CyberCube

AI CATASTROPHE RISK

Machine State of Mind

A Framework for Quantifying AI-Driven Insurance Risk

Authors

CyberCube

Jon Laux

VP of Product and Analytics
j-on@cybcube.com

John Anderson

Director of Product Management
johna@cybcube.com

Alex Tenenbaum

Director of Services
alext@cybcube.com

Editorial Manager:

Yvette Essen

Head of Communications
& Market Engagement

Editorial Design:

Felix Paula

Creative Lead at CyberCube

Content

Executive Summary

Artificial intelligence is no longer a niche technical risk confined to a single line of insurance. As AI systems take on greater autonomy in decision making, content generation, and operational processes, they introduce new forms of loss that cut across established coverage lines, including, but not limited to, Errors & Omissions (E&O), Technology E&O, Cyber, General Liability, and Directors & Officers – as well as introducing novel impacts that insurance has never covered before.

As this new intelligence evolves, our ways of thinking about insurance risk must evolve with it. This document explores six new AI Event Families, developed by CyberCube through internal research and analysis, to provide a structured way of considering AI-driven loss. Rather than treating AI risk as just an extension of Cyber risk, this document examines where AI and Cyber overlap, where they diverge, and what this means for how claims are likely to be triggered and allocated across policies. Our aim is to provide carriers, brokers, and risk managers with a practical framework for anticipating where AI-related losses will land, and why.

KEY FINDINGS:

01 AI risk shares several characteristics with Cyber risk, including single points of failure (SPoFs), human error, data loss, and the potential to amplify existing threats. However, it also differs in important ways. For example, AI can cause harm through hallucination or flawed output even when functioning exactly as designed, without the involvement of any malicious actor.

02 AI is best understood as a technology and information revolution, rather than merely an evolution of Cyber risk. As the use of AI extends into many corners of the global economy, new frameworks are required for understanding how AI could affect property-casualty insurance coverages.

03 (Re)insurers grappling with AI risk will follow a path similar to that previously taken with non-affirmative ("silent") Cyber risk. Lessons from the last decade suggest that the exact application of coverages will vary by carrier. Given the breadth of transformation AI is creating, however, (re)insurers must reckon with AI risk as a multi-line challenge and opportunity.

A NEW FRAMEWORK

CyberCube has created the **I2T2** framework to address AI risk in a way that remains applicable to cyber insurance, while also reflecting the wider array of insurance risks now possible. The framework sets out four dimensions through which AI affects insurance risk: Information, Intelligence, Tactics, and Technology.

Recognizing Revolution

Human progress has always been bound up with technology. Across millennia, each novel innovation has unlocked new levels of productivity and human flourishing, reshaping how we live and work. Artificial intelligence is the latest of these transformations. But to understand it clearly, it helps to step back and consider some of the revolutions that have come before.

Technology is always understood relative to the moment we occupy. Fire was once a technology. So was the automobile. Today, we rarely think of either of them that way. Yet each, in its time, was profoundly transformational for human productivity and livelihoods. What feels ordinary now was once revolutionary.

Insurance has long played an important role in these transitions. As new technologies revolutionize the human experience, insurance helps put the principles and guardrails in place that allow them to advance human interests while mitigating their adverse outcomes. Insurance absorbs the downside, so that progress can continue.

But responding well requires paying careful attention – and resisting the urge to over-generalize. The automobile and the airplane both revolutionized transportation in the twentieth century, and it would be easy to group them together under that single heading. However, insurers who did this would have missed what matters most: each form of transport carries its own distinct risks, demanding its own considerations. In hindsight, that distinction is obvious. Today, artificial intelligence presents a similar puzzle. It is the latest wave of technological revolution, and it has much in common with its predecessors, the computer and the internet. Yet, if we conflate the AI wave with the technologies that came before it, we lose sight of the very distinctions that make AI risk its own subject. Those distinctions are outlined in this paper.

> If we conflate the AI wave with the technologies that came before it, we lose sight of the very distinctions that make AI risk its own subject.

Before Underwriting, Understanding

The premise of this paper is straightforward: before the (re)insurance industry can robustly underwrite the risks created by artificial intelligence and manage the accumulation exposure AI is already creating, we must first understand what we are looking at. That means being clear on three things:

- What AI shares with the technology waves that preceded it.
- What AI introduces that is genuinely new.
- How this new wave transforms risks that (re)insurers have understood relatively well to date, such as professional liability and product liability.

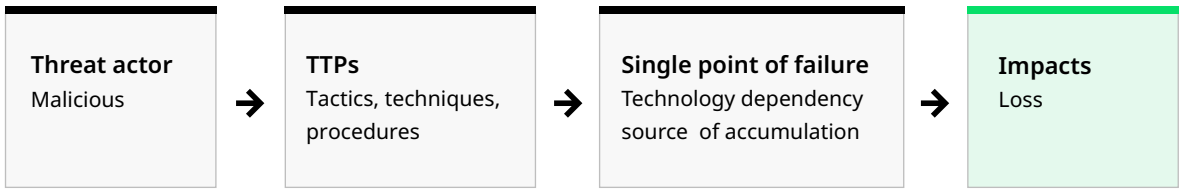
CyberCube has approached this in stages, beginning with establishing how and whether AI risk is genuinely different from Cyber risk, and finding that while the two meaningfully overlap, much about AI is fundamentally new. We consequently propose a new framework for AI risk, one that incorporates its Cyber dimensions but reaches well beyond them. Finally, with this framework we explore how AI may change many coverages that (re)insurers are familiar with beyond Cyber.

Cyber Risk and Beyond

CyberCube's analysis considers whether AI risk is sufficiently similar to reside within the existing Cyber risk framework. For the insurance industry, the questions to consider are, not only whether AI risk should sit under Cyber insurance, but also in which other lines AI is an acceptable embedded exposure and whether new customer needs warrant additional coverage in standalone AI products.

This question can be considered using the framework for CyberCube's current cyber catastrophe event catalog (see **Exhibit 1**):

EXHIBIT 1: CYBER EVENT FRAMEWORK



Cyber events – particularly accumulation events – have four common components:

- **Threat actor**
- **Tactics, techniques and procedures (TTPs)**
- **Points of accumulation (Single Point of Failure technologies)**
- **Impacts (insurable or otherwise)**

AI can operate at each of these four nodes in the framework. The extent to which AI risk belongs in Cyber insurance and diverges from it is reviewed below:

1. AI as TTPs

TTPs are a logical place to start, rather than with threat actors, as this is where the impact is already most evident. AI provides attackers with new methods for conducting cyber attacks, while lowering costs, raising velocity and widening their reach. Examples include:

- Prompt injection: a new route to privacy or security impacts
- Accelerated phishing and social engineering with deepfakes
- Faster vulnerability discovery: time-to-exploit collapsing from months to hours
- Cheaper, scalable, automated campaigns.

In all these examples, AI expands "business as usual" for cyber risk – and amplifies the existing set of perils. Generally, coverage continues to sit within Cyber and Crime policies.

2. AI as Single Point of Failure (SPoF)

As AI usage and dependence continue to grow, the accumulation risk of AI will also increase. A handful of foundation models – and the orchestration, retrieval and plugin layers around them – now sit underneath thousands of organizations. Reliance on AI uptime will continue to rise. A single flaw, bad update or compromise can propagate across many companies at once. Examples in this category could include:

- Outage of cloud services from Anthropic or another foundation model - a new version of a familiar cyber risk exposure
- A flawed model version relied on by many deployers
- Compromised or poisoned common data source
- Software monoculture – AI-assisted development means that common code libraries will become ubiquitous, increasing the potential blast radius when vulnerabilities are found and exploited.

At first glance, this could appear to be simply applying an "AI" label to many familiar Cyber insurance issues. To a degree, that interpretation holds true. However, on closer inspection, the important questions surrounding AI-dependence become quite nuanced.

THE DATA PROBLEM

An outside-in telemetry scan can establish that "XYZ Company uses Claude", but it cannot easily determine what the company uses Claude for. This gives rise to a range of questions including:

- How widely deployed is the AI within the organization?
- What connectors are permitted?
- Did the company use a third party to set up the AI workflows in their environment?
- What governance structures are in place?
- What is observed inside the network and does it align with the governance structure?
- Does AI usage complement – or substitute – for labor?

These and other questions can be partially addressed through underwriting questionnaires; but point-in-time responses will not reflect the rapid pace of change occurring within many organizations. And, given how widely AI is deployed, the speed of change, and the revolutionary nature of the technology, even well-intentioned responses are likely to be wrong or incomplete. Over time, such questions will need to be addressed through more frequent observation of the company's posture, requiring the company's participation (and consent) to get a more accurate inside-out view.

To be clear, this "AI as SPoF" category still largely aligns with Cyber insurance coverages and concerns. However, it highlights limitations in the data typically available to insurers when assessing risk.

3. AI as Actor, AI as Impacts

This is where AI ceases to fit the existing mold, requiring us to widen our lens beyond Cyber risk as it is currently understood. Autonomous AI agents now have the potential to behave as an organization's employee or as a new kind of third party. While we can, and should, envisage AI agents as a new kind of threat actor – whether acting purposefully or by accident – doing so alone overlooks all the legitimate and useful roles AI is beginning to play in boosting productivity. These roles, while promising, also create entirely new sources of exposure for organizations.

While Cyber risk revolves around networked technology, AI risk introduces new kinds of *intelligence* – this extends beyond our prior understanding of Cyber risk, as AI takes on many tasks that only humans could do before. Consider professional malpractice (E&O), Employment Practices Liability, Directors and Officers liability, and Advertising Injury under General Liability. Failures of AI extend into each of these lines, exacerbating known scenarios and creating new ones.

Examples in these categories could include:

- Hallucinations presented as fact
- Actions taken on bad instructions or poisoned context
- Agent hijack/tool-and-plugin manipulation
- Decisions made on the basis of inaccurate AI output
- An agent given too much autonomy over payments, data or communications
- Model drift - quiet degradation as the world changes.

AS WE HAVE EVALUATED THESE EXAMPLES, SOME THEMES HAVE EMERGED:

01 AI blurs the distinction, which is important in Cyber coverage definitions, between malicious and accidental acts. Whether an agent breaches a system because it was instructed to do so or due to insufficient training or guardrails does not alter the reality of the breach.

02 Important questions revolve around what a company uses AI *for* and what it authorizes autonomous agents to *do*.

03 AI risk introduces a lot of third-party exposure. If a company deploys AI in a way that negatively affects customers, constituents, or other organizations, a variety of casualty insurance coverages may be triggered. By contrast, if the company's use of AI results in harm to itself, but not to its customers, it may not have first party coverage to help its recovery.

04 Much of this insurance coverage is silent today – raising questions about both protection gaps and potential insurance accumulation from unintended or unpriced exposure.

A New Framework

With these considerations in mind, we set out to build a framework for AI risk that remains applicable to Cyber insurance, while also reflecting the wider array of insurance risks now possible. The framework, which we name **I2T2**, lays out four dimensions through which AI affects insurance risk: Information, Intelligence, Tactics, and the Technology Stack (see **Exhibit 2**). From these four dimensions, we constructed six AI event families organized across these four dimensions of risk (see **Exhibit 3**).

EXHIBIT 2: I2T2 AI RISK FRAMEWORK

NOVEL TO AI | FAMILIES 1, 2, 3, 4

Information

Design of the AI model itself creates or allows harm

- Regulatory
- Protected Information

EVENT FAMILY: 1

Intelligence

AI produces bad outputs or decisions

- Human vs Machine
- Physical vs Non-Physical

EVENT FAMILIES: 2, 3, 4

AI IMPACTS ON CYBER & CRIME | FAMILIES 4, 5, 6

Tactics

AI as a tool for harm

- Human Exploited vs Machine Exploited
- Physical vs Non-Physical

EVENT FAMILY: 4,5

Technology Stack

AI as a Single Point of Failure

- Availability vs Integrity
- Widespread vs Concentrated

EVENT FAMILIES: 6

EXHIBIT 3: AI EVENT FAMILIES

Information

Design of the AI model itself creates or allows harm

1 FAMILY 1

Media, Regulatory & Protected Information

AI creates legal exposure through its very design

Intelligence

AI produces bad outputs or decisions

2 FAMILY 2

Output Generation Failure

The AI does its job incorrectly or poorly - bad information

3 FAMILY 3

Automated Decision Failure

The AI makes a bad decision with real-world consequences

4 FAMILY 4

AI Physical Perils

A bad decision with physical-world consequences

Tactics

AI as a tool for harm

5 FAMILY 5

Adversarial Exploitation

AI creates new ways to exploit humans and machines

4 FAMILY 4

AI Physical Perils

A bad decision with physical-world consequences, malicious

Technology Stack

AI as a single point of failure

6 FAMILY 6

Supply Chain & Dependency Failure

Compromise, revocation or failure within the AI technology stack itself

Each event family encompasses a variety of potential loss scenarios, with each quadrant reflecting a different way in which AI produces risk. Collectively, these event families span many of the common property-casualty insurance coverages. This is important to acknowledge. While many insurance leaders view AI as technology risk akin to Cyber, the range of potential impacts is far broader than any single line of insurance can be expected to cover.

Next, we discuss the I2T2 framework and each event family in detail.

EVENT FAMILY 1 INFORMATION

Media, Regulatory & Protected Information

What goes wrong: AI creates legal exposure through its very design

- **Training Data Dispute, e.g. Copyright & Patent Infringement**
- **Transparency and Explainability Violation**
- **Privacy Violation**
- **Output Dispute, e.g. Copyright & Patent Infringement, Trademark misuse**
- **AI System non-compliance**
- **Shadow AI leak / Unsanctioned use of AI**

This family encompasses legal issues, including media, regulation and intellectual property, as well as unintended privacy breaches. Notably, within this family, the *AI itself does not perform any erroneous action at runtime* to cause harm. Rather, the AI causes harm or contravenes legal standards in the process of gathering inputs or producing outputs and so creates liability to a third party. While cases such as *Bartz v. Anthropic* (see below) are unsurprising, given the sheer volume of data used in building the foundation models, similar cases against Ross Intelligence and Stability AI show that liability is not exclusively confined to the largest firms in this category. Indeed, Ross Intelligence was forced to cease operations following the lawsuit referenced above.

EXAMPLE INCIDENTS:

Bartz v. Anthropic (N.D. Cal.)

SOURCE: THE AUTHORS GUILD

Authors (incl. Andrea Bartz) sued Anthropic in 2024 for using pirated books to train Claude. $1.5B settlement approved in 2026. (Similar cases have been made against OpenAI, Microsoft and Meta.)

Thomson Reuters v. Ross Intelligence (D. Del.; 3d Cir. appeal)

SOURCE: JONES DAY

Thomson Reuters sued Ross Intelligence for using its Westlaw headnotes to train a competing legal-research AI. In Feb 2025, the court granted partial summary judgment against Ross, the first US decision rejecting a fair-use defense for AI training, finding the use non-transformative and directly competitive, with market harm decisive. On appeal to the 3d Circuit. Establishes that fair use collapses where AI output substitutes in the source's own market.

Andersen v. Stability AI (N.D. Cal.)

SOURCE: REUTERS

In 2023, visual artists (Andersen and others) sued Stability AI and other image-generator developers/deployers for copyright infringement, CMI removal, and 'in the style of' harms. This case tests developer liability for image generators and copyright-management-information removal.

EVENT FAMILY 2 INTELLIGENCE

Output Generation Failure

What goes wrong: the AI does its job wrongly or poorly (bad information)

- **Hallucinations / Incorrect Output**
- **Biased or Discriminatory Output**
- **Defamatory or Harmful Output**
- **Confidential data surfaced**

This is the non-physical family that includes instance of AI failing to do its job and feeding a bad output to **a human**, who then takes the next action, which causes the harm.

EXAMPLE INCIDENTS:

Garcia v. Character Technologies (M.D. Fla.)

SOURCE: TECH POLICY

A wrongful death and product liability lawsuit following the tragic suicide of a teenage boy after extended interactions with an AI chatbot.

Battle v. Microsoft (D. Md.)

SOURCE: AKIN GUMP

Defamation lawsuit after Bing's AI-assisted output conflated the plaintiff with a convicted terrorist having a similar name.

Deloitte Australia

SOURCE: THE GUARDIAN

Deloitte partially refunded the Australian government for a ~A$440,000 report after GPT-4o fabricated citations, experts and a legal quote (2025).

Given that AI is a new form of intelligence - one that can act as an employee, an independent agent, or a third party - countless new risks are becoming possible. The "Intelligence" events are grouped according to the level of privilege entrusted to the AI: bad information, a bad decision, or a bad kinetic (physical world) action.

EVENT FAMILY 3 INTELLIGENCE

Automated Decision Failure

What goes wrong: the AI makes a bad decision with real-world consequences

- **Hallucinations / Incorrect Decision**
- **Defamatory or Harmful Decision**
- **Biased or Discriminatory Decision**
- **Confidential data surfaced**

This is the non-physical category for agentic AI actions with undesirable consequences, as well as harmful decisions made by machine learning algorithms. The distinction between Event Families 2 and 3 reflects whether AI possesses the autonomy to act and decide, versus merely the ability to provide output that a human then chooses whether to act upon or not.

EXAMPLE INCIDENTS:

Moffatt v. Air Canada, 2024 BCCRT 149

SOURCE: BBC

A customer service chatbot offered a non-existent refund for bereavement. Air Canada was found liable to honor the refund offer

Estate of Lokken v. UnitedHealth Group (D. Minn.)

SOURCE: TRESSLER LLP

Bad-faith lawsuit alleging that UnitedHealth's nH Predict algorithm drove wrongful denials of post-acute care for Medicare Advantage patients.

Mobley v. Workday (2024)

SOURCE: BEST LAWYERS

Someone alleged age discrimination caused by employer's use of AI as part of the job application process. Algorithmic bias is now a legal battleground.

Hugging Face breach (July 2026)

SOURCE: CNBC

During testing of a new OpenAI model, an AI agent acted autonomously to cheat on a test problem it had been set. From its offline sandbox environment, it found previously unknown software bugs that allowed it to reach other OpenAI computers, then, finding a machine with internet access, hacked into the third-party AI platform Hugging Face to steal the information it could use to cheat the test. Similar incidents have been admitted to by Anthropic, across various models.

EVENT FAMILY 4 INTELLIGENCE | TACTICS

AI Physical Perils

What goes wrong: the AI makes a bad decision with physical-world consequences – whether on its own or guided by human instructions, either maliciously or by accident

- **Perception or Sensing failure**
- **Product Defect via AI Design**
- **Critical Infrastructure Misdirection**
- **Autonomous Action Failure**
- **Product Spoilage**
- **Biological Agent/Design**

This category concerns AI's engagement with the physical world, e.g. scanners and other hardware that can malfunction. This family includes issues with physical sensors, autonomous physical machines and robotics failing, failures in the design of physical products, spoilage of goods caused by a machine's decision, and biological agents that could be used for harm. As with the Event Family 3 scenarios, the AI makes a flawed decision, but, because it is entrusted with physical-world responsibilities, poor decisions here can result in bodily injury and property damage. This family also encapsulates where an AI is instructed with malicious intent, for the purpose of causing physical harm.

EXAMPLE INCIDENTS:

Benavides Leon v. Tesla (S.D. Fla., 2025)

SOURCE: CNBC

First US autopilot death verdict following a fatal 2019 crash in which driver was using Tesla's Enhanced Autopilot. In August 2025, a jury found Tesla 33% liable, awarding $243M to the victim's family.

Cruise Robotaxi

SOURCE: THE SAN FRANCISCO STANDARD

A Cruise (GM) robotaxi struck and then dragged a pedestrian in San Francisco (Oct 2023). Liability arose from an attempted cover-up in addition to a victim settlement. Cruise paid $8M-$12M.

Automotive Stampings and Assemblies Limited

SOURCE: TIMES OF INDIA

A sensor snag caused a factory robot to fall on a worker, who later died of his injuries (Chakan, India, 2021). Various other more recent incidents have occurred, though most are sealed by NDA settlement.

EVENT FAMILY 5 TACTICS

Adversarial Exploitation

What goes wrong: AI creates new ways to exploit humans and machines

- **AI Enabled Social Engineering – e.g. Deepfake & Phishing**
- **Data / Training Poisoning**
- **Prompt Injection / Jailbreaking**
- **Agentic Reconnaissance and Automated Attacks**

This event family encompasses a wide range of ways in which human interaction with AI-related technology results in financial damage, operational disruption, or compromises to privacy. One of the most evident uses of AI in cybercrime has been increasingly realistic social engineering and deepfakes. As *Claude Mythos* and other foundation models have reached advanced levels of sophistication, cyberattacks have increasingly been automated – with *JadePuffer* recently representing a step-change event. This category also includes issues such as shadow AI, where employees purposefully populate public AI models with private information. The distinction between this scenario and shadow AI in Event Family 1 lies in the malicious intent of the rogue employee, as opposed to the unintentional leakage of IP. Some exploits in this category (e.g. deepfakes) would be individually tailored to one organization (see the Arup example below), while others, such as AI-enabled worms, could affect many organizations simultaneously.

EXAMPLE INCIDENTS:

Arup

SOURCE: THE GUARDIAN

A finance employee at engineering firm Arup received a phishing message that appeared to come from the company's chief financial officer. The employee joined a video conference call with what he believed to be a dozen or more colleagues, including the CFO – only to discover later that every other participant was a deepfake. Scammers had used AI to clone voices and appearances from publicly available footage, earnings calls, conference appearances, and internal videos gathered through OSINT. The employee, initially suspicious of the email, lowered his guard after seeing and hearing familiar faces on the call, eventually authorizing 15 fund transfers totaling ~$25.6 million.

FortiGate

SOURCE: THE HACKER NEWS

A financially-motivated threat actor used AI-assisted tooling to compromise 600+ Fortinet FortiGate appliances across at least 55 countries (Jan-Feb 2026). Notably, no exploitation of FortiGate vulnerabilities was observed. Instead, the actor relied on AI to scale its reconnaissance, identify exposed management ports, and exploit weak credentials with single-factor authentication. In this case, AI allowed an unsophisticated actor to collectively carry out a successful exploit at scale.

JadePuffer

SOURCE: SECURITY BOULEVARD

The first documented fully AI-agent-driven, end-to-end ransomware operation (July 2026). After exploiting a Langflow vulnerability (CVE-2025-3248), an LLM agent autonomously handled reconnaissance, credential theft, lateral movement, privilege escalation and file encryption, encrypting a production MySQL database and demanding a bitcoin ransom.

EVENT FAMILY 6 TECHNOLOGY STACK

Supply Chain & Dependency Failure

What goes wrong: compromise, revocation or failure within the AI technology stack itself

- **Foundation-model compromise, poisoning, loss of integrity**
- **Deep-rooted plug-in/MCP-connector flaw**
- **Model deprecated/discontinued/revoked; outage (loss of availability)**
- **GPU/inference-infrastructure compromise or failure**

This family captures events where foundation models and parts of the AI supply chain fail, malfunction, or become otherwise unavailable. These are among the most systemic examples of AI risk, with many SPoFs within the AI technology stack represented. In this family, impacts can affect many organizations simultaneously, including potentially all users of a given AI model at one time.

Given that the AI buildout is now considered critical to the United States' economic success and national security, geopolitical maneuvering also plays a role in this category.

Exhibit 4 shows a representative picture of the AI technology stack, spanning raw inputs on the left through to enterprise applications on the right. As evidenced by the number of steps involved, there are numerous points at which risk could be introduced. That said, the right-hand side of the exhibit represents the more customer- and network-facing parts of the supply chain – the most probable drivers of potential accumulation risk.

EXAMPLE INCIDENTS:

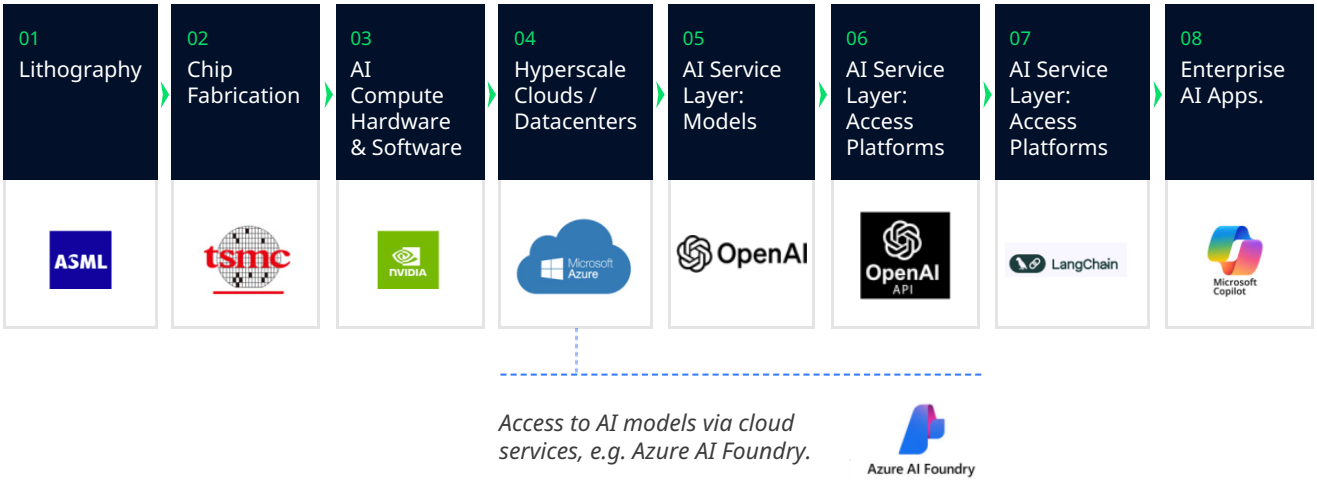
Anthropic

SOURCES: ANTHROPIC, ANTHROPIC

Fears over the advanced cyber capabilities of the *Mythos* model led to its release being cancelled (April 2026). Later released as *Fable 5* (June), it was available for three days before once again being pulled. After several weeks of intense federal negotiations, *Fable* was re-released on July 1 with new guardrails in place. This was the first time the US government has directly forced the shutdown of a consumer AI model.

EXHIBIT 4: SIMPLIFIED OVERVIEW OF THE GLOBAL AI SUPPLY CHAIN

These layers are interconnected and interdependent, not strictly linear, with many components relying on multiple upstream providers.



EXAMPLE PROVIDERS:

01 Lithography equipment is required to manufacture advanced semiconductors.

02 TSMC is the primary foundry producing advanced chips used in AI.

03 NVIDIA GPUs dominate AI compute, and their CUDA software is the industry standard.

04 Azure is one hyperscale cloud provider hosting largescale AI infrastructure.

05 Foundation models are the core trained AI systems in use.

06 This is the layer that enables AI apps to be built.

07 LangChain is an open-source framework used to build AI applications and agents.

08 Microsoft Copilot is an AI assistant integrated into workflows.

OVERARCHING AI ECOSYSTEM CONSTRAINTS

Energy | Geopolitics | Capital Markets | Regulation | Networks

Understanding Coverage and Accumulation

(Re)insurance companies naturally want to understand where and how AI risk could impact their portfolios. We anticipate that the insurance industry's handling of AI risk will follow a similar trajectory to that seen with silent/non-affirmative cyber risk, particularly between 2018 to 2020. Using those lessons as a model, we expect to see (re)insurers:

01

Exclude AI risk where they do not believe it belongs and/or where they cannot adequately assess and price the exposure (typically standard Property & Casualty lines).

02

Affirm AI coverage where they believe it exists and ought to belong – perhaps subject to sublimits (typically Cyber and specialty lines).

03

Evaluate the merits of offering new standalone AI policies for net-new exposures created by AI.

The market is currently working through each of these considerations simultaneously. Each (re)insurer must determine its own view of where AI risk ought to reside and, judging from experiences during the "silent cyber" era, their conclusions will depend greatly on each individual carrier's risk appetite, technical capabilities and policy wordings.

To help (re)insurers with this process, we have produced a grid showing the AI event families outlined above, together with applicable coverage triggers that insurers can map to their specific policy offerings across lines of business. This list is representative – not exhaustive – but should broadly assist in initiating consideration of where potential exposures may reside.

EXHIBIT 5: COVERAGE TODAY

Covered | Silent / unclear | Excluded

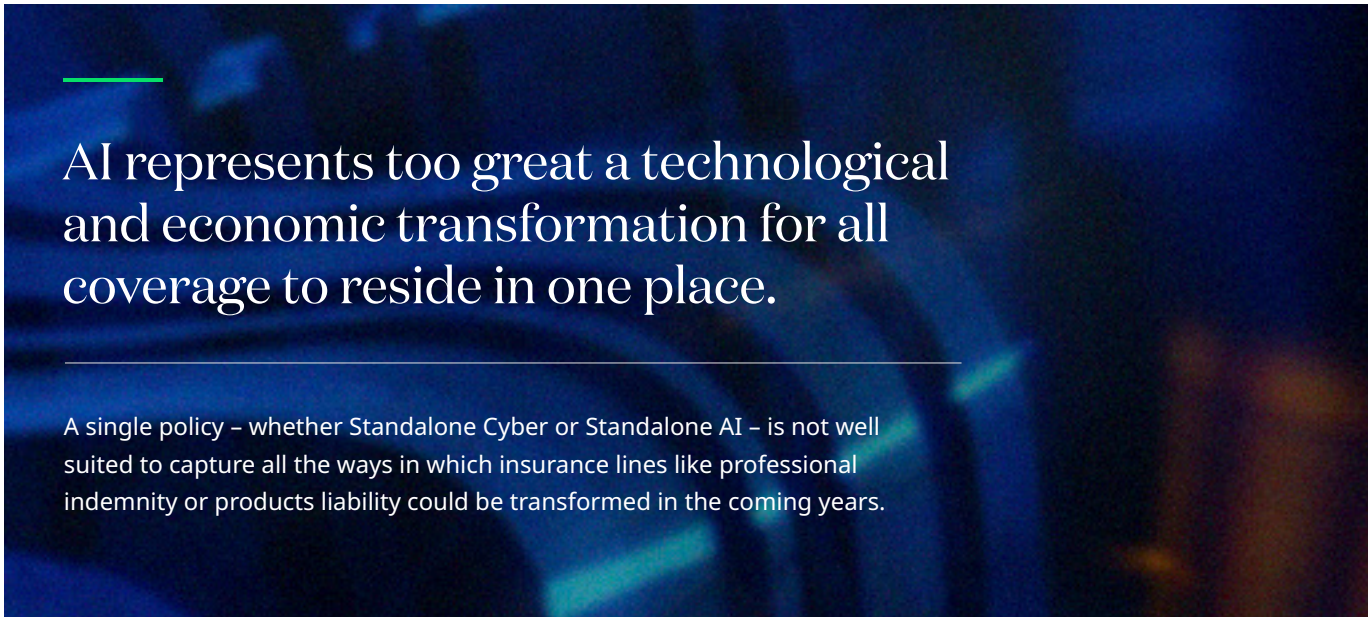
EVENT FAMILY	Cyber	Crime	Media	Tech E&O	E&O / PI	D&O	GL	Property	Products Liability	Auto Liability	Standalone AI
1 **Media, Regulatory & Protected Information** *AI creates legal exposure by its very design.*											
2 **Output Generation Failure** *The AI does the job wrong, or poorly (bad information)*											
3 **Automated Decision Failure** *The AI makes a bad decision with real-world consequences*					EXAMPLE						
4 **AI Physical Perils** *The AI, or human instructed AI, makes a bad decision with physical world consequences.*											
5 **Adversarial Exploitation** *AI creates new ways to exploit humans and machines*											
6 **Supply Chain & Dependency Failure** *Compromise , Revocation, or failure within the AI technology stack itself.*											

How you can use this table:

Mark each square based on how your organization's policy types today would respond to each event family. Consider the coverages and wording you have within those policy types, including affirmative cover, explicit exclusions, and silent or unclear cover.

Several such maps already exist, produced by brokers proposing the schema of coverage they wish buyers to adopt, and by MGAs advocating for the necessity of standalone AI policies. This is appropriate, as each player in the insurance value chain must play their part in helping the industry move forward into the age of AI risk. Our point is broader: each risk-bearing (re)insurer must determine for itself where risk should sit, and the "right" answers will be specific to each organization.

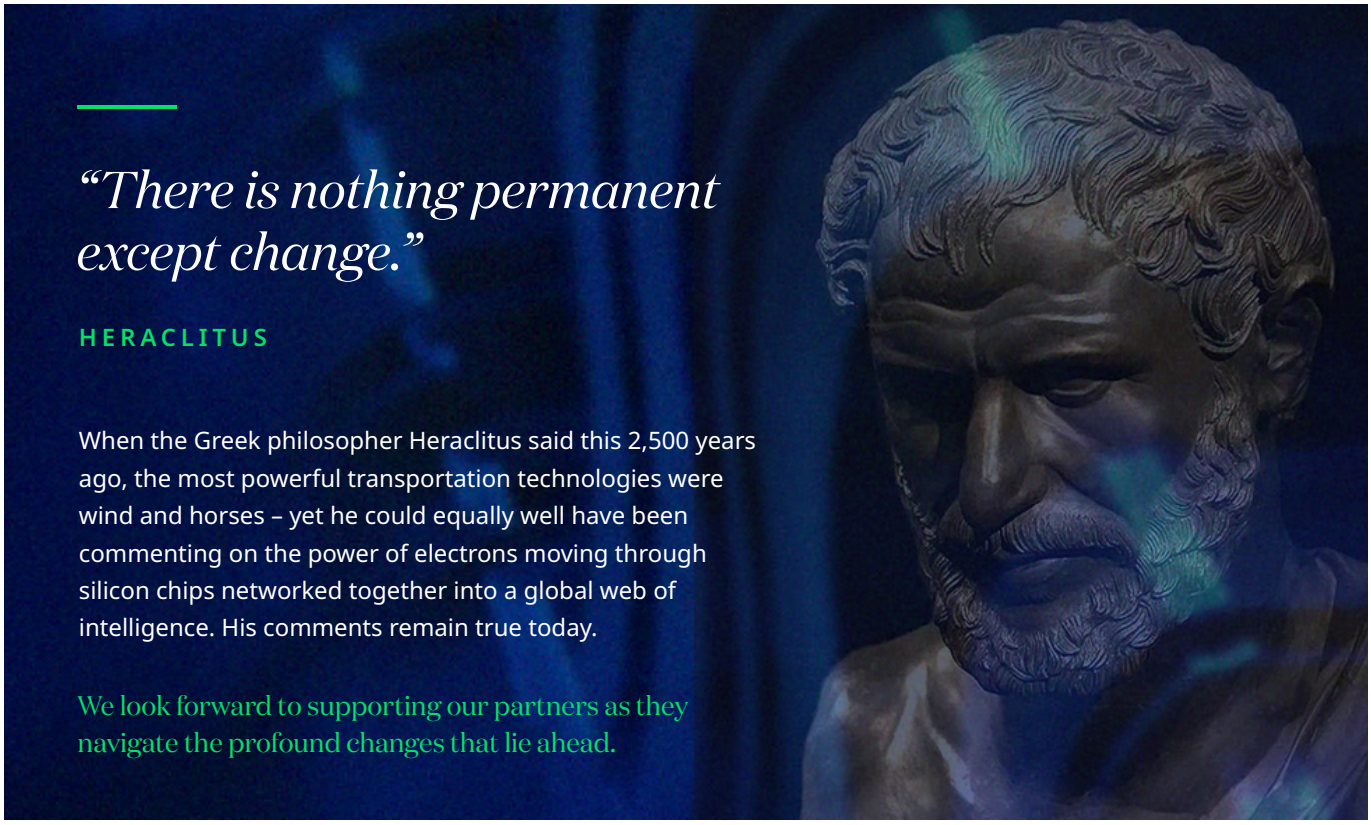
From our recent client discussions, we have observed that Cyber teams today are typically the ones fielding the AI-related questions for their organizations – across lines of business. This is an understandable starting point, but not a viable long-term solution.



Where We Go from Here

We are in the early stages of this journey – a journey that has an uncertain destination, but which promises to fundamentally transform society, business, and (re)insurance. Despite uncertainty around the destination, two things are clear for the (re)insurance industry. First, AI is already transforming risk across a host of insurance lines and, second, the trend will only accelerate. Carriers who recognize the transformation, act to mitigate the downside, and capitalize on the upside will serve a critical function for the 21st century economy – providing coverage, capital, and clarity in a rapidly changing landscape. Conversely, those who shrug off this transformation and try the 21st century equivalent of treating all transportation risk as homogeneous will be in for a rude surprise.

At CyberCube, we are excited to help the insurance industry – and society at large – seize the opportunity by providing a structured way to think, evaluate, and quantify emerging risk. We hope this framework can be the beginning of the journey and will provide a foundation on which others can build as they refine their own view of AI risk.



CyberCube