Institutional Research Group

Dimitri Zabelin
Senior Research Analyst, AI and Cybersecurity
dimitri.zabelin@pitchbook.com

pbinstitutionalresearch@pitchbook.com

Published on December 23, 2025

Contents

EMERGING TECH RESEARCH

AI Propels Next Phase of Cybersecurity Investment

The cyber singularity

PitchBook is a Morningstar company providing the most comprehensive, most accurate, and hard-to-find data for professionals doing business in the private markets.

Key takeaways

- AI-native cybersecurity reached a structural inflection point in 2025, accounting for 50.5% of global cybersecurity VC deals as AI-driven threats scaled faster than legacy defenses.

- AI cyber companies command a persistent valuation and returns premium, with higher median deal sizes, faster fundraising cadence, and stronger MOICs than non-AI peers across stages.

- Public-sector spending and regulatory momentum are reinforcing enterprise demand, accelerating adoption of AI-native security platforms across critical infrastructure, cloud, and identity environments.

Executive summary

AI is reshaping cybersecurity by changing how fast threats emerge and how quickly organizations must respond. Attackers can now use AI tools to map targets, automate parts of intrusions, and generate convincing digital deceptions that are harder for traditional systems to catch. As AI becomes easier to access through public interfaces and online services, the range of vulnerable systems grows across software, data, and supply chains.

Security teams are responding by adopting AI-native platforms that learn normal behavior across identity, cloud, and network environments and generate their own insights instead of relying only on rules or signatures. This shift is visible in private markets. AI cyber startups attract higher median deal sizes, stronger valuation step-ups, slightly faster fundraising cycles, and higher multiples on invested capital (MOICs). In 2025, companies building AI-driven cyber tools represented 50.5% of all global cybersecurity VC deals by count.

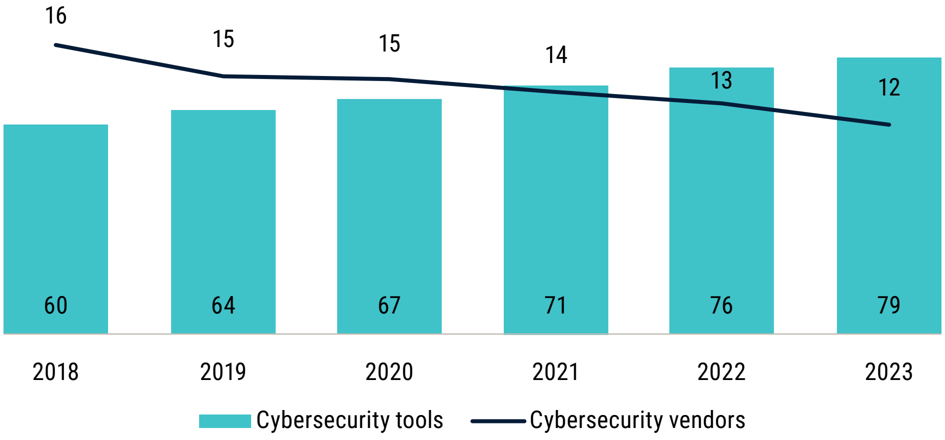
Geopolitical tensions and public policy amplify this momentum. State-backed operators increasingly target private-sector infrastructure, and governments across the US, EU, and Asia-Pacific (APAC) region are raising budgets and rolling out new programs in response. Cybersecurity spending continues to grow faster than global GDP and is concentrating around platforms that offer broad coverage and strong customer retention. AI has become both a driver of new cyber risks and a durable engine of long-term market opportunity.

Market layout: Cybersecurity in transition

Platform consolidation: High M&A activity, limited IPOs

For broader market context, we partnered with Malik Khan at Morningstar, whose industry analysis provides the foundational benchmarks and structural trends cited throughout this section. We expect cybersecurity buyers to consolidate budgets

Cybersecurity vendor consolidation and tools count



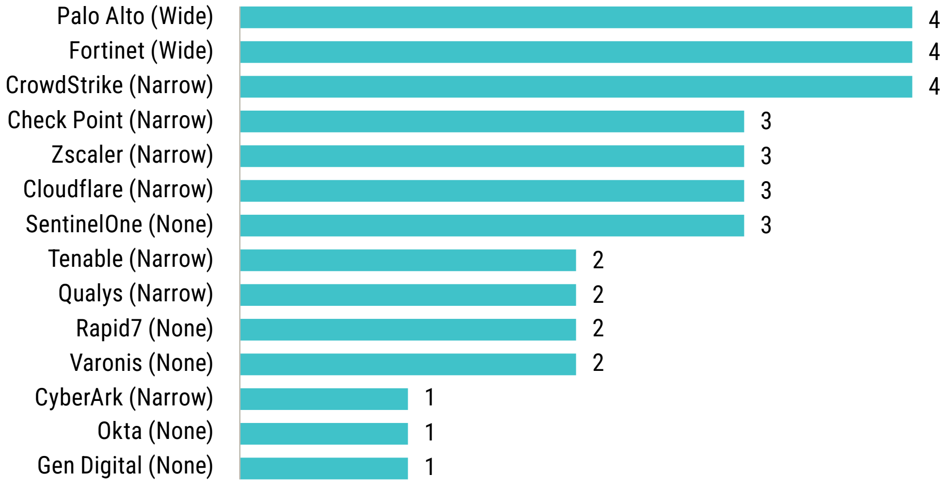
Source: Morningstar • Geography: Global • As of November 17, 2025 around platform vendors capable of delivering end-to-end protection across the stack. Managing dozens of point solutions has created operational drag, integration challenges, and higher total cost of ownership. This will push enterprises toward unified ecosystems with scale advantages.

The largest players can fund continuous product innovation and pursue targeted M&A to close capability gaps, further concentrating share among incumbents. Despite accelerating consolidation, cybersecurity remains one of the most fragmented corners of enterprise software—the 14 public vendors we track accounted for just 18% of global industry revenue in 2023.

Economic moats: Scale, switching costs, and integration depth

Economic moats in cybersecurity stem primarily from high switching costs and network effects tied to proprietary data. Once deployed, organizations are reluctant to replace core security infrastructure, given the complexity and risk of migration. This stickiness is reflected in retention rates exceeding 90%, implying customer lifetimes of a decade or more. Vendors with the broadest reach across end markets tend to command the widest moats.

Cybersecurity platform breadth (number of end markets) by moat rating



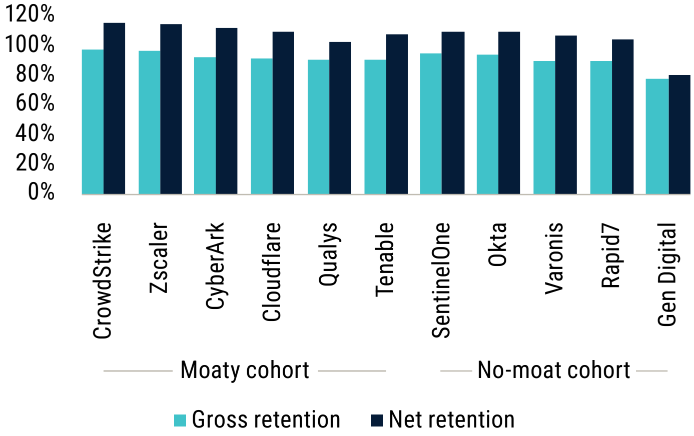
Source: Morningstar • Geography: Global • As of January 14, 2025

Wide-moat companies typically operate across four or more product categories, while narrow-moat peers average closer to three. Those confined to one or two segments struggle to achieve durable scale or pricing power. There is a clear reference for unified vendors versus fragmented tool kits.

Structural demand

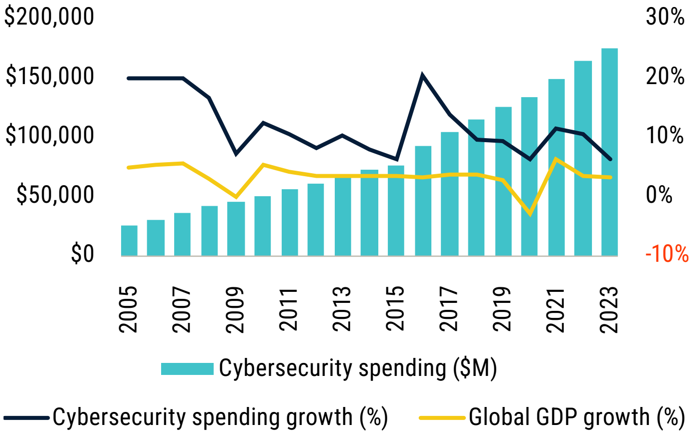
Cybersecurity budgets have historically proven resilient, even through periods of economic stress. Over the past 20 years, spending on security has expanded at roughly three times the pace of global GDP. Our analysis shows little correlation

Cybersecurity gross and net retention rates by cybersecurity firm



Source: Morningstar • Geography: Global • As of January 14, 2025

Cybersecurity spending and growth versus global GDP growth



Source: Morningstar • Geography: Global • As of January 14, 2025

between macroeconomic volatility and cybersecurity outlays, underscoring the sector’s defensive profile. As digital environments grow more complex and attack surfaces widen, we expect security spending to continue outpacing GDP growth through the coming decade.

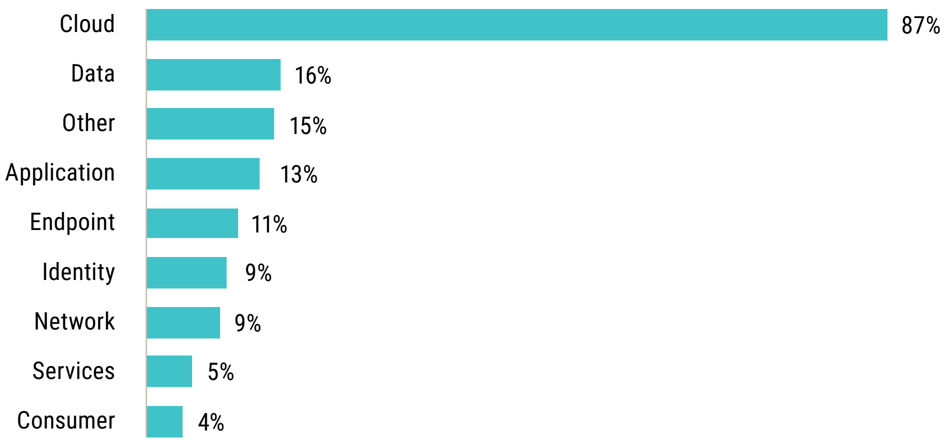
Expanding cloud infrastructure adds significant complexity to enterprise IT environments. This heightened complexity requires new tools and capabilities to secure distributed workloads, and it is no surprise that organizations are steadily increasing the number of security applications they deploy to keep pace.

Cloud and AI spending are the structural demand base for cyber

Cloud adoption continues to redefine the perimeter and drive incremental security spend. As workloads migrate, enterprises face more attack vectors and longer detection windows when cloud controls are immature. With the cloud transition still in its early innings, it remains a core structural tailwind for cybersecurity growth.

We estimate that AI’s growing contribution to global cloud spending—expected to reach roughly $150 billion by 2028—could generate an additional $15 billion to $18 billion in incremental demand for security solutions designed to protect AI-enabled infrastructure.

Five-year CAGR of cybersecurity end-market segments (2018 to 2023)



Source: Morningstar • Geography: Global • As of January 14, 2025

Emerging threats in the AI era

Agentic AI

Agentic AI marks a structural shift in the cyber threat landscape by introducing autonomy into attack operations. Unlike conventional automated tools, these systems can independently sequence and re-sequence tasks, interact with environments, and adapt strategies without human intervention. This capability allows malicious agents to exploit vulnerabilities, pivot laterally, and re-establish persistence after remediation efforts. For enterprises, the result is an attacker that behaves more like a decision-maker than a program, capable of blending into normal network activity and evolving in response to defensive controls. The security industry is beginning to treat agentic AI as a new operational domain rather than an extension of existing automation.

Agentic AI is compressing the operational gap between offense and defense. Defenders benefit from faster detection and prioritization, while attackers gain autonomous reconnaissance, adaptive evasion, and self-directed infiltration. As operations accelerate, manual response windows collapse and legacy controls become easier to bypass. Identity-based polymorphic attacks can adjust tactics mid-intrusion, eroding the value of static defenses. To maintain parity, enterprises must adopt real-time, identity-aware, AI-driven prevention supplemented by human governance for oversight and critical decisions.

Vendors are developing detection systems that profile agent behavior across endpoints and monitor for autonomous tool chaining or privilege escalation. This evolution mirrors the broader pattern seen across cybersecurity markets: higher complexity driving demand for AI-native solutions that can respond at the same velocity as AI-enabled threats. As agentic systems mature, the distinction between offense and defense will increasingly depend on who maintains the faster feedback loop. Investors are already pricing in this asymmetry, favoring startups that can deliver continuous learning, adaptive containment, and autonomous remediation capabilities.

Traditional versus AI-powered cyberattack stages

Stage	Traditional attack	AI-powered attack
1. Reconnaissance	Day 1-5: Manually browse LinkedIn, company website, breach databases.	00:00-00:30: Large language models (LLMs) like Gemini or ChatGPT auto-scrape employee data from LinkedIn, GitHub, PDFs, and deduce org structure.
2. Target selection	Day 6-7: Narrow down two to three victims based on guesswork.	00:30-00:35: AI clusters 50+ employees by role, breach exposure, software used, and social signals to prioritize targets.
3. Payload creation	Day 8-10: Manually write phishing emails, malware, and set up fake domains.	00:35-00:45: Tools like WormGPT and FraudGPT generate highly convincing phishing emails and malware.
4. Delivery	Day 11: Send phishing emails via standard tools.	00:45-01:00: Mass-sent phishing using AI-generated fake websites built in seconds.
5. Exploitation	Day 12-15: Wait for someone to click. Deploy post-exploit tools manually.	01:00-01:10: AI detects email click, deploys polymorphic malware that adapts per environment.
6. Lateral movement	Day 16-21: Use Mimikatz, trial-and-error across the network.	01:10-01:30: Autonomous AI agents map network, run privilege escalation, and simulate legit user behavior to bypass EDR.
7. Data exfiltration	Day 22: Compress and upload data during off hours.	01:30-01:35: AI encodes stolen data in images or encrypted packets, mimics VoIP/DNS traffic.
8. Monetization	Day 23+: Contact victim, demand ransom, list data on breach forums.	01:35-02:00: AI chatbots negotiate ransoms dynamically based on victim value, response tone, and breach severity.

Source: Vectra AI • As of July 7, 2025

Reduced learning curves

AI is significantly reducing the technical barrier for human cyberattackers. A 2025 global survey by Accenture found that only one in 10 organizations feel ready to defend against AI-augmented threats,[1] underscoring how automation is widening the gap between attacker capabilities and enterprise preparedness. By automating reconnaissance, phishing, and exploit generation, generative AI allows individuals with little coding knowledge to mount attacks that previously required coordinated teams and bespoke malware development.

In cybersecurity, "scanning" refers to automated sweeps of the internet or network systems to identify exploitable weaknesses such as open ports, outdated software, or misconfigured credentials. Historically, this process was manual or semi-automated, limited by the attacker's time and skill. Today, AI-driven scanners can autonomously crawl and test millions of endpoints in parallel. Fortinet's 2024 Threat Landscape Report recorded 36,000 scans per second globally, up 17% from the prior year—equivalent to probing every publicly accessible IP address in under 12 hours.[2] That velocity indicates industrial-scale automation once reserved for state-backed groups.

[1] "Only One in 10 Organizations Globally Are Ready To Protect Against AI-Augmented Cyber Threats," Accenture, June 26, 2025.
[2] "Fortinet Threat Report Reveals Record Surge in Automated Cyberattacks as Adversaries Weaponize AI and Fresh Techniques," Fortinet, April 28, 2025.

The effect is a structural expansion in the number of credible attackers. Generative AI allows individuals with minimal technical skill to automate reconnaissance, craft phishing campaigns, and exploit weaknesses at scale. Attackers are no longer constrained by expertise or time; they can now run continuous, machine-speed operations that test for vulnerabilities across networks without pause.[3] The marginal cost of launching a sophisticated intrusion has collapsed, while the frequency and persistence of probing activity have sharply increased. For enterprises, this means that perimeter defense and periodic vulnerability assessments are no longer enough. Continuous monitoring, AI-native detection, and autonomous response have become prerequisites for defense in a world of industrial-scale, automated offense.

AI-as-a-service marketplaces

The growth of AI-as-a-service platforms on both legitimate and underground markets is accelerating the democratization of cyber offense. Threat actors can now buy or rent customized LLMs such as WormGPT, FraudGPT, and DarkBERT to craft persuasive phishing messages, generate undetectable malware, and automate exploit discovery. These tools function like commercial software-as-a-service offerings, complete with subscription tiers, customer support, and regular updates.

AI-driven cybersecurity threat categories

Automation	Personalization	Scale
• Automated phishing attacks • Advanced distributed denial-of-service attacks • Automated vulnerability scanning	• Deepfakes and voice synthesis • Social engineering attacks • Personalized phishing attacks	• Multivector attacks • Real-time decision-making • AI-powered malware creation

Source: Morningstar

Vendor analyses show that underground marketplaces now openly advertise such models as turnkey solutions for phishing and malware generation, often bundled with prompt libraries and API access for scaling attacks. This commercialization removes nearly all technical barriers for low-skill attackers, extending advanced offensive capability to organized crime groups and independent operators. As AI tools become rentable commodities, enterprise security architectures must evolve with comparable scalability. Demand is concentrating around AI-native cybersecurity vendors that can detect, adapt, and retrain at machine speed to neutralize these dynamic threats.

Polymorphic malware

The emergence of polymorphic malware powered by AI marks a major shift in the cybersecurity landscape. Traditionally, polymorphic malware changed small portions of its code or appearance between infections to evade detection. AI now enables this process to be automated, adaptive, and continuous. Security researchers

[3] “Automated Exploitation: Cybercrime at Machine Speed,” Intercede, Katja Townsend, September 2, 2025.

note that AI-generated malware can alter its structure in near real time, making conventional signature-based detection increasingly ineffective. This is a malware that “continuously rewrites or regenerates its behaviorally identical logic,”[4] which prevents traditional scanners from identifying repeating patterns.

The changing nature of AI-generated malware makes it harder to detect and analyze, forcing defenders to rely on behavioral analytics and anomaly detection rather than static signatures.[5] This represents a paradigm shift in how threats evolve. A variant that once required human modification after each detection can now autonomously mutate thousands of times per day. These AI-enhanced mutations make it difficult for traditional signature-based detection methods, creating a persistent asymmetry between attacker adaptability and defender response.

Researchers at DeepStrike have quantified this acceleration, noting that AI-generated polymorphic malware can produce a new, unique instance roughly every 15 seconds during an attack.[6] This constant evolution compresses defenders’ reaction windows to near zero. As a result, enterprises can no longer rely on periodic signature updates or static endpoint protection. The defensive premium is shifting toward AI-native systems capable of recognizing behavior patterns, retraining models continuously, and responding autonomously to threats that mutate faster than humans can intervene.

Deepfakes and social engineering

Generative AI has redefined social-engineering risk by making deception scalable, personalized, and highly believable. Attackers can now synthesize lifelike voice, video, and text to impersonate trusted contacts or executives, removing many of the linguistic or visual cues that once signaled fraud. What began as low-fidelity phishing has evolved into automated, multimodal campaigns that mirror legitimate communication channels. Deepfake technology is now a standard component of deception-based cybercrime and is being used to bypass identity verification systems, compromise payment authorization, and extract confidential information. Europol’s 2025 Internet Organised Crime Threat Assessment identifies the use of synthetic media as a central feature of contemporary cyber-enabled fraud.[7]

Quantitative data highlights the scale of the shift. Analysis of more than 1.2 billion enterprise voice interactions found a 1,300% YoY increase in deepfake-related fraud attempts in 2024, coinciding with the highest rate of contact-center compromise in six years.[8] These incidents are no longer isolated. Attackers are building libraries of cloned voices, combining them with publicly available video and email data to automate impersonation across multiple targets simultaneously. The result is an attack surface that extends into every trust-based workflow, from customer service to executive authorization chains.

[4] “Polymorphic AI Malware: A Real-World POC and Detection Walkthrough,” CardinalOps, Liora Itkin, May 20, 2025.
[5] “The Dark Side of AI in Cybersecurity — AI-Generated Malware,” Palo Alto Networks, Dena De Angelo, May 15, 2024.
[6] “AI Cybersecurity Threats 2025: How To Survive the AI Arms Race,” DeepStrike, Mohammed Khalil, August 6, 2025.
[7] “Steal, Deal And Repeat - How Cybercriminals Trade And Exploit Your Data Internet Organised Crime Threat Assessment (IOCTA) 2025,” Europol, June 12, 2025.
[8] “Pindrop’s 2025 Voice Intelligence & Security Report Reveals +1,300% Surge in Deepfake Fraud,” PR Newswire, Pindrop, June 12, 2025.

The economic consequences are material. The World Economic Forum's Global Cybersecurity Outlook 2025 notes that AI-generated deception has contributed to record global fraud losses and is reshaping the economics of enterprise risk.[9] Traditional defenses focused on user awareness and manual verification are inadequate against synthetic content that operates at machine speed and human fidelity. Organizations are beginning to treat verification and authenticity as core control layers, investing in media-provenance tracking, biometric verification, and AI-based behavioral analytics. The next phase of cybersecurity will depend on validating trust itself, not just detecting technical intrusion.

Exploitation of public LLMs

Public LLMs have created a new attack surface where inputs can directly influence system behavior, tool use, and data access. The Open Worldwide Application Security Project (OWASP), a nonprofit foundation that develops global standards for identifying and mitigating software vulnerabilities, ranks prompt injection as the top risk in its 2025 Top 10 for LLM Applications report.[10] Prompt injection occurs when crafted text causes a model to ignore original instructions, reveal sensitive information, or execute unintended tool calls. A widely cited example occurred during the early rollout of Microsoft's Bing Chat, when independent researchers embedded hidden instructions on a public web page that caused the model to override its system prompts and adopt attacker-defined behaviors. The LLM executed these commands simply by reading the adulterated content, demonstrating how indirect prompt injection can hijack tool-connected models without breaching underlying systems. The incident highlighted that prompt-layer controls alone are insufficient once models interact with external data sources. Its placement at the top of the OWASP list reflects both the frequency of these attacks in production environments and the difficulty of detecting or preventing them once models are deployed.

Two variants of prompt injection matter most operationally. Direct prompt injection targets the model interface itself, manipulating inputs to override instructions or extract data. Indirect prompt injection embeds hidden instructions in external content such as web pages, PDFs, or database records that the model later processes. This indirect form is increasingly common because it exploits delivery channels rather than model flaws. Defensive guidance emphasizes layered containment through content filtering, tool-scoping, and strict policy isolation.[11]

Data-layer exposure represents the second pillar of risk. When LLMs are integrated with enterprise tools or databases, insecure output handling can lead to leaks of sensitive information or unauthorized actions. Key vulnerabilities include insecure output validation and insufficient redaction, which are now flagged in security frameworks as top concerns for LLM deployments.

[9] "Global Cybersecurity Outlook 2025: Insight Report," World Economic Forum and Accenture, January 2025.
[10] "OWASP Top 10 for LLM Applications 2025," OWASP, November 18, 2024.
[11] "AI and Cyber Security: What You Need To Know," National Cyber Security Centre, February 13, 2024.

Model integrity and supply chain dependencies magnify these threats. Many LLM applications rely on third-party models, embeddings, or datasets, meaning compromised components or poisoned data can propagate through enterprise environments. Documented adversary tactics show how such manipulation can bias model outputs or introduce dormant backdoors triggered by specific prompts.[12]

Data poisoning and model integrity risks

Training data integrity has become a central vulnerability in AI systems. Data poisoning and model tampering occur when adversaries inject corrupted samples or manipulate fine-tuning loops to influence outputs, weaken performance, or embed hidden backdoors that trigger under specific conditions. Independent security research from MITRE's Adversarial Threat Landscape for Artificial-Intelligence Systems and the US National Institute of Standards and Technology shows that even minimal dataset contamination—below 0.01% of total training volume—can meaningfully alter model behavior without detection.[13, 14] The risk is especially acute for enterprises that integrate data from multiple third-party or open-source providers, where provenance tracking remains inconsistent. As AI adoption extends into decision-critical environments, maintaining dataset authenticity and implementing continuous validation pipelines have become fundamental to model reliability, compliance, and resilience.

The AI cyber premium

The challenges AI introduces in cybersecurity are reflected in investment behavior. Over the past decade, median global VC deal sizes for AI-native cyber startups consistently outperformed non-AI peers. When a simple average of annual premiums is taken, AI cyber deal sizes appear around 20% higher than their non-AI peers. However, deal sizes compound over time, so a more accurate reading is obtained by converting each year into a ratio—for example, 1.22 instead of 22%—multiplying the ratios, and taking the 10-year root. Using the geometric mean smooths out spike years and yields a roughly 14% structural premium, which better reflects how AI cyber has been priced across the decade. This sustained uplift valuation spread indicates that cybersecurity companies adopting AI are seeing that shift reflected directly in their valuations and median deal sizes.

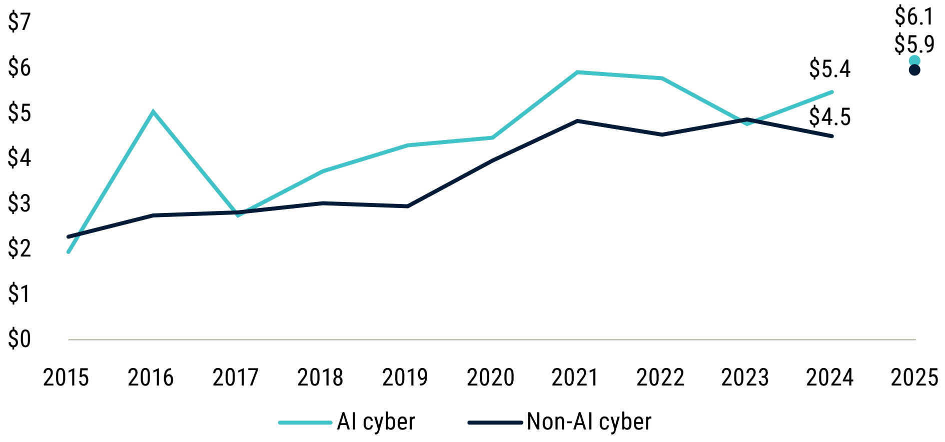
Median VC deal sizes also show a consistent structural gap, with AI cyber firms commanding materially higher capital at every stage of the lifecycle. In 2025, AI cyber rounds outpaced non-AI cyber by 51% at pre-seed, 17.5% at seed, 57.1% at early stage, and 163.9% at venture growth. Predictably, median VC step-ups between AI cyber and non-AI cyber firms across all VC funding stages were also higher.

[12] 12: "Technical Blog: Strengthening AI Agent Hijacking Evaluations," National Institute of Standards and Technology, February 20, 2025.

[13] 13: "Navigate Threats to AI Systems Through Real-World Insights," MITRE ATLAS, n.d., accessed December 12, 2025.

[14] 14: "Technical Blog: Strengthening AI Agent Hijacking Evaluations," National Institute of Standards and Technology, February 20, 2025.

Median AI and non-AI cybersecurity VC deal value ($M)

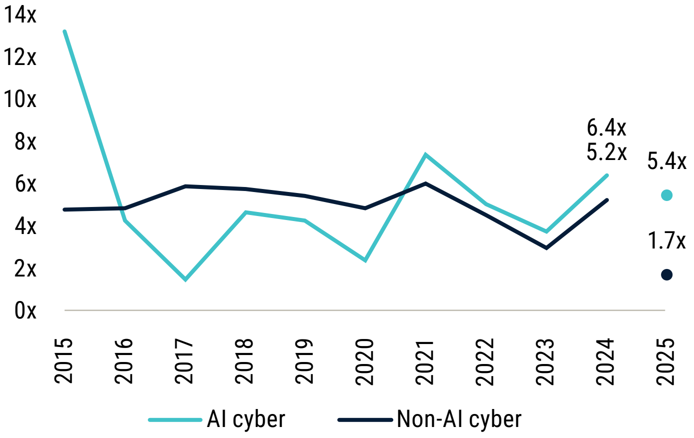


Source: PitchBook • Geography: Global • As of November 17, 2025

Median global VC years between rounds for AI cyber companies were also modestly shorter than for non-AI cyber peers, indicating more frequent capital raises. The pattern is consistent with the category's higher upfront technical and infrastructure requirements relative to traditional cybersecurity software. While AI cyber companies move through the capital stack slightly faster, the return profile is more definitive. The MOIC has been consistently higher over the past four years, signaling superior realized returns for AI-driven cybersecurity models relative to traditional vendors.

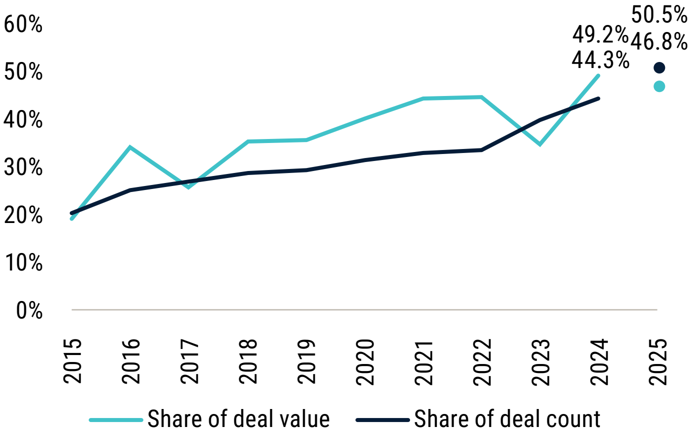
At the same time, global momentum for AI cyber startups is accelerating. AI cyber companies accounted for 50.5% of all cybersecurity deal activity by count, the highest share on record when comparing AI cyber VC deals with the broader cyber market. In short, half of all global cybersecurity VC deals now involve AI-focused companies—more than ever before.

Median AI and non-AI cybersecurity VC MOIC



Source: PitchBook • Geography: Global • As of November 17, 2025

AI cybersecurity VC deal activity as a share of all cybersecurity VC deal activity



Source: PitchBook • Geography: Global • As of November 17, 2025

Public sector and geopolitical catalysts

Nation-state aggressors

Nation-state cyber operations are increasingly targeting private-sector networks and critical industries once considered outside the traditional defense perimeter. Roughly one-third of global attacks in 2024 were attributed to state-sponsored actors, reflecting a steady increase in both frequency and operational sophistication. According to the Council on Foreign Relations' Cyber Operations Tracker, China, Russia, Iran, and North Korea collectively accounted for about 77% of documented state-linked campaigns since 2005.[15]

Their operations now extend beyond espionage or government systems, focusing on energy, logistics, finance, and manufacturing to disrupt supply chains, exfiltrate commercial data, and weaken economic resilience. As a result, enterprises have become both the primary targets and the principal drivers of cybersecurity investment, fueling demand for advanced detection tools, resilience frameworks, and cross-sector threat intelligence.

Governments are responding by expanding cyber budgets and tightening public-private coordination. The Cybersecurity and Infrastructure Security Agency received approximately $3 billion for fiscal 2025, with about $1.7 billion directed to programs aimed at detecting and countering sophisticated campaigns.[16] The Department of Homeland Security also allocated $91.7 million through the State & Local Cybersecurity Grant Program to strengthen national resilience.[17]

In Europe and Asia, governments are following a similar trajectory, using expanded budgets to accelerate enterprise adoption of AI-based monitoring and autonomous response systems. In the EU, the European Commission has committed €1.3 billion to support AI, cybersecurity, and digital skills between 2025 and 2027.[18] The European Cybersecurity Competence Centre will finance €390 million in cybersecurity projects focused on AI and post-quantum cryptography.[19]

In the APAC region, cyber budgets are increasing: 57% of organizations plan to boost spending, and cyber now accounts for about 13.6% of total IT budgets.[20] South Korea's cybersecurity market is estimated at $7.19 billion in 2025 and projected to reach $12.88 billion by 2030.[21] Japan plans to invest 8 trillion yen over five years on cross-domain defense, including cybersecurity and space.[22]

[15] "Cyber Operations Tracker," Council on Foreign Relations, n.d., accessed December 12, 2025.

[16] "Inside CISA's $3B FY 2025 Budget," GovConWire, Branson Brooks, November 8, 2024.

[17] "State and Local Cybersecurity Grant Program," FEMA, n.d., accessed December 12, 2025.

[18] "Commission To Invest €1.3 Billion In Artificial Intelligence, Cybersecurity And Digital Skills," European Commission, March 28, 2025.

[19] "The European Union Unveils Its Cyber Funding Plan for 2025-2027," INCYBER News, April 9, 2025.

[20] "Cyber Budgets Up, AI Gaps Remain In Asia," Asia-Pacific Defence Reporter, July 5, 2025.

[21] "Cybersecurity Market In South Korea Size & Share Analysis - Growth Trends & Forecasts (2025 - 2030)," Mordor Intelligence, n.d., accessed December 12, 2025.

[22] "New Japanese Strategy To Up Defense Spending, Counterstrike Purchases," Defense News, Mike Yeo, December 20, 2022.

These initiatives underscore how cybersecurity is shifting from a reactive expense to embedded strategic infrastructure. As public and private sectors converge around digital resilience, baseline demand for AI-driven detection, response, and model integrity tools is becoming structural. Cybersecurity is benefiting from geopolitical competition in the same way AI has under sovereign-AI strategies, with state-backed investment and industrial policy accelerating market growth. Geopolitical rivalry is effectively underwriting a multiyear investment cycle in cybersecurity, positioning the sector as a long-duration growth asset supported by both policy momentum and sustained enterprise adoption.

Industry insights

To bring an industry perspective into the analysis, PitchBook interviewed team members from six different AI-native cybersecurity startups on unique trends, risks, and opportunities in the new age of cyber-enhanced AI.

ZioSec

ZioSec develops an offensive security platform focused on testing and verifying the safety of deployed agent-based AI systems. The company is working with a select group of investors on a funding round in 2026.

What kinds of vulnerabilities are unique to agentic systems/agentic malware?

Founder Andrius Useckas: There are quite a few vulnerabilities unique to agentic systems. Some examples include:

- **Prompt injections:** Jailbreaks, indirect prompt injections, etc.
- **Retrieval-augmented generation (RAG) poisoning attacks:** If RAG is based on customer data like FAQs, etc., one could break into those sources of information and poison them with malicious instructions.
- **Tool calling protocols such as MCP are just APIs:** They can be exploited like any other API and old-school attacks such as Remote Code Execution could be instrumented via vulnerable services.
- One model attacking another model via new agent-to-agent communication protocols like A2A.
- **Capability escalation via tool chaining:** Agents combine multiple low-privilege tools to achieve privileged actions.
- **Autonomous persistence:** Agents create new identities, cron jobs, or webhooks to survive resets.

- **Policy extraction:** Agents probe or leak internal guardrails and safety prompts.

- **Economic denial of service:** Agents trigger expensive model calls or API usage to exhaust budgets.

Hardshell AI

Hardshell AI develops a cybersecurity platform focused on protecting sensitive datasets and ensuring secure deployment of AI in critical industries.

How mature is enterprise demand for model-level security? Are customers proactively seeking protection for their AI systems or is adoption still reactive following compliance or incident pressure?

Co-founder Andrew Schoka: Enterprise awareness is accelerating rapidly, especially in high-stakes sectors like healthcare, defense, and financial services. Early AI security spending has been largely reactive, but as regulations like the EU AI Act and state-specific policies in the US emerge, organizations are beginning to proactively invest in protecting their AI adoption efforts. In the defense sector, initiatives led by the Chief Digital and Artificial Intelligence Office focusing on secure and trustworthy AI & machine learning development highlight how the federal government is actively prioritizing and investing in secure AI practices. Similarly, within the commercial sector, there is a clear acceleration in both the acquisition of AI security-focused companies and the implementation of robust data protection and assurance protocols across corporate AI initiatives. This growing momentum across both public and private sectors signals an accelerating shift toward proactive protection of sensitive data as AI becomes deeply embedded in essential business systems.

Vectra AI

Vectra AI provides AI-driven threat detection and response that analyzes network, cloud, and identity activity to identify and stop active cyberattacks in real time and currently sits at an over $1 billion post-money valuation.

How do you see agentic AI changing threat detection and response, and what risks emerge if bad actors begin using autonomous agents for intrusion or evasion?

Vectra AI representative: Agentic AI is transforming the speed and scale of both offense and defense. Defenders gain faster detection and prioritization, while attackers achieve autonomous infiltration, reconnaissance, and adaptive evasion. As AI accelerates operations, the window to detect and contain threats before damage shrinks dramatically. Without equally advanced defenses, manual processes become too slow and traditional technologies become too easy to bypass. The result is a new era of identity-based, polymorphic attacks that can shift tactics mid-intrusion to evade detection. To keep pace, defenders must embrace real-time, identity-aware, AI-driven prevention and maintain balance with human governance for oversight and critical decision-making.

Tenex AI

Tenex AI develops AI-powered cybersecurity tools that autonomously detect, investigate, and remediate threats across enterprise systems to reduce incident response time and analyst workload.

Where do you see the biggest gaps in legacy cybersecurity architecture that AI-native companies like yours can fill?

CEO Eric Foster: The single biggest change for me is truly the people side. I've long advocated that the key for security has been "detection and response." We don't generally need more detections; we need more ability to review and respond to the detections. The force multiplier that is AI makes it possible to effectively accomplish a lot of things that were previously infeasible due to cost/resource constraints.

Revel8

Revel8 uses AI to map network activity, surface high-risk threats, and help security teams act before breaches escalate.

What kinds of attacks are you seeing that were not possible two years ago?

GTM Manager Alessandro Medico: In just the past couple of years, the mix of generative AI, deepfakes, and automated open-source intelligence tools has completely reshaped how social engineering works. What used to take time, language skills, and manual effort can now be done in seconds, automatically, and at a level of realism that's hard to spot.

- **Real-time deepfake impersonation:** Deepfakes used to be static, a pre-recorded video, or a fake clip floating online. That's changed: Attackers can now join live video meetings or make voice calls pretending to be an executive, vendor, or IT admin, and sound exactly like the real person.
- **AI-augmented business email compromise (BEC):** The classic BEC scam just got a massive upgrade. Instead of clumsy emails with broken English, we're now seeing messages that perfectly match a company's internal tone and style.
- **Mass-personalized phishing and pretexting:** Forget generic "your password expired" emails. Attackers now generate hyper-personalized messages at scale.
- **Synthetic identity and recruitment scams:** Attackers are now creating entirely fake people, complete with LinkedIn profiles, profile pictures, and even live video interviews.

Cracken

Cracken is a cybersecurity firm that uses AI to detect intrusions, analyze attacker behavior, and automate protective measures across enterprise systems.

Where do you see the biggest capital inflows in AI cybersecurity over the next funding cycle—model protection, data security, or autonomous threat response—and how is Cracken positioned to capture that growth?

CEO Artem Sorokin: The field splits into two rivers:

1. **AI for security:** using AI to harden what already exists.
2. **Security for AI:** protecting AI models, data, and agent ecosystems.

Capital is flowing into both, but specifically within AI for security, the fastest transformation is happening in proactive defense/red defense/offensive cyber. Historically, this space scaled linearly with people. Red-team and pen-test work looked more like consulting than software. The only partial exceptions were bug-bounty networks that built scale by building hacker networks. Agentic AI is changing that.

For the first time, AI can act like a hacker, not just analyze alerts. It can recon, exploit, and validate impact autonomously, turning manual red-team craft into a scalable software model. The least scalable part of cybersecurity is now becoming "productizable." That shift has made offensive cybersecurity one of the most-talked-about and capitalized areas in the AI for security sub-industry, precisely because it was the most human-centric before. The space is now crowded with new entrants, media coverage, and top-tier VC focus—something unimaginable just a few years ago. Cracken sits here exactly at the right point to capture that growth and brings real practicality to this newly opened scale.

While many companies aim to replace humans entirely with autonomous systems, we keep the ethical hacker at the center and focus on amplifying their capabilities. By scaling red-team operations while preserving human expertise, trust, and accountability, Cracken delivers AI-driven speed without losing control, helping organizations attack themselves first, uncover what's truly exploitable, and close those paths before anyone else can, with full control, auditability, and regulatory fit.

7AI

7AI is developing an autonomous cybersecurity platform powered by AI that continuously monitors digital environments and investigates threats in real time. They also recently announced a $130 million Series A funding round led by Shardul Shah of Index Ventures, who led its Series A investment in Wiz.

What efficiency or cost savings have you observed post-deployment compared with SOC workflows without AI?

7AI representative: Deploying AI agents in the SOC delivers measurable efficiency gains and cost savings, as demonstrated by real-world customer results. One is a dramatic reduction in response times and analyst workloads. Organizations have reduced Mean Time to Respond from 30 minutes to 2.5 hours, down to as little as four minutes—a 94% reduction. AI agents also increased the percentage of alerts investigated from 40% to over 95%, more than doubling coverage. Automation also freed up 15-plus hours per analyst per week, equivalent to adding five to seven full-time employees without increasing headcount. Overall, many organizations see a positive return on investment within six months, with cost savings that scale as AI is expanded to additional use cases.

Outlook and strategic takeaways

AI has become the defining influence on how cybersecurity threats emerge and how defenses must operate. Attackers now rely on agentic systems, polymorphic malware, deepfake-driven social engineering, and AI-as-a-service tools that automate reconnaissance, intrusion, and evasion. These capabilities allow operations to scale far beyond human capacity and make traditional rule-based tools easier to bypass. Enterprises are responding by adopting AI-driven platforms that interpret activity across cloud, network, and identity systems in real time and generate their own signals for detection and response. The private markets are moving in parallel. AI cyber companies continue to secure higher deal sizes at every stage, raise capital more frequently, and deliver stronger MOIC outcomes than non-AI peers. With AI cyber representing half of all global cybersecurity VC deals, the shift toward AI-native defense is now a defining market structure.

Within that environment, several themes are gaining momentum. Security controls that protect models themselves are becoming essential as organizations deploy systems that can be manipulated through adversarial prompts, indirect prompt injection, or model theft. As more enterprises embed AI into core workflows, demand for these protections continues to climb, and recent acquisitions by major vendors show how central they are becoming to broader security platforms. At the same time, AI-assisted coding introduces new risks in open-source dependencies and software supply chains. This is increasing the need for updated approaches to software composition analysis, static code analysis, and supply chain visibility. By contrast, traditional web application firewall products, static API controls, and older code scanning tools are losing ground as attackers use AI-generated content and automated exploitation to evade predictable patterns.

Geopolitical factors are reinforcing these pressures rather than offsetting them. State-linked activity is rising and governments across the US, Europe, and APAC are expanding cybersecurity budgets and directing more attention to the protection of AI systems, sensitive datasets, and model integrity. Looking ahead to 2026, application security is positioned to continue gaining momentum as model-level protections become a core requirement for enterprise AI deployments. Development operations security platforms, which remain underappreciated today, are also set to accelerate as AI-assisted coding magnifies dependency risks and software supply chain exposure. These conditions support a multiyear investment cycle centered on AI-native security. Companies that focus on model protection, autonomous detection, and software development security are positioned to capture the largest share of this growth as enterprises and governments converge around more advanced forms of defense.

PitchBook provides actionable insights across the global capital markets.

Additional research:

PitchBook

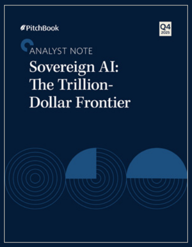
2026

OUTLOOK

Artificial Intelligence

2026 Artificial Intelligence Outlook: The Great Competition Wars Have Begun

Download the report here



PitchBook

Q4

ANALYST NOTE

Sovereign AI: The Trillion-Dollar Frontier

Q4 2025 Analyst Note: Sovereign AI

Download the report here

PitchBook Insights is an online compendium of in-depth data, news, analysis, and perspectives that shape the private capital markets.

PitchBook subscribers enjoy exclusive access to a comprehensive suite of private market insights, including proprietary research, news, data, tools, and more on the PitchBook platform.

Nizar Tarhuni
Executive Vice President of Research and Market Intelligence

Paul Condra
Global Head of Private Markets Research

James Ulan
Director of Emerging Technology Research

Report created by:

Dimitri Zabelin
Senior Research Analyst, AI and Cybersecurity

Oscar Allaway
Data Analyst

Chloe Ladwig
Graphic Designer

Learn more about PitchBook’s Institutional Research team.

Click here for PitchBook’s report methodologies.