

How Generative AI Cos. Can Navigate Product Liability Claims

By **David Kerschner, Jason Ross and Brendan Gibbons** (January 29, 2026)

In December, the California Judicial Council assigned a coordination motion judge to consider whether multiple product liability lawsuits against OpenAI Inc. over its ChatGPT artificial intelligence system should be coordinated, with a coordination hearing scheduled for Jan. 30.

The coordination request brings together a growing number of claims alleging that conversational AI systems caused psychological harm, failed to disengage in crisis scenarios or lacked adequate safeguards.

The coordination order does not address merits. It signals a familiar pattern: Plaintiffs are aggregating disputes over emerging technologies and pursuing them through mass-tort-style proceedings, borrowing tactics from litigation involving social media, pharmaceuticals and other consumer-facing products.

The protections that once insulated AI companies from civil liability have long shaped the legal landscape. For decades, technology companies operated with confidence that their AI services and products were broadly shielded from litigation tied to third-party content by Section 230 of the Communications Decency Act and the First Amendment.

Those protections are now being tested, particularly where plaintiffs are pursuing claims against AI companies that target platform design, user engagement and warnings, rather than content moderation.

In this emerging wave of litigation, plaintiffs seek to reframe generative AI as consumer products whose design and warnings create foreseeable risks.

Viewed through a mass-tort defense lens, this article explains how these claims are framed and highlights defense-side factors that may shape discovery, expert strategy and regulatory scrutiny.

From Social Media to AI: The Litigation Trajectory

In recent social media litigation, plaintiffs alleged platforms intentionally maximized engagement in ways that exploited psychological vulnerabilities and caused foreseeable harm.

Courts addressing motions to dismiss drew a critical distinction between claims treating platforms as publishers of third-party content, which remain barred by Section 230, and claims targeting a platform's own design choices or failures to warn. Courts have also declined to treat the First Amendment as a categorical bar to failure-to-warn claims.[1]

These rulings do not eliminate immunity defenses. They narrow the inquiry, so defendants



David Kerschner



Jason Ross



Brendan Gibbons

can still press traditional tort limits on duty, injury and causation.

Although arising in the online gambling context, *De Leon v. DraftKings Inc.*, decided by the U.S. District Court for the Southern District of New York in December, underscores the limits of these theories once courts move past threshold immunity questions.

There, the court dismissed addiction-based claims, notwithstanding the plaintiffs' focus on platform design, holding that New York law requires physical injury to support product liability theories.[2]

Together, these decisions create a two-phase defense landscape: initial fights over immunity, and content versus conduct, followed by conventional product liability analysis if claims survive.

Plaintiffs are now applying this framework to generative AI, arguing that anthropomorphic design, persistent memory and engagement-driven interactions fall outside Section 230.

Early AI Litigation: What Courts Have and Have Not Decided

Early AI cases show how courts are addressing these theories at the pleading stage, as in *Garcia v. Character Technologies*, in the U.S. District Court for the Middle District of Florida, where design defect and failure-to-warn claims were allowed to proceed last May while the question of whether the chatbot qualified as a "product" remained open.[3] The case later settled before the court reached the merits of those claims.

Similarly, in *Raine v. OpenAI Inc.*, in the San Francisco County Superior Court, the plaintiffs allege failures to disengage and conflicts between engagement-driven design and safety protocols, but the court has not resolved duty, defect or causation.[4]

Other cases and state enforcement actions test how far traditional product liability and consumer protection theories extend in this context. But to date, courts have not adopted a generalized duty or defect theory. Causation and proof remain central.

Product Liability Theories Being Tested Against AI Systems

Recent lawsuits test how product liability applies to novel AI technologies. While most cases remain in infancy, the claims coalesce around alleged design defects, failures to warn and causation.

Strict Product Liability — Design Defect

One of the most significant areas of exposure for AI companies arises from strict liability claims based on alleged design defects. Plaintiffs contend that AI chatbots and companion systems fail to perform as safely as an ordinary consumer would expect, or that the risks inherent in their design outweigh their benefits.

A recurring theme is "sycophancy." Plaintiffs allege that some large language models used in AI chatbots are trained to maintain conversational flow by affirming user statements.

When a user expresses distress or suicidal ideation, a model optimized for engagement may validate those thoughts rather than challenge them or disengage. Plaintiffs argue this behavior is the foreseeable result of engagement-driven design choices.

Plaintiffs also focus on the alleged weakening or removal of safety features, asserting that safer alternative designs, such as hard refusals for self-harm content or automatic disengagement during crisis scenarios, were available but not implemented, or were rolled back for business reasons.

From a defense perspective, these claims are likely to rise or fall on evidence of feasible alternative designs, not generalized rhetoric about engagement or harm.

Strict Product Liability – Failure to Warn

Plaintiffs also assert strict liability failure-to-warn theories. They allege AI companies failed to disclose foreseeable risks, including hallucinations, confident presentation of inaccurate information and the potential for emotional reliance through prolonged interaction.

In cases involving minors, plaintiffs further contend companies failed to warn parents about risks to developing brains or emotional displacement.

Plaintiffs now argue that anthropomorphic design masks AI-related risks, leading users to believe they are receiving empathetic, reliable advice rather than probabilistic outputs. For defendants, however, these claims often turn on the visibility, timing and repetition of warnings, not merely their existence in terms of service.

Failure-to-warn claims present several threshold pressure points on plaintiffs. Defendants are likely to challenge whether the alleged risk was sufficiently known or knowable at the relevant time, whether the warning proposed by plaintiffs would have altered user behavior and whether the claimed injury falls within the scope of risk that an additional warning would remedy.

In cases involving sophisticated users or repeated interactions, defendants may further argue that warnings were adequate as a matter of law, or that the risk was open and obvious in context, particularly where the system expressly disclaimed providing medical or professional advice.

Negligence and Voluntary Undertaking

Beyond strict liability, plaintiffs are pursuing negligence claims based on alleged design defects and failure to warn. A developing theory is negligent or voluntary undertaking.

Plaintiffs argue that by establishing safety teams and publicly emphasizing user well-being, AI companies assumed a duty to protect users, and breached that duty by inadequately implementing safeguards.[5]

But courts have not embraced a broad duty theory based on generalized safety commitments, and early voluntary undertaking claims are likely to turn on the specificity of the alleged representations and reliance.

Generalized safety statements and aspirational policies should not be conflated with enforceable undertakings, and defendants are likely to challenge these claims by emphasizing the absence of specific promises, reasonable reliance, or a causal link between the alleged undertaking and the claimed harm.

Causation and Proof

While plaintiffs have succeeded in pleading the above discussed theories in at least one case, proving them will require extensive expert testimony. Plaintiffs must show that a specific feature of an AI system caused a specific harm.

In addiction-based claims, experts may opine that certain design features contributed to compulsive use or dependence. In suicide and psychosis cases, experts may address whether validation and reinforcement were a substantial contributing factor to the alleged harm.

Defendants are likely to counter with alternative causation theories, including preexisting mental health conditions, external stressors and intervening conduct. In appropriate cases, defendants may also assert comparative fault or misuse as defenses, arguing that even if a causal link were assumed, user conduct or safeguard circumvention limits or bars liability.

These disputes mirror long-standing causation disputes in tobacco and pharmaceutical litigation, where evidence of risk does not establish causation. As in other mass torts, these cases may survive early motions, yet ultimately turn on expert admissibility and individualized proof, rather than rulings at the pleading stage.

Regulatory and Legislative Context

Proposed federal legislation, including the AI LEAD Act, would explicitly classify certain AI systems as products for purposes of product liability law, and authorize claims against developers and deployers of those systems for negligent design, failure to warn and strict liability.

Although passage remains uncertain, plaintiffs may cite such proposals as evidence of foreseeability and evolving standards of care.

California has moved further, with legislation addressing companion chatbots, disclosure obligations and developer accountability. At the same time, federal preemption of state AI laws has been discussed but not adopted, leaving companies to plan against the backdrop of continuing state-law claims.

Even where legislation does not pass, regulatory scrutiny can shape litigation by informing expectations about safeguards and documentation.

Defense Takeaways for AI Companies

Immunity defenses remain critical. But design- and warning-based claims may proceed past early motions.

Meanwhile, traditional product liability requirements — injury, causation and feasible alternative design — continue to provide meaningful defenses.

In many cases, expert admissibility and individualized causation will be the real inflection points, not pleading-stage rulings.

Early coordination on design decisions, warnings and internal documentation can materially shape litigation exposure years later.

How AI Companies Can Prepare

Several recurring issues drive risk and leverage in AI product liability cases.

Early decisions about product design, safety features, warnings and internal governance can materially shape how claims are evaluated on dispositive motions, on expert challenges, and before juries. Companies that address these issues proactively are better positioned to narrow claims and manage long-term exposure.

Design Choices and Safety Architecture

AI companies should approach these risks through a product safety lens. Plaintiffs increasingly focus on system prompts, engagement objectives and guardrails as evidence of design defect.

Design choices that encourage prolonged interaction or fail to disengage in sensitive scenarios may later be scrutinized for foreseeability and reasonableness. Even where no defect is ultimately found, companies should expect design tradeoffs to be examined closely through discovery and expert testimony.

Warnings and Disclosures

Failure-to-warn claims often turn on the visibility, timing and repetition of warnings, not merely their existence.

AI companies may wish to consider whether warnings are delivered at moments of heightened risk, reinforced during prolonged sessions, and tailored appropriately for minors and parents.

User Controls and Safeguards

Age verification, parental controls, usage limits and escalation protocols can serve both safety and litigation defense objectives. While the adoption of such measures does not concede liability, plaintiffs frequently cite their absence as evidence that risks were foreseeable and unaddressed.

How these tools are implemented, and whether they are consistently enforced, may affect both causation arguments and jury perception.

Documentation and Communications

AI companies should record why safety decisions were made, and how risks were evaluated. In product liability litigation, the ability to explain the reasoning behind design and safety choices often matters as much as the choices themselves.

Contemporaneous documentation can provide critical context when plaintiffs argue that alternatives were ignored or risks were discounted. In long-running product cases, poorly contextualized internal documents often drive punitive narratives more than underlying design decisions themselves.

Public statements, including marketing materials, blog posts, investor communications and litigation responses, can quickly become evidence, underscoring the importance of disciplined communications practices.

Why This Moment Matters for AI Companies

Plaintiffs are adapting long-standing product liability tactics to generative AI, testing the limits of immunity doctrines that technology companies have long relied upon. While those defenses remain critical, they may not resolve design- and warning-based claims at the pleading stage.

AI companies that approach AI litigation with the same rigor applied to other complex consumer products, focusing on design choices, warnings, documentation and expert strategy, may be better positioned to manage risk as this litigation landscape evolves.

David Kerschner and Jason Ross are partners, and Brendan Gibbons is a senior associate, at Arnold & Porter Kaye Scholer LLP.

Arnold & Porter counsel Rachel Forman contributed to this article.

The opinions expressed are those of the author(s) and do not necessarily reflect the views of their employer, its clients, or Portfolio Media Inc., or any of its or their respective affiliates. This article is for general information purposes and is not intended to be and should not be taken as legal advice.

[1] See, e.g., *In re: Social Media Cases*, JCCP 5255 (Nov. 5, 2025); *In re: Social Media Adolescent Addiction Personal Injury Prods. Liab. Litig.*, No. 4:22-md-03047-YGR, Dkt. No. 1730 (N.D. Cal. Feb. 28, 2025); see also *In re: Social Media Adolescent Addiction Personal Injury Prods. Liab. Litig.*, 702 F. Supp. 3d 809 (N.D. Cal. 2023).

[2] *De Leon v. DraftKings Inc.*, No. 25-cv-00644, Dkt. No. 70 (S.D.N.Y. Dec. 11, 2025).

[3] *Garcia v. Character Techs. Inc.*, No. 24-cv-01903, Dkt. No. 115 (M.D. Fla. May 20, 2025).

[4] *Raine v. OpenAI Inc.*, No. CGC-25-214237 (Cal. Super. Ct. S.F. Cnty.).

[5] See, e.g., *Raine*, No. CGC-25-214237; *Brooks v. OpenAI Inc.*, No. 25STCV32386 (Cal. Super. Ct. L.A. Cnty.)