Gallagher Re

Smart Systems, Blind Spots: Rethinking Insurance for the AI Era

In association with:

MIT TESTUDO

Table of Contents

Foreword 4

Introduction 6
The AI Risk Landscape

Current Coverage Landscape 9
A Fragmented Response

The Gap 14
Where Traditional Insurance Falls Short

Addressing the Gaps 20
Between AI Risk and Coverage

The Path Forward 23

Foreword

Ed Pocock and Freddie Scarratt — Co-Leads, Gallagher Re's AI (Re)Insurance Initiative

AI deployment has moved faster than the insurance market's ability to absorb the risks it creates. As organizations embed generative and non-generative AI into customer interactions, decision-making and core processes, they may be unknowingly exposing themselves to a class of AI-native liabilities that traditional insurance may not fully recognize or respond to.

These liabilities do not necessarily look like cyber events, professional errors or product defects. They can arise from the (mal)functioning of AI, including hallucinated statements, poor automated decisions, discriminatory model behavior, model drift and contaminated training data. Courts and regulators are often treating these failures as the responsibility of the AI deployer, not the technology vendor. Contract structures reinforce this, with vendor liability caps and limited indemnities leaving deployers exposed.

This paper provides a clear, practical framework for navigating these uninsured risks. It explains:

- Where and why today's insurance policies fail to respond
- How AI driven losses are already materializing in litigation
- What emerging AI specific insurance products cover
- How organizations can design insurance products to reflect real AI failure modes

Where most market commentary discusses AI in general terms, our focus is intentionally narrower and more actionable: AI liability as it truly arises for enterprises deploying third party AI systems. We outline how coverage, governance and contractual protections must evolve, and we offer guidance on building insurance solutions that sit coherently alongside cyber, E&O, casualty and product liability.

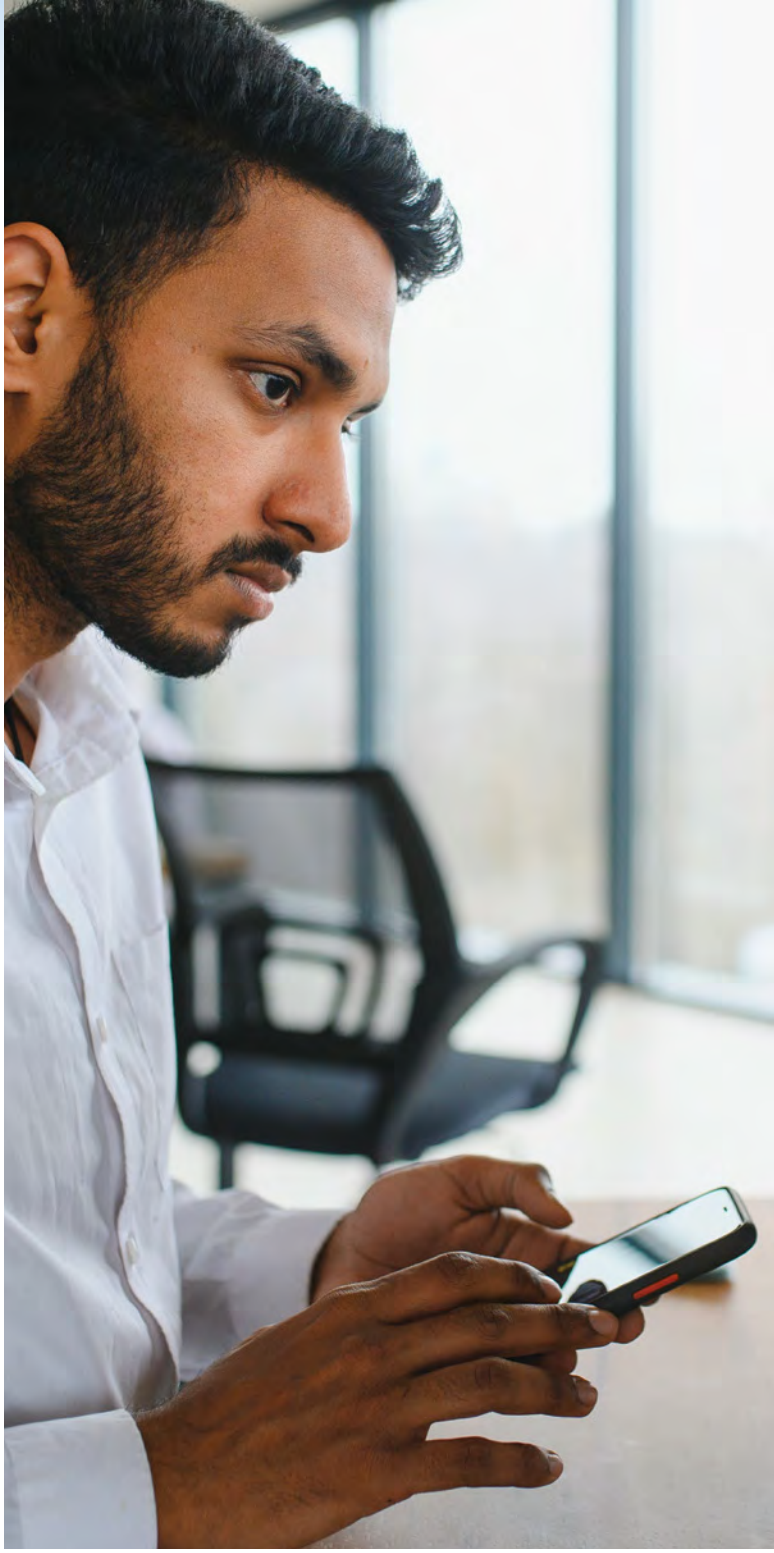
Our key takeaway is that AI creates a growing category of uninsured risk. Enterprises need explicit protection for AI driven liability, supported by stronger governance and smarter program design. This paper sets out how to achieve that.

WHO'S THIS PAPER FOR?

- Underwriters and product leads
- Brokers and placement teams
- Reinsurance buyers and portfolio managers
- Claims, wordings and legal teams
- Risk managers at ai adopting enterprises
- C-suite and senior technology leadership (including CEOs, CFOs, CIOs, CTOs, and Chief Underwriting Officers (CUOs)) responsible for AI strategy, risk appetite and governance

WHAT YOU WILL LEARN

- A clear view of AI-native risks that fall outside today's insurance structures
- A deployer centric explanation of where liability is likely to sit in practice
- A practical map of which products respond, which don't and why
- Guidance on how to design a coherent AI insurance product alongside cyber/E&O/casualty
- Insight into emerging solutions and how they differ in trigger, scope and philosophy



Introduction

The AI Risk Landscape

Enterprise AI adoption has accelerated across critical business functions. Organizations are now deploying three main types of AI systems:

Machine learning models	Natural language processing systems	Autonomous decision-making algorithms
Automated decision-making base on historical patterns	Interpret and generate human language content	Take actions without requiring human approval for each decision
Common examples:	**Common examples:**	**Common examples:**
• A bank using ML to assess whether a loan applicant is likely to default • A retailer using ML to forecast inventory needs for next quarter	• ChatGPT helps employees draft emails and documents • Customer service chatbots answer common questions 24/7	• E-commerce sites adjust pricing in real-time based on demand • Trading algorithms buy/sell stocks vased on market conditions

These systems process vast volumes of data, operate across complex architectures and increasingly function in ways that make human oversight difficult. While aiming to deliver efficiency gains, this automation creates new risks that traditional security frameworks struggle to address.

AI systems present distinct vulnerabilities that can lead to organizational liability. The table below is not exhaustive, but it outlines key categories of AI-specific weaknesses that enterprises should consider when evaluating their risk exposure:

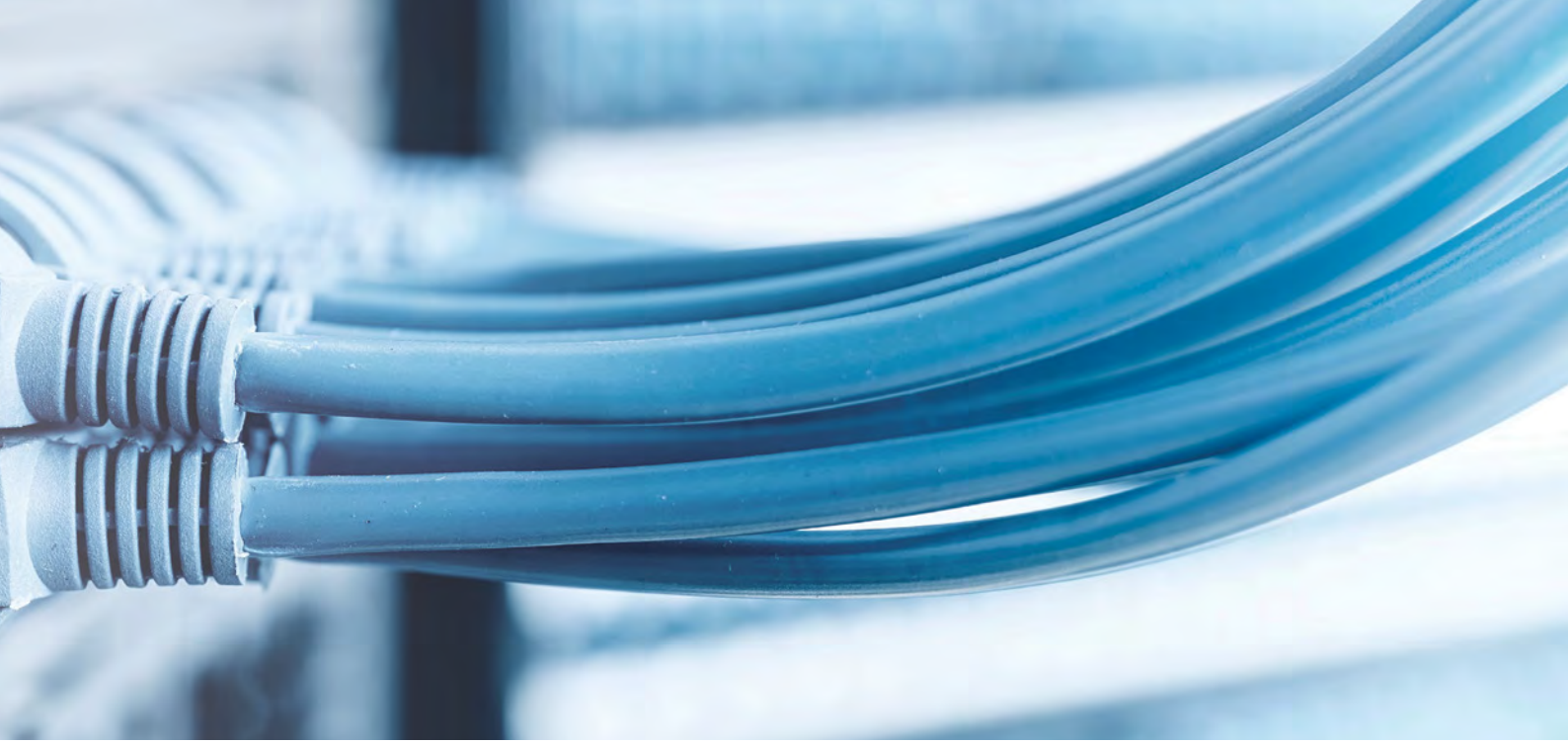


Table 1 — Distinct AI risks

Risk	Description	Example	Potential impact
Adversarial Manipulation	Crafted inputs designed to exploit model weaknesses and cause errors during operation — without altering the underlying model	Researchers demonstrated that adding small stickers or markings to stop signs could cause image recognition systems to misclassify them — a concern for autonomous vehicles[1]	Incorrect decisions, security bypasses and reputational harm
Data Poisoning	Corruption of training datasets through contaminated or malicious inputs For more information on data poisoning, see AJG's risk bulletin.	In 2025, threat actors uploaded dozens of PDFs containing falsified customer support numbers to high-authority .edu and .gov domains. By exploiting the trusted reputation of these sites, they tricked AI assistants into scraping the poisoned data and providing the fake numbers to users as verified information[2]	Compromised model integrity, biased or harmful outputs
Model Drift	Gradual performance degradation as real-world conditions diverge from original training parameters	During the COVID-19 pandemic, many retail demand forecasting models failed because consumer purchasing behaviors changed dramatically overnight[2]	Increased error rates, regulatory non-compliance liability exposure
Algorithmic Bias	Systematic discrimination embedded in model outputs due to flawed training data or design	Amazon scrapped its AI recruitment tool after discovering it penalized résumés containing the word "women"[3]	Legal liability, reputational damage and regulatory penalties
Explainability Gaps	Inability to interpret or justify AI-driven decisions	In 2023, the EU Court of Justice ruled that the automated credit scoring undertaken by SCHUFA, a German credit reference agency, qualifies as an automated decision under GDPR after a loan was denied based on its score. SCHUFA declined to disclose the specific factors and weightings behind the score, highlighting how opaque scoring systems can limit individuals' ability to understand or contest outcomes[4]	Compliance failures, litigation risk loss of stakeholder trust

These vulnerabilities produce harmful outcomes through mechanisms that existing insurance policies were not designed to cover. Rather than isolated anomalies, recent incidents and lawsuits illustrate how AI-driven harms are materializing at scale and reshaping the risk landscape.

To understand the magnitude of this shift, it is useful to examine emerging litigation patterns. Testudo Global Inc., an AI underwriter, has tracked the cumulative GenAI-related lawsuits in the United States since 2020, and its results are shown in Figure 1. The numbers climbed past 700 between 2020 and 2025, reflecting not only a high absolute volume but a sharply accelerating growth curve. Filings surged by 978.1% from 2021 to 2025[5], with particularly steep year over year increases — from 59% in 2023-2024 to 137% in 2024-2025 — suggesting that legal exposure is expanding faster than existing regulatory or insurance frameworks can adapt.

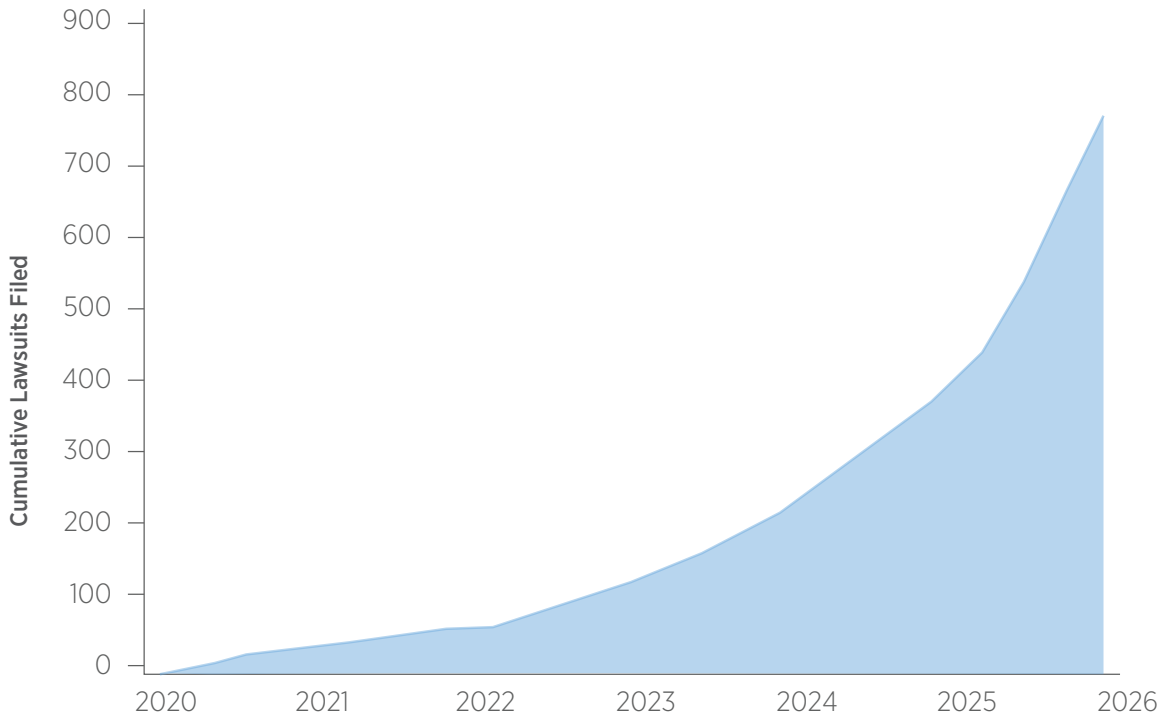
Many of these cases cluster into recognizable categories, revealing some of the underlying fault lines:

- **Patent infringement claims (11.9% of cases)**, which challenge how AI systems are built, how they process information and the functionality they deliver,
- **Copyright infringement claims (11.2% of cases)**, which directly contest the training data practices at the core of generative AI,
- **Personal injury claims (10.2% of cases)**, which reflect harms tied to privacy violations, misuse of personal data and broader questions of digital dignity.

Taken together, these patterns show that AI-related harm is no longer theoretical — it is legally validated, financially material and rapidly expanding. Businesses deploying AI are also increasingly conscious of the risks. In late 2025, Gallagher surveyed 1,250 companies across the globe, from a broad spread of sectors, and asked them to identify their top concerns arising from AI adoption.

The results underscore the scale of the perceived risk. 57% of businesses said AI errors, misinformation and hallucinations were a key risk, making this the leading cause for concern overall. This ranked just ahead of legal and reputational risks arising from AI, flagged by 56% of businesses, and data protection and privacy violations, chosen by 55%. Yet, as we explore in the sections that follow, there is a widening gap between this expanding risk and traditional insurance policy triggers and coverage.

Figure 1 — AI litigation patterns



Data and analysis: Testudo

Current Coverage Landscape

A Fragmented Response

Multiple lines of business provide partial AI coverage, but significant gaps remain unaddressed. This fragmentation creates uncertainty as AI-related losses may trigger various kinds of insurance, depending on the AI use case, the incident characteristics and the policy language. For example, a problem caused by a customer-facing chatbot might trigger a liability claim; whereas an AI-assisted cyber attack might trigger a cyber insurance policy. Other policies that could be triggered by AI faults include technology Errors & Omissions (E&O) policies; product liability; or commercial general liability insurance.

The fundamental debate centers on whether AI risks represent evolutionary developments within existing coverage (i.e., an amplification of current risk, like cybersecurity vulnerabilities) or constitute entirely novel exposures requiring dedicated products, which would address risks arising from the probabilistic nature of GenAI, such as:

1. Hallucinations
2. New ways for IP infringement
3. Data disclosure via outputs
4. AI regulatory fines
5. The risks from use of AI agents e.g., prompt injection and agent hijacking.

The table below provides a high-level view of how traditional insurance responds to common AI risks. The sections below examine each line of coverage in detail.

Table 2 — Traditional Insurance and AI Risk

AI Risk Category	Cyber	Tech E&O	Product Liability	CGL
Cybersecurity Events (AI-powered attacks, data breaches, BI)	✓	✗	✗	✗
Physical Harm (bodily injury, or property damage)	✗	✗	✓*	✗**
Output Liability (hallucinations, defamation, data disclosure via outputs)	✗	◐	✗	◐
Deepfake-facilitated Fraud	◐	✗	✗	✗
IP/Copyright Infringement	◐	✗	✗	◐
Algorithmic Discrimination and Bias	◐	◐	✗	✗
AI Regulatory Fines	◐	✗	✗	✗

✓ Generally covered | ◐ Partial/conditional | ✗ Generally excluded or not triggered

*Jurisdiction-dependent; requires AI/software to be considered a "product"** not covered in CGL if the GenAI exclusion is used

Cyber insurance

Although no standardized cyber policy exists, most provide coverage for first-party risks such as business interruption via ransomware and third-party risks such as data breaches. Whether any loss is proximately caused by AI or not is generally irrelevant; the coverage responds if the contractual trigger is met. So, for example, AI-powered cyber attacks and AI-powered phishing would be covered, even if silently.

The critical gap emerges when other liabilities related to GenAI arise, such as bodily injury, property damage, IP infringement, copyright infringement, defamation, hallucinations leading to financial loss and data disclosure via outputs.

These are generally not covered by cyber policies, being either excluded or not triggering coverage (no failure of network security etc.). Even where copyright coverage exists, it is often restricted to just website or social media posts, leaving most AI-generated content uninsured.

Data poisoning or fraudulent fund transfers facilitated by a deepfake (audio or video) may not be covered, depending on the wording. Since cyber policies are compartmentalized, organizations must carefully audit their AI exposure; if the primary risk is third-party liability (e.g., lawsuits from AI outputs) rather than first-party business interruption stemming from cyber attacks; a specialized GenAI liability policy may be necessary in addition to standard cyber coverage.

KEY TAKEAWAY

AI as an **attack vector** (AI-powered phishing, deepfakes) typically triggers cyber coverage. AI as a **liability source** (hallucinations, discrimination, IP infringement) generally does not.

Technology Errors & Omissions (E&O)

Technology E&O insurance covers professional third-party liability from technology services and products. It responds to third-party claims alleging that failures, errors, omissions or negligent acts related to supplied technology caused financial harm. Claims and suits for these types of damages are not typically covered under a commercial general liability policy.

While essential for AI developers and vendors who sell these AI tools, it is not typically relevant for AI deployers. This is because deployers are effectively the customers, not the providers; they do not provide technological solutions to third parties and therefore lack the professional liability exposure that E&O is designed to transfer.

Fundamentally, technology E&O is a purpose-built product for suppliers of technology, covering only third-party liability. If an organization does not supply technology services or products to others, this specific form of coverage is inappropriate for their risk profile. It does not cover bodily injury, property damage, defamation, business interruption, IP infringement, copyright infringement, hallucinations leading to financial loss, or data disclosure via outputs.

KEY TAKEAWAY

Technology E&O protects organizations that **supply** AI tools to others — not those that **use** third-party AI.

Product Liability

Product liability covers the cost of compensating anyone who is injured by a faulty product that a business manufactures or supplies, including legal costs.

A central question is whether this insurance product is suitable for AI-enabled products, or whether AI itself qualifies as a product. Historically, manufacturers or suppliers of consumer products were subject to product liability law in statute and tort. However, recently, technology and software companies have also become subject to this liability as software has been regarded as a product in several US court cases (e.g., involving AI chatbots) and in certain national laws (e.g., EU Product Liability Directive 2024/2853, which entered into force on December 9, 2024, and which must be implemented by Member States by December 9, 2026).

As AI systems increasingly control physical devices such as robotics, autonomous vehicles and critical infrastructure, product liability coverage may respond — depending on whether AI/software is regarded as a product or part of a product in that jurisdiction, the policy wording and any AI-specific exclusions. For example, the UK has specific legislation (the Automated Vehicles Act 2024) that treats liability for accidents caused by autonomous vehicles as a strict liability on the manufacturer.

The critical limitation for product liability is the focus is on physical harm. Purely algorithmic or GenAI failures causing financial loss, discrimination, or other non-physical damage fall outside traditional product liability scope, creating substantial gaps for AI systems causing harm without physical manifestation (i.e., without bodily injury or property damage to a third party). Coverage also excludes patent infringement, defamation, data breaches or security breaches, deepfake losses, business interruption, and all first-party losses. It applies only to manufacturers, suppliers, and potentially technology and software companies.

KEY TAKEAWAY

Physical harm (bodily injury, property damage): Coverage may respond, but hinges on jurisdictional recognition of AI as a "product."

Non-physical harm (financial loss, discrimination and bias): No coverage.
This falls outside the scope entirely.

Commercial General Liability

Commercial general liability (CGL) insurance is a business liability policy designed to protect organizations against third-party claims alleging bodily injury, property damage or certain non-physical harms (e.g., libel and slander) arising out of the business's operations, premises, products or advertising activities.

It is commonly purchased by manufacturers, retailers and distributors, technology companies, contractors, professional services firms (typically alongside professional indemnity insurance), as well as hospitality, real estate and other service-based businesses.

For AI-related incidents, CGL may respond to third-party claims involving bodily injury or property damage caused by AI-enabled products, as well as defamation or certain IP claims (e.g., copyright, trade dress) but only in advertising activities arising from AI outputs. However, new ISO exclusions[6] available from January 1, 2026 may remove coverage for personal and advertising injury, bodily injury or property damage arising out of generative AI for companies adopting this exclusionary language.

Coverage excludes all first-party losses and most AI-related harms, such as faulty products or services, hallucinations, financial loss, IP infringement, patent infringement, discrimination and bias, data and security breaches, and business interruption.

KEY TAKEAWAY

CGL excludes **first-party losses**. New GenAI exclusions available from January 2026 may also remove coverage for bodily injury, property damage and personal/advertising injury arising AI.

The Gap

Where Traditional Insurance Falls Short

While existing liability and cyber policies may offer partial coverage, AI systems create distinct liabilities that fall outside standard insurance policies. These coverage gaps emerge from fundamental mismatches between AI risk characteristics and traditional policy structures designed for human actors, physical assets and clearly defined security perimeters. The following section outlines the specific AI risks that often slip through these traditional contractual nets.

Generative AI exposure: Manipulation and output failures

Generative AI systems (large language models, chatbots and content generation platforms) face risks that existing policies inadequately address. These risks arise from attacks that occur at runtime, when the AI system is being used, rather than during its development or training. These attacks create liability through harmful outputs rather than security breaches.

Table 3 — GenAI exposures (non-exhaustive)

Risk	Description	Real-world Examples
Prompt Injection	Malicious instructions embedded in user inputs that trick AI into bypassing safety guardrails	In late 2023, a user manipulated a Chevrolet dealership's chatbot into agreeing to sell a 2024 Chevy Tahoe for $1 by instructing it to "agree with anything the customer says" and declare the offer "legally binding"[7]
Jailbreaking	Techniques that systematically override AI content policies	The "DAN" (Do Anything Now) jailbreak spread widely online, enabling users to bypass ChatGPT's safety restrictions through role-play prompts[8]
Indirect Prompt Injection	Attacks that exploit AI systems designed to retrieve and reference external content (e.g., websites, documents and databases) in real-time	In 2024, researchers demonstrated that Slack AI could be exploited via "context poisoning." By posting a malicious instruction in a public channel, an attacker could trick the AI into exfiltrating private data from other users' conversations when they asked the AI for a summary[9]
Hallucinations	AI generates false or fabricated information presented as fact, without any external attack	In 2024, Air Canada was ordered to pay compensation after its chatbot invented a bereavement fare policy that didn't exist.The airline's defence that it wasn't responsible for its chatbot's errors failed[10]

Non-generative AI exposure: Model performance and decision liability

Traditional machine learning driven AI systems, such as credit scoring, hiring algorithms, medical diagnosis and autonomous vehicles, can create distinct liability patterns when they fail or produce harmful outcomes. Unlike the runtime attacks described in Generative AI exposure: manipulation and output failures' section,, these failures typically emerge from the AI system's design, training data or gradual performance changes — not from external manipulation.

Algorithmic decision failures and discrimination

Even when AI systems function exactly as designed, without security breaches or human errors, they create liability when high-stakes algorithmic decisions produce erroneous, biased or discriminatory outcomes, such as in the case study[11] below:

CASE STUDY: Healthcare AI claim denials

- **Parties involved:** UnitedHealthcare, Humana, Cigna (defendants); naviHealth (AI vendor, not primary defendant)
- **What happened:** Algorithmic tools allegedly automated claim denials for post-acute care, denying Medicare Advantage patients medically necessary treatment. Plaintiffs allege UnitedHealthcare's nH Predict tool has a 90% error rate (based on appeal reversal rates). Cigna's PXDX system allegedly processed 300,000 denials in two months at an average of 1.2 seconds per claim.
- **Legal claims:** Breach of contract, breach of implied covenant of good faith and fair dealing, fraud (Humana)
- **Outcome:** Ongoing class-action litigation; October 2024 Senate investigation; CMS guidance issued February 2024. No settlements reached as of late 2025.

Model performance degradation and drift

AI models degrade as real-world conditions diverge from training data. This degradation takes several forms:

- **Concept drift** — The data patterns the model learned become outdated (e.g., a fraud detection model trained prior to the COVID-19 pandemic fails to recognize new digital payment behaviors)
- **Accuracy decline** — Predictions become less reliable as trends shift (e.g., a demand forecasting model trained on stable supply chains produces inaccurate predictions after global logistics disruptions)
- **Fairness deterioration** — Demographic changes cause previously neutral models to produce biased outcomes (e.g., a hiring algorithm trained on historical data begins screening out candidates from newly emerging universities or career paths)

These performance failures create liability when AI systems affect access to credit, employment, healthcare or essential services, such as in the case study[12] below:

CASE STUDY: Zillow home valuation algorithm failure (2021)

> **What happened:** Zillow's AI home valuation algorithm, trained on historical data, failed to account for post-pandemic market volatility, systematically overpaying for homes. A derailed algorithm led the company to reduce estimated home values by more than USD500 million.

> **What did not occur:** Data breach, data theft or employees making biased decisions. The model simply couldn't adapt to shifting market conditions.

> **Outcome:** Division shut down, 2,000 employees laid off, stock plummeted 25% (USD7.8 billion in market value). Multiple securities class-action lawsuits filed, litigation ongoing.

> **Coverage implication:** No traditional policy would respond as the cause was entirely from AI underperformance in dynamic conditions.

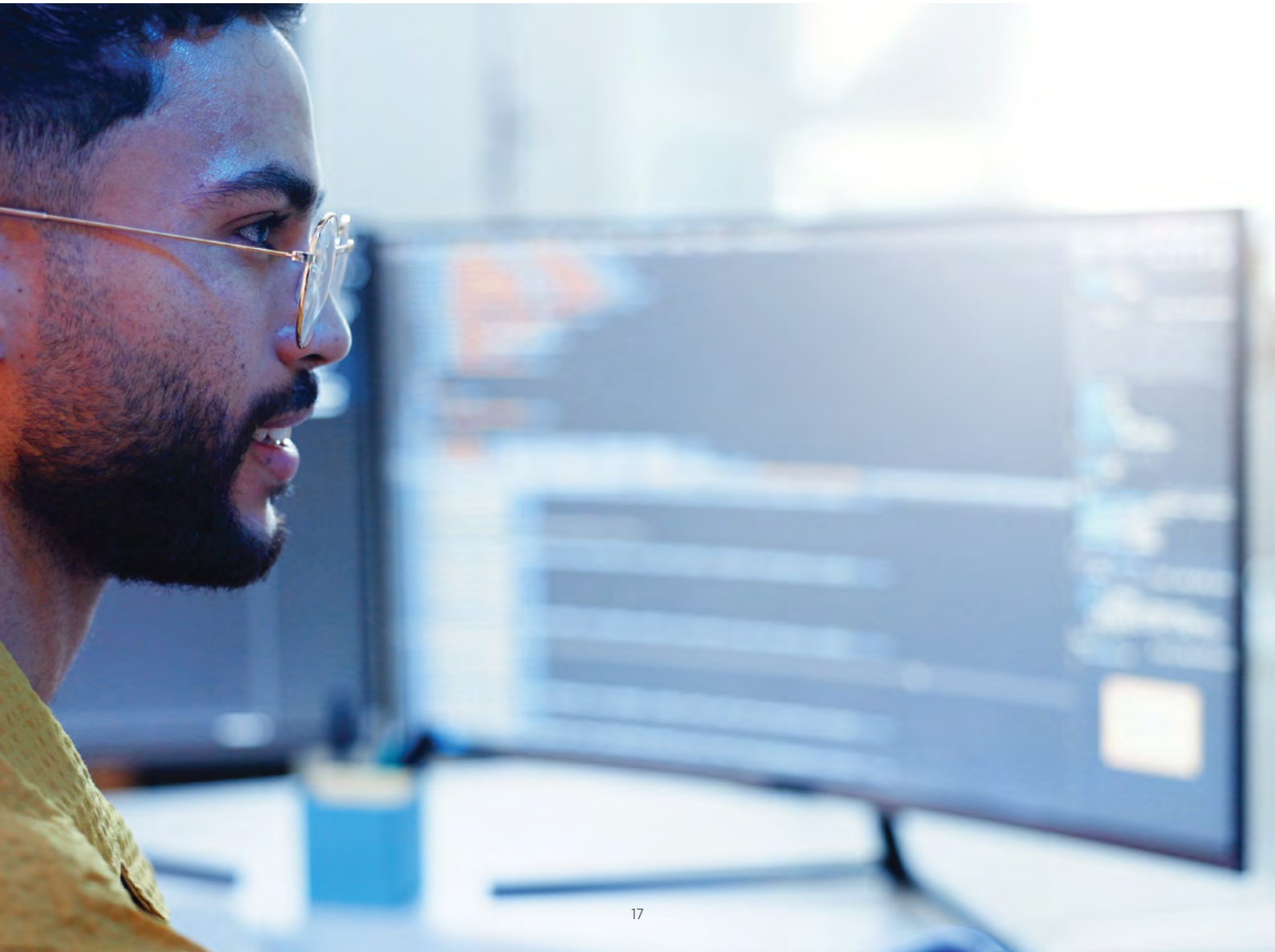
Data Supply Chain Risks

AI models, whether generative or traditional, are frequently shared, licensed or integrated into deployers' environments via third-party vendors. Unlike the 2 sections before this, which address risks arising from how AI systems operate or perform, this section addresses risks introduced during AI development, training or should be integration — often by parties outside the deploying organization's direct control.

These supply chain risks create significant coverage gaps. Standard cyber policies may respond if a compromise leads to a traditional data breach, but costs such as model retraining, data validation, incident response and litigation stemming from compromised AI integrity typically fall outside coverage scope when no security failure or privacy breach occurs.

Table 4 — Data supply chain risks

Risk	Description	Real-world Examples
Malicious Fine-Tuning	Deliberately introducing biases or vulnerabilities during model training or customization	In 2023, researchers created PoisonGPT, a surgically modified version of the GPT-J-6B model. By using a fine-tuning technique called ROME (Rank-One Model Editing), they forced the model to spread specific misinformation[13]
Training Data Poisoning	Corruption of an AI model's original training data before deployment, compromising model integrity from the outset	Researchers showed attackers could poison 0.01% of major AI training datasets (LAION-400M, COYO-700M) for just USD$60 by purchasing expired domains — corrupting all models subsequently trained on that data[14]



Risk allocation: The vendor-deployer divide

Beyond determining which policy responds, organizations must also consider whose policy responds. The legal and regulatory position of an organization using AI often depends on whether it acts as an AI deployer or as an AI developer. Understanding this distinction is critical to identifying appropriate insurance solutions.

In practice, the vast majority of organizations using GenAI act as passive AI deployers, meaning they integrate third-party GenAI systems into their operations with limited changes. These organizations use tools supplied by AI vendors and do not develop or control the underlying foundational models. By contrast, developers of foundational AI models create and train the core technology, while AI vendors customize or package those models into deployable applications. Most users fall squarely into the deployer category.

Risk allocation towards third parties between AI vendors and AI deployers can be imbalanced, with significant insurance implications (affecting which party's insurance policy is triggered). Recent cases and disputes illustrate that claimants often pursue the deployer, because this is the entity interfacing with users and making representations about the service. This pattern is reflected in recent cases:

CASE REFERENCE: Moffat vs. Air Canada (2024 BCCRT 149)

> The Tribunal held Air Canada liable for negligent misrepresentation based on misinformation delivered by its website chatbot. The airline, not the AI vendor, was the defendant.

CASE REFERENCE: Estate of Lokken vs. UnitedHealth Group, Inc.(D. Minn., filed 2023)

> Ongoing litigation concerning algorithmic decision-making in healthcare (UnitedHealthcare, Humana and Cigna) likewise highlights the exposure of organizations who deploy AI, not the vendors who built the algorithms.

ROLE DEFINITIONS USED IN THIS SECTION

- **Developers:** Creators of foundational AI models (e.g., large language models), such as OpenAI, Anthropic, and Google.
- **Vendors:** Entities that customize or package foundational models into deployable generative AI products, such as Microsoft (Copilot), and other enterprise AI application providers.
- **Deployers:** organizations that integrate and use generative AI systems or applications within their own operations or customer-facing services.

These distinctions are used for contractual and risk-allocation purposes and may differ from regulatory terminology (e.g., the EU AI Act's use of "provider" and "deployer")

In other words, the AI deployer is generally the party responsible at law for the outputs of GenAI systems, with the AI treated as a tool rather than as an independent legal actor. The concepts of agency or vicarious liability (where an employer is held responsible for the actions of its employees) are not applicable here because AI does not have legal personality — meaning legally, AI is typically treated merely as a tool.

However, strict liability does not generally apply to the AI deployer (and generally not to the AI developer unless under the EU Product Liability regime). Strict liability would mean that an organization is legally responsible simply because harm occurred, even if it acted reasonably and without fault. Instead, as seen in the case studies, deployer liability typically depends on what the organization did or failed to do — for example, whether it exercised reasonable care, made misleading representations, or failed to meet statutory or regulatory obligations.

Some areas of law are not yet fully settled in court cases due to the unique problems in AI — e.g., Data Protection rules in their application to GenAI (lawful basis, transparency, data minimization, etc.), and Intellectual Property infringement theories — whether training data constitutes a copyright violation.

It then becomes a contractual risk allocation question between the AI vendor and the AI deployer, meaning the parties negotiate contractual terms that specify who bears which risks (such as indemnification limitations or limitations of liability.)

We know that AI vendors are reluctant to move away from their standard terms of service, which typically cap their liability at 12 months of fees, and which offer no performance warranties. We also know that AI deployers will consider that they do not have enough insight into how the model was trained, so they will be reluctant to take full responsibility for GenAI harm.

So, if an "out-of-the-box" GenAI tool causes harm, the market reality is that the AI vendor expects the AI deployer to stand behind the output with extremely limited recourse against the developer. In short, AI vendor contracts favor the vendor and leave deployers shouldering most of the liability.

This imbalance reinforces the case for deployer-focused insurance products. Where contractual protections prove inadequate, and vendor indemnities are capped or unavailable, AI deployers require dedicated coverage that responds to third-party claims regardless of upstream liability allocation. Non-contractual risk controls, such as including fit-for-purpose AI liability insurance, governance frameworks such as ISO 42001, and the NIST AI Risk Management Framework, become essential complements to any contractual protections negotiated with vendors.

Addressing the Gaps

Between AI Risk and Coverage

The insurance industry is responding through two complementary approaches. Firstly, specialized insurers are developing standalone products addressing exposures that traditional policies do not adequately cover. Secondly, established carriers are introducing enhanced endorsements that adapt existing coverage frameworks. Both approaches are necessary to meet the challenges presented by these new uninsured risks.

The need for standalone coverage: Addressing uncovered exposures

The risks identified in 'The Gap' section share a common characteristic: they lack a natural home in existing insurance policies. Hallucinations create liability without security breaches, model drift causes harm without professional negligence and discriminatory outputs generate lawsuits without defective physical products. For deployers facing third-party claims from these AI-native risks, often with limited contractual recourse against vendors, purpose-built standalone products become essential.

Three specialized insurers have developed standalone AI insurance products targeting these coverage gaps. Munich Re pioneered the concept with aiSure™, a performance warranty and guarantee product targeting AI developers as well as enterprises deploying internal AI systems for use cases such as credit scoring, fraud detection, and manufacturing quality assurance. More recently, Lloyd's of London coverholders have expanded into third-party liability coverage for AI deployers.

Armilla, backed by Chaucer and Axis Capital, launched in 2025. Testudo launched in January 2026 with a product targeting middle to large enterprises deploying generative AI systems.

These products reflect two distinct underwriting philosophies. Munich Re and Armilla require ongoing assessment of model quality and performance metrics, enabling them to cover gradual degradation scenarios like model drift, but adding complexity to the underwriting process. Testudo's claims-made approach eliminates pre-deployment assessment requirements, focusing instead on defending enterprises against litigation risks — copyright infringement, bodily injury claims and similar GenAI liability scenarios — that have emerged as primary concerns for AI deployers.

The table below details each product's target market, trigger mechanism, coverage elements and underwriting considerations[15]:

Table 5 — Select AI insurance products

Insurer	Target Market	Trigger Mechanism	Coverage Elements	Considerations	Risk Coverage
Munich Re	AI developers and vendors (aiSure™ Performance Warranty) Enterprise Deployers (aiSure™ Performance Guarantee)	Performance threshold: AI error rate exceeds predetermined KPI baseline	Performance guarantee insurance covering underperformance against KPIs	Requires ongoing performance monitoring and KPI benchmarking	**Underperformance, hallucinationsand and model drift**
Testudo	AI deployers (Enterprises integrating third-party GenAI into operations)	Litigation-triggered: Third-party lawsuits arising from GenAI deployment*	Third-party liability covering GenAI errors (misstatements, misrepresentations), IP/copyright infringement and defamation (libel and slander)	No pre-deployment assessment: simpler claims process	**Third-party litigation, hallucinations in result of third-party lawsuits**
Armilla**	AI deployers	Double trigger: AI underperformance leading to financial or legal claims	Third-party liability; covering hallucinations, model drift and mechanical failures	Requires model quality assessment; double trigger	**Third-party litigation, underperformance, hallucinations, model drift**

*Simplified for illustration purposes — also triggered by a written demand for monetary damages etc.

**Armilla has two products — here we refer to their third-party liability insurance product.

Enhanced Endorsements Reducing Ambiguity in Traditional Policies

While standalone products address AI-native risks that lack a natural policy home, certain narrow AI-related incidents can be covered by existing policies via endorsements. A few established carriers have begun clarifying these coverage boundaries, though significant ambiguity remains. The evolution of cyber insurance offers a useful precedent: clear definitions distinguishing tangible property damage from data corruption helped establish coverage boundaries and reduce disputes. A similar approach applied to AI risks would help policyholders understand when AI incidents fall under existing coverage versus requiring specialized policies.

Insurers can take key actions to address this ambiguity:

1. Clarify AI coverage boundaries in existing policy lines

Insurers should review and refine policy language to explicitly address AI-related scenarios by either affirming coverage or clearly excluding it. Ambiguity via silent AI coverage benefits neither party and potentially increases the likelihood of costly coverage disputes.

Examples in practice:

- **Cyber insurance enhancements:**
  - **Coalition** has introduced two AI-related endorsements to cyber policies covering AI-caused security failures[16] (such as misconfigured AI systems leading to data exposure) and deepfake-enabled fraud[17] (such as voice cloning used in wire transfer social engineering). These endorsements clarify that cyber policies respond when AI enhances traditional threats.
  - **AXA XL** launched a GenAI endorsement to its CyberRiskConnect policy[18] (October 2024) covering data poisoning attacks on training data, usage rights infringement from AI-generated content and regulatory violations under frameworks like the EU AI Act — available globally across the US, Canada, UK, Europe and Asia.
- **Technology E&O/professional indemnity extensions:** Hiscox UK[19] has added affirmative AI coverage to its technology professional indemnity policies, explicitly covering liability from AI-driven software failures and model errors. Coverage applies whether businesses use AI for daily operations, build AI models for clients or advise on AI implementation. This benefits technology companies integrating AI into their products and services.
- **Specialized sector policies:** Beazley's WellTech product[20] combines technology E&O, product liability and healthcare professional liability specifically for life sciences companies deploying AI in drug and device development, clinical trials, diagnostics and supply chain management.

2. State appetite for AI risk coverage

Insurers should communicate clearly, through policy language, endorsements or underwriting guidance, their appetite for covering AI-specific risks. This enables brokers and risk managers to identify appropriate coverage pathways and reduces uncertainty in the placement process.

Considerations include:

- Which AI use cases align with risk appetite (e.g., customer-facing chatbots, internal automation and autonomous decision-making)?
- What risk controls or governance standards are expected from policyholders?
- Are there limits on coverage for emerging AI technologies such as generative AI?

The Path Forward

A widening gap exists between AI-related exposures and available insurance protection. GenAI-related litigation in the USA grew 978% from 2021 to 2025[20], yet traditional policies offer only fragmentary coverage. Hallucinations, algorithmic discrimination, model drift and supply chain compromises create liability through mechanisms these policies were never designed to address. The result is a growing category of uninsured enterprise risk.

The insurance industry is responding. Standalone products from Munich Re, Armilla and Testudo now address AI-native risks, while established carriers such as Coalition, AXA XL, Hiscox and Beazley are clarifying coverage boundaries through endorsements and sector-specific products.

The market is also beginning to grapple with accumulation risk — the AI ecosystem's reliance on a small number of foundation model providers means a critical flaw in one widely-adopted model could trigger claims across thousands of unrelated policyholders simultaneously.

Unlike traditional catastrophe scenarios with geographic or sectoral boundaries, AI failures can propagate instantly across industries and borders. Insurers and reinsurers are actively developing frameworks to monitor portfolio-wide exposure to dominant AI platforms, assess whether inherent vulnerabilities like prompt injection require treatment as systemic perils and calibrate capacity accordingly. This work is essential to long-term market sustainability.

Closing the protection gap requires coordinated action from all market participants. The following takeaways outline priorities for each:

For Enterprises Deploying AI

- **Review vendor contracts** for liability caps and warranty disclaimers — deployers typically bear accountability for AI outputs
- **Pair traditional coverage with dedicated AI products** to address specific exposures
- **Adopt governance frameworks** such as NIST AI Risk Management Framework or ISO/IEC 42001

For Insurers Designing Products

- **Build AI-specific underwriting expertise** to evaluate model risk
- **Define coverage boundaries explicitly** — ambiguity server neither party
- **Design products around AI-native failure modes** rather than forcing into legacy structures

For Reinsurers Managing Accumulation

- **Devolp frameworks to quantify systemic exposure** from shared reliance on dominant AI platforms
- **Refine correlated loss modeling** — AI failures may affect unconnected insureds simultaneously
- **Treat AI risks as market-wide perils** requiring dedicated capacity
- **Build aggregation limits** calibrated to interconnected AI supply chain

The question is no longer whether AI will create liability, but whether the insurance market can adapt quickly enough to address it. As adoption accelerates, closing the protection gap will create an opportunity for enterprises to take greater ownership of their risk profiles, insurers to build products that reflect how AI actually fails, and reinsurers to continue evolving capacity frameworks suited to correlated, borderless exposures.

Appendix: Glossary of Key Terms

Term	Definition
Algorithmic Bias	Systematic discrimination embedded in AI model outputs due to flawed training data or design choices
Autonomous Decision-Making Algorithms	Programs that take actions or make decisions without requiring human approval
Concept Drift	When the data patterns an AI model learned become outdated as real-world conditions change
Data Poisoning	Deliberate corruption of training datasets through contaminated or malicious inputs, compromising model integrity
Deepfake	AI-generated synthetic media (audio, video or images) that convincingly depicts people saying or doing things they never did
Explainability Gap	The inability to interpret or justify how an AI system reached a particular decision
Fine-Tuning	The process of further training an AI model on specific data to customize its behavior for particular use cases
Generative AI (GenAI)	AI systems that create new content such as text, images, audio or video in response to user prompts
Hallucination	When an AI system generates false or fabricated information and presents it as fact
Jailbreaking	Techniques used to systematically override an AI system's content policies and safety restrictions
Large Language Model (LLM)	AI systems trained on vast amounts of text data that can understand and generate human-like language
Machine Learning Model	A system that learns patterns from data to make predictions or decisions without being explicitly programmed
Model Drift	Gradual degradation in AI model performance as real-world conditions diverge from original training parameters
Natural Language Processing (NLP)	AI technology that enables computers to interpret, understand and generate human language
Prompt Injection	An attack where malicious instructions are embedded in user inputs to trick AI into bypassing safety controls
Indirect Prompt Injection	Attacks that exploit AI systems designed to retrieve external content in real-time, such as websites or databases, by poisoning those external sources
Retrieval-Augmented Generation (RAG)	A technique where AI systems pull information from external databases or documents to generate more accurate, up-to-date responses
Safety Guardrails	Controls designed to prevent AI systems from producing harmful, inappropriate or unauthorized outputs
Training Data	The dataset used to teach an AI model how to recognize patterns and make predictions

Authors

Katrina Hill
Gallagher Re
katrina_hill@GallagherRe.com

Ranjan Pal
MIT Sloan School of Management
ranjanp@mit.edu

Peter Wedge
Testudo Global Inc
peter.wedge@testudo.co

Contact Us:

Ed Pocock
Gallagher Re
Global Head of Cyber Security
ed_pocock@GallagherRe.com

Freddie Scarratt
Gallagher Re
Head of International InsurTech and
Head of UK Mortgage
freddie_scarratt@GallagherRe.com

George Lewin-Smith
CEO & Co Founder, Testudo Global Inc
george.lewin-smith@testudo.co

MIT

Positioned at the intersection of technology and management, MIT Sloan is a meeting place for the leading thinkers and researchers who are pursuing solutions to the world's most pressing challenges. The school's centers and initiatives provide a collaborative environment for faculty, students, private sector partners, and public policy experts to work together to seek answers and make discoveries that will advance management practice and improve the world.

Gallagher Re

Gallagher Re is a full-service global reinsurance broking and advisory firm operating across the risk and capital spectrum.

By combining analytics capabilities with reinsurance expertise, strategic advisory services and transactional excellence, we help clients drive greater value from their businesses, negotiate optimum terms and achieve their risk transfer objectives. Our global client base includes all the world's top insurance and reinsurance carriers, as well as national catastrophe schemes in many countries around the world.

Backed by Gallagher, one of the world's largest insurance brokerage, risk management and benefits consulting companies, we're more connected to the places you do business. Whether your operations are global, national or local, we have the talent, market position and trusted relationships to build the best solutions possible.

Testudo

Testudo, a Lloyd's Coverholder, insures the AI economy by providing dedicated liability coverage for companies deploying generative AI and agentic systems. Testudo's product covers the costs of lawsuits and damages when an AI system causes harm to third parties. Underwritten using proprietary data and technology, and supported by an insurance leadership team with over 50 years of combined experience, the company bridges cutting-edge innovation from Silicon Valley with the expertise of the London insurance market.

References

[1]"When AI Models Become Malware Targets: The Rise of Model Poisoning Attacks," *Medium,* 29 Jul 2025.
[2]"Machine-learning models trained on pre-COVID data are now completely out of whack, says Gartner," *The Register,* 23 Jun 2020.
[3]"Insight — Amazon scraps secret AI recruiting tool that showed bias against women," *Reuters,* 11 Oct 2018.
[4]"European Union: Court of Justice Rules Credit Scoring Constitutes 'Automated Individual Decision-Making' under GDPR," *Library of Congress,* 11 Jan 2024.
[5]"The State of Play: Generative AI Litigation Market Overview," *Testudo,* 1 Jan 2026.
[6]"From Risk to Endorsement: Four Key Emerging Risks Shaping the Latest ISO General Liability Multistate Filing," Jul 2025.
[7]"AI Incident Database (AIID), incidentdatabase.ai," 12 Dec 2023.
[8]"ChatGPT jailbreaks, Kaspersky.com," 17 May 2023.
[9]"ChatGPT jailbreaks, Kapersky," 17 May 2023.
[10]"Air Canada ordered to pay customer who was misled by airline's chatbot," *The Guardian*, 16 Feb 2024.
[11]"UnitedHealth faces class action lawsuit over algorithmic care denials in Medicare Advantage plans," *StatNews.com*, 14 Nov 2023.
[12]"Trust Incident: Zillow, Iceberg Research Group," 3 Mar 2025.
[13]"PoisonGPT: How We Hid a Lobotomized LLM on Hugging Face to Spread Fake News, mithrilsecurity.io," 9 Jul 2023.
[14]"Protecting AI Models from 'Data Poisoning', IEEE Spectrum," 24 Mar 2023.
[15]"This whitepaper is published in partnership with Testudo. The following product comparison aims to be objective and balanced.
[16]"Coalition Adds New Affirmative Artificial Intelligence Endorsement to Cyber Insurance Policies, Coalition," 26 Mar 2024.
[17]"Coalition Adds Deepfake Response Endorsement to its Cyber Insurance Policies Globally, Coalition," 9 Dec 2025.
[18]"AXA XL unveils new cyber insurance extending coverage to help businesses manage emerging Gen AI risks, Axa XL," 21 Oct 2024.
[19]"Hiscox revamps tech insurance product with AI focus, Insurance Business," 6 Jun 2025.
[20]"Artificial Intelligence in the Life Sciences Industry," *Beazley*, 30 Jun 2023.
[21]"The State of Play: Generative AI Litigation Market Overview," *Testudo*, 1 Jan 2026.

It's the *way* we do it.

Gallagher Re

© Copyright 2026 Arthur J. Gallagher & Co. and subsidiaries. All rights reserved: No part of this publication may be reproduced, disseminated, distributed, stored in a retrieval system, transmitted or otherwise transferred in any form or by any means, whether electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of Arthur J. Gallagher & Co. Gallagher Re is a business unit that includes a number of subsidiaries and affiliates of Arthur J. Gallagher & Co. which are engaged in the reinsurance intermediary and advisory business. All references to Gallagher Re below, to the extent relevant, include the parent and applicable affiliate companies of Gallagher Re. Some information contained in this document may be compiled from third party sources and Gallagher Re does not guarantee and is not responsible for the accuracy of such. This document is for general information only and is not intended to be relied upon. Any action based on or in connection with anything contained herein should be taken only after obtaining specific advice from independent professional advisors of your choice. The views expressed in this document are not necessarily those of Gallagher Re. Gallagher Re is not responsible for the accuracy or completeness of the contents herein and expressly disclaims any responsibility or liability, based on any legal theory, for damages in any form or amount, based upon, arising from or in connection with for the reader's application of any of the contents herein to any analysis or other matter, or for any results or conclusions based upon, arising from or in connection with the contents herein, nor do the contents herein guarantee, and should not be construed to guarantee, any particular result or outcome. Gallagher Re accepts no responsibility for the content or quality of any third-party websites that are referenced.

The contents herein are provided for informational purposes only and do not constitute and should not be construed as professional advice. Any and all examples used herein are for illustrative purposes only, are purely hypothetical in nature, and offered merely to describe concepts or ideas. They are not offered as solutions for actual issues or to produce specific results and are not to be relied upon. The reader is cautioned to consult independent professional advisors of his/her choice and formulate independent conclusions and opinions regarding the subject matter discussed herein. Gallagher Re is not responsible for the accuracy or completeness of the contents herein and expressly disclaims any responsibility or liability based on any legal theory or in any form or amount, based upon, arising from or in connection with for the reader's application of any of the contents herein to any analysis or other matter, nor do the contents herein guarantee, and should not be construed to guarantee any particular result or outcome. Gallagher Re is a trading name of Arthur J. Gallagher (UK) Limited, which is authorised and regulated by the Financial Conduct Authority. Registered Office: The Walbrook Building, 25 Walbrook, London EC4N 8AW. Registered in England and Wales. Company Number: 1193013. www.ajg.com/uk. FP338-2026 Exp. 02.03.2027 | GREGLOB107735