A Guide to AI Risk Management and Insurance for Modern Companies

AUTHORED BY

GUNDERSON
DETTMER

Authors

Katie Gardner is a Strategic Licensing and Transactions Partner in Gunderson Dettmer's New York office. Katie has over a decade of experience in the EC/VC space, and specializing in strategic transactions, intellectual property and technology licensing and protection, AI use and implementation, product counseling, data and privacy strategy, and the commercialization of intellectual property, software, data and other technology assets.

Aaron Rubin is a Strategic Licensing and Transactions Partner in Gunderson Dettmer's New York office. Aaron counsels emerging growth and startup companies on strategic IP and technology matters in all aspects of their businesses, from day-to-day operations to financings to M&A activities. He focuses on strategic partnerships, IP and technology licensing, AI use and implementation, product counseling, data privacy strategies, and the commercialization of software, data, and other tech assets.

Table of Contents

Introduction

Artificial intelligence (“AI”) tools have quickly become an essential element of modern business operations, transforming how companies innovate and compete.

At the same time, the rapid pace of AI adoption has introduced a unique set of risks that commercial business insurance and existing standard practices and contractual provisions were not designed to protect, and thus is fast evolving.

This white paper aims to help founders and executives understand and mitigate AI-related risks, including by demystifying the evolving AI insurance landscape. We begin by outlining the key risk categories facing companies that use, develop or deploy AI technologies, followed by a discussion of the limitations of traditional insurance products in addressing these exposures. We then examine how AI-specific coverage can close critical gaps, breaking down the distinct types of coverage available, the nature of claims they address, and the operational factors companies should weigh in evaluating their insurance needs. We close with a roadmap for securing appropriate coverage and implementing legal best practices to proactively manage AI-related business risks.

Emergent Risks for Companies Using or Offering AI Products and Services

Companies developing or deploying AI products and services must navigate a complex and rapidly evolving landscape of regulations, industry standards, and contractual requirements.

These legal and compliance requirements vary widely depending on several factors, including geographical location, industry, risk profile and use case. Key areas of AI-related exposure include:

Intellectual Property Infringement

- Use of copyrighted material (e.g., images, text, code, datasets) in training or outputs without proper rights or licenses
- Risk of misappropriating trade secrets or confidential information used in model development or fine-tuning
- Lack of clarity around IP ownership in generated outputs, co-developed models, or fine-tuned versions of foundation models

Contractual Liability

- Breach of contract due to inaccurate representations or warranties about AI functionality or compliance
- Indemnification obligations triggered by downstream misuse or harm caused by AI tools
- Misalignment between contractual commitments (e.g., service levels, use restrictions) and actual capabilities or limitations of the AI system
- Failure to flow down key terms in subcontractor or vendor agreements (e.g., data usage, security requirements)

Liability Under Applicable Laws

- Violations of data privacy laws (e.g., GDPR, CCPA, HIPAA) through the collection, processing, or output of personal or sensitive data
- Discrimination or bias in automated decision-making, especially in regulated areas like hiring, credit, housing, or healthcare (e.g., violations of the EU AI Act, Title VII, or the Fair Credit Reporting Act)
- Non-compliance with transparency, explainability, or labeling obligations under emerging AI-specific regulations (e.g., EU AI Act, NY AEDT Law, FTC guidance)
- Export control and sanctions risks related to training data, compute resources, or deployment locations

Tort and Criminal Liability

- Negligence claims arising from failure to adequately monitor or test AI systems before deployment
- Product liability for physical or financial harm caused by autonomous or semi-autonomous AI systems
- Defamation, libel, or false light arising from AI-generated content
- Exposure to criminal liability for use or misuse of AI in ways that violate laws (e.g., fraud, impersonation, wiretapping, surveillance)

Intellectual Property Infringement

Multimodal generative AI applications enable users to create content, such as images, text, video, or audio based on input prompts. These systems work by identifying patterns within extensive training datasets and generating outputs that reflect those patterns. However, because many models are trained on datasets that may include unlicensed or copyrighted material, both developers and users face the risk of inadvertently infringing third-party intellectual property rights.

Copyright Infringement

COPYRIGHT INFRINGEMENT RISKS

Businesses may face copyright liability from two key aspects of generative AI: (1) the use of copyrighted materials in training datasets, and (2) the generation of outputs that resemble or reproduce protected works. These risks generally fall into two categories:

- **DIRECT INFRINGEMENT:** Direct liability can arise where (a) copyrighted works are used during the training of an AI model without authorization, potentially constituting unlawful reproduction or creation of derivative works; or (b) the AI system generates outputs that are substantially similar to a copyrighted work, thereby infringing the original author's exclusive rights.

- **CONTRIBUTORY OR VICARIOUS INFRINGEMENT:** An AI provider may be held secondarily liable if it facilitates or benefits from infringing activity, especially where it knew or should have known of the infringement, or had the ability to control it but failed to act.

Companies that knowingly use AI technology trained on unlicensed data to generate outputs may be found to have created infringing derivative works, exposing them to significant liability under U.S. copyright law – including statutory damages of up to $150,000 per infringed work.

LEGAL LANDSCAPE: DERIVATIVE WORKS AND THE FAIR USE DEFENSE

A growing number of copyright owners, including artists and authors, have filed lawsuits against AI companies, alleging that their copyrighted works were used without authorization to train AI models and generate infringing "derivative works." In response, many AI providers assert that such use is protected under the **fair use doctrine**, a key limitation on copyright infringement under U.S. law.

These lawsuits are shaping the legal boundaries of AI training practices and often center on whether large-scale scraping and use of copyrighted content for model development constitutes fair use. Courts are examining issues such as the lack of express consent from rights holders, the degree to which the models or outputs are "transformative," and the economic impact on the original creators.

- **DERIVATIVE WORK:** Under the U.S. Copyright Act (17 U.S.C. § 101), a derivative work is a new creation based on one or more preexisting works, including adaptations such as translations, dramatizations, and editorial revisions. A derivative work may still infringe the original author's rights if created without authorization — even if it includes new elements. The exclusive right to prepare derivative works remains with the original copyright holder, and unauthorized derivations may give rise to liability.

- **FAIR USE DOCTRINE:** Fair use (17 U.S.C. § 107) is a statutory defense to copyright infringement, allowing limited use of copyrighted materials without permission in specific circumstances. Courts evaluate fair use based on four non-exclusive factors:

Purpose and character of the use, including whether it is transformative and whether it is for commercial or nonprofit purposes;	**Nature of the copyrighted work, with creative works typically receiving greater protection than factual ones;**	**Amount and substantiality of the portion used in relation to the whole work; and**	**Effect of the use on the market for the original work.**

- **DIGITAL MILLENNIUM COPYRIGHT ACT (DMCA):** Initially enacted in 1998, the DMCA was designed to address copyright issues in the digital age and established rules for online service providers for the use, distribution, and protection of copyrighted content. The DMCA contains two key provisions: (1) a "safe harbor" that shields online service providers from liability for copyright infringement arising from user-generated content hosted on their platforms, and (2) a prohibition on the circumvention of technological measures (such as encryption or digital rights management tools) designed to protect copyrighted works. While the DMCA's safe harbor could apply if AI providers operate or host platforms that allow end users to post content, AI providers must demonstrate they lack direct control over the creation of the infringing content, promptly respond to valid takedown notices, and implement a repeat infringer policy. Additionally, when training AI tools with copyrighted works, companies risk violating the DMCA by removing copyright management information (such as the author's name, the title of the work and any copyright notices) from the data used as training inputs.

The outcome of pending cases may significantly shape how courts apply these principles to generative AI, with implications for both developers and downstream users.

EMERGING LITIGATION AND INDUSTRY IMPACT

Ongoing lawsuits highlight the complex and unsettled legal landscape surrounding ownership, originality, and liability in the context of AI-generated content. These cases have amplified calls for clearer statutory guidance to define the boundaries of copyright protection for both creators and AI systems. In the absence of such clarity, companies leveraging generative AI tools remain exposed to potential claims. Recent high-profile cases include:

- **THE NEW YORK TIMES V. MICROSOFT AND OPENAI:** The Times filed suit in December 2023, alleging that OpenAI and Microsoft used millions of its copyrighted articles without authorization to train large language models such as ChatGPT. The complaint includes claims of direct and vicarious copyright infringement, violations of the Digital Millennium Copyright Act (DMCA) for removal of copyright management information, and unfair competition.

- **ANDERSEN V. STABILITY AI:** A group of visual artists brought a putative class action against Stability AI, DeviantArt, and Midjourney, asserting that the defendants scraped billions of copyrighted images from online sources to train their AI model, Stable Diffusion. The plaintiffs claim that the resulting AI-generated images are unauthorized derivative works, and that the defendants engaged in direct, contributory, and vicarious copyright infringement, as well as violations of DMCA provisions and unfair competition laws.

- **RIAA V. SUNO AND UDIO:** On behalf of Sony Music, Universal Music Group, and Warner Music Group, the RIAA filed copyright infringement lawsuits against Suno and Udio, alleging that the companies used copyrighted music recordings in their training datasets without permission. The complaints include claims of direct infringement, vicarious infringement, and willful misappropriation of copyrighted sound recordings.

- **CONCORD MUSIC V. ANTHROPIC:** Universal Music Group, Concord Music, and ABKCO Music sued Anthropic in October 2023, alleging that its AI model, Claude, generated song lyrics substantially similar to hundreds of copyrighted works, including ***I Will Survive*** by Gloria Gaynor. The plaintiffs asserted direct copyright infringement and sought statutory damages of up to $150,000 per work, totaling over $75 million.

- **BARTZ V. ANTHROPIC:** Three authors sued Anthropic in August 2024, alleging its Claude AI model was trained on copyrighted books without authorization, including both lawfully acquired works and pirated digital copies. In June and July 2025, Judge William Alsup issued a split decision by ruling that: (1) it was “exceedingly transformative” and fair use for Anthropic to use legally obtained books, including digitized copies of purchased books in print, for training Claude AI models; however, (2) the court also found that using pirated titles from sites like LibGen was not fair use, regardless of whether such works were later bought legally. The court

consequently certified a nationwide class action for authors whose works were included in Anthropic's "pirate library," exposing Anthropic to significant potential damages if found liable for infringement of these pirated materials.

Ownership Uncertainty of AI-Generated Outputs

The question of who owns AI-generated content remains a significant legal gray area, with important implications for companies building or using generative AI tools. Under current U.S. law, copyright protection is only available for works that are the product of **human authorship**. The U.S. Copyright Office has repeatedly stated that content generated solely by an AI system without meaningful human input is not eligible for copyright protection. This position introduces several practical and legal challenges:

- **LACK OF COPYRIGHT PROTECTION:** If a work is determined to be generated entirely by an AI system without sufficient human creativity, it may fall into the public domain — leaving it without enforceable IP rights. This creates risk for companies relying on such content for competitive advantage or commercialization.

- **COLLABORATIVE INPUTS AND UNCLEAR AUTHORSHIP:** Many AI-assisted works involve varying levels of human involvement, e.g., crafting prompts, editing outputs, or combining generated content with human-authored material. In such cases, ownership may be uncertain or shared, and assessing whether the human contribution is substantial enough to qualify for copyright protection is highly fact-dependent.

- **THIRD-PARTY MODEL USE COMPLICATIONS:** When outputs are generated using third-party or open-source models, additional layers of complexity emerge. Model licenses may be silent or ambiguous on output ownership, or may assert that all rights remain with the model provider — raising commercial and IP concerns for downstream users.

- **IMPLICATIONS FOR LICENSING AND COMMERCIAL USE:** If a company cannot claim exclusive rights in its AI-generated outputs, it may be unable to prevent others from copying, modifying, or reusing that content. This affects the ability to monetize, license, or enforce rights in AI-assisted works.

Risk of Misappropriating Trade Secrets or Confidential Information

The development and fine-tuning of AI models often involve the ingestion of large datasets — sometimes provided by third parties, obtained from public or scraped sources, or derived from user interactions. If these datasets contain trade secrets or confidential business information, companies may face legal exposure for misappropriation, even if the inclusion was inadvertent.

- **UNINTENTIONAL INGESTION OF PROTECTED INFORMATION:** Ingesting data from unvetted sources — such as scraping public websites, using user-uploaded content, or incorporating third-party datasets without proper diligence — raises the risk of incorporating proprietary information protected under trade secret laws or NDAs. This can occur even if the information is not labeled as confidential.

- **TRADE SECRET MISAPPROPRIATION:** Under U.S. law, liability for trade secret misappropriation can arise where a company knew or had reason to know that it acquired or used confidential information through improper means. Use of a model trained on misappropriated data, whether internally or through a vendor, can give rise to both direct and vicarious liability.

- **RESIDUAL DISCLOSURE THROUGH OUTPUTS OR MODEL BEHAVIOR:** Even if training data is no longer accessible in its original form, there's growing scrutiny around whether models can memorize or reproduce sensitive data points in outputs, especially in the case of fine-tuned models with narrow datasets. This is particularly concerning in regulated industries like healthcare, finance, and life sciences, where inadvertent disclosure of sensitive information could also trigger regulatory penalties.

- **VENDOR AND OPEN MODEL RISK:** When using third-party vendors or open-source foundation models, companies should assess whether proper rights were obtained for the training data. Absent clear contractual representations and audit rights, businesses may inherit liability for downstream use of improperly sourced data.

Contractual Liability

While much of the legal focus around AI centers on compliance with evolving laws and regulations, a significant source of risk arises from the contractual commitments companies make when developing, licensing, or deploying AI systems.

These obligations may be found in customer agreements, vendor contracts, partnership terms, or even publicly posted terms of service. Because AI systems are inherently probabilistic and often trained on third-party data, overbroad representations, indemnities, or performance guarantees can expose companies to liability well beyond what the law would otherwise impose. Carefully navigating and negotiating these contractual terms is essential to managing risk in the commercial use of AI.

Misrepresentations and Warranties

Contractual liability often arises from inaccurate or overly broad representations and warranties concerning an AI system's functionality, compliance, or intended use. Because generative and predictive models are inherently probabilistic, making absolute claims about their accuracy, legality, or applicability to sensitive use cases can create significant legal exposure.

- **OVERPROMISING CAPABILITIES:** AI systems may be marketed as "accurate," "safe," or "bias-free"—terms that can be difficult to defend if outputs are inconsistent or flawed. Companies that oversell system capabilities, especially in contracts or sales materials, risk breach of warranty claims or allegations of misrepresentation. This risk is heightened when performance metrics (e.g., error rates, accuracy thresholds) are not clearly defined or caveated.

- **COMPLIANCE REPRESENTATIONS:** Some contracts include representations that an AI system is "compliant with applicable law," which can be problematic given the evolving regulatory landscape and uncertainty around how existing laws apply to AI. Such representations may give rise to indemnification obligations or breach claims if downstream users face legal consequences.

- **USE CASE RESTRICTIONS:** Even if a system performs adequately in general use, liability can arise when it is used in high-risk or regulated contexts—such as employment screening, credit scoring, insurance underwriting, or healthcare decision-making. Without clear disclaimers or contractual limitations, a provider may be held responsible for outputs used in prohibited or unintended ways. In some cases, regulators may treat the provider as a participant in the decision-making process, potentially triggering legal obligations under sector-specific laws.

Indemnification Obligations

AI-related contracts often include indemnification clauses that allocate responsibility for third-party claims. These provisions can create significant liability if not narrowly tailored to account for the unique risks of AI systems. Companies developing, licensing, or deploying AI tools may find themselves exposed to unanticipated indemnity obligations, especially in areas such as IP infringement, data misuse, or downstream harm caused by generated outputs.

- **INFRINGEMENT INDEMNITIES:** Customers frequently seek broad indemnification for third-party claims arising from alleged intellectual property infringement by the AI system or its outputs. This is particularly risky where the model was trained on data of uncertain provenance or where outputs may resemble protected works (e.g., copyrighted images or proprietary content). Without proper exclusions or limitations, AI providers may be on the hook for costly litigation—even when claims are speculative or outside their control.

- **DATA AND PRIVACY CLAIMS:** Indemnity obligations may also be triggered by the inclusion of personal data, protected health information, or other regulated content in training or output data. This includes claims under laws like the GDPR, CCPA, or HIPAA, especially where the AI tool unintentionally memorizes or reveals sensitive information. If vendors don't adequately vet training data sources, customers may demand indemnification for resulting privacy violations or data breach claims.

- **BROAD OR AMBIGUOUS TRIGGERS:** Some indemnity clauses are drafted to cover any claim "arising out of" or "relating to" the use of the AI tool, without limiting the types of claims or requiring fault. These sweeping provisions can effectively turn AI vendors into insurers of downstream use, regardless of the user's conduct or the unpredictability of model behavior. This is particularly concerning for generative AI, where outputs may be shaped by user prompts or external context beyond the provider's control.

Performance and Availability Commitments

AI systems—particularly generative and large language models—are inherently probabilistic and may exhibit inconsistent behavior depending on inputs, prompts, context, or system load. Despite this, enterprise customers often seek performance guarantees that are difficult or impossible for AI providers to meet with reliability. Overcommitting in service-level agreements (SLAs) or other contractual provisions can expose providers to breach claims, service credits, or termination rights.

- **SERVICE LEVEL AGREEMENTS (SLAS):** Traditional SLAs, such as those used in SaaS or infrastructure agreements, may not align with the realities of AI performance. Uptime commitments or deterministic output guarantees may not be feasible where the system's functionality depends on evolving training data, user prompts, or third-party APIs. SLAs that fail to account for the non-deterministic nature of AI can result in chronic underperformance, contractual disputes, or demands for remedies the provider cannot support.

- **SUPPORT AND RESPONSE OBLIGATIONS:** AI tools may experience unique failure modes, including hallucinations, prompt injection attacks, or model drift over time. If support obligations do not address these risks specifically, vendors may be contractually required to remediate issues that are not technically resolvable or may require costly retraining or reengineering. Vague "bug fix" or "error correction" language may inadvertently encompass these AI-specific anomalies.

- **FAILING TO MEET AI-SPECIFIC KPIS:** Some customers attempt to impose accuracy thresholds, bias mitigation standards, or explainability requirements as contractual deliverables. While well-intentioned, these terms can be difficult to define, measure, or control—especially where outputs vary with use case or rely on third-party infrastructure. Providers that accept strict performance metrics without clear exclusions or caveats may find themselves in breach even when the system operates as designed.

Use Restrictions and Acceptable Use Violations

AI providers may face risk when customers or end users deploy their systems in ways that exceed intended or authorized use cases. Without clear contractual limitations and robust enforcement mechanisms, companies may be held liable for misuse — particularly when outputs cause harm or implicate legal and regulatory frameworks.

- **END USER MISUSE:** Generative AI systems can be used to create harmful, offensive, or unlawful content, including hate speech, deepfakes, discriminatory outputs, or misinformation. If contractual terms do not clearly prohibit such use — and if monitoring or enforcement is lacking — providers may face reputational harm, regulatory scrutiny, or even tort liability for enabling dangerous applications. This risk is amplified in high-stakes sectors like employment, healthcare, financial services, and education.

- **BREACH OF MODEL OR DATA USE RESTRICTIONS:** Many foundation models and training datasets are subject to license terms that limit commercial use, redistribution, or retraining. Downstream customers who unknowingly or carelessly breach these restrictions may expose the original provider to liability — particularly where the provider failed to impose "flow-down" obligations or monitor compliance. Similarly, using outputs to train new models may violate upstream data or model licenses if not explicitly permitted.

- **UNAUTHORIZED OUTPUTS OR DERIVATIVE USES:** Companies may assume they have broad rights to use AI-generated content for commercialization, product development, or redistribution. However, unless the contract specifies permissible use, disputes can arise over ownership, scope, or rights of reuse. Additionally, some customers may feed outputs into other systems or use them in unintended ways (e.g., medical diagnosis, automated hiring decisions), potentially triggering legal obligations or risks that were never contemplated by the provider.

Flow-Down Risk in Partner and Supply Chain Agreements

As companies increasingly rely on complex AI supply chains — including third-party model providers, data licensors, and infrastructure partners — contractual misalignment across the chain can create significant risk. A failure to properly flow down critical terms or reconcile obligations between upstream and downstream agreements can expose providers to liability they did not anticipate or intend to assume.

- **LACK OF ALIGNMENT ACROSS CONTRACTS:** Companies may make representations, warranties, or commitments to customers that exceed or conflict with the rights they have obtained from upstream vendors or licensors. For example, a provider may promise unrestricted commercial use of an AI-generated output while relying on a model or dataset that is licensed only for research or non-commercial use. This mismatch can result in breach of contract, indemnity claims, or license termination.

- **FAILURE TO FLOW DOWN KEY TERMS:** Contracts with customers often require compliance with data protection laws, use restrictions, audit rights, or content moderation obligations. If these requirements are not passed down to vendors, model providers, or subcontractors involved in model training, hosting, or integration, the company may be unable to fulfill its obligations or control key risks. This is particularly relevant where sensitive data is processed or outputs are generated that could trigger regulatory obligations.

- **LACK OF AUDIT AND OVERSIGHT RIGHTS:** Without appropriate audit or reporting rights, companies may not be able to verify that third parties are complying with contractual requirements, such as data minimization, content filtering, or license compliance. This creates blind spots in risk management and makes it difficult to defend against customer or regulatory claims tied to third-party behavior.

- **INDEMNITY AND LIABILITY GAPS:** Even if a company is required to indemnify a customer for a third-party claim, it may not have back-to-back indemnification rights from its vendors. This can result in uncovered losses, particularly in cases involving IP infringement, privacy violations, or misuse of training data.

Limitation of Liability and Remedy Limitations

Given the unpredictable and evolving nature of AI systems, it is critical for companies to carefully negotiate limitations on liability and available remedies in AI-related contracts. Without tailored protections, providers can find themselves exposed to disproportionate financial and legal risk, especially where AI outputs are used in sensitive or high-stakes environments.

- **DISPROPORTIONATE EXPOSURE:** AI vendors may be asked to accept liability for harms stemming from the use or misuse of their systems, including errors, bias, hallucinations, or downstream decisions based on AI outputs. If standard limitation of liability clauses are omitted or insufficiently scoped, companies may be exposed to uncapped or high-dollar claims, particularly where contract damages include consequential, incidental, or special damages — such as reputational harm, loss of business, or regulatory fines.

- **MISALIGNMENT WITH PRODUCT REALITIES:** AI systems are often non-deterministic and probabilistic in nature, meaning that outputs may vary over time or differ based on input phrasing. Agreeing to traditional commercial remedies, like refunds, service credits, or re-performance, may not be practical or effective for addressing failures like hallucinated outputs, unpredictable behavior, or prompt injection attacks. These issues often cannot be "fixed" in the conventional sense.

- **INSUFFICIENT CARVE-OUTS OR OVERBROAD EXCEPTIONS:** Many contracts attempt to exclude limitations of liability for certain types of claims (e.g., IP infringement, privacy violations, or breaches of confidentiality). Without careful drafting, these carve-outs can unintentionally gut the liability cap or extend to areas of high, unbounded risk, particularly where IP or data issues are tied to third-party models or training datasets.

For more information on AI commercial contracting best practices, please refer to the "AI Terms in Commercial Deals: Negotiating Realistic Terms and Aligning Expectations" webinar, which is also available on the Gunderson Dettmer Generative AI Resources hub.

Compliance with Applicable Laws

Companies deploying or offering AI tools must navigate a growing web of legal and regulatory obligations. In addition to existing privacy, consumer protection, anti-discrimination, and sector-specific rules, AI-specific laws are rapidly emerging across jurisdictions. This section outlines the key areas of legal exposure under U.S. and international laws.

Consumer Protection and Fair Competition

AI technologies that generate content, automate decisions, or provide recommendations can expose companies to liability under longstanding consumer protection and competition laws. U.S. regulators—including the Federal Trade Commission (FTC), Consumer Financial Protection Bureau (CFPB), and Department of Justice (DOJ)—are actively applying existing statutes to AI-powered tools, particularly where they may cause consumer harm, unfair competitive advantage, or discrimination.

FEDERAL TRADE COMMISSION (FTC)

The FTC enforces Section 5 of the FTC Act, which prohibits unfair or deceptive acts or practices and unfair methods of competition. In recent guidance and enforcement actions, the FTC has emphasized that AI does not excuse companies from longstanding legal responsibilities. Key areas of regulatory focus include:

- **DECEPTIVE AI MARKETING CLAIMS:** Businesses must substantiate performance claims about AI tools. Unverified statements like “bias-free” or “human-level accuracy” can be deemed deceptive.

- **DISCRIMINATORY OR HARMFUL OUTPUTS:** Companies may violate the law if they use or deploy AI tools that result in unfair or discriminatory outcomes without appropriate testing, safeguards, and oversight.

- **IMPROPER USE OF TRAINING DATA:** The FTC has penalized companies for using data collected without consent to train AI models and has required the deletion of both the data and resulting algorithms.

- **DARK PATTERNS AND MANIPULATIVE DESIGN:** AI-powered interfaces that deceive, manipulate, or exploit users (e.g., via misleading prompts or choices) may be considered unfair practices.

Further, notable enforcement actions include:

- **OPENAI INVESTIGATION:** Focused on potential consumer harm from false or misleading outputs generated by ChatGPT and related models.

- **RING (2023):** $5.6 million settlement for privacy violations, including unauthorized use of customer video footage to train algorithms; required deletion of affected data and models.

- **AMAZON (2023):** $25 million fine for violating COPPA by retaining and using children's voice data collected through Alexa to train algorithms.

CONSUMER FINANCIAL PROTECTION BUREAU (CFPB)

The CFPB regulates consumer financial products and services, with a growing focus on algorithmic decision-making in lending and credit. Key risks include:

- **ADVERSE ACTION NOTICE FAILURES:** Even when decisions are made using complex algorithms, creditors must provide clear and specific reasons for adverse decisions under the Equal Credit Opportunity Act (ECOA).

- **DISCRIMINATORY LENDING MODELS:** The CFPB has warned that using AI does not insulate companies from liability under fair lending laws, particularly when inputs or proxies result in disparate impacts on protected classes.

DEPARTMENT OF JUSTICE (DOJ)

The DOJ enforces civil rights laws across sectors such as housing, education, and employment, and has increasingly scrutinized the use of AI tools that may result in discriminatory outcomes.

- **FAIR HOUSING ACT GUIDANCE:** The DOJ has cautioned that algorithm-based tenant screening tools must not result in discriminatory exclusions based on race, disability, or other protected characteristics.

- **CORPORATE COMPLIANCE EXPECTATIONS:** In its guidance on evaluating corporate compliance programs, the DOJ highlights the importance of managing AI risks—including monitoring outputs, limiting uses to intended purposes, and integrating AI governance into enterprise risk management.

Employment and Anti-Discrimination Laws

The use of AI in employment decisions, such as hiring, promotion, termination, or performance evaluation, raises significant legal risks under anti-discrimination laws. Regulators at the federal, state, and local levels have prioritized enforcement in this area, particularly as employers increasingly rely on Automated Employment Decision Tools (AEDTs) to assess candidates or employees.

EQUAL EMPLOYMENT OPPORTUNITY COMMISSION (EEOC)

The EEOC enforces federal anti-discrimination laws, including:

- **Title VII of the Civil Rights Act**, prohibiting discrimination based on race, color, religion, sex, or national origin;
- **The Americans with Disabilities Act (ADA)**;
- **The Age Discrimination in Employment Act (ADEA)**; and
- **Genetic Information Nondiscrimination Act (GINA)**.

Key EEOC guidance and actions include:

- **ADA TECHNICAL ASSISTANCE (2022):** Clarifies that employers using AI or algorithmic tools to evaluate applicants must ensure reasonable accommodations for individuals with disabilities and avoid practices that disproportionately screen them out.
- **ITUTOR GROUP CASE:** Online education company paid $365,000 to settle EEOC allegations that its AI-driven hiring system automatically rejected female applicants over 55 and all applicants over 60, violating the ADEA and Title VII.

The EEOC has emphasized that employers cannot outsource liability to third-party vendors: if an AI tool results in unlawful discrimination, the employer may still be held responsible.

STATE AND LOCAL EMPLOYMENT AI LAWS

A growing number of jurisdictions have enacted laws regulating AI in employment contexts, with a focus on transparency, fairness, and bias mitigation.

- New York City Local Law 144 (AEDT Law):
  - Requires employers using AEDTs for hiring or promotion decisions to conduct an independent **bias audit**.
  - Requires **notice and disclosure** to job candidates about the use of such tools.

  - Defines AEDTs broadly to include any algorithmic or statistical model that "substantially assists" in decision-making.
  - Enforcement began on **July 5, 2023**.
- Illinois HB 3773 (Effective Jan. 1, 2025):
  - Requires employers to notify employees when AI tools are used to make employment-related decisions (e.g., hiring, promotion, training, discipline, termination).
  - Builds upon the **Illinois AI Video Interview Act**, which mandates transparency in AI-assisted interview evaluations.
- Other State Developments:
  - Several states (e.g., California, Maryland, New Jersey) have introduced legislation addressing AI-driven hiring practices, some of which would require audits, disclosures, or consent.

KEY LEGAL RISKS IN EMPLOYMENT CONTEXTS

- **DISPARATE IMPACT LIABILITY:** AI tools may unintentionally screen out protected classes, triggering liability even absent discriminatory intent.
- **LACK OF TRANSPARENCY:** Failure to disclose how AI is used in employment decisions may violate state or local law.
- **VENDOR OVERSIGHT GAPS:** Employers must ensure that third-party vendors using AI in recruiting or HR tools are compliant with applicable legal requirements.

For more information on forthcoming AI-specific privacy regulations and requirements, please refer to Gunderson Dettmer's "AI Regulatory Update" publication series.

Data Privacy and Data Protection

AI systems that ingest, generate, or process personal data are subject to a complex and evolving body of privacy laws. These laws vary by jurisdiction but generally impose strict requirements around transparency, consent, data minimization, security, and data subject rights. Regulatory focus is increasing on how personal and sensitive data is used in training, fine-tuning, and real-time inference—particularly where outputs may expose that data or where the collection lacked a lawful basis. AI providers and users may face liability under major frameworks such as the General Data Protection Regulation (GDPR) (EU/UK), California Consumer Privacy Act (CCPA) and California Privacy Rights Act (CPRA) (U.S.), the Health Insurance Portability and Accountability Act (HIPAA) (U.S.), and numerous other state, federal, and international laws.

KEY PRIVACY RISKS IN THE AI CONTEXT

There are several key risks, including:

- **INCLUSION OF PERSONAL DATA IN TRAINING OR OUTPUTS:** Personal or sensitive data may be embedded in training datasets without proper authorization or consent. If training data includes information about identifiable individuals, this may constitute unlawful processing under laws like the GDPR, CCPA, or HIPAA—especially where the data was scraped from public sources or acquired without clear notice or user control.

- **MODEL MEMORIZATION AND DATA LEAKAGE:** Regulatory authorities are increasingly concerned with AI models' tendency to memorize training data and later reproduce it in outputs. This can lead to inadvertent disclosure of personal or sensitive information—especially when models are prompted to recall specific names, facts, or phrases. This risk is particularly acute for large language models (LLMs) trained on vast corpora with minimal human curation.

- **TRANSPARENCY AND CONSENT REQUIREMENTS:** Most privacy laws require companies to inform individuals about the use of their data, including for AI model training, fine-tuning, or decision-making. Where AI is used in ways that materially affect individuals (e.g., employment, credit, or healthcare), additional disclosures—and often affirmative opt-in consent—may be required. Failure to clearly communicate that AI is being used or that personal data is involved may be deemed deceptive or unlawful.

CORE PRIVACY OBLIGATIONS

Across jurisdictions, privacy frameworks typically impose the following key obligations:

- **TRANSPARENCY AND NOTICE:** Businesses must publish clear privacy notices describing their data collection, use, sharing, and retention practices, especially where AI is involved. This includes identifying if data is used to train models, generate content, or make automated decisions.

- **LEGAL BASIS FOR PROCESSING:** Companies must establish a lawful basis for processing PII. Under the GDPR, this includes:

  - Consent (particularly for sensitive or high-risk uses),
  - Contract performance,
  - Legal obligation,
  - Vital interests,
  - Public interest, or
  - Legitimate interests (balanced against individual rights).

- **DATA SUBJECT RIGHTS:** Individuals are typically granted rights to:
  - Access, correct, and delete their data;
  - Opt out of profiling or automated decision-making;
  - Object to certain types of processing;
  - Port their data; and
  - Know whether AI is involved in decisions about them.
- **SECURITY AND GOVERNANCE:** Businesses must implement safeguards to protect personal data throughout the AI lifecycle. This includes:
  - Data encryption,
  - Access controls,
  - Breach detection and notification procedures,
  - Governance protocols for model training and inference,
  - Data retention and deletion mechanisms,
  - Periodic audits of model behavior and data leakage risks.

SPECIAL CATEGORIES OF DATA

AI systems often interact with sensitive personal data, which carries heightened regulatory obligations and enforcement risk. Examples include:

- **BIOMETRIC DATA:** Regulated by laws like Illinois BIPA, Texas's biometric statute, and others requiring prior consent, strict use limitations, and detailed retention policies.
- **HEALTH DATA:** Governed by HIPAA for covered entities and state laws like Washington's My Health My Data Act. Applies even to consumer-facing apps if health-related inferences are made.
- **CHILDREN'S DATA:** Covered under COPPA and state-level laws, often requiring verifiable parental consent and restricting behavioral profiling and targeted advertising.
- **FINANCIAL DATA:** Regulated by the Gramm-Leach-Bliley Act (GLBA), Fair Credit Reporting Act (FCRA), and related state statutes, which require opt-in rights, notice, and security measures.
- **EMPLOYMENT DATA:** Subject to state privacy laws (e.g., CPRA), federal regulations (e.g., ADA, FCRA), and agency oversight (e.g., EEOC). Includes data from employee monitoring tools, productivity tracking, and algorithmic evaluation systems.

NOTABLE ENFORCEMENT ACTIONS

Regulators are increasingly enforcing privacy laws in the context of AI systems:

- **GOODRX & BETTERHELP (FTC):** Fined $1.5M and $7.8M, respectively, for sharing sensitive health data with advertisers in violation of their public-facing privacy policies.
- **CLEARVIEW AI:** Settled lawsuits under BIPA for scraping and processing billions of facial images without consent. Required to delete collected data and cease sales to private entities.
- **TILTING POINT MEDIA:** Fined $500,000 for unlawfully collecting and sharing children's data without parental consent in violation of COPPA and California's privacy laws.

AI-Specific Regulations

In addition to applying existing laws to AI systems, regulators around the world are enacting AI-specific legislation that directly governs the development, deployment, and use of AI technologies. These new laws reflect growing concerns about algorithmic discrimination, misinformation, lack of transparency, and systemic risk. The regulatory landscape is rapidly evolving, and companies must proactively track developments to ensure compliance across jurisdictions.

EUROPEAN UNION: E.U. AI ACT

The E.U. AI Act is the world's first comprehensive AI-specific regulatory framework. It imposes detailed obligations on providers and users of AI systems based on a risk-based classification, with requirements scaling according to the risk posed to health, safety, and fundamental rights. Key features include:

- **RISK CATEGORIES:**
  - **Unacceptable Risk**: Banned systems (e.g., social scoring, real-time biometric surveillance).
  - **High-Risk**: Includes AI used in employment, credit scoring, biometric ID, education, and critical infrastructure.
  - **Limited Risk**: Transparency requirements apply (e.g., chatbots, emotion recognition).
  - **Minimal Risk**: No obligations (e.g., AI in spam filters or video games).

- **CORE OBLIGATIONS FOR HIGH-RISK SYSTEMS:**
  - Risk management and impact assessments
  - Robust data governance and record-keeping
  - Human oversight and transparency
  - Accuracy, robustness, and cybersecurity measures
  - Conformity assessments and CE marking
- **GENERAL-PURPOSE AI (GPAI):** Providers of large-scale foundation models (e.g., LLMs) must comply with additional obligations around documentation, risk mitigation, transparency, and model testing—especially if the model poses **systemic risk**.
- **PENALTIES:** Up to **€35 million or 7%** of global annual turnover for serious violations, underscoring the need for proactive compliance.
- **EXTRATERRITORIAL REACH:** The Act applies to any company introducing AI systems into the E.U. market or using outputs that affect E.U. citizens, even if located outside the E.U

For more information on compliance obligations under the E.U. AI Act, please refer to Gunderson Dettmer's "Demystifying the E.U. AI Act" publication series.

UNITED STATES: FEDERAL AND EXECUTIVE ACTION

Unlike the E.U., the U.S. does not yet have a comprehensive federal AI law. Instead, AI regulation is emerging through a **patchwork of agency guidance, executive action, and state-level legislation**.

- **EXECUTIVE ORDER 14179 (JAN. 2025) — *"Removing Barriers to American Leadership in Artificial Intelligence"***
  - Rescinds previous AI risk-mitigation policies issued under President Biden, including Executive Order 14110 (Oct. 2023) which was focused on ***"Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence."***
  - Promotes innovation over regulation, calling for an **"AI Action Plan"** (published July 2025) to enhance U.S. leadership and reduce regulatory burdens.
  - Seeks stakeholder input to guide development of a light-touch national AI strategy.
  - The final plan is expected in **July 2025**.

- **AI ACTION PLAN (JULY 2025) — *"Advancing American Innovation and Leadership in Artificial Intelligence"***
  - Replaces and reduces federal oversight and compliance requirements for AI companies, aiming to accelerate commercial and governmental AI adoption, and prioritizing support for AI research, development, and deployment.
  - Emphasizes voluntary standards, industry self-regulation, and public-private partnerships over prescriptive mandates.
  - Sets objectives to grow the domestic AI workforce, modernize federal procurement of AI, and boost investments in foundational AI infrastructure.
  - Commits to annual reviews and stakeholder consultations to adapt the plan as AI technology evolves.
- **FEDERAL AGENCY ACTIVITY (PRE-2025 RESCISSION):**
  - FTC, EEOC, CFPB, and HHS issued guidance on AI-related risks under existing laws (many of which remain relevant).
  - Pending legislation (e.g., the **Algorithmic Accountability Act**) has not yet passed but could resurface in future sessions.

UNITED STATES: STATE-LEVEL AI LEGISLATION

In the absence of a federal framework, multiple states have moved forward with **AI-specific laws**, creating a complex and fragmented regulatory environment.

- **CALIFORNIA:** California passed several AI-specific laws, including AB 2013 and the California AI Transparency Act (SB 942), which require developers to make public disclosures about AI inputs and outputs.
  - **AB 2013** requires companies to disclose the source, number of data points, description of data types (including whether such data contains PII, is licensed, or is in the public domain), date of use, and purpose of processing for underlying training datasets. There are limited exemptions for AI systems developed for security and integrity purposes (e.g., detection of security incidents and illegal content), or otherwise developed for national security, military, or defense purposes and made available only to federal entities.
  - The **California AI Transparency Act** also requires "covered providers" (i.e. any entity that "creates, codes, or otherwise produces a generative artificial intelligence system that has over [1 million] monthly visitors or users and is publicly accessible" within California) to provide disclosures about the use of AI systems, and provide AI detection tools that allow users to assess whether content has been created or altered using a AI system. Both laws go into effect on January 1, 2026.

- **COLORADO:** Colorado passed the Colorado AI Act (SB 24-205) earlier this year, which adopts the same definition of "AI system" as the E.U. AI Act and applies to developers and deployers (i.e. AI system users) that do business in Colorado. SB 24-205 imposes obligations on "high-risk" processing activities that make consequential decisions related to education, employment, financial or lending services, essential government services, healthcare, housing, or legal services. These obligations range from transparency disclosures about AI system inputs and outputs, adopting risk management policies, conducting impact assessments, and imposing guardrails against algorithmic discrimination. The law goes into effect on February 1, 2026.

- **UTAH:** Utah enacted the Utah AI Policy Act ("UAIPA"), which became effective on May 1, 2024. UAIPA imposes disclosure requirements on businesses using AI tools with their customers in "regulated occupations" (i.e. occupations that require a license or state certification to practice, such as accountants, architects, and healthcare professionals) or that engage in sales, telemarketing, and other consumer solicitation activities involving Utah-based consumers.

- **OTHER STATES TO WATCH:** New York, Maryland, and Massachusetts are actively considering AI-specific laws focused on consumer protection, algorithmic discrimination, and transparency.

Tort and Criminal Liability

The use of AI systems creates significant new risks of tort and criminal liability. Unlike traditional tools, AI can act autonomously, make unpredictable decisions, replicate biases at scale, and expose companies to cybersecurity threats.

These risks can result in serious legal consequences, including claims for negligence, product liability, discrimination, fraud, and even criminal charges. As regulators and courts move quickly to address the challenges AI presents, companies must navigate an evolving legal landscape where liability may attach not just for how AI is designed and deployed, but also for how it is monitored and controlled. This section examines how core tort and criminal law principles apply to AI and where new forms of exposure are emerging.

Tort Liability

Under U.S. and common law principles, tort liability may arise when a company's use of AI causes foreseeable harm due to negligence, design defects, or failure to warn.

NEGLIGENCE

Companies may be held liable for failing to act with reasonable care in the design, deployment, or oversight of AI systems. Common fact patterns include:

- Failing to test or validate models before use;
- Relying on flawed or biased data;
- Ignoring known risks associated with specific use cases;
- Failing to implement sufficient safeguards or human-in-the-loop controls.

Negligence claims are most likely to arise where AI is used in **safety-critical applications**, such as autonomous vehicles, healthcare diagnostics, industrial automation, or public infrastructure.

PRODUCT LIABILITY

When AI systems or outputs are embedded in physical products or software-as-a-product offerings, companies may face **strict liability** for harms caused by:

- **DESIGN DEFECTS** (e.g., flawed training logic or unsafe outputs),
- **MANUFACTURING DEFECTS** (e.g., corrupted model parameters), or
- **FAILURE TO WARN** (e.g., inadequate instructions, disclaimers, or user safeguards).

These claims do not require proof of negligence, meaning that even careful providers can be held liable if the product causes injury.

DEFAMATION, FALSE LIGHT, AND EMOTIONAL DISTRESS

Generative AI systems can produce text, images, or audio that defames individuals, places them in a false light, or causes reputational or emotional harm. Examples include:

- Generating false criminal accusations or offensive impersonations;
- Fabricating images or videos that imply misconduct;
- Publishing harmful content about real people without fact-checking or editorial review.

In a first-of-its-kind defamation suit, ***Mark Walters v. OpenAI***, radio host Mark Walters sued OpenAI alleging that ChatGPT generated defamatory and libelous output that suggested Mr. Walters was involved in an embezzlement scheme. Similarly, in ***Battle Enterprise v. Microsoft,*** Jeffrey Battle filed a lawsuit against Microsoft alleging that Microsoft's AI-powered search engine, Bing, produced a libelous AI-generated search result that conflated the plaintiff with an individual with a similar name that was previously convicted for seditious conspiracy.

LIABILITY FOR USER-GENERATED CONTENT

The rise in AI is also challenging traditional application of federal safe harbors for online service providers that may host infringing and illegal user-generated content, including intermediary immunity under Section 230 of the Communications Decency Act ("CDA 230").

- **CDA 230:** Initially enacted in 1996, CDA 230 has long provided immunity to online platforms for content posted by third parties, shielding such platforms from being treated as "publishers" of such third-party content. CDA 230 has allowed the internet to thrive as a space for user-generated content, enabling the growth of social media platforms, online marketplaces, forums, and countless other services that rely on user participation.

The application of CDA 230 becomes increasingly complex when providers deploy AI systems that generate or moderate content autonomously using AI technology. Unlike traditional platforms that passively host third-party content, AI systems actively produce outputs in response to user prompts, often blurring the lines between hosting and creating content. This raises critical questions about whether AI-generated output should be treated as third-party content under CDA 230, which is potentially protected by the statute, or as original content attributable to the AI provider, which would not be covered by the statute's grant of immunity. Courts and regulators are beginning to scrutinize these issues, especially in cases involving defamation, misinformation, or bias in AI-generated content. While no decisions on the merits have been issued as of the date of this article, previous decisions on the scope of CDA 230 immunity may provide insight as to how a court may rule on this issue in the AI context. For example, in ***Fair Housing Council of San Fernando Valley v. Roommates.com***, the court found that providers lose CDA 230 immunity when they are "responsible, in whole or in part, for the creation or development" of unlawful content. In this case, Roommates.com was denied immunity for content it helped develop through mandatory questionnaires, but retained immunity for user-generated "additional comments". Similarly, if an AI system materially contributes to the creation of illegal or harmful content, courts may decline to grant CDA 230 to the operator of that system.

Recent years have seen intensified calls for reform or repeal of CDA 230, particularly under the Trump administration. Key officials at the DOJ, FTC, and FCC have advocated for a narrower interpretation of the statute, arguing that it should not shield providers from liability arising out of their content moderation decisions or for removing user content. This is a significant departure from longstanding judicial interpretations that broadly protected both hosting and moderation activities. Legislative proposals have also emerged, including bipartisan bills that would sunset CDA 230 unless Congress enacts new liability standards. These efforts reflect growing bipartisan concern that the current liability shield may be ill-suited for the age of generative AI and the evolving role of online platforms in shaping public discourse.

Criminal Liability

While criminal prosecution of AI developers or deployers remains rare, several emerging scenarios may give rise to criminal enforcement—particularly where harm results from reckless disregard, intentional misuse, or failure to mitigate foreseeable risks.

FRAUD, IMPERSONATION, AND DEEPFAKES

AI tools used to impersonate individuals—such as voice cloning, fake customer service bots, or political disinformation campaigns—can expose perpetrators to criminal liability under statutes governing fraud, wiretapping, identity theft, and impersonation. Advances in AI now enable bad actors to bypass traditional account verification and authentication methods with alarming sophistication, increasing the vulnerability of organizations to cyberattacks and social engineering scams. For instance, in one high-profile case, a finance employee at a Hong Kong-based company was deceived into transferring $25.6 million to cybercriminals who used AI-powered deepfake technology to convincingly impersonate the company's Chief Financial Officer and other executives during multiple video calls.

States have enacted new legislation criminalizing AI tools used for digital fraud and other such harms caused by AI. Such cybercrimes include:

- Online impersonation created with the intent to intimidate, bully, threaten, or harass a person through electronic or online communications (e.g., social media sites).
- Use of another person's likeness to create a digital replica or commit digital identity theft.
- Malicious use or commercial dissemination of manipulated audiovisual content that falsely depicts others without consent.
- Nonconsensual "intimate image" production, such as the creation of sexually explicit or pornographic content of another person without consent.
- Deceptive media or falsified electioneering communications created with the intent to injure a candidate's reputation (e.g., election disinformation).
- Granting individuals a property right (and right to assert criminal action for trespass) to their name, image, voice, and likeness.

Further, even if a third party misuses the tool, companies may face liability if they:

- Facilitated the misuse knowingly,
- Failed to enforce terms of service,
- Ignored obvious signs of abuse.

Companies can protect against the risk of deepfakes and other cybercrimes by obtaining adequate cyber insurance and developing internal procedures and policies to educate employees about the emergent AI-powered cybercrimes.

SURVEILLANCE AND UNLAWFUL DATA COLLECTION

Using AI tools for mass surveillance, facial recognition, or audio analysis may implicate:

- Wiretap laws,
- Eavesdropping statutes,
- Biometric privacy laws (e.g., BIPA),
- Computer trespass laws.

In some jurisdictions, unlawful access to user data—even without malicious intent — can lead to criminal penalties.

RECKLESS ENDANGERMENT OR NEGLIGENT HOMICIDE

In extreme scenarios, such as fatal failures of autonomous systems (e.g., self-driving vehicles or medical AI tools), prosecutors may pursue charges where companies:

- Had notice of defects or failures,
- Ignored expert warnings or safety concerns,
- Deployed the technology without proper testing or safeguards.

While still novel, these theories are being explored in multiple jurisdictions.

Types and Cost of AI Insurance Coverage

Are AI Risks New and How Does That Impact What Insurance I Need?

The vast majority of risks that AI-driven companies face are not new, and therefore don't require new types of insurance, but the use of AI tends to increase a company's exposure to their pre-existing risks (eg more data), while also moving them into new territory that they've not worried about before (e.g., IP breaches in LLM training data).

For example, two of the most likely risks that the average AI company has is:

- negligence in performing their obligations to customers; and
- exposure to a cyber security breach.

They're covered by E&O and Cyber insurance policies, respectively, and the risks are not new. However, the fact patterns that would trigger the use of those policies ***are*** new. In this new era, negligence might be driven by an AI hallucination, while a cyber security breach could be driven by deep fake fraud.

As a result, companies need to:

- work with carriers who are keeping pace with emerging AI risks;
- spend time mapping out their key risk exposures and pressure test how their policies would respond, with the help of their broker and insurance carriers. Their broker should be a great resource as they'll have access to claims data and anecdotes.
- negotiate with carriers to define policy terms more clearly where possible (this is more possible for Series B+ companies). Broadly defined terms can be positive in an emerging legal area but they can also cause confusion and lead to drawn out disputes in the event of a claim.

What Insurance do AI-Driven Companies Need and What is the Cost?

The costs and types of AI-specific insurance are shaped by the coverage scope, industry dynamics, and the unique exposures of each company. Here's a breakdown of the primary coverages available to AI companies, their focus areas, and why they're essential:

ERRORS AND OMISSIONS (E&O) INSURANCE

E&O insurance protects against financial losses caused by errors, omissions, or negligence in delivering AI-based services or products. In short, it covers a company from financial loss caused to its customer due to its service or product not performing that way it should have. In many instances, the damage that a company can cause to a customer far exceeds the value of the contract. This coverage is crucial for businesses whose AI solutions are integrated into critical business operations or are operating in highly regulated industries, as even minor failures can lead to significant client losses and legal disputes. By covering legal fees, settlements, and judgments, E&O insurance enables companies to focus on innovation and growth without being derailed by costly claims. Lawyers will carefully draft indemnity language and place liability caps in the contract, but those caps do not apply to acts of negligence (one cannot contract out of negligence), which is one of the most shocking realities for non-lawyers to discover.

- **WHY IT'S IMPORTANT:** Failure to meet contractual obligations or performance standards could result in lawsuits for breach of contract or negligence.
- **REAL-WORLD EXAMPLE:** An AI-powered logistics platform provides inaccurate delivery schedules, leading to missed deadlines and financial losses for a retail client.
- **COST:** Premiums vary annually for companies depending on factors like annual revenue, the criticality of the AI service, and prior claims history. Industries like healthcare and fintech, where operational risks are higher, often see increased premiums.

BIAS AND DISCRIMINATION COVERAGE

Bias and discrimination coverage covers claims stemming from biased AI decisions that result in discriminatory outcomes, whether intentional or not. Companies often face challenges in auditing and mitigating bias within AI systems, particularly when algorithms are trained on incomplete or unrepresentative datasets. This coverage protects against financial losses and reputational damage arising from lawsuits, ensuring companies can address these risks while maintaining client and stakeholder trust. A good rule of thumb is that Companies should consider bias and discrimination if their AI solution is used for **identification or decision making** in some way (e.g., lending decisions (credit cards, mortgages), healthcare decisions (recommended treatment plans), public safety (analysis of footage) and recruitment decisions (analyzing job applications and making recommendations)). AI models are probabilistic,

meaning that outputs will not be the same every time and accuracy is assessed in terms of likelihood, so it's very hard for companies to get comfortable that their models will not provide biased and discriminatory outputs.

- **WHY IT'S IMPORTANT:** Violations of anti-discrimination laws may lead to costly legal actions, including class-action lawsuits and penalties.
- **REAL-WORLD EXAMPLE:** A mortgage company's AI model is accused of systematically offering higher interest rates to minority applicants, triggering regulatory scrutiny and legal action.
- **COST:** Costs vary annually and are influenced by the complexity of the AI model, regulatory environment, and historical complaints. Companies deploying AI in sensitive areas like hiring or credit assessments may face higher premiums.

INTELLECTUAL PROPERTY (IP) CLAIMS COVERAGE

IP claims coverage safeguards against claims of copyright, trademark, or patent infringement related to AI technologies. Companies often face heightened exposure to IP disputes due to the competitive nature of AI development and the reliance on proprietary datasets and algorithms. This coverage helps mitigate the financial and operational risks of lawsuits, enabling companies to defend their innovations while continuing to scale. IP disputes are particularly prevalent in industries like healthcare, fintech, and autonomous vehicles, where cutting-edge technology and high stakes drive intense competition.

- **WHY IT'S IMPORTANT:** Accusations of IP theft can lead to injunctions halting product distribution or monetary damages that exceed initial investment costs.
- **REAL-WORLD EXAMPLE:** An ed-tech company faces a lawsuit after its AI-powered learning tool is alleged to have used copyrighted training datasets without authorization.
- **COST:** Premiums vary annually depending on the technology's proprietary nature, the industry, and the likelihood of infringement claims. AI companies working in competitive sectors, like entertainment or software development, often pay more.

REGULATORY INVESTIGATION COVERAGE

Regulatory investigation coverage provides financial and legal support during investigations by regulatory bodies over non-compliance with data protection or AI transparency laws. Companies often face unique challenges in keeping pace with evolving regulations while scaling their operations, making this coverage critical. It also covers costs like legal counsel and audit readiness, helping companies address inquiries under frameworks such as GDPR, CCPA, or AI-focused legislation without disrupting growth efforts.

- **WHY IT'S IMPORTANT:** Penalties for violations, such as under GDPR or CCPA, can include substantial fines and damage to brand reputation.

- **REAL-WORLD EXAMPLE:** A marketing firm using AI for personalized campaigns is investigated for non-compliance with GDPR's data privacy requirements.

- **COST:** Coverage costs vary annually depending on factors like the volume of data processed, geographical regions of operation, and regulatory environment. Industries dealing with high volumes of consumer data, such as e-commerce and advertising, often pay more.

CYBER INSURANCE COVERAGE

Cyber insurance coverage covers financial losses from cyberattacks, data breaches, or fraud enabled through AI technology. While cyber insurance covers risks like data breaches, ransomware, and social engineering attacks, AI-specific coverage focuses on liabilities unique to AI systems, such as algorithmic errors, AI-enabled fraud (e.g., invoice manipulation), and compliance issues related to AI decision-making. These policies complement each other, addressing both traditional cyber risks and emerging vulnerabilities tied to AI technologies.

- **WHY IT'S IMPORTANT:** Liability for AI-related breaches and fraud can result in lawsuits from clients, as well as costs for system recovery and customer notifications.

- **REAL-WORLD EXAMPLE:** A smart manufacturing system is hacked, causing production delays and exposing sensitive supplier contracts, leading to legal disputes.

- **COST:** Premiums vary annually depending on company size, data sensitivity, and cybersecurity measures. Companies in industries like finance or healthcare, which handle sensitive data, face higher premiums.

Navigating AI Insurance

Securing the right AI insurance is critical to protecting your company from unforeseen risks while enabling sustainable growth.

The process of procurement, claims, and underwriting may seem complex, but with the right approach and questions, it can be streamlined and effective. Navigating these processes requires thoughtful planning and the right partnerships. By proactively engaging with insurers and brokers experienced in your specific technology and business needs, companies can secure tailored coverage, streamline claims, and mitigate risks effectively.

Procurement: Asking the Right Questions

Navigating the procurement process begins with understanding your company's specific risks and selecting a broker or insurer who truly understands AI.

▸ ASSESSING YOUR COMPANY'S INSURANCE NEEDS

- **Identifying Key Risk Areas**: Start by evaluating where your AI systems could pose financial, operational, or reputational risks. For example, consider potential liabilities related to biased decision-making, regulatory compliance, or intellectual property disputes.

- **Determining Appropriate Coverage Levels**: Work with your broker to calculate the coverage limits that align with your risk exposure, factoring in company size, industry standards, and contractual obligations.

▸ SELECTING THE RIGHT INSURANCE PROVIDER AND BROKER

- **Evaluating Expertise in AI Risks**: Ask brokers or insurers about their experience with AI-related businesses. How do they address specific risks like AI-driven errors, algorithmic failures, or regulatory non-compliance?

- **Comparing Policy Options and Premiums**: Ensure that policy language is clear and comprehensive, avoiding excessive exclusions. Compare premiums across insurers while factoring in coverage quality and claims support.

Key Questions to Ask Your Broker

- What industries and AI technologies does your insurer specialize in?
- How do these policies address emerging risks, such as AI-enabled fraud or compliance with new regulations?
- Are there bundled policies or cost-saving measures that fit my company's stage and needs?

Claims and Underwriting Process

Understanding how insurance claims and underwriting work is vital for maintaining a robust risk management strategy. Companies leveraging AI must provide detailed insights into their risk management practices, including robust cybersecurity protocols, model validation processes, and compliance with data protection laws. Documentation demonstrating these safeguards often helps in securing more favorable terms.

Strategies for Mitigating Similar Risks

Mitigating risks associated with AI technologies requires a proactive approach that combines technical safeguards, operational policies, and strategic planning. Here are five key strategies:

CONDUCT REGULAR ALGORITHM AUDITS

- **What to Do:** Periodically evaluate your AI models to identify biases, errors, and vulnerabilities. Use third-party audits for an impartial perspective and ensure model decisions are explainable.
- **Why It Matters:** Transparent algorithms reduce the likelihood of biased outcomes or unintentional errors, minimizing reputational and legal risks.

IMPLEMENT STRONG DATA GOVERNANCE POLICIES

- **What to Do:** Establish clear protocols for data collection, storage, and usage, ensuring datasets are representative and compliant with privacy laws. Monitor for data drift that can impact model performance.
- **Why It Matters:** Proper data governance prevents regulatory violations and improves the reliability of your AI systems, safeguarding against data misuse claims.

- **INVEST IN ROBUST CYBERSECURITY MEASURES**
  - **What to Do:** Deploy advanced security measures like multi-factor authentication, encryption, and continuous monitoring. Conduct regular penetration testing to identify vulnerabilities.
  - **Why It Matters**: Protecting your systems from breaches reduces exposure to AI-enabled fraud, data theft, and costly downtime.
- **ESTABLISH A CLAIMS RESPONSE PLAN**
  - **What to Do:** Create an internal protocol for addressing incidents, including clear roles for notifying insurers, gathering evidence, and managing communication with affected stakeholders.
  - **Why It Matters:** A well-defined response plan accelerates claims processing and minimizes the impact of incidents on your operations and reputation.
- **STAY AHEAD OF REGULATORY CHANGES**
  - **What to Do:** Monitor updates to AI-specific regulations (e.g., EU AI Act, GDPR) and industry standards. Engage legal counsel or compliance experts to stay aligned with evolving requirements.
  - **Why It Matters:** Proactive compliance reduces the risk of fines, investigations, and reputational harm, particularly as AI regulations continue to develop globally.

Common Pitfalls and How to Avoid Them

Navigating the claims and underwriting process for AI insurance can be challenging, especially for companies unfamiliar with its complexities. Below are a few common missteps that companies should proactively address to secure comprehensive coverage while avoiding costly surprises.

INADEQUATE OR MISALIGNED COVERAGE

- **PITFALL:** Startups often underestimate their exposure to AI-specific risks, opting for standard cyber insurance or general liability policies that don't fully address the unique liabilities of AI systems.
- **HOW TO AVOID IT:** Conduct a detailed risk assessment that considers algorithmic failures, intellectual property disputes, and regulatory requirements. Work with a broker experienced in AI-related risks to ensure your coverage matches your company's specific needs.

AMBIGUOUS OR OVERLY RESTRICTIVE POLICY TERMS

- **PITFALL:** Policies with unclear definitions or excessive exclusions can leave critical gaps in coverage, particularly around algorithmic errors or failures to meet evolving AI regulations.

- **HOW TO AVOID IT:** Scrutinize policy language, particularly exclusions related to AI-specific risks like model errors, data misuse, or outdated software patches. Engage your broker to clarify terms and advocate for adjustments if necessary.

FAILURE TO COMMUNICATE RISK MANAGEMENT PRACTICES

- **PITFALL:** Insurers rely heavily on underwriting criteria, and insufficient documentation of risk management protocols can result in higher premiums or even coverage denials.

- **HOW TO AVOID IT:** Provide detailed documentation of your company's risk mitigation practices, such as regular audits of AI models, cybersecurity measures, and compliance with data protection laws. Demonstrating strong internal controls often leads to more favorable underwriting outcomes.

Best Practices for Companies Using or Offering AI Products and Services

As AI technologies become integral to business operations, companies must adopt comprehensive best practices to mitigate emergent risks and ensure compliance.

This section identifies best practices and provides guidance for companies using or offering AI products and services, including:

- Implementing clear policies and procedures for AI usage and development is essential to uphold ethical standards and regulatory requirements.
- Monitoring and tracking the use of training data to address compliance concerns (e.g., bias, accuracy, copyright, privacy, etc.).
- Implementing robust technical guardrails that reduce the likelihood of AI misuse or errors.
- Conducting thorough vendor diligence ensures third-party AI tools meet legal and operational standards.
- Securing sufficient IP rights (in, e.g., data licensing, end user, customer, vendor, partner and other such commercial agreements) to enable the company to collect, process, and otherwise use data necessary for the company's business.
- Management of potential liability through appropriate risk allocation mechanisms, including contractual indemnities and limitations of liability provisions.
- Obtaining AI-specific insurance offers financial protection against unforeseen risks.

As companies use, develop and deploy AI technologies, they must navigate an increasingly complex regulatory and commercial landscape. It is critical to develop compliance mechanisms with guidance of legal counsel to meet obligations under evolving applicable laws, guidance, and market standards.

Adopt and Implement Internal Policies

In addition to regulatory scrutiny, potential investors and acquirers expect target companies to adopt, implement, and maintain commercially reasonable internal policies and procedures to enable AI compliance. External use, publication, or distribution of AI outputs opens the door to potential exposure to infringement claims, so companies should implement a system for internal review and escalation, especially for higher-risk use cases (e.g., product development, marketing, etc.).

Companies should adopt an AI Usage Policy that outlines employee requirements and restrictions on prompting and using AI outputs, including when to seek approval for higher-risk use cases. Companies will need to tailor their AI Usage Policy to address their unique business risks, which can range from a strict prohibition of using AI for any business purpose to a more permissive Policy that sets guidelines and parameters for prompts and outputs for approved uses. Companies must also provide regular employee training and notify employees of any changes to the AI Usage Policy to ensure that all company personnel (e.g., employees, consultants, and other such service providers) understand the guidelines and responsibilities associated with AI use.

In many ways, these AI internal compliance policies may mirror the procedures companies use to track licenses and monitor usage of open-source software in back- and front-end operations—such as including robust monitoring programs (e.g., regular code scans and audits), requiring human review and oversight, and tracking internal organization-wide usage of AI outputs. At minimum, a AI Usage Policy should include the **prohibitions on the use of AI tools that are made available to employees on company-authorized accounts or otherwise accessed by employees on non-company accounts**, such as:

- Do not use AI tools to conduct illegal activities (e.g., fraud, phishing, etc.); create illegal or unethical content; or manipulate or deceive another person.
- Do not use AI tools to invade the privacy of individuals; violate data protection and privacy laws; impersonate another person; or generate misrepresentations or falsehoods regarding another person.
- Do not use AI tools that infringe upon third-party IP rights.

- Do not use AI tools to disrupt, harm, or gain unauthorized access to systems or networks of the company or any third party.

- Do not use AI tools to create discriminatory content; or make decisions that have unfair or adverse impacts on people.

- Do not use AI tools to create content that could harm the reputation or interests of the company or its stakeholders.

Unless the employee is using an enterprise account managed by the company, do not ingest company confidential or proprietary information (e.g., customer or vendor lists, source code, product development details, presentations, user information, consumer personal information, etc.) as a prompt in any AI tools.

Furthermore, in addition to adopting and implementing a AI Usage Policy, companies must also maintain strict internal policies for data usage, specifying permissible purposes for processing, retention timelines, deletion protocols, and business continuity/recovery procedures. The company's internal data usage policies should be reviewed by legal counsel to ensure it adheres to applicable privacy laws and any AI- or industry-specific regulations, and the company should regularly audit and review its compliance with such policies.

Monitoring and Tracking Use of Data

Effective monitoring and tracking of data are critical for companies using, developing, or offering AI products and services. A robust data governance framework ensures compliance while minimizing risks related to data misuse or infringement. Key practices include:

- Data Usage Policies: **See "Adopt and Implement Internal Policies" above.**

- **DATA MAPPING AND INVENTORY:** Conduct data mapping and inventory exercises to track what personal data is used in training, inference, and feedback loops. Maintain a detailed "map" of all data used in training or deploying AI models, include source/origin (e.g., end user, customer, vendor, partner, data broker, etc.), data type (e.g., categories of PII), and nature of data (e.g., anonymized, de-identified, aggregated, etc.). Clearly distinguish between proprietary, licensed, unlicensed, and publicly available data to assess risks relating to compliance, ownership, and licensed rights.

- **TRANSPARENCY AND DOCUMENTATION:** To demonstrate accountability, facilitate audits, and maintain an audit trail, companies must document the provenance of AI training datasets, modifications or preprocessing steps taken, approvals/authorizations for AI Usage Policy exceptions, recordation of acknowledging and facilitating data subject requests, retention and destruction of sensitive categories of data in accordance with applicable laws, etc.

- **PERIODIC AUDITS:** In addition to regular data security and integrity audits, companies should also conduct regular legal compliance audits to verify adherence with applicable requirements, including licensing agreements, regulatory requirements, or flow-down usage limitations or business restrictions.

Companies should take steps to guard against data leakage and ensure compliance with laws, such as:

- Adopting AI governance frameworks that include:
  - Human review of automated decisions,
  - Model auditability and documentation,
  - Secure training pipelines,
  - Periodic testing for data leakage or bias.
- Providing clear notice and obtain any required consents from applicants and employees;
- Establishing internal accountability for AI tools, including HR and legal review of third-party vendor tools;
- Ensuring privacy notices reflect actual data practices, especially regarding AI usage and third-party data sharing; and
- Implementing robust consent mechanisms, including granular, opt-in workflows for high-risk processing.

Companies may benefit from outsourcing these compliance functions instead of performing them in-house. For earlier-stage companies that may not have legal operations/compliance teams, there are many service providers that offer commercially available "responsible AI" and data compliance support services.

Vendor Diligence and Management

Conducting comprehensive vendor and partner diligence is a critical step for companies using or offering generative AI products and services. Ensuring that third-party entities supplying data or AI tools adhere to rigorous legal, ethical, and technical standards helps mitigate downstream risks associated with non-compliance, data misuse, or poor-quality outputs. Key elements of vendor and partner diligence include:

- **BACKGROUND CHECKS AND REPUTATION ASSESSMENT:** Evaluate the vendor or partner's track record, including prior compliance issues, data breaches, or legal disputes. Favor partners with a proven history of ethical data sourcing and AI development that adheres to applicable current industry standards.

- **DATA VERIFICATION:** Require vendors to provide detailed documentation about the provenance of their data. Confirm that data sources comply with all applicable laws, including sector-specific and comprehensive privacy regulations (e.g., GDPR, CCPA), and do not include unauthorized or sensitive information without consent for sub-processing.

- **CONTRACTUAL SAFEGUARDS:** Negotiate agreements that clearly outline data ownership, permissible uses, and liability for breaches or non-compliance. Include indemnification clauses and ensure vendors warrant that their data is legally and ethically sourced. ***See "Contractual Risk Allocation" below.***

- **COMPLIANCE CERTIFICATIONS AND AUDITS:** Insist on relevant certifications (e.g., SOC 2 Type II report, ISO 27001 data security certification, etc.) and conduct regular audits or reviews of vendor data practices. Verify adherence to contractual obligations and legal standards. Review the vendor's protocols for identifying and addressing risks, including those related to data misuse, privacy violations, and security breaches. Require robust incident response and remediation plans.

- **TECHNICAL AND ETHICAL OBLIGATIONS:** Assess the vendor's data curation and AI development processes for compliance with technical quality benchmarks and ethical guidelines, such as reducing bias, avoiding discrimination, and ensuring transparency. Establish mechanisms for continuous oversight—such as periodic reporting, independent audits, or use of automated tracking tools—to ensure vendors maintain compliance throughout the partnership.

When procuring AI tools, such as AI-powered coding companions or notetaking tools, carefully assess how your company will use these third-party products and identify strategies to mitigate potential infringement risks. During the review of vendor agreements, including arrangements governed by a vendor's standard clickthrough terms, consider the rights your company needs and the safeguards the vendor provides for its AI offerings.

Secure Sufficient IP Rights

Companies must ensure that their inbound and outbound IP license agreements secure sufficient IP rights to data necessary for the operation and conduct of business, while also insulating the company from potential down- or up-stream liability. Companies should: (1) clearly define ownership rights to inputs (e.g., user-provided data/prompts) and outputs (e.g., AI-generated content); (2) specify whether users retain ownership of their inputs, and whether outputs are proprietary to the user, shared, or owned by the company; (3) specify whether the company or vendor retains rights to use AI-generated outputs for further model training, improvement, or other purposes; (4) ensure that agreements explicitly outline the scope of IP rights granted to users or obtained from vendors (e.g., whether the use of outputs is restricted to personal or commercial purposes, whether sublicensing is allowed, etc.); and (5) clearly define any limitations or exclusions to these rights.

To guard against uncertainty related to ownership of AI-generated outputs, companies should take steps to:

- Document and preserve evidence of meaningful human contributions to outputs;
- Review model terms of use to confirm rights over outputs;
- Consider incorporating human-in-the-loop design where copyright ownership is a priority; and
- Avoid relying exclusively on AI-generated assets for content or branding that requires IP protection.

Further, given the variability in how courts and regulators may interpret the fair use defense and federal safe harbors for online service providers, it is essential for companies to:

- Conduct a thorough assessment of their legal rights concerning the training data used to develop AI models (e.g., whether such data is licensed, unlicensed, or in the public domain); and
- Evaluate potential exposure to claims alleging that their use or distribution of AI technologies that facilitate illegal activity, uses harmful moderation practices, or generates outputs that are infringing or otherwise cause harm (e.g., defamation, bias, or misinformation).

Finally, to minimize the risk of misappropriating trade secrets or company proprietary information, companies can:

- Conduct diligence on training and fine-tuning datasets, particularly when sourced externally;
- Implement technical safeguards to monitor for data leakage or memorization in model outputs;
- Include strong representations, warranties, and indemnities in vendor contracts regarding data sourcing;
- Train teams on the handling of proprietary information in model development workflows; and
- Design user sign-up processes to enhance the enforceability of clickthrough agreements to protect proprietary data from unauthorized data scraping (e.g., implementing enforceable website terms, restricting public access to sensitive information, incorporating risk-shifting provisions in applicable commercial contracts, and prohibiting the use of proprietary company data for unauthorized or risky purposes).

Contractual Risk Allocation

Companies can also manage potential liability through appropriate contractual risk allocation, including use of indemnification, limitation of liability, disclaimers, and representation and warranty provisions informed by and tailored to business-specific exposure (e.g., intellectual property disputes, data misuse, and harm arising from AI-generated outputs, etc.). Companies should develop a contract playbook for engagement of customers, vendors, and partners that covers at least the following:

Stakeholder	Indemnities	Limitations of Liability	Outputs; Disclaimers	Data Governance
Customers	Limited indemnification for claims related to AI tools; require customer indemnification for improper use of outputs.	Cap liability tied to fees paid; exclude indirect or consequential damages with carve-outs for gross negligence or fraud.	Clarify customer responsibility for use of outputs and include disclaimers about limitations and potential biases.	Limit liability to areas under company control; require customers to follow agreed data security standards.
Vendors	Require vendor indemnity for IP infringement, misuse of data, and claims caused by their AI tools.	Negotiate caps sufficient to cover risks like third-party claims; exclude caps for gross negligence or fraud.	Ensure liability for vendor-provided outputs aligns with agreed warranties and compliance standards.	Establish vendor responsibility for data compliance and breaches; require adherence to privacy laws.
Partners	Mutual indemnification for breaches, IP violations, or misuse of shared AI tools.	Define proportional caps based on each party's control over AI products or services.	Define responsibilities and rights for outputs, or jointly-created IP.	Define clear responsibilities for data governance and liability for misuse of shared datasets.

Companies can manage risk more effectively by carefully crafting risk-shifting contractual provisions. Legal counsel should regularly review contract templates, negotiate material customer and provider arrangements, and align the company's contract negotiation and review playbook to align with new regulatory requirements, industry standards, and market trends. Contractual best practices and recommendations include:

- **TO MITIGATE RISKS ARISING FROM MISREPRESENTATIONS AND WARRANTIES:**
  - Align contractual language with the actual capabilities and limitations of their AI tools;
  - Avoid sweeping representations of legal compliance; and
  - Explicitly restrict or disclaim high-risk use cases where appropriate.

- **WHEN DRAFTING AND REVIEWING INDEMNIFICATION OBLIGATIONS:**
  - Narrow indemnity triggers to specific, high-risk areas (e.g., IP infringement directly caused by the provider's training data);
  - Exclude or limit coverage for user misuse, third-party inputs, or unauthorized downstream use;
  - Include indemnity caps, notice requirements, and control-of-defense provisions to manage exposure; and
  - Where possible, require upstream indemnities from model providers or dataset licensors.
- **WHEN DRAFTING AND REVIEWING PERFORMANCE AND AVAILABILITY COMMITMENTS:**
  - Use disclaimers that clarify the probabilistic nature of AI outputs;
  - Define performance standards in terms of reasonable efforts or representative use cases, rather than absolute guarantees;
  - Tailor support obligations to AI-specific issues, including retraining limitations and latency variability; and
  - Avoid rigid KPIs unless the model is purpose-built and validated for a narrow, well-controlled use case.
- **WHEN PREPARING USE RESTRICTIONS AND ACCEPTABLE USE POLICIES:**
  - Include detailed acceptable use policies (AUPs) in customer agreements and enforce them through technical and contractual means;
  - Restrict high-risk use cases contractually, with clear language and default prohibitions for sensitive domains;
  - Incorporate flow-down provisions to ensure downstream parties comply with upstream license terms for models and data; and
  - Reserve audit and suspension rights for misuse, and consider requiring customer indemnification for unauthorized or harmful uses.
- **TO ADDRESS FLOW-DOWN RISK IN PARTNER AND SUPPLE CHAIN AGREEMENTS:**
  - Ensure all key rights, obligations, and restrictions are flowed down contractually to vendors, model providers, and data licensors;
  - Align upstream and downstream terms before committing to customer-facing representations or indemnities;

- Include audit, cooperation, and indemnity clauses in vendor agreements to close liability gaps; and
- Maintain a centralized register of third-party dependencies, licenses, and obligations to enable contract harmonization.

- **WHEN DRAFTING AND REVIEWING CONTRACTUAL LIABILITY LIMITATIONS:**
  - Include clear liability caps tied to fees paid or a fixed dollar amount, with tightly scoped exceptions.
  - Exclude indirect, consequential, and special damages unless specifically negotiated;
  - Align remedy provisions with what is realistically achievable given the nature of the AI product (e.g., use reasonable efforts language rather than guarantees of re-performance); and
  - Consider disclaimers or limitations specific to AI outputs, such as lack of fitness for a particular purpose or no guarantee of accuracy or legality.

Implement Technical Guardrails

One of the most effective tools companies can use to limit liability associated with AI models is the implementation of robust internal technical guardrails such as:

- For companies developing AI models, it is essential to establish controls that minimize the risk of incorporating copyrighted works. This can include internal restrictions on permissible training data and clear guidelines for employees when utilizing third-party AI tools.
- When using data to train AI models or feed inputs into them, companies must ensure they have obtained the necessary rights to use the data, particularly for purposes beyond providing services to customers. This is especially critical when handling PII, where compliance with privacy laws and principles is imperative to avoid costly financial penalties.
- For businesses integrating generative AI outputs into operations, employee policies and oversight procedures should clearly define how these outputs can be used. Restrictions on incorporating third-party generated content into products and services can help prevent unintentional violations of copyright or IP laws.

Additionally, challenges such as inherent biases in AI algorithms, which can result in discriminatory outcomes, and the generation of inaccurate outputs, commonly referred to as “hallucinations,” can result in significant financial penalties. Courts are increasingly scrutinizing whether companies have exercised reasonable care in the training, deployment, and ongoing monitoring of AI models to mitigate foreseeable risks. To reduce legal and regulatory exposure, organizations should establish robust

AI governance frameworks. These include transparent documentation of AI model design, rigorous testing for bias and inaccuracies, and clear disclaimers regarding AI-generated content. Regular independent audits, adherence to established ethical guidelines, and maintaining meaningful human oversight, especially in high-risk applications, are additional best practices to demonstrate proactive risk management and compliance with evolving legal standards.

Finally, given that AI data practices often conflict with traditional privacy principles, companies should establish technical guardrails in accordance with the latest commercial industry standards and regulatory guidance to protect their ability to collect, use, and store the data essential to their operations. These measures ensure compliance with privacy regulations while supporting the company's data-driven objectives. By proactively implementing these safeguards, organizations can reduce liability, promote ethical AI use, and align with legal and regulatory standards.

Obtain AI Insurance

As AI continues to transform how enterprise operations, the rise in potential liabilities associated with developing and utilizing AI models has exposed companies to unprecedented risks. For example, to reduce the risk of tort and criminal liability, companies can:

- Conduct thorough **model testing and validation**, especially for high-risk use cases;
- Implement **human oversight and failsafes**, particularly where outputs affect health, safety, or legal rights;
- Include clear **disclaimers, instructions, and limitations** in product materials and user interfaces;
- Monitor and **enforce acceptable use policies** to prevent misuse;
- Maintain audit logs, impact assessments, and internal reviews to document responsible development.

One of the most effective safeguards against claims related to AI is the procurement of specialized AI insurance that provides financial protection for businesses using or providing AI tools, covering liabilities that may arise from their use or deployment. Ensuring that AI insurance policies align with a company's specific practices and needs is critical. Tailored coverage can address risks unique to AI operations, such as intellectual property disputes, algorithmic errors, or misuse of AI-generated outputs.

Conclusion

As AI continues to transform business operations and redefine sources of competitive advantage, it is more important than ever for companies to proactively leverage AI capabilities while carefully managing associated risks. Traditional insurance products and outdated contractual protections are increasingly inadequate in addressing the complex and evolving landscape of AI-related exposures. By adopting a forward-thinking approach to risk management—one that integrates customized insurance solutions with strong legal and operational best practices—founders can more effectively protect their organizations from unnecessary liabilities when developing, deploying, or distributing AI technologies.

This white paper explores the varied and multifaceted risks inherent in AI adoption, including those related to liability, regulatory compliance, intellectual property, and operational challenges. These considerations highlight the critical need to understand both the limitations of conventional insurance coverage and the emergence of innovative, AI-specific risk management solutions. This white paper aims to provide companies with practical guidance for securing appropriate coverage and implementing effective strategies to mitigate immediate AI-related risks, while also positioning themselves for long-term resilience and success.