and that the "cause of K.G.M.'s harms is a disputed factual question that must be resolved by the jury."[142] After six weeks of evidence, the jury found for the plaintiff, determining that the defendants knew that the design or operation of their platforms was dangerous or was likely to be dangerous when used by a minor, and that they failed to adequately warn of the danger.

The federal MDL (MDL No. 3047, N.D. Cal.) has expanded dramatically.[143] By February 2026, approximately 2,325 claims were active in the MDL, up from approximately 1,745 in April 2025. Twenty-nine state attorneys general have asked the court to consolidate their claims into a single trial. Nearly 800 school districts have filed suits. Six school district bellwether cases—from Maryland, Georgia, Kentucky, New Jersey, South Carolina, and Arizona—have been selected for federal MDL trials expected in late 2026.

---

[136] *Id.* ¶ 12.
[137] See Press Release, *Tech Justice Law Project and Social Media Victims Law Center lawsuits accuse ChatGPT of emotional manipulation, supercharging AI delusions, and acting as a "suicide coach,"* Tech Justice Law (Nov. 6, 2025), https://techjusticelaw.org/2025/11/06/social-media-victims-law-center-and-tech-justice-law-project-lawsuits-accuse-chatgpt-of-emotional-manipulation-supercharging-ai-delusions-and-acting-as-a-suicide-coach/.
[138] *Id.*
[139] See Complaint ¶ 52 (available at https://www.ag.ky.gov/Press%20Release%20Attachments/CTI%20Complaint%20Motion%20and%20Order%20Filed.pdf)
[140] *P. F. v. Meta Platforms, Inc.*, No. 22STCV21355 (Cal. Super. L.A. County, filed Oct. 24, 2022).
[141] *P. F. v. Meta Platforms, Inc.*, No. 22STCV21355, 2025 WL 4112819 (Cal. Super. Nov. 05, 2025)
[142] *Id.* at *2.
[143] *In re Social Media Adolescent Addiction/Personal Injury Prods. Liab. Litig.*, MDL No. 3047 (N.D. Cal.).

The design-defect theories advancing in these cases build on a line of authority allowing plaintiffs to target a platform's design features without triggering Section 230 immunity. In *Lemmon v. Snap, Inc.*, the Ninth Circuit permitted a claim premised on Snap's "speed filter," reasoning that the platform could have eliminated the allegedly dangerous feature "without altering the content that Snapchat's users generate."[144] Similarly, in *A.M. v. Omegle.com, LLC*, the court distinguished the platform's random pairing of users—a design choice that occurred before any third-party content existed—from the subsequent publication of their conversations.[145] The Northern District of California systematized this approach in the social media MDL, *In re Social Media Adolescent Addiction/Personal Injury Products Liability Litigation*, 702 F. Supp. 3d 809 (N.D. Cal. 2023), distinguishing features that may support liability outside Section 230—such as inadequate parental controls, the absence of time-restriction options, barriers to account deletion, insufficient age verification, and appearance-altering filters without labeling—from features reflecting protected editorial functions, such as endless content feeds and algorithmic content promotion for engagement. The court's line turned on whether addressing the alleged defect would require the platform to "publish less third-party content" or alter its editorial decisions about "what to publish and how." The *KGM* court applied precisely this distinction at summary judgment, and it is the framework against which AI design-defect claims are likely to be tested.

A common thread connects these cases: the question of whether AI-powered inference models—recommendation algorithms, generative AI chatbots, and other systems that shape user experience through algorithmic decision-making—can be treated as "products" subject to traditional product liability principles. Two primary categories of AI "product" claims are now emerging across the litigation landscape. *First*, recommendation algorithms: the social media MDL targets AI-driven systems that select and promote content to maximize engagement, alleging they are defectively designed because they exploit the neurobiology of minors. *Second*, generative AI chatbots: the *Garcia* and *Raine* cases target AI systems that produce novel conversational output, alleging they are defectively designed because they foster psychological dependency, fail to detect self-harm, and lack adequate safety interventions. Across both categories, plaintiffs allege that AI companies knew of the risks their products posed and failed to adequately warn users or implement available safety features. Indeed, the *KGM* jury specifically found that Meta and YouTube failed to adequately warn of the dangers of their platforms to minors. The *Garcia* complaint's failure-to-warn claims survived the motion to dismiss. And the *Raine* complaint centers the allegation that OpenAI failed to warn users and parents about the risks of psychological dependency. For AI companies, the failure-to-warn theory is particularly significant because it is a familiar product liability theory that may require less doctrinal innovation than design defect claims and may prove more straightforward for juries to evaluate.

Driven in part by the litigation developments described above, a wave of legislative and regulatory activity targeting AI chatbot safety and social media design has emerged at both the federal and state level. At the federal level, the Kids Online Safety Act (KOSA) was reintroduced in May 2025 with bipartisan support and would require platforms to implement safeguards for minors, including "limit by default design features that encourage or increase the frequency, time spent, or activity of minors on the covered platform, such as infinite scrolling, auto playing, rewards for time spent on the platform, notifications, and other design features that result in compulsive usage" of the platform.[146] The House advanced the Kids Internet and Digital Safety (KIDS) Act in early 2026, which includes provisions specifically addressing AI chatbots—requiring disclosure to minors that they are interacting with an AI, crisis intervention resources when minors express suicidal ideation, and mandatory break reminders after three hours of use. The GUARD Act would prohibit minors from using AI companions entirely and impose criminal penalties for chatbots that

[144] 995 F.3d 1085 (9th Cir. 2021).
[145] 614 F. Supp. 3d 814 (D. Or. 2022).
[146] Kids Online Safety Act, S. 1748, 119th Cong. § 103(C) (2025), (available at https://www.congress.gov/bill/119th-congress/senate-bill/1748/text#toc-id541c6e8b-6708-4d97-83e1-44abf7a184b2); *see also* GUARD Act, S. 3062, 119th Cong. §91 (2025).

promote certain harms. The Federal Trade Commission has also launched an inquiry under its Section 6(b) authority into AI chatbots acting as companions, examining what actions companies are taking to mitigate alleged harm and compliance with the Children's Online Privacy Protection Act.[147]

At the state level, California enacted SB 243, which requires operators of companion chat platforms to provide clear notice that the chatbot is AI, send reminders every three hours encouraging minors to take a break, and implement measures to prevent sharing sexually explicit content with minors.[148] A separate ballot initiative, the California Kids AI Safety Act, is gathering signatures for the November 2026 ballot and would impose stricter requirements including independent safety audits for AI products directed at minors.

Internationally, the European Union's revised Product Liability Directive, adopted in October 2024, extended the definition of a "product" to include digital files, online platforms, and all types of software, including AI systems.[149] This Directive applies the EU's strict (non-fault) product liability framework to consumer-facing generative AI systems, including chatbots. It will apply to products placed on the market as of December 2026. The EU Artificial Intelligence Act classifies AI systems used as products, or as safety components of products, as "high-risk" systems subject to third-party conformity assessments.[150] As the December 2026 implementation date approaches, AI companies operating in or selling into the EU market should be preparing for compliance with these requirements.

In sum, AI companies should evaluate their products and design choices against this rapidly developing legal framework, with particular attention to safety features for vulnerable users, adequacy of warnings and disclosures, internal processes for detecting and responding to foreseeable harms, and defensibility of design decisions that prioritize engagement or growth over user safety. For instance, any AI system that interacts with consumers (especially minors) without adequate disclosures about its limitations, risks, or potential for generating outputs that cause harm may face failure-to-warn claims under existing products liability frameworks. Further, a conduct-versus-content distinction may extend to design decisions, such as prioritizing engagement maximizing responses, choosing not to implement intervention protocols, or failing to implement age-verification or parental controls, so as to preclude the safe harbors of the First Amendment or Section 230.

IX. AI Algorithm Tools & Section 230 Immunity

The Fourth Department's recent decision in *Patterson v. Meta Platforms, Inc.*, -– N.Y.S.3d ----, 2025 WL 2092260 (4th Dept. July 25, 2025), highlights a growing circuit split over whether social media platforms' content-recommendation algorithms transform third-party content into first-party speech, potentially stripping companies of their Section 230 immunity. This decision, rejecting the Third Circuit's approach in *Anderson v. TikTok*, creates uncertainty for technology companies relying on algorithmic content curation.

a. Section 230's Traditional Protection

Section 230 of the Communications Decency Act, 47 U.S.C. § 230, provides that "no provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider." This provision creates federal

[147] Press Release, FTC Launches Inquiry into AI Chatbots Acting as Companions, Fed. Trade Comm'n (Sept. 11, 2025), https://www.ftc.gov/news-events/news/press-releases/2025/09/ftc-launches-inquiry-ai-chatbots-acting-companions.
[148] Cal. SB 243 (2025).
[149] Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024.
[150] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024.

immunity to any cause of action that would make service providers liable for information originating with third-party users.[151]

Courts have historically interpreted Section 230 broadly to protect social media companies from tort liability, including product liability claims. In *Force v. Facebook, Inc.*, 934 F.3d 53 (2d Cir. 2019), the Second Circuit held that using algorithmic tools to match third-party information with consumer interests does not strip platforms of Section 230 protection. The court reasoned that "[m]erely arranging and displaying others' content to users of Facebook through such algorithms—even if the content is not actively sought by those users—is not enough to hold Facebook responsible as the 'develop[er]' or 'creat[or]' of that content."[152]

The Fourth Circuit recently reinforced this interpretation in *M.P. By & Through Pinckey v. Meta Platforms, Inc.*, 127 F.4th 516 (4th Cir. 2025), affirming dismissal of strict products liability claims against Facebook. Despite allegations that Facebook's algorithms "radicalized" a shooter through targeted content recommendations, the court held that "decisions about whether and how to display certain information provided by third parties are traditional editorial functions of publishers."[153]

b. Strategy One: Targeting Platform Design Features

Plaintiffs have developed two primary strategies to circumvent Section 230 immunity. The first focuses on allegedly dangerous features of social media platforms that exist independently of third-party content publication, and relies on traditional product liability doctrines. *See supra* Section VIII.

c. Strategy Two: Algorithms as First-Party Speech

The second strategy emerged from the Supreme Court's decision in *Moody v. NetChoice, LLC*, 603 U.S. 707 (2024), which held that content-moderation algorithms constitute expressive activity protected by the First Amendment. Justice Kagan explained that "deciding on the third-party speech that will be included in or excluded from a compilation—and then organizing and presenting the included items—is expressive activity of its own."[154]

The Third Circuit seized on this reasoning in *Anderson v. TikTok, Inc.*, 116 F.4th 180 (3d Cir. 2024). The court accepted plaintiffs' argument that TikTok's algorithm "amalgamates third-party videos" to create "an express product that communicates to users that the curated stream of videos will be interesting to them."[155] Crucially, the court concluded: "Given the Supreme Court's observations that platforms engage in protected first-party speech under the First Amendment when they curate compilations of others' content via their expressive algorithms . . . it follows that doing so amounts to first-party speech under § 230, too."[156] Because Section 230 only immunizes "information provided by another," TikTok's algorithmic recommendations constituted unprotected first-party speech.

d. The Circuit Split: *Patterson* Rejects *Anderson*

The Fourth Department in *Patterson* explicitly rejected the Third Circuit's reasoning, creating a clear circuit split. The case arose from the May 14, 2022 Buffalo mass shooting, with

[151] *Zeran v. America Online, Inc.*, 129 F.3d 327, 330 (4th Cir. 1997).
[152] *Id.* at 66.
[153] *Id.* at 526.
[154] *Id.* at 731.
[155] *Id.* at 183.
[156] *Id.* at 184.

plaintiffs alleging that social media platforms' recommendation algorithms fed the shooter "a steady stream of racist and violent content," leading to his radicalization and isolation.

The *Patterson* court held that Section 230 continues to protect platforms using content-recommendation algorithms, finding *Anderson* unpersuasive. The court warned that accepting *Anderson*'s logic would mean "Internet service providers using content-recommendation algorithms (including Facebook, Instagram, YouTube, TikTok, Google, and X) would be subject to liability for every defamatory statement made by third parties on their platforms."[157] This result, the court concluded, would be "contrary to the express purpose of section 230."

Most significantly, *Patterson* articulated what it called a "Heads I Win, Tails You Lose" proposition favoring social media companies. Either platforms receive Section 230 immunity because their algorithms do not deprive them of publisher status (following *Force* and *M.P.*), or they receive First Amendment protection because the algorithms create protected first-party speech (following *Moody* and *Anderson*). The court concluded: "Of course, section 230 immunity and First Amendment protection are not mutually exclusive, and in our view the social media defendants are protected by both. Under no circumstances are they protected by neither."[158]

The court expressed concerns about *Anderson*'s implications, warning it would "gut the immunity provisions of section 230 and result in the end of the Internet as we know it." Platforms using algorithms would face "unlimited liability" for all tort claims, including defamation, ultimately causing the Internet to "devolve into mere message boards."[159]

e. Looking Forward: Uncertainty and Supreme Court Review

The current landscape presents significant uncertainty for technology companies and plaintiffs alike. While certain platform features—particularly those unrelated to content publication—may support viable product defect claims outside Section 230's protection, the status of algorithmic content recommendation remains contested.

The Second and Fourth Circuits maintain that recommendation algorithms are traditional editorial functions protected by Section 230. The Third Circuit, conversely, views algorithmic curation as first-party speech falling outside Section 230's scope. This split creates forum-shopping opportunities and inconsistent liability exposure for platforms operating nationally.

Supreme Court review appears increasingly likely given the fundamental disagreement about Section 230's scope and its interaction with First Amendment principles. However, it is less certain how the Supreme Court will decide these issues. As Justice Barrett noted in her *Moody* concurrence, the "complexities" involved include considerations of "rapidly evolving" AI technology.[160]

Until the Supreme Court resolves this split, technology companies should consider both defensive strategies: maintaining robust Section 230 documentation while simultaneously preparing First Amendment defenses for their algorithmic systems. Plaintiffs, meanwhile, will likely focus on identifying platform features that can be characterized as defects independent of content publication functions, while monitoring developments in circuits following the *Anderson* approach.

---

[157] *Patterson*, 2025 WL 2092260, at *4.
[158] *Id.* at *5.
[159] *Id.* at *6.
[160] *Moody*, 603 U.S. at 746 (Barrett, J., concurring).

X. Libel & Defamation

A widespread adoption of a new technology for expression, such as social media platforms like X, is often followed by an increase in libel suits. The broad use of ChatGPT and other large language models (LLMs) is no exception. Given the myriad contexts for LLM-generated statements that form the bases for defamation claims that plaintiffs are continuing to assert across multiple jurisdictions, AI companies, users of generative AI tools, and potential defamation plaintiffs should consider how the defamation principles that were developed for human speakers can apply to made-to-order AI speech. Libel claims based on AI-generated fictitious statements are starting to work through the courts, but with results less favorable than the plaintiffs likely had expected. For example, in May 2025, a Georgia state court granted summary judgment to OpenAI on a defamation claim concerning indisputably false statements in ChatGPT output. The court held that under the circumstances present—where the user knew ChatGPT was likely to provide a fictitious response to his prompt—no reasonable person could have understood that the output was communicating facts.

a. AI-Generated Statements with Potential Defamatory Meaning

LLM-based AI programs are not designed to generate outputs by replicating particular existing expression. Rather, these models employ probabilistic and predictive algorithms to generate a logical sequence of words that are responsive to each new prompt. Inherent in this technology is the software's capacity to create—or even hallucinate—content, which may contain fictitious statements that the software convincingly presents as factual assertions. Courts are starting to grapple with the different types of AI statements that may convey defamatory meaning:

Creating False Statements from False Attribution of True Facts: AI programs can present facts that are true in isolation, but, through omission or misattribution, convey false and harmful statement about an individual or a company. This kind of false statements formed the basis for *Battle v. Microsoft Corp.*, No. 1:23-cv-01822 (D. Md. filed July 7, 2023). The plaintiff, Jeffery Battle, alleged that Microsoft defamed him when Microsoft's search engine Bing returned the search results of his name with an AI-generated "summary" that conflated him with another individual with a similar name—Jeffrey Battle—and attributed facts about this convicted terrorists to the plaintiff. That summary stated:

> Jeffrey Battle, also known as The Aerospace Professor, ... ***However***, ***Battle*** was sentenced to eighteen years in prison after pleading guilty to seditious conspiracy and levying war against the United States. He had two years added to his sentence for refusing to testify before a grand jury.

The use of the "however" connector linked the facts about someone else with the plaintiff, creating a false association that imply harmful facts about the plaintiff. Unfortunately, this case will not result in an order that could provide guidance in future AI defamation cases: In October 2024, the court granted Microsoft's motion to compel arbitration.

Creating False Statements that Could Imply Defamatory Meaning: Defamation claims may be based on AI-generated statements that are not literally false but imply a defamatory fact. In November 2024, Asian News International (ANI), an Indian news agency, sued OpenAI in India for copyright infringement and for reputational harm caused by ChatGPT's "hallucinations" that incorrectly attributed fictitious sources to ANI. And in the United States, *The New York Times*, *The Wall Street Journal*, and others have asserted trademark dilution claims, alleging that AI output's false attribution to the newspapers of the content they did not write harms the journalism's brand. We may soon see other AI defamation by implication cases, particularly as image and

video generating tools create content that can imply false statements through visual or audiovisual juxtapositions. A company selling solar energy products recently filed suit against Google concerning false statements that appear in the "AI overview" section on top of Google search results. Those statements are linked to several articles and other sources that, though they do not support the false statements at issue, may suggest that the AI-generated statements convey facts that are capable of defamatory meaning. *See LTL LED, LLC v. Google LLC*, Case No. 25-cv-02394 (D. Minn.), Dkt. No. 1-1 (Compl.).

Creating False Statement Hallucinations: An AI output—particularly in response to engineered prompts—can fill in the gaps in training with made-up facts and sources. A case pending in Georgia state court, *Walters v. OpenAI*, involves such AI "hallucinations." Case No. 23-A-04860-2 (Ga. Super. Ct. Gwinnett County, filed July 17, 2023). A talk show host with a syndicated radio program, Mark Walters, alleged that when a journalist researching a real lawsuit asked OpenAI's ChatGPT to summarize the complaint that the Second Amendment Foundation ("SAF") filed against Washington Attorney General, ChatGPT responded that Walters was a defendant in a made-up lawsuit. Even worse, in response to the journalist's further prompting, ChatGPT elaborated on the fake lawsuit, stating that the complaint accused Walters of defrauding and embezzling funds from the SAF. In May 2025, the court granted OpenAI's motion for summary judgment. It found that, under the circumstances present in that case, a reasonable reader in the journalist's position could not have concluded that the output was communicating actual facts.

b. Applying Defamation Law to AI-Generated Statements

To prevail on a defamation claim, it is not enough to show that the LLM software hallucinated harmful, false statements. The plaintiff must establish that there was: (1) a publication to a third party, (2) of a false statement of fact about the plaintiff, (3) made with the requisite mental state, (4) that caused harm to the plaintiff's reputation. We take these elements in turn to discuss how traditional defamation elements may be applied to AI-generated false speech.

Publication of a statement that is capable of a defamatory meaning. Defamation law does not require the publication of the allegedly defamatory statement to a broad public audience. If AI companies that program LLMs are deemed to be the speakers, rather than the user entering the prompt, then the publication element may be easily established when the AI software communicates a response containing the allegedly defamatory statement to the user who entered the prompt.

The more difficult question is whether the context of the speech, an LLM generated answer to a user prompt, can convey statements of fact (and not fiction) to a reasonable user. In *Walters*, OpenAI has argued that the journalist ***knew*** ChatGPT was giving fictitious responses, and yet, continued to ask ChatGPT to provide additional gap-filling details about a lawsuit he knew to be fictitious. There was no publication, OpenAI argued, because the user knew the output information was made up. This argument relies on the well-established principle that if the reader knows that the speech was not conveying facts—for example, statements made in a heated argument, as a joke, or puffery—the statement is not capable of defamatory meaning and therefore not actionable.

The Georgia state court agreed with OpenAI. The court noted that: (1) ChatGPT had notified the journalist–user that the software could not access the complaint at the provided link, (2) ChatGPT also had provided multiple warnings that the output may contain "factually inaccurate information,"(3) the journalist–user had a copy of the complaint to verify the truth of the output's statement, and (4) the journalist–user in fact had realized that the output contained

"fantasized" statements. Under these facts, the court found that no reasonable user reading the output in the journalist's position could have concluded that ChatGPT was communicating ***facts***.

But that case did not present the question of what happens if the user does not know that the statements are hallucinations, *i.e.*, fake. Other cases may do so. For example, a manufacturer of solar panels and systems sued Google in March 2025, alleging that the "AI Overview" section on top of Google search results had falsely stated that the Minnesota Attorney General had sued the company for violating state law. *See LTL LED, LLC v. Google LLC*, Case No. 25-cv-02394 (D. Minn.), Dkt. No. 1-1 (Compl.). Here, the AI-generated summaries were directed to the general public, not a particular user with knowledge of what had happened. Even worse, the AI overviews linked to articles and other sources that, although they did not confirm that the state attorney general had sued the solar company, allegedly suggested to the public that the statements in the AI overviews are intended to communicate facts. The case is in its early stages. The courts will also need to decide whether the public's general awareness that AI applications, at times, hallucinate or AI companies' express notice that AI applications can make up gap-filling information means that the model's response to a prompt is not capable of defamatory meaning. That finding would exclude AI-generated statements from defamation liability altogether. Courts may not be willing to go that far. Further, despite the occasional hallucinations, to the public, AI tools are increasingly recognized as powerful, accurate, and reliable research tools.

Falsity. An argument that the allegedly defamatory statement is, in fact, true is one of the most powerful and common defenses. But in most AI defamation cases, the falsity of the statements is unlikely to be disputed.

Fault (Mental State). The bigges-t challenge for a plaintiff asserting a defamation claim against an AI company will be establishing the requisite state of mind with respect to the falsity of a spontaneously generated AI statement. For a public official, public figure, or limited purpose public figure plaintiff, the required degree of fault is actual malice—*i.e.*, that the defendant subjectively knew that the statement was false or entertained serious doubts about its truth. A private citizen, too, must show falsity, but only under a negligence standard.

These falsity standards, addressing the mental state of a human speaker, are not easily applicable to AI-generated speech. It seems improbable to say that probabilistic LLM ***models*** knew the truth or falsity of the next sequence of words it was predicting, much less entertained serious doubts about its truth. Are LLMs even capable of having sufficient "consciousness" to meet the standard? And how can the AI companies—which do not control either the user's prompt or the spontaneous expression AI generates in response to that prompt—possess knowledge of falsity or entertain doubts about the truth of statements they did not know the LLMs would generate?

In the OpenAI case, the court rejected the plaintiff's argument that merely releasing an AI software that was ***capable*** of hallucinations can establish fault under either the negligence or the actual-malice standard.

OpenAI's summary judgment motion asked the Georgia state court to find that Walters cannot establish fault. OpenAI argued that Walters, self-designated as the "loudest voice in America" on Second Amendment issues, is a public figure who must establish actual malice (not negligence). The court agreed. The court also agreed with OpenAI that Walters could not meet the high standard of actual malice, because the evidence OpenAI had offered (on its extensive mitigation efforts to decrease the likelihood of hallucinations) showed that OpenAI did not act with actual malice.

It is unclear, even for private individuals, how defamation plaintiffs could establish that AI companies were negligent about the false speech. In non-AI defamation cases, courts have

considered how a reasonable publisher would have acted. This suggests that AI companies may be able to negate negligence by showing that they have kept up with the safeguards that other companies in the industry generally adopt. In the OpenAI case, the court noted that, even under the negligence standard, OpenAI's efforts to reduce the likelihood of hallucination can show that OpenAI acted with reasonable care.

The question of AI companies' mental state may ultimately turn on whether the plaintiff asked the AI companies to retract the harmful speech and what the AI companies did in response. Some have posited that AI companies' refusal to remove or disable false statements that were flagged to them could establish the requisite knowledge of falsity. But under traditional defamation principles, mental state is usually measured at the time that the statement is made. We will need to wait for a court to weigh on this. OpenAI's summary judgment motion tried to seize on the fact that neither Walters nor the journalist notified OpenAI about the allegedly defamatory statement before filing the lawsuit, but the court did not address it. Other courts may soon have an opportunity to do so. For example, a case filed in April 2025 against Meta concern LLAMA outputs that asserted that the Plaintiff had participated in the January 6th Capitol attack and was arrested and charged with a misdemeanor for his supposed involvement. *See Starbuck v. Meta Platforms, Inc.*, Case No. N25C-04-283 (Del.), Complaint. His allegations detail his efforts to ask Meta to stop publishing these statements and Meta's failure to comply with his requests. The case tees up how an AI company's response to notice of allegedly false statements can affect its defamation liability. The case is still in its early stages.

Harm. Even though damages for reputational harm can be difficult to quantify, defamation plaintiffs could recover presumed damages in certain circumstances. That remedy, however, may be out of reach for many AI defamation plaintiffs because, even for private individuals, a showing of actual malice is required for the recovery of presumed damages, if the speech relates to a matter of public concern. Because it will be challenging for AI defamation plaintiffs to establish actual malice, the damages in many cases would be limited to actual (or special) damages for the harm that is proximately caused by the publication of the defamatory statement. The limited potential for recovery may discourage some plaintiffs from filing defamation lawsuits for AI-generated statements.

c. Liability for AI-Generated Statements? The Key Takeaways.

Whether AI-generated false speech creates a significant risk for defamation liability will depend on some key questions the courts will continue to address. In addition to the critical guidance courts are starting to offer on how different defamation principles apply to speech in AI outputs, the courts will need to decide another key question about immunity: Are AI companies exempt from defamation liability under section 230 of the Communications Decency Act, which protects the owner of a web site against liability for content placed there at the request of a user? Given that the statute exempts a "provider or user" from being "treated as the publisher or speaker of any information ***provided by another information content provider***," it is unclear whether the courts will deem the companies' own software outputs as information provided by someone else. Here are some current takeaways for both those harmed by and those that could face liability for AI-generated speech:

Current Takeaways for Potential Defamation Plaintiffs: Even though a plaintiff in an AI defamation lawsuit could face some challenges convincing the courts to impose liability based on the principles developed for human speakers, there are some advantages in AI lawsuits that a potential plaintiff should take into account when evaluating a defamation claim.

First, unlike most defamation cases, AI hallucinations are likely to be indisputably false. As noted above, this will deprive the defendants of one of its most powerful defenses—that the

statements are true. Further, where falsity is not disputed, the defendants also lose the ability to take extensive, embarrassing discovery aimed at establishing the truth of the harmful statements. The prospect of facing that discovery, often on sensitive topics, can deter the filing of even meritorious claims. Without that hurdle, plaintiffs may be more inclined to file suit.

Second, even though it may be difficult to establish fault with respect to AI companies, there may be some situations where it becomes easier to establish fault in AI defamation suits, as compared to traditional lawsuits. For example, if a user publishes AI-generated false statements (and becomes a defendant), there will be a contemporaneous record of that user's state of mind. Discovery should yield the series of prompts the user entered to generate the false statement and deposition testimony about those prompts. That discovery would give the plaintiff a tool to inquire into whether the user knowingly misused AI tools to engineer or design prompts that are likely to pressure LLM models to generate false statements. Such documentation of the speaker's state of mind is rarely available in a traditional defamation lawsuit.

Third, to create the best record on the difficult element of fault, the plaintiff should give prompt notice to the AI company about the LLM model's false statements and ask for a retraction and mitigation efforts to prevent additional outputs with those statements. AI companies aim to provide accurate research tools and are incentivized to avoid defamation liability. They may help avoid further republication of the harmful statement without a costly litigation.

Current Takeaways for AI Companies: On the flipside, AI companies should consider implementing a protocol for responding to complaints about false and harmful statements. Regardless of the effect on the plaintiff's ability to establish the fault element, it will not be a good fact if it turns out that the plaintiff notified the AI company about the defamatory statement, before filing suit, but the company did nothing.

AI companies should design the software to provide user notices or warnings in situations where the output may hallucinate or provide inaccurate information. The May 2025 decision from the Georgia state court suggests that providing such warnings to the user—particularly when the prompt is forcing the model to respond to questions based on information it cannot access or falls outside of its training set—may offer protection against the argument that the user understood the output to contain provably false statements of fact.

AI companies should also continue to educate the courts and the public about the nature of probabilistic modelling. Explaining that LLMs as not creating a mosaic of existing text or information but crafting new predictive responses will provide an accurate technological framework on which the courts can apply traditional defamation principles to a new context.

Current Takeaways for Individuals and Companies Using AI Tools: The risk for libel liability is not limited to AI companies. As non-AI companies and media organizations increasingly incorporate generative AI tools to conduct research and create public-facing content, they should consider the legal risk of potentially republishing AI-generated false statements. As the law develops, companies using these powerful tools should continue to engage in common-sense best practices. Before publishing a factual statement identified by an LLM model, one should verify those facts and document that process. As much as possible, companies should also use enterprise, not free versions, of LLM software that contain the up-to-date safeguards that AI companies have put in place. That could help avoid a finding of negligence with respect to the falsity of AI-generated statements.

XI. AI and Employment: Federal and State Regulation

As AI tools have become more advanced, they are beginning to see applications in a wide

variety of employment contexts, from generating job postings to initial resume screenings to conducting fully automated interviews. This shifting AI landscape has been met with a shifting regulatory landscape. In 2023, the Biden Administration issued an executive order outlining rules for the use of artificial intelligence in many areas.[161] However, the Trump Administration withdrew that executive order.[162] Although some litigants have argued that federal employment laws such as Title VII of the Civil Rights Act, the Americans with Disabilities act, and the Age Discrimination in Employment Act all apply to the use of AI in hiring, there is limited federal guidance on how they do so.

Some cases have started to directly litigate claims of AI discrimination, but not enough for a clear framework to emerge. *Mobley v. Workday* involved claims that Workday's hiring screening tools discriminate against "job applicants who are African American, over the age of forty, and/or disabled."[163] *EEOC v. iTutorGroup Inc.,* involved claims that iTutorGroup used application software that automatically rejected female applicants over age 55 and male applicants over age 60.[164] The limited guidance these cases offer are discussed below—but many issues regarding AI and its role in employment discrimination litigation remain uncertain.

In *EEOC v. iTutorGroup*, a woman named Wendy Pincus applied for an English-tutor position with iTutorGroup.[165] The only stated requirement for the role was a bachelor's degree. Upon applying with her full biographical information, she was immediately rejected.[166] However, Ms. Pincus applied the next day with identical information except for a more recent date of birth and was offered an interview.[167] Ms. Pincus then filed a complaint with the EEOC, which sued iTutorGroup on Ms. Pincus's behalf under the ADEA.[168] The case settled with iTutorGroup agreeing to pay compensatory damages and remove the software that automatically rejected older applicants.[169] The case produced no court opinions or meaningful precedent because of the settlement. But the case provides an early example of how regulators may argue that the ADEA and other antidiscrimination laws apply to a company's use of AI tools to make hiring decisions. The case also provides a template future plaintiffs may follow in alleging AI discrimination and seeking evidence to support those claims. If slight changes in information pertaining to a protected class result in different hiring outcomes, employers may face discrimination claims. Employers should consider whether AI screening tools will lead to claims of discrimination when similar applicants see different outcomes where the sole difference between applicants is the applicant's race, sex, age, religion, disability status, or any other protected feature.

Mobley v. Workday, Inc.[170] provides more substantial guidance. There, plaintiff Derek Mobley claims to have applied to over 100 different positions that use Workday's AI screening tools.[171] Workday provides these tools to other companies to help aid their hiring and screening procedures.[172] Mobley is an African American male over the age of forty who struggles with anxiety and depression.[173] He claims that the Workday platform required him to take a "Workday-branded assessment and/or personality test" which are more likely to discriminate and screen out those with anxiety and depression.[174] After his application was rejected from all of the over

161 Exec. Order No. 14110 88 FR 75191 (Nov. 1, 2023).
162 Exec. Order No. 14179 90 FR 8741 (Jan. 23, 2025).
163 *Mobley v. Workday*, 740 F. Supp. 3d 796, 803 (N.D. Cal. 2024).
164 *EEOC v. iTutorGroup*, 2022 WL 20699029 (E.D.N.Y. Aug. 3, 2022).
165 *EEOC v. iTutorGroup*, JVR No. 2310200016 (E.D.N.Y. 2023).
166 *Id.*
167 *Id.*
168 *Id.*
169 *Id.*
170 740 F. Supp. 3d 796 (N.D. Cal. 2024).
171 *Id.* at 802.
172 *Id.*
173 *Id.*
174 *Id.*

100 positions that used Workday's platform, Mobley sued Workday under Title VII of the Civil Rights Act of 1964, the ADEA, and the ADA for intentional discrimination and disparate impact discrimination.[175]

The court granted in part and denied in part Workday's motion to dismiss. The court found that Workday may not be sued as an "employment agency" because Workday does not "bring[] job listings to the attention of those looking for employment."[176] However, the court allowed Mobley to proceed on a theory that Workday acted as an agent of the various employers.[177] According to the court, by embedding artificial intelligence and machine learning into its decision making tools, Workday is "participating in the decision-making process" rather than simply "implementing in a rote way the criteria that employers set forth."[178] Therefore, Workday was acting as an agent of the employer, who "may be held independently liable" under Title VII, the ADA, and the ADEA.[179]

This ruling is significant, because it holds that both employers who use AI tools, as well as the company that provides the AI tools, may each face claims under federal discrimination law. The case also suggests defendants may face challenges when claiming the decisional logic of an AI tool is a "black box" to the user.[180] This may support the idea that discrimination by AI is the same as any other manner of discrimination under federal discrimination law. However, the ruling does limit some liability to AI companies, by holding that they are not "employment agencies," although they can still be categorized as an agent of employers.

The *Mobley* court went on to dismiss claims based on intentional discrimination while allowing claims based on disparate impact discrimination.[181] The court held that Workday's knowledge of the discriminatory effects of its screening tools was not enough to rise to the level of intentional discrimination.[182] But the court allowed disparate impact claims because Mobley was able to "(1) show a significant impact on a protected class or group; (2) identify the specific employment practices or selection criteria at issue; and (3) show a causal relationship between the challenged practices or criteria and the disparate impact."[183] Importantly, Mobley was able to sufficiently plead disparity and causation because he applied to over 100 jobs rather than a single job, had a zero percent success rate, and presented academic literature demonstrating how bias can affect algorithms.[184] As AI hiring tools become more prevalent, more plaintiffs may be able to allege disparities across large sample sizes like Mobley.

On May 29, 2026, the *Mobley* court denied Mobley's motion to compel production of Workday's bias-testing data, finding it was privileged because Workday's attorneys compiled the bias-testing data and used the results to provide legal advice.[185] On the other hand, the court granted Mobley's motion to compel Workday's reports to federal agencies about its workforce demographics.[186] In doing so, the court recognized that, although these reports concern

---

175 *Id*.
176 *Id.* at 808.
177 *Id.* at 806.
178 *Id.* at 807.
179 *Id.* at 804.
180 *Id.* at 807 ("Nothing in the language of the federal anti-discrimination statutes or the case law interpreting those statutes distinguishes between delegating functions to an automated agent versus a live human one. To the contrary, courts applying the agency exception have uniformly focused on the 'function' that the principal has delegated to the agent, not the manner in which the agent carries out the delegated function.").
181 *Id* at 811-813.
182 *Id.*
183 *Id.* at 809. (Citing *Bolden-Harge v. Off. Of Cal. State Controller*, 63 F.4th 1215, 1227 (9th Cir. 2023)).
184 *Id.* at 811.
185 *Mobley v. Workday, Inc.*, No. 23-CV-00770-RFL (LB), 2026 WL 1510537, at *4 (N.D. Cal. May 29, 2026).
186 *Id.* at *7.

Workday's own hiring processes—not those of their clients—the reports nonetheless show Workday's knowledge of potential demographic disparities when using AI tools.[187]

Another recently-filed class action against Eightfold AI alleges violations of the Fair Credit Reporting Act, the Investigative Consumer Reporting Agencies Act, and the Unfair Competition Law for the deployment of software that uses AI to evaluate job applications and offer employers recommendations based on applicant data.[188] Plaintiffs allege that Eightfold AI's tools engage in AI-assisted collection of job applicants' personal information—including social media profiles, location data, cookies, and internet and device activity—to provide employers using the product with "final decisions" and "likelihood of success" scores for job applicants.[189] Eightfold AI removed the case to federal court in the Northern District of California, and it is still in the pleadings stage.[190]

Moreover, federal courts have applied various state privacy laws to the use of AI tools in hiring. In *Baker v. CVS Health Corporation*, plaintiffs brought a class action alleging that CVS's use of AI in the application process violated Massachusetts state laws against lie detector tests.[191] The court denied CVS's motion to dismiss, noting that the HireVue interview tools are advertised as being able to detect "whether an applicant 'has an innate sense of integrity and honor,'" and can help with "lie detection" and "screening out embellishers."[192] Therefore, because the plaintiffs did not receive notice of a lie detector test, they may have suffered an injury under the state statute as pled.[193]

In *Deyerler v. HireVue, Inc.*, plaintiffs accused HireVue interview software of violating the Illinois BIPA.[194] The plaintiffs claimed that HireVue's analysis of video interviews captures biometric data from applicants without providing notice or describing guidelines for destroying said data pursuant to the state law.[195] HireVue disputed that their software does in fact capture or store any biometric data and filed a motion to dismiss. The court denied HireVue's motion and held that the plaintiffs had adequately pled a claim under state law. This string of cases demonstrates that plaintiffs may be able to pursue claims that various state privacy laws apply to the use of AI in screening and hiring procedures.

While the federal government appears unlikely to regulate the use of AI in hiring in the near term, several state and local governments have passed or are considering legislation that restricts and places conditions on the use of AI in employment and hiring decisions. Illinois passed the Artificial Intelligence Video Interview Act that requires employers to disclose their use of AI to analyze video interviews and obtain consent to evaluate applicants with AI.[196] Under the amended Illinois Human Rights Act, Illinois employers cannot use AI systems that generate discriminatory outcomes based on protected characteristics, regardless of intent.[197] Employers must notify applicants and employees when they use AI in hiring or employment decisions.[198] Maryland has a facial recognition technology law that prohibits the use of facial recognition services during job interviews without applicant consent.[199] Texas passed the Texas Responsible AI Governance Act, a more employer-friendly law that bars *intentional* AI-based discrimination

---

187 *Id.* at *6.
188 Compl. ¶ 1, 125-155, *Kistler v. Eightfold AI Inc.*, 26-cv-01768 (N.D. Cal. Mar. 2, 2026).
189 *Id.* at ¶¶ 1, 8.
190 Notice of Removal, *Kistler v. Eightfold AI Inc.*, 26-cv-01768 (N.D. Cal. Mar. 2, 2026).
191 717 F. Supp. 3d 188 (D. Mass. 2024).
192 *Id.* at 191.
193 *Id.* at 192-193.
194 2024 WL 774833 (N.D. Ill. Feb. 26, 2024); 740 ILCS 14/1, *et seq.*
195 *Id.* at *6.
196 820 ILL. COMP. STAT. 42/5 (2020).
197 775 ILL. COMP. STAT. 5/2-102. § (L)(1) (2025).
198 *Id.* § (L)(2).
199 Md. Code Ann. Lab. & Empl. § 3-717.

and grants employers sixty days upon notice to cure violations.[200]

Connecticut passed the Connecticut Artificial Intelligence Responsibility and Transparency Act, which regulates the use of AI-driven tools in employment decision-making.[201] Developers of the AI tools must provide deployers with adequate information to ensure legal compliance.[202] The law also requires deployers to notify employees or applicants when they use AI-driven tools unless it would be reasonably obvious to a person that they are doing so.[203] In the event an AI-driven tool generates employment decisions, deployers must issue a written notice to the impacted individual before a decision is made.[204]

Colorado repealed and replaced the 2024 Colorado AI Act, showing that legislation in this space is quickly evolving.[205] Effective January 1, 2027, the new law focuses on "automated decision-making technology" (ADMT) used to materially influence a consequential decision in areas such as employment.[206] An ADMT materially influences a consequential decision when it affects the outcome of the decision by constraining, ranking, scoring, recommending, or meaningfully altering how the decision is made.[207] When a consequential decision from an ADMT (which *does not* include summaries of a candidate's profile for human review)[208] results in an adverse outcome, an individual designated by the deployer who has authority to approve, modify, or override a consequential decision, and who satisfies another four criteria, must conduct "meaningful human review."[209] This new law also includes a sixty-day notice and cure period unless the attorney general can show that the employer knowingly and repeatedly violated the law.[210]

Proposed legislation in New York would also place restrictions around an employer's use of AI to screen candidates, including by conducting an impact assessment prior to the system being put into use.[211] This proposed state legislation lags behind New York City, which in 2021 passed a local law prohibiting employers from using an automated employment decision tool to assess candidates unless an independent auditor completes a bias audit of the tool within one year of its use, information about the bias audit is made publicly available, and the employer provides notice to the candidates of the use of the tool.[212]

California has taken significant steps to regulate the use of automated decision-making in employment contexts. In October 2025, California amended the Fair Employment and Housing Act (FEHA), prohibiting employers from deploying "Automated-Decision Systems" (ADS) that produce discriminatory hiring or employment outcomes.[213] Under FEHA, covered employers must retain ADS-related data for a minimum of four years.[214] The California Privacy Protection Agency issued regulations on "Automated Decision-Making Technology" (ADMT), requiring advance notice to employees and candidates if ADMT is used, disclosure of how this technology functions, and notice to individuals on their right to opt-out.[215]

---

200 Tex. Bus. & Com. Code Ann. §§ 552.056(b), 552.104 (2024).
201 Conn. S.B. 5, Act 26-15 (2026) (effective Oct. 1, 2026).
202 *Id.* at § 8(a).
203 *Id.* at § 9.
204 *Id.* § 10.
205 CO Senate Bill 26-189.
206 *Id.* at § 6-1-1701(5).
207 *Id.* at § 6-1-1701(13)(a)(II).
208 *Id.* at §§ 6-1-1701(2)(b)(II), 6-1-1701(3)(b)(IV).
209 *Id.* at §§ 6-1-1701(15), 6-1-1705(1)(a)(II).
210 *Id.* at § 6-1-1706(3)(a)-(c).
211 N.Y. Senate Bill S9401; Senate Bill S7623A.
212 N.Y.C. Admin. Code §§ 20-870, 20-871, 20-872, 20-873.
213 Cal. Admin. Code tit. 2 §§ 11017(e), 11028(m), 11039(1)(J).
214 Cal. Admin. Code tit. 2 § 11013(c).
215 Cal. Admin. Code tit. 2 § 7220(a), (c).

XII. Federal Securities Regulations: "AI Washing"

The term "AI washing"—derived from "greenwashing" (cases in which companies were alleged to make misleading claims about their environmental impact)—describes companies making alleged materially false or misleading claims about their artificial intelligence capabilities.[216] What may begin as marketing hyperbole can result in regulatory penalties, private civil/class action exposure, and, at the extremes, potential criminal exposure.

Most commonly, public and private companies that misrepresent their artificial intelligence capabilities in connection with the offer or sale of securities face potential legal exposure under long-established anti-fraud provisions of the federal securities laws.[217] Such alleged misrepresentations may be challenged by the SEC or private litigants, although, critically, the SEC and private litigants are subject to different pleading requirements.[218]

This section discusses the emergence of federal securities claims based on AI washing, highlighting patterns in the SEC's recent AI washing enforcement actions, and then focusing on the application of the elements of the key anti-fraud provision, and related defenses, to AI washing claims in the context of the private securities class actions against Apple regarding its representations about its "Apple Intelligence" features.[219]

a. SEC "AI Washing" Enforcement Actions

MMC Ventures documented the phenomenon's scope in its foundational study, finding that 40% of 2,830 European startups claiming to be "AI companies" showed no evidence of AI material to their value proposition.[220] The regulatory response crystallized in 2023 when SEC Chair Gary Gensler cautioned: "Don't do it . . . One shouldn't greenwash, and one shouldn't AI wash. I don't know how else to say it."[221]

The SEC brought its first enforcement actions regarding AI washing on March 2024, settling charges against investment advisers Delphia Inc. ($225,000 penalty) and Global Predictions Inc. ($175,000 penalty) for false claims about using client data in investment algorithms that never existed.[222]

---

[216] *See* "Businesses Should Scrub Public Statements, as the SEC Focuses on 'AI Washing.'" Quinn Emanuel Government & Regulatory Litigation Update (May 2024), https://www.quinnemanuel.com/the-firm/publications/government-regulatory-litigation-update-may-2024.

[217] We note that the FTC may be able target the same kinds of misrepresentations under its consumer protection framework and announced in September 2024 the creation of Operation AI Comply, a "law enforcement sweep" targeting "multiple companies that have relied on artificial intelligence as a way to supercharge deceptive or unfair conduct that harms consumers." FTC, *FTC Announces Crackdown on Deceptive AI Claims and Schemes*, https://www.ftc.gov/news-events/news/press-releases/2024/09/ftc-announces-crackdown-deceptive-ai-claims-schemes. But that framework is independent of the federal securities laws and regulations, and there is no evidence to date that the SEC and FTC are coordinating regarding "AI washing."

[218] For example, to establish intentional fraud, a private litigant must plead additional elements that the SEC is not required to plead, and the SEC, but not private litigants may pursue claims based on negligent fraud.

[219] *Tucker et al. v. Apple Inc., et al.*, Case 3:25-cv-05197 (N.D. Cal.), https://assets.law360news.com/2355000/2355921/https-ecf-cand-uscourts-gov-doc1-035125880637.pdf.

[220] Will Knight, *About 40% of Europe's "AI Companies" Don' Use Any AI At All*, MIT Technology Review (Mar. 5, 2019), https://www.technologyreview.com/2019/03/05/65990/about-40-of-europes-ai-companies-dont-actually-use-any-ai-at-all/.

[221] *See* Richard Vanderford, *SEC Head Warns Against 'AI Washing,' the High-Tech Version of 'Greenwashing'*, WSJ (Dec. 5, 2023, at 5:15 ET), https://www.wsj.com/articles/sec-head-warns-against-ai-washing-the-high-tech-version-of-greenwashing-6ff60da9; Hailey Konnath, *SEC Chair Warns Businesses Against AI Washing: 'Don't Do It'*, Law360 (Dec. 5, 2023, 23:35 ET), https://www.law360.com/articles/1773759.

[222] Press Release, SEC, *SEC Charges Two Investment Advisers with Making False and Misleading Statements About Their Use of Artificial Intelligence*, No. 2024-36 (Mar. 18, 2024), available at https://www.sec.gov/newsroom/press-releases/2024-36.

1. What *Saniger* Means for AI-Disclosure Fraud

The Securities and Exchange Commission has been bringing fraud cases against tech companies for lying to investors about their core technology for years. The sequence is well established: materially misrepresent what your product does, raise capital on the back of those claims, and face the consequences when the truth emerges. The SEC has applied that framework across industries and technology types alike, treating false statements about a company's core technology as securities fraud regardless of how novel the technology may be.

Now, given AI's explosive growth and its allure in the capital markets, the SEC has its eyes on the sector. But the Commission is not writing new law. It is applying familiar theories to a new class of companies. This update surveys the SEC's recent AI-disclosure enforcement actions in the context of that longer pattern, and draws out the practical implications for companies that build or market AI-related products.

2. The Enforcement Pattern

The Commission's technology-fraud cases span across a variety of industries and stretch back well before the current interest in artificial intelligence. Perhaps most famously, in 2018 the SEC took action against the blood-testing company, Theranos, and its founder, Elizabeth Holmes. In *SEC v. Holmes*, the Commission alleged that Theranos raised more than $700 million on the representation that its proprietary analyzer could run a broad array of diagnostic tests from a few drops of blood, when in truth the device performed only a limited number of tests and most samples were processed on modified, commercially available machines. Holmes and the company settled without admitting or denying the findings, with Holmes agreeing to a $500,000 penalty, a ten-year officer-and-director bar, the return of 18.9 million shares, and the surrender of her voting control.[223]

Three years later, the Commission turned to the electric- and hydrogen-fuel-cell vehicle maker Nikola Corporation. In a settled order, the SEC alleged that, before Nikola had produced a single commercial product, the company's founder, Trevor Milton, gave investors the false impression that Nikola had reached key technological and production milestones. According to the SEC, Milton misled investors in tweets, media appearances, and even one promotional video depicting a prototype seemingly driving down a road when it was, in fact, rolling down an incline using gravity. Nikola resolved the matter for a $125 million penalty on a neither-admit-nor-deny basis, and the SEC pursued the founder separately in *SEC v. Milton*.[224]

Later, in *SEC v. Perryman*, the Commission charged the former chief executive of medical-device startup Stimwave with defrauding investors of roughly $41 million by misrepresenting the company's nerve-stimulation device. According to the complaint, one of the device's implanted components was a non-functioning piece of plastic. The executive falsely told investors the device had been approved by the FDA. Related conduct was also the subject of a parallel criminal

---

[223] Press Release, SEC, *Theranos, CEO Holmes, and Former President Balwani Charged With Massive Fraud*, No. 2018-41 (Mar. 14, 2018), available at https://www.sec.gov/newsroom/press-releases/2018-41; Complaint, *SEC v. Holmes*, No. 5:18-cv-01602 (N.D. Cal. Mar. 14, 2018), available at https://www.sec.gov/files/litigation/complaints/2018/comp-pr2018-41-theranos-holmes.pdf.

[224] Press Release, SEC, *Nikola Corporation to Pay $125 Million to Resolve Fraud Charges*, No. 2021-267 (Dec. 21, 2021), available at https://www.sec.gov/newsroom/press-releases/2021-267; *In re Nikola Corp.*, Securities Act Release No. 11018, Exchange Act Release No. 93838, Admin. Proc. File No. 3-20687 (Dec. 21, 2021), available at https://www.sec.gov/files/litigation/admin/2021/33-11018.pdf; Complaint, *SEC v. Milton*, No. 1:21-cv-06445 (S.D.N.Y. July 29, 2021), available at https://www.sec.gov/files/litigation/complaints/2021/comp-pr2021-141.pdf. Milton was also indicted for the same conduct in the Southern District of New York, but he received a full and unconditional presidential pardon in March 2025. *Executive Grant of Clemency to Trevor Milton* (Mar. 27, 2025), available at https://www.justice.gov/pardon/media/1395001/dl.

prosecution by the U.S. Attorney's Office for the Southern District of New York.[225]

Though these cases span different industries and technologies, they follow a familiar pattern: in each, the company claimed its technology could do something it either could not do at all, or could only do through means it was not disclosing—manual processes or third-party equipment. The SEC treated those disclosure failures as securities fraud, regardless of the technology involved.

3. Application To Artificial Intelligence

Against that backdrop, the SEC's AI-related cases are best understood not as a new body of law but as the same principles applied to a new, fast-growing sector. The SEC's actions to date have ranged from settled civil proceedings against companies to parallel civil and criminal charges against individual corporate officers.

The earliest arrived in March 2024, when the Commission settled proceedings premised on an AI-washing theory[226] against two investment advisers, Delphia (USA) Inc. and Global Predictions, Inc. Both firms were alleged to have marketed AI-driven investment capabilities they did not possess. According to the SEC's orders,[227] Delphia claimed to use machine learning and artificial intelligence to "predict which companies and trends are about to make it big," and Global Predictions billed itself as the "first regulated AI financial advisor" offering "[e]xpert AI-driven forecasts." The SEC found those statements false and misleading and that neither entity implemented appropriate compliance procedures to safeguard against those types of misstatements, charging both under the Investment Advisers Act. Without admitting or denying the findings, the firms were censured, agreed to cease and desist, and paid civil penalties of $225,000 and $175,000, respectively. Announcing the settlements, then-Chair Gary Gensler warned that advisers "should not mislead the public by saying they are using an AI model when they are not," because "[s]uch AI washing hurts investors."[228]

The government has also pursued parallel criminal actions against a corporate officer for alleged misstatements concerning AI-related products. In *SEC v. Roberts*, the Commission alleged that the former chairman and CEO of advertising-technology firm Kubient fabricated results for its flagship product. That product, KAI, was marketed as artificial-intelligence software that detected when advertisement views occurred, not by humans, but by software programs designed to artificially inflate views, and therefore the advertising price. According to the SEC, Kubient's IPO Offering Materials stated that KAI "was able to successfully ingest hundreds of millions of rows of data in real-time" which prevented "approximately 300% more digital ad fraud than" competing products despite there being "no actual data" analyzed by the company. The U.S. Attorney's Office for the Southern District of New York separately charged the former CEO, who pleaded guilty and was sentenced to one year and one day in prison.[229]

---

[225] Press Release, SEC, *SEC Charges Former CEO of Medical Device Startup Stimwave with $41 Million Fraud*, No. 2023-255 (Dec. 19, 2023), available at https://www.sec.gov/newsroom/press-releases/2023-255; Complaint, *SEC v. Perryman*, No. 1:23-cv-10985 (S.D.N.Y. Dec. 19, 2023), available at https://www.sec.gov/files/litigation/complaints/2023/comp-pr2023-255.pdf; Indictment, *United States v. Perryman*, No. 23-cr-00117 (S.D.N.Y. Mar. 9, 2023), available at https://www.justice.gov/d9/press-releases/attachments/2023/03/09/u.s._v._perryman_indictment_0_0.pdf.

[226] The term "AI washing" derives from "greenwashing," which describes false or exaggerated claims about the environmental benefits of a company's products or practices.

[227] *In re Delphia (USA) Inc.*, Investment Advisers Act Release No. 6573, Admin. Proc. File No. 3-21894 (Mar. 18, 2024), available at https://www.sec.gov/files/litigation/admin/2024/ia-6573.pdf; *In re Global Predictions, Inc.*, Investment Advisers Act Release No. 6574, Admin. Proc. File No. 3-21895 (Mar. 18, 2024), available at https://www.sec.gov/files/litigation/admin/2024/ia-6574.pdf.

[228] Press Release, SEC, *SEC Charges Two Investment Advisers with Making False and Misleading Statements About Their Use of Artificial Intelligence*, No. 2024-36 (Mar. 18, 2024), available at https://www.sec.gov/newsroom/press-releases/2024-36.

[229] Press Release, SEC, *SEC Charges Former Chairman and CEO of Tech Co. Kubient With Fraud and Lying to Auditors*, No. 2024-131 (Sept. 16, 2024), available at https://www.sec.gov/newsroom/press-releases/2024-131; Complaint, *SEC v. Roberts*, No. 24-cv-06990 (S.D.N.Y. Sept. 16, 2024), available at https://www.sec.gov/files/litigation/complaints/2024/comp-pr2024-131a.pdf; Information, *United States v. Roberts* (S.D.N.Y. Sept. 16, 2024), available at https://www.justice.gov/usao-sdny/media/1368126/dl;

Perhaps the most aggressive government enforcement of AI-washing theory specifically came in April 2025, when the SEC charged Albert Saniger, the founder and former chief executive officer of the e-commerce startup Nate, Inc., in a fraud action filed in the Southern District of New York,[230] and the U.S. Attorney's Office for the same district indicted Saniger in a parallel criminal action.[231] According to the complaint and the indictment, Saniger raised more than $40 million by representing that Nate's mobile shopping app used proprietary AI to complete online purchases autonomously, "without human intervention." Saniger allegedly contrasted competitors' "dumb bots" with Nate's purported intelligence, assuring one investor that "[N]ate is an intelligent machine," "not a bot, or a combo of bots," and claiming a "success" rate of 93% to 97%. In reality, the government alleges, Nate's automation rate was "effectively zero": orders were placed manually by contractors in a Philippines call center, later supported by a center in Romania. Saniger allegedly concealed the manual operation—restricting access to an automation dashboard and prioritizing investors' orders so they would appear seamless—and when human operation alone proved insufficient, directed his team to build the type of bots he had previously disclaimed. Even after a June 2022 news article questioned Nate's use of AI, Saniger purportedly represented that Nate "only relied on 'humans-in-the-loop' for payments risk, data labeling, reinforcement learning training, and purchase completion for certain 'edge cases.'" A year later, Nate, Inc. collapsed. The criminal indictment charges Saniger with one count of securities fraud and one count of wire fraud; the SEC's civil complaint charges him with violations of Section 17(a) of the Securities Act, Section 10(b) of the Exchange Act, and Rule 10b-5, seeking an injunction barring further violations, an officer-and-director bar, disgorgement, and civil penalties.[232]

The *Saniger* case is worth watching closely. It is the first AI-disclosure matter proceeding toward trial rather than settling, which means it will be the first to test these theories in adversarial litigation. Because prior cases have been resolved on a neither-admit-nor-deny basis, the outer limits of AI-washing liability are left ill-defined. If *Saniger* goes to verdict, the result will either validate the enforcement approach or expose its limits—and in either case, it will give companies and practitioners their first real look at how these cases are argued and decided.

4. Implications for AI Companies

For companies that build or market AI, the through-line of this enforcement history is more instructive than any single case. Across every matter discussed here—from *Theranos* to *Nate*—the underlying conduct follows the same pattern: the company claimed its technology could do something autonomously, at scale, or with a degree of sophistication it either did not possess or was achieving through undisclosed means. The gap between the claimed capability and the reality is what drew enforcement. AI companies are squarely within that pattern, and the SEC's prior leadership made the point plainly, describing one AI case as "old school fraud using

---

Press Release, U.S. Att'y's Off., S.D.N.Y., *Former CEO Of Kubient, Inc. Sentenced To Prison In Connection With Accounting Fraud Scheme*, No. 25-061 (Mar. 20, 2025), available at https://www.justice.gov/usao-sdny/pr/former-ceo-kubient-inc-sentenced-prison-connection-accounting-fraud-scheme; *see also* Indictment, *United States v. Chidambaran*, No. 26-cr-97 (E.D.N.Y. Apr. 15, 2026), available at https://www.justice.gov/usao-edny/media/1436506/dl?inline (criminal indictment alleging corporate officers made materially misleading statements regarding finances of technology company claiming to provide AI-driven business automation solutions).

[230] Complaint, *SEC v. Saniger*, No. 1:25-cv-02937 (S.D.N.Y. Apr. 9, 2025), available at https://www.sec.gov/files/litigation/complaints/2025/comp26282.pdf; SEC Litig. Rel. No. 26282, *SEC Charges Founder and Former CEO of Artificial Intelligence Startup with Misleading Investors* (S.D.N.Y. Apr. 11, 2025), available at https://www.sec.gov/enforcement-litigation/litigation-releases/lr-26282.

[231] Indictment, *United States v. Saniger*, No. 25-cr-00157 (S.D.N.Y. Apr. 9, 2025), available at https://www.justice.gov/usao-sdny/media/1396131/dl?inline; *see also* Press Release, U.S. Att'y's Off.S.D.N.Y., *Tech CEO Charged In Artificial Intelligence Investment Fraud Scheme*, No. 25-082 (Apr. 9, 2025), available at https://www.justice.gov/usao-sdny/pr/tech-ceo-charged-artificial-intelligence-investment-fraud-scheme.

[232] Both cases remain in their early stages. In his criminal case, Saniger—a dual citizen of Spain and the United States—voluntarily waived extradition, pleaded not guilty, and was released on a $250,000 bond. Pretrial motions are due in September 2026 and a status conference is set for November 5, 2026. The SEC's civil case advanced more slowly because Saniger had to be served in Spain under the Hague Convention. With service now effected, the district court has scheduled an initial conference for June 26, 2026.

new school buzzwords."[233] On this front, the current Commission's view is little changed. The SEC's current Director of the Division of Enforcement has described his philosophy as a "back to basics" approach, centered on offering, accounting, and disclosure fraud—the same theories implicated when a company misstates material facts about its products or technology.[234] Companies should expect this enforcement pattern to continue, with the Commission treating AI-washing cases consistent with any other material misstatement.[235]

The practical lessons are familiar. Companies should ensure that any public statements about their use of AI are accurate and supported by a reasonable basis before they raise capital, report to the markets, or market their services, and should take particular care with concrete, verifiable metrics. They should maintain policies and procedures reasonably designed to ensure the accuracy of those statements, recognizing that the SEC has faulted not only the claims themselves but the absence of adequate controls. And they should appreciate that exposure is not limited to the entity: founders, executives, and gatekeepers have been named in these matters, in civil and criminal proceedings alike. The potential costs for individuals—monetary penalties, disgorgement, officer-and-director bars, and in some cases imprisonment—are substantial.[236]

b. Private "AI Washing" Litigation

1. The Apple Litigation: The Latest Test Case

The Apple litigation presents a critical test case for how courts will evaluate AI-related representations in the federal securities context generally and apply the elements of Rule 10b–5 claims by private litigants to alleged AI-related misrepresentations specifically . The consolidated actions against Apple, brought by Apple investors, center on its June 2024 announcement of "Apple Intelligence" features, particularly enhanced Siri capabilities that would purportedly enable "hundreds of new actions" across applications and leverage "personal context" for tailored user experiences.

The complaints allege a fundamental disconnect between Apple's public representations and its internal development reality. According to the pleadings, when Apple showcased these features at its 2024 Worldwide Developer Conference (WWDC)—demonstrating Siri's supposed ability to access personal information, take contextual actions across apps, and understand on-screen content—the company lacked a functional prototype. The significance of this alleged misrepresentation is underscored by internal statements attributed to Robby Walker, senior director responsible for Siri, who purportedly acknowledged that Apple promoted the technology "before it was ready" and characterized the situation as "ugly and embarrassing."

The market impact allegations follow a clear pattern: Apple's initial representations drove investor expectations for the iPhone 16 cycle, with the advanced Siri features positioned as a "key selling point" for devices that otherwise lacked major hardware innovations. The subsequent

---

[233] Press Release, SEC, *SEC Charges Founder of AI Hiring Startup Joonko With Fraud*, No. 2024-70 (June 11, 2024), available at https://www.sec.gov/newsroom/press-releases/2024-70.

[234] David Woodcock, Dir., Div. of Enf't, SEC, Remarks at the MFA Legal & Compliance 2026 Conference (May 13, 2026), https://www.sec.gov/newsroom/speeches-statements/woodcock-remarks-mfa-legal-compliance-2026-conference-051326.

[235] The Commission's retooling of its specialized units—replacing its crypto-and-cyber unit with a Cyber and Emerging Technologies Unit whose remit expressly includes fraud involving artificial intelligence—further signals institutional attention to the AI space, even if the broader program scales back on overall enforcement. Press Release, SEC, *SEC Announces Cyber and Emerging Technologies Unit to Protect Retail Investors*, No. 2025-42 (Feb. 20, 2025), available at https://www.sec.gov/newsroom/press-releases/2025-42.

[236] More broadly, SEC enforcement is not the only consequence of inaccurate AI disclosures: the same representations that draw regulatory scrutiny increasingly give rise to private securities class actions. Such private AI-related securities class actions have proliferated, with fifteen such actions in 2024 and sixteen in 2025—both more than double the seven filed in 2023. Cornerstone Research, *Securities Class Action Filings: 2025 Year in Review* at 5 (2026), available at https://www.cornerstone.com/wp-content/uploads/2026/01/Securities-Class-Action-Filings-2025-Year-in-Review.pdf.

delays—first acknowledged in March 2025 when Apple stated the features were "taking longer than [the Company] thought"—allegedly caused material stock price declines. Morgan Stanley's analysis that approximately 50% of consumers who didn't upgrade to iPhone 16 attributed their decision to the delayed rollout provides empirical support for the materiality allegations.

2. How Courts Evaluate AI Washing Allegations Under Rule 10b-5 in Private Securities Litigation

While the Apple litigation remains pending, recent decisions provide guidance on how courts analyze AI-related representations under Rule 10b-5 in private securities litigation.

Rule 10b–5, which implements the anti-fraud provisions of section 10(b) of the Securities Exchange Act, makes it unlawful for any person, directly or indirectly, to misstate or omit a material fact in connection with the purchase or sale of securities. 17 C.F.R. § 240.10b–5. To state a claim for securities fraud, a private litigant must allege:

> (1) a material misrepresentation or omission by the defendant;
>
> (2) scienter;
>
> (3) a connection between the misrepresentation or omission and the purchase or sale of a security;
>
> (4) reliance upon the misrepresentation or omission;
>
> (5) economic loss; and
>
> (6) loss causation.

Halliburton Co. v. Erica P. John Fund, Inc., 134 S.Ct. 2398, 2407 (2014) (citations omitted).

Defenses like puffery go to Section 10(b)'s materiality requirement, which is intended "to filter out essentially useless information that a reasonable investor would not consider significant, even as part of a larger 'mix' of factors to consider in making his investment decision." *Basic, Inc. v. Levinson*, 485 U.S. 224, 234 (1988) (citation omitted). Thus, statements involving material facts—i.e., facts where there "is a substantial likelihood that a reasonable shareholder would consider it important in deciding how to vote" or invest, *id*. at 231 (quoting *TSC Indus., Inc. v. Northway, Inc.*, 426 U.S. 438, 449 (1976))—are actionable, but "[i]mmaterial statements [like] vague, soft, puffing statements or obvious hyperbole" are not. *In re K-tel Int'l, Inc. Sec. Litig.*, 300 F.3d 881, 897 (8th Cir. 2002); *see also In re Ford Motor Co. Sec. Litig.*, 381 F.3d 563, 570–71 (6th Cir. 2004) ("Courts everywhere 'have demonstrated a willingness to find immaterial as a matter of law a certain kind of rosy affirmation commonly heard from corporate managers and numbingly familiar to the marketplace—loosely optimistic statements that are so vague, so lacking in specificity, or so clearly constituting the opinions of the speaker, that no reasonable investor could find them important to the total mix of information available.' " (quoting *Shaw v. Digit. Equip. Corp.*, 82 F.3d 1194, 1217 (1st Cir. 1996))).

The Private Securities Litigation Reform Act of 1995 ("PSLRA") also includes a safe harbor provision for forward-looking statements, like revenue projections, future objectives, future economic performance, etc. 15 U.S.C. § 78u-5(c); *see also id.* § 78u-5(i)(1) (defining "forward-looking statement"); H.R. Rep. No. 104-369, at 45 (1995) (Conf. Rep.). If a "forward-looking statement" is accompanied by meaningful cautionary language, "a defendant's statement is protected regardless of the actual state of mind." *Miller v. Champion Enters. Inc.*,

346 F.3d 660, 672 (6th Cir. 2003). If it is "not accompanied by meaningful cautionary language, actual knowledge of [its] false or misleading nature is required." *Id*. (citing 15 U.S.C. § 78u-5(c)(1)(B); *Helwig v. Vencor, Inc.*, 251 F.3d 540, 552 (6th Cir. 2001) (en banc), *cert. denied*, 536 U.S. 935 (2002). To be considered meaningful, cautionary language must be specific and substantive enough that it "realistically could cause results to differ materially from those projected in the forward-looking statement." *Helwig*, 251 F.3d at 558–59 (citation omitted); *see also Asher v. Baxter Int'l, Inc.*, 377 F.3d 727, 732–33 (7th Cir. 2004) (deeming boilerplate warnings like "all businesses are risky" not meaningful, while noting that "the cautions need not identify what actually goes wrong and causes the projections to be inaccurate").

The cases below provide a snapshot of the types of statements a court may find actionable and those that a court may view as mere puffery or forward-looking statements.

a.) Key Decisions: Misrepresentation and Puffery

Bond v. Clover Health: In *Bond v. Clover Health Investments, Corp.*, 587 F.Supp.3d 641 (M.D. Tenn. 2022), plaintiffs alleged that defendants cast their company as a "paradigm-disrupting tech firm" powered by AI software called Clover Assistant. The court found these were not mere boasts because defendants made factual claims that:

- The Clover Assistant actually worked at scale
- Software-enabled successes drove company growth

According to the complaint, those claims were false or actionably misleading because the company's growth was "actually simply the result of kickbacks and exploiting [] connections in the New Jersey insurance market." *Id*. at 689.

Defendants argued that "their general statements attributing Clover's growth to the quality of its products were simply too anodyne and vague to be materially misleading." *Id*. This argument was rejected. According to the court, "The defendants' assertions about the causes of Clover's growth . . . were built around a core contention that was factual and definite in nature: the claim that the Clover Assistant actually worked—and did so on a scale capable of translating into more attractive Medicare Advantage plans." *Id*. at 670. Crucially, the Court found that the "proposition that the Clover Assistant worked well and drove growth was the core of Clover's pitch to investors, and there [was] no reason for the court to assume that a reasonable investor would have ignored it." *Id*.

The court therefore found that plaintiffs "have pleaded, with particularity, that the statements about the Clover Assistant's adoption were at least misleading, which, in light of the Assistant's stated centrality to Clover's business, is all that is required to prevail on the elements of falsity and materiality." *Id*. at 672.

In re Upstart: in *In re Upstart Holdings, Inc. Securities Litigation*, 2023 WL 6379810 (S.D. Ohio Sept. 29, 2023), plaintiffs alleged that defendants "touted the strength of [their] AI underwriting model," which was the "centerpiece" of defendants' business: "it is what improves upon the traditional FICO-based system for determining a potential borrower's creditworthiness and makes Upstart's platform appealing for lenders who think that the old approach prices loans inaccurately." *Id*. at *12. The court identified "a few sub-categories of statements," but importantly issued the following findings regarding statements about Upstart's AI model.

The court first acknowledged that certain statements about the "superiority of the AI model" were "inactionable puffery." *Id*. Specifically, statements that "turbulent times presented

an opportunity for a new platform to shine, that the platform offered compelling and valuable benefits, and that the AI model was a fairly magical thing capable of doing the right things" were all "loosely optimistic statements that cannot be objectively verified." *Id*. According to the court, "there is no measuring stick by which the veracity of these optimistic statements can be evaluated—no objective way for reasonable investors (or [the] Court) to assess whether Upstart's AI model was, in fact, 'magical' or whether it 'shined.'" *Id*. Those statements, therefore, were inactionable.

More "specific, material, and verifiable" statements made by Upstart about the superiority of its AI model were, however, found to be actionable. For example, statements about the "significant advantage of the AI model over traditional FICO-based models," such as "higher approval rates and lower interest rates at the same loss rate" and "superiority . . . in elevated risk environments" were statements that could be "objectively assessed by looking to" the metrics. *Id*. at *13. According to the court, plaintiffs had adequately pled that the AI model did not provide the verifiable advantages, which is why the "statements [were] actionable material misstatements under the first step of the Rule 10b-5 analysis." *Id*.

The court also found statements that the AI model had the ability to "respond very dynamically to macroeconomic changes" as material and not mere puffery. *Id*. Plaintiffs alleged that defendants had represented that one of the main advantages of the AI model is "its ability to very quickly adjust itself to handle the macro situation that is evolving out there, even when macroeconomic trends in the broader economy are going through substantial disruptions and dislocations" such as "rising interest rates." *Id*. The court admitted that "whether these statements are inactionable puffery presents a close call." *Id*. First, the court acknowledged that the statements were "difficult to verify": how quickly must a model react for the statement to be true? However, the court gave more weight to the fact that the statements were "more specific than vague affirmations about the benefits of the model or its ability to do the right thing" and were "about a specific attribute of the model—its speed." *Id*. Furthermore, the "purported ability of the AI model to react quicker and more accurately than traditional FICO-based models [was] at the core of Upstart's pitch to lenders and investors." *Id*. And because "statements about the speed and adaptability of the model . . . would likely be considered essential to a reasonable investor," the statements were "material and not mere puffery." *Id*.

3. Implications for the Apple Litigation

Based on the complaint, Apple's statements about Siri's AI capabilities present a complex matrix of potentially protected and actionable elements that courts must parse under the framework discussed above.

a) The Safe Harbor Analysis

It is uncertain if the PSLRA's safe harbor provision will apply to Apple's forward-looking statements. On one hand, the timeline projections ("over the next year," "in the coming months") constitute classic forward-looking statements, but their protection under the PSLRA's safe harbor provision depends on whether they were accompanied by meaningful cautionary language that identified the specific risks that could cause actual results to differ materially.

The mixed present/future tense construction of Apple's statements complicates the analysis. When Apple stated that Siri "will be able to take hundreds of new actions," this forward-looking component might receive protection. However, statements that Apple Intelligence "combines the power of generative models with personal context" suggest present capability.

b) The Puffery Defense: Specificity as the Dividing Line

Apple's puffery defense will likely require the court to evaluate each statement for their specificity and verifiability:

Potentially Puffery:

- General claims about Siri becoming "more natural" or "more personal"
- Assertions about "breakthrough" technology or "revolutionary" capabilities
- Marketing language about "transforming" user experiences

Potentially Actionable:

- The specific metric of "hundreds of new actions"—a quantifiable claim similar to the "higher approval rates" in *Upstart*
- Demonstrations of specific functionalities (finding flight details, adding contact information)—concrete capabilities that can be objectively verified
- Claims about understanding "personal context"—while somewhat vague, this suggests specific technical functionality

The *Bond v. Clover Health* analysis may be helpful for plaintiffs. Just as Clover's statements about its AI assistant "working at scale" were deemed actionable because they formed the "core pitch" to investors, plaintiffs will likely argue that Apple's Siri enhancements were positioned as the primary driver for iPhone 16 upgrades. Plaintiffs' theory will be that the centrality of these claims to the investment thesis weighs against puffery protection.

c) The "Vaporware" Problem

If proven true, John Gruber's characterization of the features as "vaporware" and Apple's 2024 WWDC presentation as a "concept video" may raise issues similar to those in *Helo v. Sema4*. There, statements about Centrellis's capabilities were actionable when former employees revealed the platform "never existed" as described. If discovery reveals that Apple lacked even basic prototypes of the demonstrated features, the parallel to *Sema4* strengthens.

Alleged admission that Apple showed the technology "before it was ready" could also prove damaging. Unlike *Alich v. Opendoor*, where disclosed human involvement defeated AI washing claims, Apple's issue isn't about undisclosed human oversight but about the existence of the technology itself. This distinction matters: transparency about limitations (protected under *Alich*) differs fundamentally from demonstrations of non-existent capabilities (actionable under *Sema4*).

4. Practical Guidance: Avoiding AI Washing Liability

Although the federal securities case law on AI washing is still emerging, companies can draw some lessons from the early decisions:

- Avoid conflating aspirational development goals with product roadmaps.

- Avoid demonstrating or advertising capabilities without underlying functional prototypes.
- Reconsider boilerplate disclosures about technological or other risks.
- Ensure relevant business teams review AI-related representations for accuracy.

a) The Path Forward

The Apple litigation represents an inflection point for AI accountability. As courts grapple with applying the federal securities anti-fraud framework to AI representations, several trends are emerging that will influence corporate disclosure practices:

- *The End of "Conceptual" Marketing.* The era of marketing AI capabilities based on theoretical potential may be closing. Companies can no longer assume AI announcements will be treated as aspirational vision statements—courts increasingly view them as concrete commitments requiring substantiation.
- *The Verification Imperative.* The case law reveals a consistent pattern: courts apply rigorous scrutiny to specific, measurable AI claims while permitting general quality descriptors. Accordingly, companies should carefully vet any specific, measurable AI claims they make.
- *The Competitive Disclosure Dilemma.* On one hand, the AI race creates pressure to announce capabilities early to maintain competitive positioning. On the other, the Apple litigation demonstrates the severe consequences of premature disclosure. Companies must consider and balance this tension.

While the next few months and years will likely provide helpful clarification on the application of the federal securities anti-fraud provisions to AI representations, it is already clear that companies trying to create AI hype and excitement among investors can run afoul of those provisions. Any companies considering a "fake it till you make it" strategy would be well served to adopt a new mantra: "prove it before you promote it."

XIII. Prospects for New Regulation

If history is any guide, we should not be optimistic about regulation of AI coming from Washington, D.C. Despite years of hand-wringing over important issues like privacy and social media, we still do not have any national laws on those subjects in the United States. Leaders of some of the major U.S. technology companies have sought to establish close ties to President Trump. Yet within that group, their views about regulating AI stretch to the ends of the spectrum. Notwithstanding the lack of action at the national level, about half the states have some regulation of AI, the most common subjects being the use of AI tools in making employment decisions and deep fakes/false information.

As with privacy, it appears that the bellwether for regulation throughout the world will come from Europe. The world's first comprehensive legal framework for regulating AI systems emerged in Europe with the passage of the EU AI Act. The EU AI Act will likely become a default international standard for AI regulation, as companies will not create separate compliance programs for Europe and other continents but will instead apply their European compliance program across the globe—an example of what has been termed the "Brussels Effect."

AI regulation stands poised to impact not only companies developing AI technologies, but

also companies that deploy AI tools. Producers and consumers of AI must each be aware of the various regulatory frameworks going into effect and prepare accordingly. Below, we explore proposed and enacted regulations around several active areas of regulation or proposed regulation: (1) "deepfake" technology, (2) AI in the employment and human resources context, (3) election-related use of AI, and (4) privacy and transparency regulation for AI platforms.

a. AI-Generated Deep Fakes

The term "deep fake" can appear in a variety of contexts, and while it has no fixed definition the term generally refers to the use of AI to create or alter media content that is then passed off as genuine. The EU AI Act defines deep fakes as an "AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful." (EU AI Act, Art. 3 § 60.)

Deep fakes have drawn scrutiny from their use in news content, entertainment media, and even pornography. Some professionals in the entertainment industry, for example, are concerned about how their digital likenesses are used and whether such usage will interfere with their livelihoods. More broadly, deep fakes may result in reputational damage, or aid in the dissemination of false or misleading information.

To address these concerns, legislatures in the US and around the world are considering ways to regulate the use of deep fakes. At the simplest level, this regulation does not ban deep fakes outright, but requires a disclosure that AI was used to generate the content. This is the approach taken in the EU AI Act, which subjects providers of AI systems that generate large quantities of synthetic content to transparency obligations, in particular, the implementation of sufficiently reliable techniques to ensure the content can be easily identified as AI-generated rather than human-created.[237] Limited-risk AI systems, including those that interact with natural persons or generate potentially deceptive content, must comply with disclosure obligations from August 2, 2026. This is also the approach taken by legislation that was pending in last Congress, the DEEPFAKES Accountability Act, which would have established disclosure requirements for deep fakes (including video and audio) and criminal and civil penalties for willful violations of those disclosure obligations.

California has taken regulation of deep fakes a step further with the passage of two laws in 2024 geared to protect entertainers (and their estates) from the misappropriate of their likenesses. One creates a cause of action for beneficiaries of deceased celebrities to recover damages for the unauthorized usage of a "digital replica" of that celebrity. (AB-1836.) The second, which pertains to living performers, voids contract provisions that permit the use of an entertainer's "digital replica" in lieu of an actual performance from the entertainer themselves, subject to exceptions. (AB-2602.)

b. Employment-Related AI Regulations

In Europe, the EU AI Act regulates the use of AI in the workplace in a variety of ways. Article 5 concerning AI systems posing "unacceptable" risk prohibits outright the use of AI systems in the workplace that "infer emotions of a natural person" (unless they serve a medical or safety purpose).[238] Article 6 classifies as "high risk", *inter alia*, systems that use AI for recruiting or to make other workplace decisions, including promoting and terminating employees.[239] The high-risk classification subjects those AI systems to significant regulation, including the implementation of a risk management system, quality controls, record-keeping and

[237] EU AI Act, Art. 50 § 4, and Art. 99.
[238] EU AI Act, Art. 5, § 1(f)
[239] *Id.*, Art. 6, § 2, and Annex III, § 4.

transparency obligations, and some level of human oversight under Articles 8-15.[240] For instance, employers must inform workers and their representatives of the deployment of AI systems in the workplace. Non-compliance with these transparency obligations can lead to fines.[241] The EU AI Act has been partially implemented, with Articles 1-5 concerning AI systems posing "unacceptable" risk already in effect since February 2, 2025. Obligations concerning these "high-risk" AI systems within the meaning of Annex III of the AI Act will likely start applying from December 2, 2027.[242] AI systems such as spam filters used in the workplace that present minimal risks for individuals are not subject to any additional obligations and are required simply to comply with other legislation already in force such as the General Data Protection Regulation (the "GDPR").[243]

c. Election-Related AI Regulations

Deep fakes have also made an impact in the political world, with concerns over misinformation and election integrity spawning regulations of deep fakes and other AI-generated content used in political advertisements. Colorado, for example, enacted legislation in 2024 that requires political advertisements featuring AI-generated deep fakes to contain disclaimer language. (HB24-1147.) And while there is no comparable regulation at the federal level, two proposed bills (the REAL Political Advertisements Act and the Protect Elections from Deceptive AI Act) would require similar disclaimer language for AI-generated political ads.

California also joined the fray by enacting three pertinent laws in 2024. The Deceptive Media in Advertisements Act contains a broad prohibition on the knowing dissemination of any election advertisement that includes "materially deceptive content" (defined to include deep fakes). This prohibition is in place in the 120 days leading up to an election and, in some cases, 60 days after an election. California additionally amended its Political Reform Act of 1974 to require that AI-generated political advertisements from political committees contain disclaimers that the content was created or substantially altered through the use of AI. Going one step further, California's Defending Democracy from Deepfake Deception Act takes a novel and more aggressive approach to AI regulation by imposing affirmative obligations on "large online platforms"—defined broadly as any website, app, or social media platform with at least "1,000,000 California users during the preceding 12 months"—to implement procedures to identify deep fakes that portray candidates for elected office in a false or misleading manner. Platforms must remove these deep fakes within 120 days of an election and, in some circumstances, 60 days after an election; outside these time periods the Act only requires that deceptive deep fakes be labeled as such.

d. Transparency and Data Privacy

Current and anticipated legislation seeks to impose broad regulatory obligations concerning transparency, record keeping, and reporting obligations.

---

[240] *Id.*, Arts. 8-15. On 19 May 2026, the Commission published draft Guidelines on the classification of "high-risk" AI systems, clarifying the applicable provisions and providing practical examples to support assessments across various sectors and use cases.
[241] *Id.*, Art. 99.
[242] The European Parliament has already approved the corresponding amendment to the AI Act. Its entry into force is, however, subject to formal adoption by the Council of the European Union. See European Parliament Press Release, *AI Act: EP approves simplification measures and "nudifier" app ban* (June 16, 2026), available athttps://www.europarl.europa.eu/news/ro/press-room/20260611IPR45207/ai-act-ep-approves-simplification-measures-and-nudifier-app-ban. In addition, rules on General-Purpose AI ("GPAI") systems that need to comply with transparency requirements, governance, confidentiality, and penalties provisions are becoming applicable on August 2, 2025. European Commission, *Timeline for the Implementation of the EU AI Act*, available at https://ai-act-service-desk.ec.europa.eu/en/ai-act/timeline/timeline-implementation-eu-ai-act.
[243] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

Transparency is a core tenet of the EU AI Act, requiring all AI systems that interact with natural persons to inform users they are interacting with an AI system.[244] (EU AI Act, Art. 50.) The EU AI Act also imposes a series of broad obligations for "high risk" systems. This includes AI systems used in connection with infrastructure (e.g., traffic and utilities), education, employment, public services, law enforcement, and biometric identification. (EU AI Act, Annex III.) These obligations require deployers of "high-risk" systems to implement, among other things, (i) risk management processes that continuously evaluates potential risks associated with the system (id. Art. 9); (ii) the use of data sets for training that meet certain quality criteria (id. Art. 10); (iii) automatic event logging (id. Art. 12); (iv) mechanisms that allow oversight by natural persons (id. Art. 14); (v) and general quality control standards that ensure the accuracy and security of the system (id. Art. 15).

Specific obligations apply to providers of general purpose AI ("GPAI") models, namely, AI models that display significant generality and are capable of competently performing a wide range of distinct tasks, such as GPT-4, Claude Sonnet 4.5, Gemini 3 Pro or Llama 4.[245] GPAI model providers have an obligation to (i) maintain information relating to the technical documentation of their model (including on how they train and test it, as well as the results of the evaluation); (ii) make available certain information to other AI providers that plan to integrate with their model; (iii) establish a policy to comply with EU copyright law; and (iv) publish a detailed summary of the content used for training their model. (EU AI Act, Art. 53.) The two first obligations do not apply to providers of GPAI models that are released under a free and open-source license allowing third parties to access, use, modify and distribute the model, and whose parameters, including on the weights (i.e., learnable numerical parameters within a machine learning model that define its internal structure and decision-making logic, enabling the system to transform input data into patterns used to generate predicted outputs), the model architecture, and model usage, are made publicly available.

Providers of GPAI models that pose systemic risk must comply with an additional set of burdensome obligations, including (i) the performance of model evaluations to assess the model's capabilities, propensities, affordances and effects;[246] (ii) the assessment and mitigation of systemic risks; (iii) the report of serious incidents to the EU AI Office, which is tasked with supervising, investigating, enforcing and monitoring the obligations for providers of GPAI models;[247] (iv) the adoption of corrective measures; and (v) the implementation of an appropriate level of cybersecurity for the model and its physical infrastructure. (EU AI Act, Art. 55.) Moreover, GPAI models that pose systemic risk cannot benefit from the "open-source" exemptions mentioned above.

---

[244] On June 10, 2026, the European Commission published the final Code of Practice on the marking and labelling of deepfakes and AI-generated or AI-manipulated text related to matters of public interest. The Code is structured in two parts. The first aims to assist providers in ensuring that AI-generated or AI-manipulated audio, images, video, and text are marked in a machine-readable format, enabling their identification as artificially generated or manipulated. The second outlines the obligations of deployers to clearly label deepfakes and AI-generated or AI-manipulated text made available to the public on matters of public interest, particularly where no human review or editorial oversight has taken place. While adherence to the Code is voluntary, it can support providers and deployers in complying with the relevant obligations under the EU AI Act. The Code will be complemented by guidelines clarifying the scope of the legal obligations and addressing aspects not covered by the Code. In parallel, the Commission published draft Guidelines on the implementation of the transparency obligations for certain AI systems under Art. 50 of the EU AI Act, aimed at clarifying the scope of the legal obligations and addressing aspects not covered by the code.

[245] A GPAI model, often developed by foundational model providers, is the result of training an algorithm on data. The outcome is then integrated into AI systems (i.e., in software and infrastructure that will be capable of generating outputs) and placed on the market to accomplish varied objectives without requiring the model to be rebuilt or extensively retrained for each new task. According to the guidelines on the scope of the obligations for GPAI models published on 18 July 2025, the European Commission presently considers that a model is likely to qualify as GPAI model if (i) the amount of computational resources used to train it is superior to 10^{23} floating point operations per second; and (ii) it is capable of generating language (text or audio), text-to-image, or text-to-video outputs.

[246] In practice, this means using some of the following methods: Q&A sets, task-based evaluations, benchmarks, red-teaming and other methods of adversarial testing, human uplift studies, model organisms, simulations, and/or proxy evaluations for classified materials.

[247] The EU AI Act mandates each Member State to designate a market surveillance authority, as well as a notifying authority.

To help GPAI model providers comply with the obligations contained in Articles 53 and 55 of the EU AI Act, the Commission issued on 10 July 2025 a GPAI Code of Practice. According to the Commission, voluntary compliance with the code will reduce the administrative burden on businesses, and increase legal certainty in comparison to situations where compliance through other methods is chosen and "trust from the Commission and other stakeholders." Moreover, commitments given in the context of the Code may be taken into account as mitigating factors in determining the amount of fines ensuing from potential future violations. Adherence to the Code is unconditional, as "[a]ny opt-out from chapters of the code of practice results in losing the benefits of facilitating the demonstration of compliance in that respect." (Guidelines on GPAI models, paras 94-100.) Several leading tech companies including Amazon, Anthropic, Google, IBM, Microsoft, OpenAI, Aleph Alpha, and Mistral AI have already signed up to the Code. By contrast, Meta indicated it does not intend to sign the Code, while xAI only signed up to the chapter of the Code named "Safety and Security" which contains the measures relating to the additional obligations contained in Art. 55 of the EU AI Act (e.g., implementing technical safeguards or conducting ongoing post-market monitoring of the model's use and potential unintended consequences). Given that companies choosing not to sign up to, or opting-out of certain parts of, the Code must rely on other methods to demonstrate compliance, the Commission is likely to scrutinise them more rigorously.

In the United States, Colorado's AI Act imposes similar transparency obligations, requiring developers to publish the types of high risk systems that are deployed, as well as a statement concerning how it manages the risk of algorithmic discrimination. (6-1-1702 § (4).) California has also opted for maximum transparency through the AI Training Data Transparency Act, which requires deployers of AI systems in California to publish summaries of data sets used for training. Additionally, California's Health Care Services: Artificial Intelligence Act requires health care providers to provide disclaimers to patients in certain circumstances where communications are generated by AI.

Importantly, the EU AI Act also contains several provisions concerning personal privacy that are not mirrored in existing US legislation. The EU AI Act prohibits the use of AI systems that: (a) engage in biometric categorization based on race, politics, religion, and certain other beliefs; (b) employ real-time biometric identification systems in public spaces for use by law enforcement (absent limited exceptions); (c) create or expand facial recognition databases through scraping; (d) make risk assessments of the likelihood an individual will engage in a criminal offense; and (e) classify persons through social scoring in a manner that leads to unfavorable treatment. (EU AI Act, Art. 5.) Pending formal adoption by the Council of the European Union, the EU AI Act is set to be amended to also ban AI systems that generate child sexual abuse material or create images, videos and audio depicting an identifiable person's intimate parts, or sexually explicit activities, without their consent.[248]

Finally, the Commission's Guidelines on prohibited AI practices under Art. 5 of the EU AI Act, published on February 4, 2025,[249] clarify that emotion recognition systems that are not deemed to pose "unacceptable risk" will nevertheless be classified as "high-risk" AI systems, subjecting them to the stringent compliance obligations referred to above, including the implementation of a risk management system, quality controls, record-keeping and transparency obligations, and some level of human oversight. The same applies to certain AI-based scoring systems, such as those used for credit-scoring or risk assessment by health and life insurance companies, which do not fulfil the conditions for outright prohibition provided by Article 5(1)(c) of

[248] European Parliament Press Release, *AI Act: EP approves simplification measures and "nudifier" app ban* (June 16, 2026), available at https://www.europarl.europa.eu/news/ro/press-room/20260611IPR45207/ai-act-ep-approves-simplification-measures-and-nudifier-app-ban.

[249] Annex to the Communication from the Commission - Approval of the content of the draft Communication from the Commission - Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), C(2025) 884 final ("Guidelines on prohibited AI practices").

the regulation.[250] The Guidelines also explain how, in the Commission's view, the EU AI Act will intersect with other related or overlapping EU statutes,[251] notably, the GDPR, the Law Enforcement Directive,[252] and Regulation (EU) 2018/1725.[253]

XIV. AI Entrants and Data-Controlling Incumbents

As discussed in Section IV, a range of technical and contractual measures are available to moderate access to data sets. How incumbents exercise control over such data raises questions under the antitrust laws: could an incumbent's control over data constitute an anticompetitive barrier to entry, or is it a legitimate exercise of property or contractual rights? In particular, several forms of control, such as blocking API access, enforcing terms of service, or asserting trade secret rights, may be characterized as lawful self-protection by one party and as exclusionary monopolization by another. The law in this area remains at an early stage of development as these questions are only starting to be raised in litigation and regulatory proceedings. We address some of the principal issues that are beginning to emerge in these disputes below.

a. Antitrust Claims by the AI Entrant

Data-Access Aftermarket Monopolization. At the threshold, courts must resolve whether data itself constitutes a relevant input market, or whether the relevant market is defined downstream at the product or service level. Where a dataset reflects years of accumulated network effects, an entrant may argue that it is not practically replicable within a competitive timeframe, and that the incumbent is monopolizing a separate aftermarket for data access distinct from the primary product market. For example, in *Celonis SE v. SAP SE*, the court recognized a distinct "data access" market in evaluating the plausibility of a monopolization claim under Section 2 of the Sherman Act.[254]

Broken Open-Platform Promises. Some cases involve allegations that incumbents made affirmative data openness commitments—to regulators, customers, or the market—and subsequently restricted access after developing or acquiring a competing product. Prior openness representations may bear on whether a later restriction reflects a legitimate business justification or pretextual foreclosure under Section 2. In *Ardent Labs, Inc. v. Applied Systems, Inc.*, Ardent Labs, doing business as Comulate, alleges that Applied made representations that it would maintain open access to its insurance data platform, but then later restricted access to the detriment of Comulate.[255] A similar argument is invoked in *Google LLC v. SerpApi, LLC*, with SerpApi alleging that Google's restrictions on access to search data are inconsistent with its prior openness representations, including in describing the data as "open source."[256]

---

[250] EU AI Act, art. 5(1)(c). See also *id.*, recitals 37-38, 58 and Annex III.
[251] Commission Guidelines on prohibited artificial intelligence practices, paras 42-52. See also, *id.*, paras 135-145, 178-183, 219-221, 238 and 287-288. The Commission approved the content of the draft guidelines on July 18, 2025.
[252] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA.
[253] Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC.
[254] 2025 WL 3013158, at *1 (N.D. Cal. Oct. 27, 2025) (denying defendant's motion to dismiss claim because plaintiff "adequately alleged for now that the data access market is separate . . . from either the ERP software market or the process mining market . . . and that [defendant's] restrictive data access policies have harmed competition").
[255] Compl. at ¶¶ 2-3, 145-46, No. 26-cv-00591 (N. D. Ill. Jan. 19, 2026).
[256] Mot. to Dismiss at 3–4, No. 4:25-cv-10826-YGR (N.D. Cal. Feb. 20, 2026).

Sham Litigation and Predatory Enforcement. A recurring theory in data-access disputes holds that an incumbent's IP enforcement actions—e.g., trade secret claims, DMCA takedowns, copyright litigation—may themselves constitute exclusionary conduct under Section 2 where the litigation is objectively baseless and filed as a means of interfering directly with a competitor's business rather than as a genuine effort to vindicate IP rights. For example, in *Ardent Labs*, Ardent Labs characterizes Applied's trade secret lawsuit as sham litigation filed as part of a coordinated antitrust campaign.[257] CREXi raised analogous arguments in response to CoStar's intellectual property and contract claims in the commercial real estate data market.[258]

The Customer-Data-Hostage Theory. A distinct antitrust theory posits that the incumbent does not own the data it is restricting—its customers do—and that the incumbent's blocking of third-party access interferes with customer rights rather than protects the incumbent's own property. This framing is structurally distinct from a *Trinko*-type refusal-to-deal claim, as the entrant does not argue that the incumbent must share its own assets, but that the data at issue belongs to the incumbent's customers rather than to the incumbent itself, placing the claim outside much of the doctrinal apparatus governing refusals to deal.[259]

In a complaint filed against Applied by Ardent Labs, for instance, Ardent Labs, alleges that insurance brokers' (Applied's own customers) own policy and transaction data is effectively held hostage within Applied's Epic platform, with access also restricted for third-parties.[260] In the healthcare context, AI entrants have asserted that Epic Systems' restrictions on patient data access conflict with patients' and providers' rights under applicable federal interoperability rules.[261]

IP Rights and Antitrust. The exercise of valid intellectual property rights does not automatically immunize a data restriction from antitrust scrutiny, but courts have long held that IP rights are a relevant factor in assessing whether a challenged restriction has a legitimate business justification.[262] The same dataset may thus simultaneously be asserted as a protectable trade secret by the incumbent and as a source of alleged Defendant market power by the entrant.[263]

b. The Mirror-Image Problem: Incumbent IP Claims and Entrant Antitrust Claims

As evidenced by the preceding discussion, the same facts will typically lend themselves to both antitrust and intellectual property claims and/or arguments. Conduct the incumbent characterizes as legitimate enforcement of its data rights—asserting trade secret misappropriation, invoking contract or terms of service restrictions, pursuing CFAA or copyright claims—may also give rise to antitrust counterclaims by the entrant under the exclusionary conduct or sham litigation theories described above. Alternatively, such arguments may simply be invoked as a defense. And, similarly, defendants facing antitrust claims may argue that they are merely protecting their intellectual property. *Applied Systems* (Illinois case), described above, is

---

257 *Ardent Labs*, Compl. at ¶¶ 7, 10, 16.

258 Mot. To Dismiss at 2, *CoStar Grp., Inc. v. Com. Real Est. Exch., Inc.*, No. 2:20-cv-8819-CBM-AS (C.D. Cal. Nov. 23, 2020) ("[I]t is CoStar, not CREXi, that has engaged in unfair and unlawful competition by blocking CREXi from accessing publicly-available data on CoStar's platforms.").

259 *Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, LLP*, 540 U.S. 398 (2004).

260 *Ardent Labs, Inc. v. Applied Systems, Inc.*, Compl. at ¶¶ 20, 122, No. 1:26-cv-00591 (N.D. Ill. Jan. 19, 2026).

261 Compl. at ¶¶ 65–66, *Particle Health Inc. v. Epic Sys. Corp.*, No. 1:24-cv-07174-NRB (Sept. 23, 2024).

262 *Image Tech. Servs., Inc. v. Eastman Kodak Co.*, 125 F.3d 1195, 1218 (9th Cir. 1997).

263 *See, e.g.*, *Thomson Reuters Enter. Centre GMBH v. Ross Intel. Inc.*, 2025 WL 1488015, at *1 (D. Del. May 23, 2025) (certifying for appeal the "hard" questions of whether the challenged data is under copyright and whether defendant's use of such data for their AI algorithm constitutes "fair use"); *see also Thomson Reuters Enter. Centre GMBH v. Ross Intel. Inc.*, 765 F. Supp. 3d 382, 390-91 (D. Del. 2025) (providing background facts).

one example: Applied's trade secret and fraud claims and Ardent Labs' antitrust counterclaims arise from the same data access dispute and are being litigated in parallel proceedings.[264] *IQVIA v. Veeva* presented the same structure over eight years before settling without judicial resolution of either set of claims.[265]

These disputes reflect an unsettled area of law in which restricting data access can be characterized either as lawful protection of intellectual property rights or as exclusionary conduct under Section 2, with courts yet to converge on a consistent framework for resolving this tension.

***

If you have any questions about the issues addressed in this memorandum, or if you would like a copy of any of the materials mentioned in it, please do not hesitate to reach out to:

Viola Trebicka
Partner - Privacy, Products Liability & Regulation
Los Angeles
violatrebicka@quinnemanuel.com, Tel: 213-443-3243

Patrick Curran
Partner – AI IP
Boston
patrickcurran@quinnemanuel.com, Tel: 617-712-7103

Sean Pak
Partner – AI IP
San Francisco
seanpak@quinnemanuel.com, Tel: 415-875-6320

Robert Schwartz
Partner – Copyright, Defamation, and Section 230 Safe Harbor
Los Angeles
robertschwartz@quinnemanuel.com, Tel: 213-443-3675

Hope Skibitsky
Partner – Scraping
New York
hopeskibitsky@quinnemanuel.com, Tel: 212-849-7535

Corey Worcester
Partner – Scraping
New York
coreyworcester@quinnemanuel.com, Tel: 212-849-7471

To view more memoranda, please visit www.quinnemanuel.com/the-firm/publications/

[264] 1:25-cv-14251 (N.D. Ill. Nov. 11, 2025); 1:26-cv-00591 (N.D. Ill. Jan. 19, 2026).
[265] 2021 WL 12319551, at *1–3 (May 7, 2021).

To update information or unsubscribe, please email updates@quinnemanuel.com