CyberCube

H1 2026 Global Threat Briefing

AI Risk Landscape: Implications for Cyber (Re)Insurance

Authors

Author:

William Altman
Director of Cyber Threat Intelligence Services

Editoral Manager:

Yvette Essen
Head of Communications & Market Engagement

Editorial Design:

Felix Paula
GTM Creative Lead at CyberCube

Introduction

Dear Readers,

CyberCube was founded with a clear purpose: to help the cyber (re)insurance industry quantify and manage cyber risk in financial terms. That purpose takes on renewed urgency as we enter a defining moment shaped by artificial intelligence (AI). AI is advancing rapidly, from frontier models to increasingly autonomous systems, and is now being embedded across enterprise infrastructure, security operations, and core business processes. As a result, enterprise cyber risk is evolving in parallel. The years ahead may prove to be an inflection point in how cyber risk behaves and how it must be viewed by (re)insurers.

AI is simultaneously amplifying both offensive and defensive capabilities, creating an environment in which defender automation competes directly with attacker automation. The central question is how could AI adoption shift cyber risk toward greater volatility and increased potential for correlated losses, particularly as shared dependencies on AI infrastructure and frontier models could introduce new pathways for systemic disruption?

These developments raise fundamental capital questions for (re)insurers. For example: Does AI increase loss frequency within established patterns? Does it heighten correlation across portfolios? Does it create new aggregation pathways that test current reinsurance structures? This Global Threat Briefing addresses these issues directly.

We examine how AI-enabled threats are developing, where systemic concentrations may be forming, and what broking, underwriting, and portfolio management actions are most relevant today. Our goal is not to speculate distant technological outcomes, but to ground the discussion in observable developments and measurable exposure.

At CyberCube, we remain committed to partnering with the global cyber (re)insurance industry through analytics, modeling, and threat intelligence that translate cyber risk complexity into actionable insight. As AI reshapes the digital economy, disciplined growth will depend on understanding both the opportunities and the structural risks that accompany it. We look forward to continuing that work together.

Chris Methven, CEO, CyberCube

Executive Summary

AI is reshaping cyber risk by increasing the speed, scale, and coordination of attacks while potentially introducing new dependency structures across the technology ecosystem. This report finds that AI is currently treated by (re)insurers as a force multiplier within existing loss frameworks rather than a distinct risk class. However, emerging developments - including compressed attack timelines, identity-layer propagation, and the rise of AI agents - are beginning to challenge this view.

At the same time, concentration across compute, cloud, and model providers creates the potential for aggregation risk, while expanding enterprise adoption embeds AI deeper into critical operations. As a result, the severity and correlation of losses are increasingly driven by recovery capability, identity security, and dependency management, suggesting that underwriting and modeling approaches may need to evolve.

CyberCube's analysis shows:

Cyber threats are changing with AI as:

- Threat actors are exploiting common security gaps more quickly, particularly identity misconfigurations and unpatched systems.
- The time from initial breach to operational disruption is shrinking, reducing opportunities for detection and containment.
- AI agents introduce a new privileged execution layer with direct access to data and systems, expanding the internal attack surface.

Cyber underwriting can adapt by:

- Focusing on identity security and patch latency as primary drivers of how attacks propagate and convert into loss.
- Evaluating recovery capability as a key determinant of business interruption (BI) severity, not just detection and containment.
- Underwriting directly responding to the governance of AI agents, including permissions, API scope control, logging, and segregation of duties.

Cyber catastrophe and aggregation risk modelers should consider:

- The global AI supply chain is a tightly coupled, interdependent ecosystem spanning semiconductor manufacturing, compute infrastructure, hyperscale cloud platforms, foundation models, and downstream applications. Cyber risk concentration could be structurally embedded at multiple critical layers of the AI supply chain, particularly where market dominance or technological centralization exists (e.g., GPUs, model training and compute, hyperscale cloud providers).
- If AI shifts from an augmentative capability to core operational infrastructure, including decision-making, automation, and control planes, failures at concentrated points in the AI supply chain could be more likely to translate into systemic, cross-sector loss events.

(Re)insurers Treat AI as a Force Multiplier Within Existing Loss Patterns

AI is viewed as increasing the speed, scale, and efficiency of known attack dynamics, while underlying loss triggers remain unchanged, leading to evolutionary adjustments within traditional risk frameworks rather than structural repricing or capital reallocation.

Accordingly, AI is embedded within existing pricing assumptions. Underwriting generally remains focused on assessing critical cyber risk controls with emerging scrutiny of AI-specific risks such as evaluating AI agent security. Meanwhile, catastrophe models treat AI as an accelerator of established scenarios. Accumulation efforts are beginning to incorporate AI dependency mapping, although this is not yet standardized, and policy language is evolving gradually, with some generative AI exclusions but continued reliance on traditional cyber loss triggers.

Adapting (re)insurance to AI-driven cyber risk requires focus across several key dimensions. These include understanding how AI technologies are developed, controlled, and deployed across enterprises; assessing how attackers are operationalizing AI in real-world campaigns and the resulting impact; and evaluating whether AI introduces new single points of failure (SPoF) or concentrated technology dependencies within the supply chain.

The Consequences of AI

Exhibit 1

Examples of AI in Real-World Cyber Attacks: 2024 – 2026 YTD (3/11/26)

Attack Type:	**AI Credential-Phishing Campaign**	**"Vibe Hacking" Ransomware** (aka. GTG 2002)
Victim's Country:	**United States, 2025**	**Global, 2025**
Brief Attack Description:	Microsoft identified a campaign that used AI-generated code to obfuscate payload behavior and evade defenses.	Anthropic observed criminals misusing Claude in an agentic way; AI helped plan, execute, and automate attacks on 17 orgs.

"Remote IT Worker" Infiltration	**AI-Enabled Scaling of Tradecraft**
United States, 2024 - 2025 **(and other countries)**	**Global, 2026**
Microsoft states North Korea's remote IT workers leveraged AI to improve operational scale and sophistication.	Amazon observed a "lower-skill" threat actor leveraging commercial AI to attack 600 FortiGate devices across 55 countries.

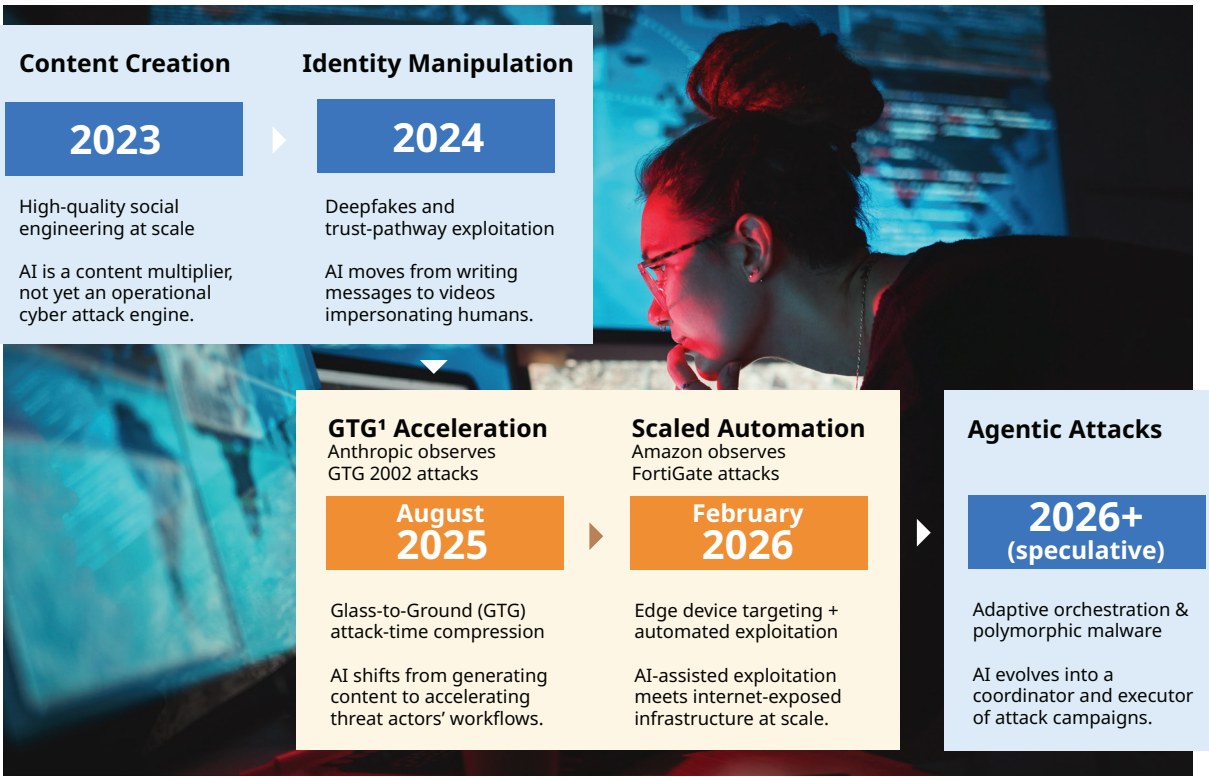
Deepfake Video-Meeting Fraud	**Deepfake Video-Meeting Fraud**
Hong Kong, 2024	**Singapore, 2025**
Scammers used AI deepfake video of the CFO and other colleagues in a video call to induce a fraudulent $26M transfer.	Victim instructed to perform a transaction of over $499k from the company account; the transaction was flagged and stopped.

Publicly disclosed examples of AI-enabled attacks show AI being used to enhance traditional tradecraft by improving impersonation realism, automating reconnaissance, scaling phishing, and accelerating operational workflows (see **Exhibit 1**). The shift is not theoretical; AI is already embedded in active campaigns across fraud, data breach, and ransomware.

Early use of AI focused on content generation, such as scaling social engineering, followed by identity manipulation through deepfakes and impersonation. However, two observed attack campaigns in August 2025 and February 2026 highlighted the evolution of AI-enabled cyber threats toward faster exploitation of common defense gaps (**Exhibit 2**). Attacks have progressed to "glass-to-ground" acceleration, where AI compresses the time between initial access and operational impact by coordinating and accelerating attacker workflows.

Exhibit 2

Evolutionary View of AI-Enabled Ransomware Attacks



Content Creation

2023

High-quality social engineering at scale

AI is a content multiplier, not yet an operational cyber attack engine.

Identity Manipulation

2024

Deepfakes and trust-pathway exploitation

AI moves from writing messages to videos impersonating humans.

GTG[1] Acceleration

Anthropic observes GTG 2002 attacks

August 2025

Glass-to-Ground (GTG) attack-time compression

AI shifts from generating content to accelerating threat actors' workflows.

Scaled Automation

Amazon observes FortiGate attacks

February 2026

Edge device targeting + automated exploitation

AI-assisted exploitation meets internet-exposed infrastructure at scale.

Agentic Attacks

2026+ (speculative)

Adaptive orchestration & polymorphic malware

AI evolves into a coordinator and executor of attack campaigns.

[1] In security contexts, a "GTG threat" can refer to scenarios where a model materially accelerates or lowers the barriers to executing harmful real-world actions, rather than merely providing abstract information. i.e., making it faster and easier to go from interacting with an AI system ("glass," the screen) to causing real-world effects ("ground"). Note, Anthropic has not publicly defined "GTG" as an acronym; it functions for Anthropic as a label for clustered threat activity.

By February 2026, AI-assisted exploitation was being applied at scale to internet-exposed infrastructure, enabling rapid targeting and compromise of vulnerable systems. This trajectory suggests a shift from AI as a supporting tool to an operational engine at-scale, with future developments pointing toward agentic attacks capable of adaptive orchestration and further autonomous execution across target environments.

The GTG 2002 attacks that occurred in 2025 demonstrate that even highly coordinated, AI-enabled ransomware operations still rely on familiar and preventable weaknesses. The campaign did not hinge on exotic zero-days or novel technical breakthroughs. It succeeded by exploiting exposed systems, unpatched VPNs, weak credential controls, excessive privileges, insufficient monitoring, and inadequate backups. Advanced orchestration amplified the attack, but basic hygiene gaps enabled it.

The February 2026 FortiGate breaches mark an inflection point in AI-driven threat evolution, combining AI-assisted operational acceleration with internet-scale targeting of widely deployed technology. While large-scale scanning and exploitation are not new, AI reduced human bottlenecks in prioritizing and coordinating attacks, compressing timelines, and enabling synchronized exploitation of unpatched systems. More than 600 devices across 55+ countries were reportedly compromised, with post-breach activity consistent with ransomware, although no confirmed cyber insurance losses have been reported.

Source(s): *Anthropic Threat Intelligence, https://www-cdn.anthropic.com/b2a76c6f6992465c09a6f2fce282f6c0cea8c200.pdf?, Amazon Threat Intelligence, https://aws.amazon.com/blogs/security/ai-augmented-threat-actor-accesses-fortigate-devices-at-scale/?utm*

AI is also compressing the cyberattack lifecycle by reducing the time between stages and enabling parallel execution across targets, as illustrated in **Exhibit 3**.

Exhibit 3

Evolution of the Cyber Attack Lifecycle: Before vs. After AI

Simplified Cyber Kill Chain	Before AI: Sequential, Human-Limited Attacks	After AI: Parallel, Automated, and Compressed Attacks
Step 1: Reconnaissance and Targeting	**Manual and time intensive target selection** *Before AI, reconnaissance was largely automated, but target prioritization, analysis, and execution remained manual, creating a bottleneck between discovery and exploitation.*	**AI scans for the presence of technologies and vulnerabilities, and prioritizes targets at scale**
Step 2: Initial Access	**Phishing or vulnerability exploitation**	**Continuously scans and prioritizes any exploitable weakness across the observable attack surface**
Step 3: Lateral Movement	**Step-by-step, hands-on, operator-driven**	**Automates identity-layer abuse, discovers and prioritizes attack paths based on success probability**
Step 4: Impact	**Slower, fewer simultaneous victims**	**AI enables parallel actions such as data theft and ransomware across multiple victim environments**

Before AI, attacks were sequential and human-limited, with manual targeting and step-by-step execution that provided more time for detection and containment. AI-enabled attacks now continuously scan for vulnerabilities, prioritize exploitable weaknesses, and automate lateral movement, enabling coordinated, parallel actions across multiple environments.

As a result, CrowdStrike found that attack timelines have compressed significantly, with average breakout times falling to 29 minutes in 2025 and, in some cases, occurring in seconds.[1]

This compression reduces the effectiveness of detection and containment, making disruption more likely once access is gained. Consequently, the severity and duration of business interruption losses are increasingly driven by recovery capability rather than the ability to detect and stop intrusions.

[1] *https://www.crowdstrike.com/explore/2026-global-threat-report?utm_medium=org*

AI Agents - Rise of A New Privileged Execution Layer

The market for AI agents is expected to grow rapidly as organizations increasingly deploy systems capable of autonomously executing tasks across data, applications, and workflows. As adoption expands, AI agents introduce a new privileged execution layer within enterprise environments, enabling direct interaction with critical systems. This shift creates new risk pathways, where misconfigurations, vulnerabilities, or misuse can lead to data exposure or operational disruption, making the secure deployment and governance of AI agents an emerging priority for cyber (re)insurers.

AI agents introduce new enterprise cyber risk pathways by enabling direct interaction with systems and data, creating opportunities for both manipulation and unintended actions. Agents can be influenced by malicious inputs, perform harmful actions while appearing to follow instructions, or operate with excessive permissions that allow errors or manipulation to trigger outages or data loss. Even without an external attacker, autonomous mistakes can lead to operational failures, while the automation and interconnected nature of agentic systems can amplify the impact of a single failure, allowing issues to propagate rapidly across enterprise environments.

Mitigating cyber risk may increasingly depend on how agents are permissioned, what controls govern their actions, and how agentic activity is monitored (**Exhibit 4**).

Exhibit 4

Three Key Control Areas for Cyber Underwriters to Assess AI Agent Risk

Permissions	Controls	Monitoring
Key Question to Ask: *Do AI agents operate with privileged access and are their permissions restricted using least-privilege?* **Key Actions to Take:** Underwriters can evaluate agent permissions, identity architecture, and access controls.	**Key Question to Ask:** *Are there approval gates before agents execute high-impact actions such as production changes or data deletion?* **Key Actions to Take:** Underwriters can assess whether controls exist that slow, detect, or contain machine-speed failures.	**Key Question to Ask:** *Which enterprise systems can AI agents interact with, and are those interactions logged and monitored?* **Key Actions to Take:** Underwriters can assess agent integrations, tool access, and monitoring controls.

First, underwriters can scrutinize permissions, determining whether AI agents operate with privileged access and whether their permissions adhere to least-privilege principles, with a clear understanding of identity architecture and access controls.

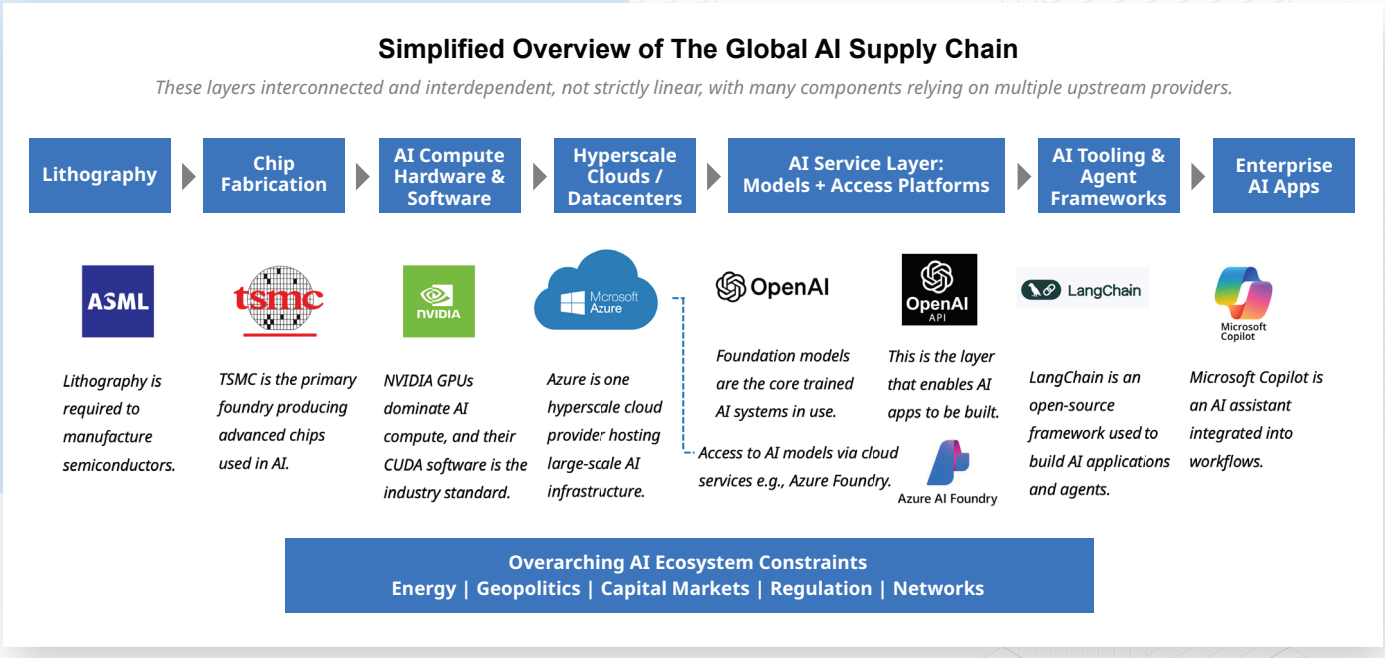
Second, underwriters can evaluate controls, specifically whether approval gates and safeguards exist before agents can execute high-impact actions such as production changes or data deletion, and whether these controls are capable of slowing, detecting, or containing failures that can propagate at machine speed.

Third, underwriters should assess monitoring, examining which enterprise systems AI agents can access, how those interactions are logged, and whether sufficient visibility and oversight mechanisms are in place.

AI Supply Chain Risk: The emerging risk of AI single-points-of-failure

(Re)insurers should also consider the supply chain risk presented by AI, which is built on interconnected technologies spanning hardware, clouds, and foundation models, to developer frameworks and finally AI applications.

Exhibit 5



The global AI supply chain is a tightly coupled, interdependent ecosystem spanning semiconductor manufacturing, compute infrastructure, hyperscale cloud platforms, foundation models, and downstream applications.

For cyber catastrophe and aggregation risk modelers, the key implication is that risk concentration could be structurally embedded at multiple critical layers, particularly where market dominance or technological centralization exists (e.g., lithography, advanced chip fabrication, GPU compute, foundation models, and hyperscale clouds). Disruption or compromise at any of these nodes, whether through cyber intrusion, operational failure, or geopolitical constraint, has the potential to propagate downstream across a wide set of insureds simultaneously, creating correlated loss scenarios rather than isolated events.

The extent to which AI-driven single points of failure should be explicitly modeled within cyber catastrophe scenarios that are not currently contemplated in existing frameworks depends on how deeply AI becomes embedded in business-critical functions across industries. If AI shifts from an augmentative capability to core operational infrastructure, including decision-making, automation, and control planes, failures at concentrated points in the AI supply chain could be more likely to translate into systemic, cross-sector loss events.

Tracking Key Developments as AI And Cyber Risk Evolve

As AI continues to evolve, several developments are likely to shape the future of cyber risk and (re)insurance. Improvements in model capabilities are enabling more complex and autonomous tasks, while growing enterprise adoption is embedding AI deeper into core business operations.

At the same time, the integration of AI into critical infrastructure and identity systems is expanding potential attack surfaces and creating new dependency risks. The use of AI to automate privilege escalation, lateral movement, and attack execution further increases the speed and scale of cyber incidents.

Together, these trends suggest a shift toward more interconnected, faster moving, and potentially systemic cyber risks, requiring (re)insurers to closely monitor where exposure is increasing and how attack dynamics are evolving.

AI has attracted an unprecedented level of attention, capital, and discourse in a relatively short period of time. Through our **Concierge** service, CyberCube's expert analysts and consultants will continue to monitor the ever-changing threat landscape and the implications for the cyber (re)insurance industry.

CyberCube