ARMILLA

Navigating AI Litigation and Risk

October 15, 2025 | White Paper

AI incidents and disputes are increasingly leading to legal action, turning AI-related litigation into a valuable source of information for understanding AI risk, liability trends and mitigation opportunities. This white paper provides an introduction to Armilla's AI Litigation Database, a live resource providing key insights into where AI liability risk is concentrating and why. The Database has been developed by Armilla as a complement to its core technical evaluation and underwriting capabilities for AI models, its monitoring of the regulatory and contractual landscapes, and other proprietary AI risk analytics. This paper highlights five key trends identified by Armilla upon analysis of the database, which include: the rise of corporate plaintiffs, the emergence of AI litigation activity in all sectors of the economy, more frequent and significant enforcement actions by regulators, an increase in certified classes, and a shift from broad to precise claims.

Courts and regulators no longer consider AI an experimental technology deserving of special forbearance. Rather, they expect that enterprises deploying AI will have strong AI governance controls in place. Moreover, Armilla expects that enhanced clarity on the liability regimes applicable to powerful AI systems will further catalyze AI-related litigation. Given these dynamics, Armilla recommends a comprehensive approach to AI risk management, including documented governance and oversight, continuous evaluation of AI systems, proactive with evolving regulations and industry standards, and reliance on affirmative AI insurance to address exposure of complex, sensitive AI applications.

To provide companies with additional tools for understanding and mitigating AI risk, Armilla is making its database available to select partners, enabling them to monitor the evolving AI risk landscape, anticipate emerging areas of exposure, and identify potential sources of liability.

Introduction

AI systems now power critical functions across every sector of the global economy. What began as isolated deployments led by technology companies has become pervasive integration into the core operations of enterprises worldwide.

Rapid advances in AI capabilities and adoption are transforming the enterprise risk landscape. AI systems introduce new failure modes unique to algorithmic decision-making, while also amplifying traditional risks, especially when AI is implemented at scale. These risks vary depending on the source of the AI system: first-party models (developed in-house) present different risk profiles and control opportunities as opposed to third-party vendor models, which may introduce opacity, dependency and integration challenges.

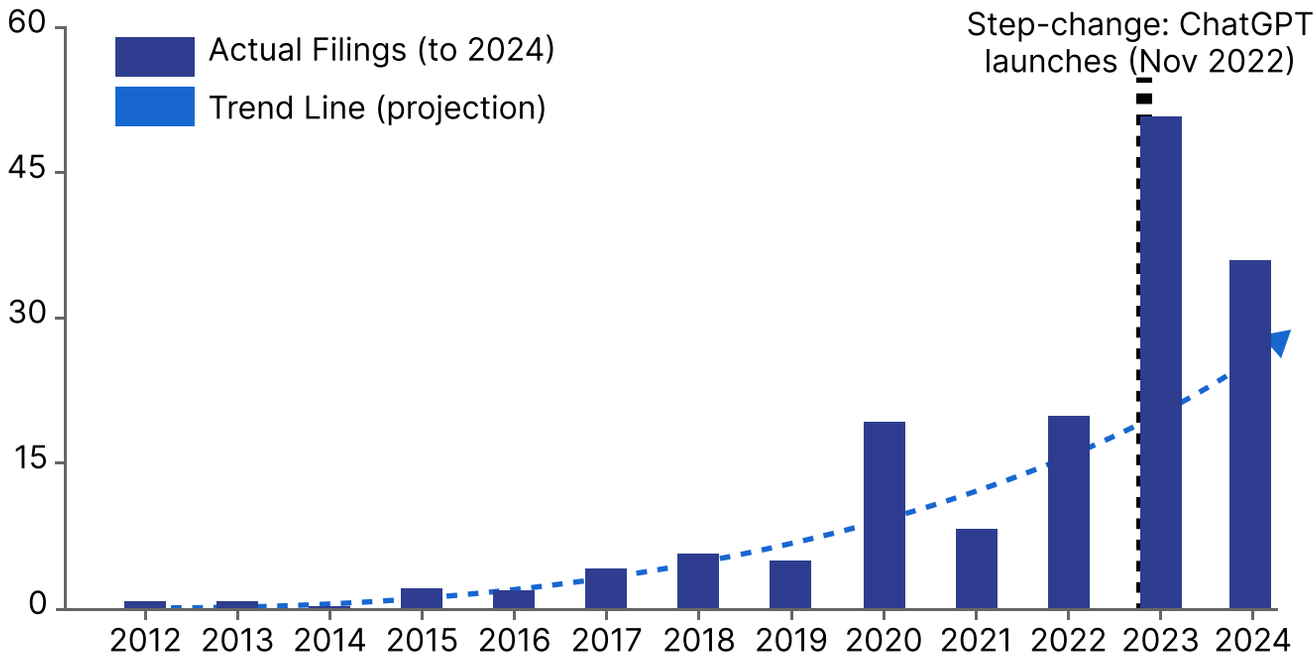
Litigation involving AI systems has shifted from scattered blips to a sustained surge, as evidenced by a new, elevated baseline for such litigation.

New Filings YTD
14

Decisions YTD
19

Total Matters (Since 2012)
189

AI Litigation Accelerates After ChatGPT



Notes: Causation not implied; 2025 incomplete, graph shows filings up to 2024.

For the insurance industry, visibility into how courts, regulators, and sophisticated plaintiffs are actually scrutinizing AI deployments can help them answer relevant questions like: Which attributes of AI implementation matter most when systems face legal challenges? Which industries show concentrated exposure? How are damages calculated and liability allocated across AI supply chains?

To explore these topics and provide an early warning system for emerging risks, Armilla has developed a proprietary AI Litigation Database, which tracks cases and enforcement actions across U.S. jurisdictions. Combined with Armilla's core technical evaluation capabilities, this database enables a more complete view of AI risk, from system-level vulnerabilities to market-level liability trends. Armilla is now making the database available to key partners.

Armilla's AI Litigation Database

Armilla’s Approach

Since 2020, Armilla has pioneered methods to evaluate, quantify, and ultimately price AI risk—establishing itself as a trusted partner across the insurance ecosystem. This includes Armilla’s proprietary technical assessment platform, which standardizes model evaluation, integrates custom datasets tailored to specific risks, and incorporates advanced adversarial testing techniques alongside agentic AI for dynamic and context-aware model testing. Working closely with brokers, underwriters, reinsurers, enterprises, and vendors, Armilla has developed a comprehensive framework for understanding and quantifying the risks inherent in AI deployment.

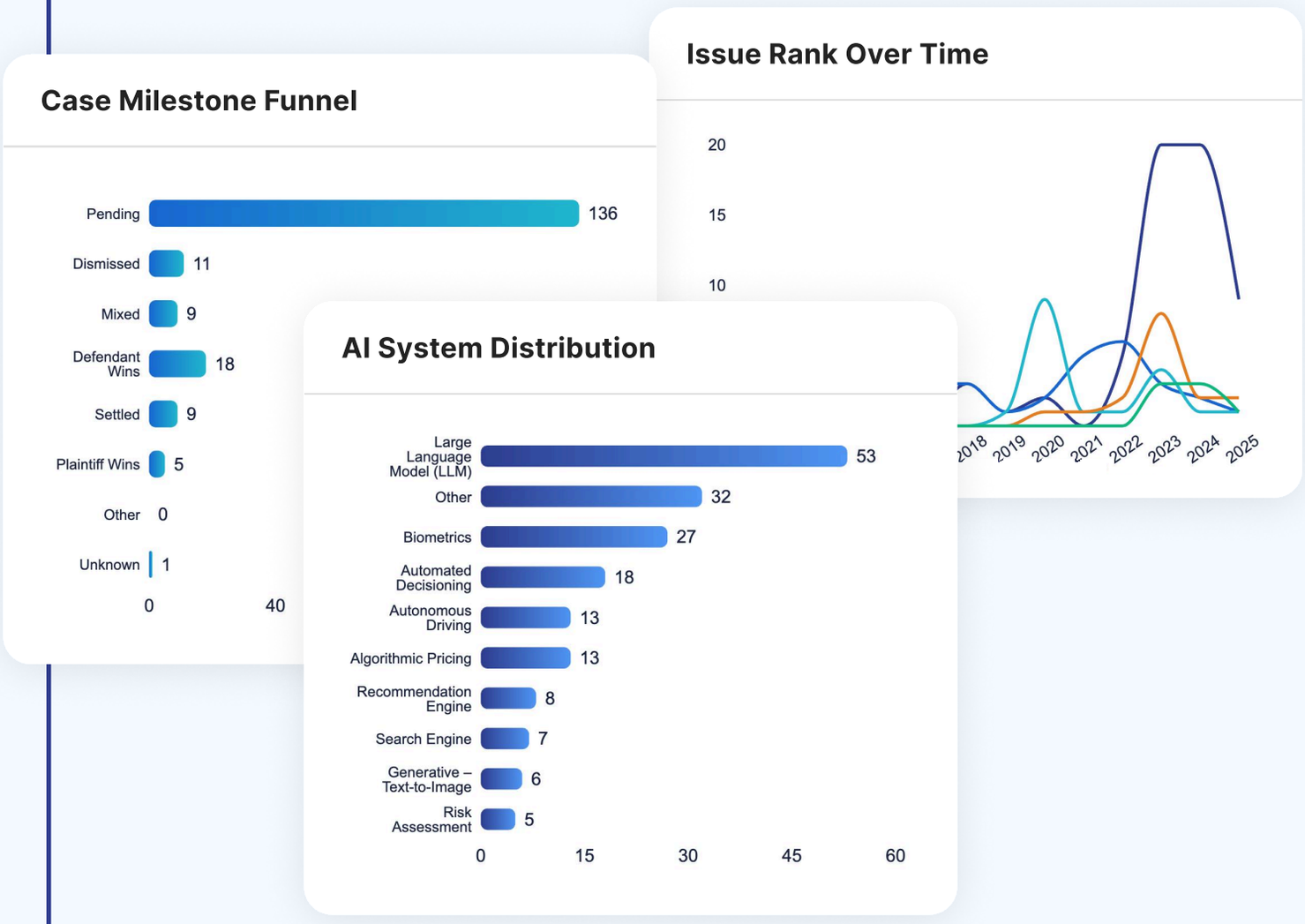
At the core of Armilla's methodology are rigorous technical evaluations of AI systems. These assessments employ advanced techniques including red teaming exercises that simulate adversarial attacks, comprehensive testing protocols, and systematic validation of model outputs against ground truth data. By examining AI systems from multiple angles, including their architecture, training data, deployment context, and operational controls, Armilla can help identify, measure, and price AI risk. In addition to technical evaluation, Armilla's approach is further informed by careful analysis of three critical external signals: emerging litigation trends, evolving regulatory frameworks, and shifting contractual practices in AI procurement and deployment. This multi-faceted perspective enables Armilla to better understand and anticipate AI risks.

Traditionally, insurance relies on litigation outcomes and historical claims data to determine potential loss scenarios for actuarial, underwriting, and rating models. While this approach has proven effective for established perils, it is inherently backward-looking and cannot fully capture the evolving and forward-facing nature of AI risks. Given the absence of historical data on AI-related failures, liabilities, and claims, system- and model-level evaluations offer a valuable, predictive view of how these risks are likely to manifest and where exposures may concentrate.

The combination of both perspectives, therefore—litigation intelligence to understand how risk has materialized and technical evaluation to anticipate how it will—forms the foundation of Armilla’s risk-intelligence platform. This integration allows Armilla to build a continuous feedback loop between emerging litigation signals and system-level evaluations, strengthening its ability to credibly underwrite and price AI risk across sectors and use cases.

AI Litigation Database

Armilla's AI Litigation Database is an actively maintained and monitored database of AI cases and enforcement actions from the U.S. federal government, fifty states and the District of Columbia. Continuously updated and normalized across the attributes that drive exposure—system type, use case, plaintiff profile, industry, legal theory, and outcome—the database allows users to see case details and to analyze where and why risk is concentrating.



By analyzing which organizational practices, AI system characteristics, and supply chain relationships are scrutinized by courts and enforcement bodies, Armilla identifies the attributes that matter most when AI systems are subject to legal review. These insights have profound implications for how enterprises should govern AI, how vendors should design and document their systems, and how insurers may structure coverage and price risk.

The database also enables a critical meta-analysis: determining whether insurance products are keeping pace with the actual risk trends revealed in litigation and enforcement data, or whether gaps are emerging between affirmative policy language and real-world exposures.

Trends and Takeaways

After a quantitative analysis of 200 AI-related cases from the database, 18 high-profile cases were selected for deeper, qualitative analysis. Our analyses revealed five major trends that are shaping the AI risk landscape:

Trend 1

Rise of Corporate Plaintiffs

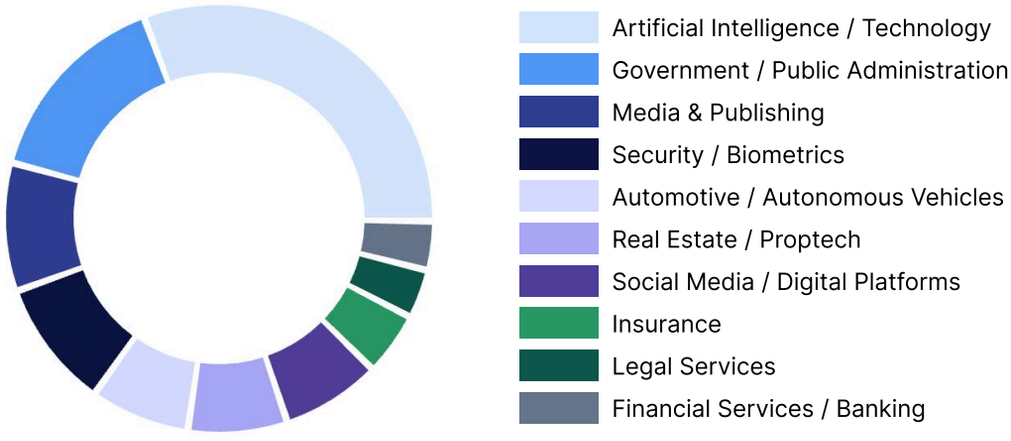
Early landmark AI cases were brought by individuals, such as artists concerned about IP infringement and consumers worried about privacy violations. More recently, major corporations with substantial resources and sophisticated legal teams have become plaintiffs in high-profile AI cases. This shift has dramatically increased both the financial stakes, with potential damages in the hundreds of millions or billions, and the strategic implications, as corporate plaintiffs can sustain prolonged litigation and set precedents that reshape entire industries.

Example

New York Times v. OpenAI and Microsoft

New York Times v. OpenAI and Microsoft illustrates how institutional plaintiffs can pursue claims on a scale far beyond early disputes, alleging systematic copyright infringement across vast training datasets and seeking remedies that could fundamentally alter AI development practices.

AI Litigation by Industry Sector



Trend 2

Economy-wide Effects

AI-related litigation once concentrated in technology companies and HR software providers has now spread into the broader economy. Cases now cut across health care, financial services, media/publishing, real estate/proptech, automotive/AV, and employment. Within the health care industry alone, major cases are underway related to AI use by industry-leaders, including UnitedHealth Group, Cigna, and Humana. This trend is in line with secular adoption of AI across different industries to drive critical decisions. No industry can assume immunity from AI-related liability.

Example

Estate of Gene B. Lokken v. UnitedHealth Group

Estate of Gene B. Lokken v. UnitedHealth Group highlights how AI-driven medical utilization review decisions can create direct liability exposure in healthcare, a sector that historically faced malpractice risk but is now confronting algorithmic accountability as well.

Trend 3

Increasing Enforcement

Following the 2021 Everalbum enforcement action by the US Federal Trade Commission, widely considered the first explicitly AI-related FTC action, the number and scope of regulatory enforcement actions has accelerated. Regulators have moved decisively from issuing guidance documents related to AI use to bringing landmark cases and imposing substantial fines. This transition signals that AI is no longer in a grace period of regulatory forbearance.

Example

Operation AI Comply

In 2024, the FTC announced *Operation AI Comply*, an enforcement initiative targeting enterprises for allegedly making deceptive claims about their AI use. It also shared details of a settlement with *DoNotPay, Inc.*, which included a civil penalty of $193,000 and stringent injunctive relief.

Trend 4

Rise of Class Actions

When courts certify classes in AI-related disputes, enterprise exposure multiplies dramatically. Class certification transforms individual complaints into systemic challenges, multiplying potential damages and creating settlement pressure even when defendants believe they have strong defenses. The increasing willingness of courts to certify AI-related class actions reflects growing judicial comfort with algorithmic harm as a legal theory.

Example

Mobley v. Workday

In *Mobley v. Workday (2023)*, the court allowed claims of algorithmic hiring discrimination to proceed as a collective action, a decision reaffirmed in 2025. This opens the door for potentially thousands of affected individuals to join the litigation.

Trend 5

Shift from Broad to Precise Claims

Earlier cases often alleged broad, somewhat exploratory harms, such as general concerns about data scraping. Recent cases demonstrate increasing legal sophistication, with plaintiffs alleging specific violations, such as material misrepresentations in SEC filings or direct vendor liability for autonomous agent failures. This precision makes cases harder to dismiss on procedural grounds.

Example

Global Predictions Inc.

The March 2024 *Global Predictions Inc.* action exemplifies this trend toward specificity, with regulators citing concrete misrepresentations about AI system capabilities rather than making general assertions about consumer harm.

Projections

As noted above, AI litigation data offers valuable insight into emerging liability risks while presenting distinct analytical challenges. Litigation operates retrospectively: courts adjudicate harms from models deployed years earlier, creating a natural gap between observable case law and the risk profile of current and future systems. This temporal offset is particularly pronounced in AI given its pace of technical development and the evolving science related to its control and alignment. Effective risk assessment therefore requires dual analysis: extracting doctrinal patterns from existing litigation to understand how courts are interpreting traditional liability frameworks in AI contexts, while simultaneously evaluating current and anticipated system capabilities to project forward-looking risk trends relevant to insureds.

These trends collectively indicate a maturing AI litigation landscape characterized by increased claim frequency, technical specificity, plaintiff sophistication, and financial exposure. This evolution supports the development of AI risk as a distinct insurable category.

In 2026 and beyond, Armilla expects to see the following trends play out:

1 Cases involving specific AI systems will have higher stakes

Remarkably, the plaintiff in Mobley v. Workday brought suit against Workday, not the enterprises using its AI-powered employment screening platform. The court is allowing the case to proceed under a novel **agent theory of liability**, whereby Workday can potentially be held liable as an agent of those enterprises. This theory of liability, combined with trends related to class actions and enforcement, will raise the stakes in lawsuits and actions related to specific AI systems.

Clarifying Liability Regimes and Implications for AI Litigation
AI is reconfiguring traditional liability frameworks by challenging core doctrines of negligence, strict liability, vicarious liability, and agency law.

- **Tort and the law of negligence** holds parties liable for failing to exercise reasonable care that a prudent person would use to prevent foreseeable harm to others.
- **Strict liability** imposes responsibility for certain inherently dangerous activities or defective products regardless of fault or intent, focusing solely on causation of harm.
- **Vicarious liability** makes one party (typically an employer) legally responsible for the wrongful actions of another (typically an employee) performed within the scope of their relationship.
- **Respondeat superior** is a concept within agency law, whereby principals are liable for their agents' conduct. It traditionally assumes human-to-human relationships with clear control and authority structures.
- **Agency law**, under a new theory of liability currently being considered in Mobley v. Workday, would treat a software provider as an agent subject to liability in the course of its performance of functions on behalf of multiple employer-principals.

For helpful commentary and analysis of these topics, see Americans for Responsible Innovation, *The Stick, the Carrot, and the Net: Policy Approaches for Addressing AI Agent Harms*, August 18, 2025; and Gabriel Weil et al., *Insuring Emerging Risks from AI*, November 14, 2024, available online.

2 Clearer AI liability regimes will catalyze AI-related litigation

Together, AI-related case law, legislative reforms, and standardized industry practices are establishing clearer AI liability rules and lines of enterprise responsibility. Clearer AI liability regimes will result in easier identification of defendants and enforceable pathways and remedies for AI-related claims. **Armilla expects this newfound clarity to catalyze AI-related litigation, as theoretical liability gives way into unambiguous accountability.**

3 Enterprise AI risk will concentrate in specific use cases

AI risk is concentrated where three conditions overlap: high-volume end-user, consumer, or patient interactions; statutory damages or established anti-discrimination frameworks; and system-level decisions that affect thousands or millions of people at once. This will place enterprises in financial services, health care, employment technology, and consumer platforms at the sharpest edge of AI risk exposure.

4 Beyond hallucinations, enterprises will be held responsible for unchecked model errors and performance issues

While the providers of generative AI models will continue to grapple with copyright and other IP-related claims, courts will expect enterprises deploying these generative AI models to play a role in assuring AI model performance and reliability. Hallucinations are a well-known phenomenon related to generative AI use. 5% of the cases examined within Armilla's AI Litigation Database relate to hallucinations or misrepresentations, a category that Armilla expects to grow significantly within the near future. However, in addition to human validation of outputs, enterprises can now draw upon well-known technical methods to control hallucinations. So, Armilla expects courts to find that, when integrating generative AI outputs into their products and services, enterprises are responsible for taking steps to reduce model errors and performance issues.

> **"Over 90% of business customers surveyed by the Geneva Association indicate they need coverage for Gen-AI-related losses."**
>
> Based on a 2025 survey of 600 respondents

Implications for Enterprises

The trends revealed in Armilla's AI Litigation Database paint a clear picture: AI risk is accelerating, broadening, and becoming more legally sophisticated.

Enterprises must adopt a comprehensive, multi-layered strategy that translates AI litigation intelligence into actionable risk management. It should include the following elements.

1 Governance: Build Defensible Decision Clarity

AI governance is now an evidentiary necessity. In cases such as *Mobley v. Workday* and *Estate of Lokken v. UnitedHealth Group*, plaintiffs are winning the rights to discovery into how AI decisions were made, validated, and overseen. Enterprises must document their AI governance programs, including clear policies defining oversight, model inventories, data quality standards, approval workflows for high-risk applications and system-level validation processes. Robust documentation and audit trails can be critical in limiting liability—as shown in the 2021 investigation by the New York Department of Financial Services, which found no violations in discrimination claims related to the Apple Card and its underwriter, Goldman Sachs.

2 Evaluation and Assurance: Technical Rigor into Legal Resilience

Courts are increasingly scrutinizing whether enterprises have performed meaningful testing and monitoring of AI systems. For high-risk AI systems, courts now expect enterprises to conduct rigorous technical evaluations including systematic testing, adversarial red-teaming, and independent third-party audits. Occasional audits are not enough. Rather, courts expect continuous monitoring systems that track model performance, detect data drift, and alert stakeholders when systems deviate from expected behavior. Over half of the cases analyzed from the AI Litigation Database included allegations of insufficient model validation or deceptive claims about performance. Documentation of rigorous technical evaluations has become an important liability shield and a form of legal and financial risk mitigation.

Maintaining stakeholder trust, from customers and employees to investors and regulators, depends on demonstrable commitment to responsible AI deployment using all available tools: governance that creates evidentiary trails, evaluation that proves system control, compliance that anticipates enforcement, and insurance that affirmatively protects against AI-related exposures.

3 Compliance: Anticipate Enforcement, Don't Chase It

Enterprises must align with legal requirements under the EU AI Act, the Colorado AI Act, California's Automated Decision making Technology Regulations and other applicable laws and regulations. Some laws—such as the Illinois Biometric Information Privacy Act (BIPA) and, in certain cases, the California Consumer Privacy Act (CCPA)— even include a private right of action, allowing individuals to sue directly for violations. At the same time, regulators are increasingly bringing high-profile enforcement actions, as evidenced by the FTC's Operation AI Comply as well as actions brought by state Attorneys General. Though enforcement intensity may vary with political leadership, it is now generally accepted that false or opaque AI claims can constitute deception. Enterprises should map AI deployments against existing industry-specific and jurisdiction-specific regulations and resource compliance teams and specialized legal counsel. Litigation data indicates that documentation of proactive compliance efforts can also help reduce legal and financial risk.

4 Insurance: Translate Residual Risk into Financial Protection

Even the most mature AI governance programs cannot eliminate residual exposure. The distributed nature of AI supply chains means exposure extends to training data providers, model developers, cloud vendors, and integration partners. Since courts have also allowed evolving theories of liability, enterprises, vendors, or both could be drawn into litigation related to an algorithmic event. This is why enterprises use insurance to transfer unpredictable AI-related exposure into quantifiable, transferable risk.

Armilla's AI liability coverage responds directly to these evolving theories, insuring not just negligence but also performance degradation, hallucination-induced harm, and agentic liability.

> **"The emergence of affirmative AI insurance products marks a key shift in the industry's approach to managing AI-driven risks. With companies like Armilla leading the charge, insurers are beginning to address perceived coverage gaps that traditional policies may overlook."**
>
> Bracken, Levine and Pappas, Hunton LLP

> **"For AI-forward clients, putting affirmative AI coverage on the table isn't optional—it's part of my professional responsibility. I owe them the opportunity to compare wording, coverage, limits, and pricing—and make an informed decision with the best options in front of them."**
>
> EVP, Top 5 US Brokerage, Armilla Appointed

Next Steps

Armilla will make the AI Litigation Database available through a portal to key partners across the insurance ecosystem, providing early visibility into emerging AI liability trends.

Armilla is committed to continuously improving the database by expanding coverage to include cases and jurisdictions beyond the U.S., developing more granular classification systems that capture specific AI failure modes and legal theories, and implementing an agentic interface that enables deep dives into specific cases for interactive research and pattern analysis. Though monitoring and analyzing litigation signals is not a replacement for developing deeper understanding and evaluation of AI risks, it can provide early warning of emerging AI risks and supplement other sources of regulatory and contractual intelligence.

By maintaining the AI Litigation Database as a resource for partners, Armilla aims to help enterprises navigate AI risk with greater confidence and foresight.

ARMILLA

Ready to Navigate the AI Risk Landscape? Request Partner Portal Access

Contact us at

info@armilla.ai

www.armilla.ai

This report is based on analysis of 197 AI litigation cases and proprietary market research.
Prepared with support from the council on AI governance.