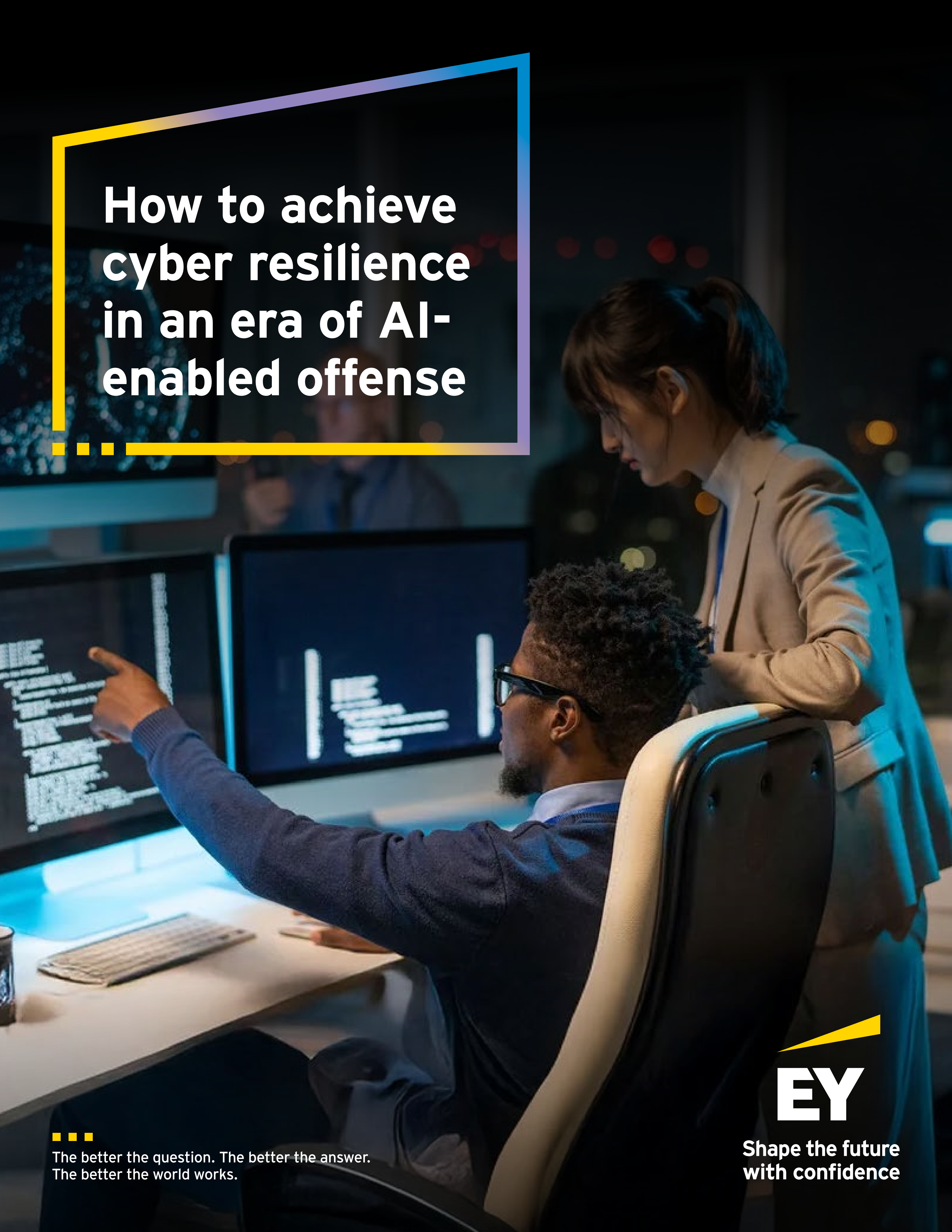
How to achieve cyber resilience in an era of AI-enabled offense



The better the question. The better the answer.
The better the world works.

Shape the future
with confidence

Organizations with AI-driven cyber resilience boost trust and competitive advantage; those that don’t risk their IP and consumer confidence.

In brief

- The rise of AI in cybersecurity is transforming attack methods, making threats more automated and adaptive, challenging traditional defense mechanisms.
- Organizations must prioritize foundational security measures like zero trust and proactive risk governance to effectively manage evolving cyber threats.
- Strategic cyber resilience integrates security into all aspects of operations so organizations can swiftly adapt to rapid changes in the threat landscape.

Cyber security is in a period of accelerated change. Artificial intelligence (AI) is reshaping how attacks are executed, how vulnerabilities are identified and how adversaries scale their campaigns. What was once a human-driven threat landscape is evolving into something more automated, more adaptive and significantly harder to defend.

AI’s effectiveness depends on the quality and volume of training data it learns from. AI-powered attacks scale and automate at a level that overwhelms traditional defensive models. This is why our clients must establish and strengthen foundational security building blocks, particularly zero trust, proactive security, strong risk governance, and rapid detection and response. These fundamentals remain the priority, even as the threat landscape accelerates. Resilience requires a shift in how organizations design, govern and operate their security capabilities.

An AI safety and research company, Anthropic, recently disclosed that a state-sponsored threat group used an AI platform to coordinate simultaneous cyber espionage intrusions across multiple global companies and government agencies.[1] The AI system assisted with vulnerability identification, live exploitation, escalation, lateral movement and the creation of new attack paths. It also attempted to solve technical problems during active operations, reducing the time and knowledge traditionally required. Similar disclosures from OpenAI and Microsoft, having documented state-affiliated misuse of large language models (LLMs),[2,3] along with assessments and guidance from the UK National Cyber Security Centre and the Cybersecurity and Infrastructure Security Agency (CISA),[4] reinforce this trajectory.

[1] “Disrupting the first reported AI-orchestrated cyber espionage campaign,” *Anthropic*.

[2] “Staying ahead of threat actors in the age of AI,” *Microsoft*.

[3] “Disrupting malicious uses of AI by state-affiliated threat actors,” *OpenAI*.

[4] “Impact of AI on cyber threat from now to 2027,” *National Cyber Security Centre*.

Although some outputs needed manual validation, the broader pattern was clear. AI increased the speed and enabled constant iteration of the intrusion cycle. Notably, the attack relied on existing exploits and was only partially automated. It was not a fully agentic operation. This follows the usual pattern: a new attack emerges, triggering an iterative arms race between attackers and defenders, with successful intrusions validating techniques that become more automated over time. This reinforces why organizations always need to constantly advance core cyber practices to close critical gaps before the next wave of automation matures.

This acceleration mirrors the broader environment described in the 2025 EY Global Risk Transformation Study, where risks evolve in nonlinear, accelerated, volatile and interconnected (NAVI) ways, collectively captured in the NAVI framework. AI-enabled cyber threats reflect this reality by escalating quickly, crossing organizational boundaries and challenging static controls.

The changing nature of cyber threats

Cyber threats are progressing through a significant shift. Earlier attacks relied heavily on the skill of human operators, manual reconnaissance and phased exploitation. Intrusions required time to plan and execute.

Today's attackers use AI inside live environments. They help craft malicious code, refine payloads, generate synthetic communications and troubleshoot technical barriers. Public threat intelligence has already confirmed that attackers use AI tools to support decisions once they are inside a network. Malware has been observed calling out to LLMs to generate evasive commands in real time.

Tomorrow's threat landscape is likely to involve semi-autonomous intrusion ecosystems that operate continuously across cloud, identity, data and application layers. These systems may test defenses, shift tactics based on detection and remain persistent without continuous human oversight. AI compresses the intrusion lifecycle into a rapid, adaptive sequence, reducing the distinction between reconnaissance, exploitation and persistence.

In a NAVI environment, this pace and interdependence requires new cyber resilience approaches. Traditional assumptions that threats move predictably or linearly are no longer valid.

Threats to AI and threats through AI

As organizations adopt AI to streamline operations and improve decision-making, they face two categories of risk that influence both security posture and enterprise resilience:

- Threats to AI arise when attackers target the data, models and infrastructure that power AI. Compromised training data can distort model behavior. Theft of proprietary models undermines competitive advantage and may reveal sensitive information. Attacks on the underlying environment, including data pipelines and machine learning operations (MLOps) workflows, can degrade performance or introduce malicious logic. Even subtle manipulations can affect how a model interprets inputs which influences downstream business decisions.
- Threats through AI occur when attackers exploit AI systems as tools or enablers. AI with elevated privileges can be manipulated into actions that users did not intend. Attackers can use AI to analyze environments, identify weaknesses, generate exploit paths and refine attacks while inside an organization's network. Synthetic audio, video and text make social engineering more effective. Public-facing AI applications may reveal sensitive information if probed in certain ways. As attackers automate these capabilities, they reduce the time and skill needed to run sophisticated campaigns.

The two threat categories are interdependent. Weak controls around AI create opportunities for misuse, while attackers who use AI increase the pressure on organizations to secure their own models, data and workflows. Resilience requires a holistic view of the entire AI ecosystem. These risks are amplified when baseline security practices are inconsistent. Applying zero trust principles to development processes, data environments and production systems limits the impact of a compromise and reduces the opportunities for attackers to manipulate or misuse AI assets.

How Ernst & Young LLP (EY US) can help

Risk consulting services

Discover how EY Risk Consulting team can help your organization embrace disruption and turn risk into a competitive advantage. Read more

Responsible AI

AI boosts business but presents challenges. A Responsible AI framework allows leaders to harness its transformative potential while mitigating risks. Read more

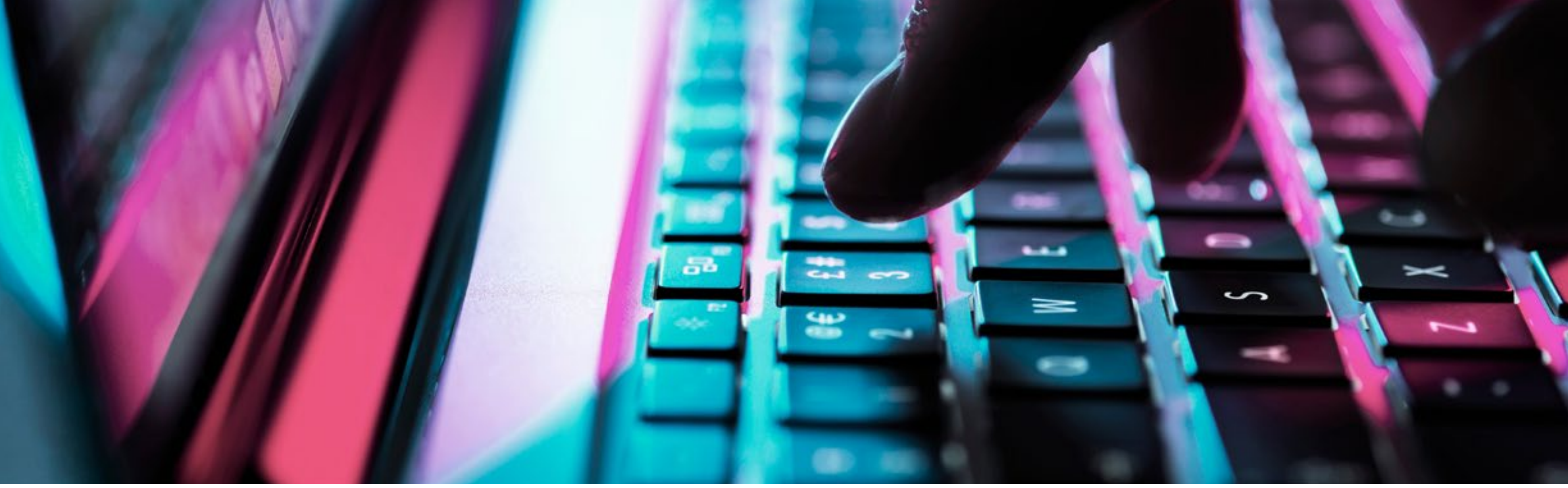
Cyber services

Evaluate how effective and efficient your organization's cybersecurity and resiliency programs are in driving business growth and operational strategies – helping you determine how, where and why to invest in cyber risk management. Read more

Why traditional controls struggle with AI-enabled threats

Foundational controls remain essential, but they were built for a world with slower attack cycles. Identity governance, network segmentation, secure development and security monitoring retain value, yet they cannot manage threats that adapt in real time. AI compresses the time between reconnaissance, exploitation and escalation. Intrusions progress too quickly for manual analysis, periodic assessments or static signatures to keep pace.

This mismatch requires organizations to rethink their defensive architecture and shift toward models that anticipate rapid change and support automated detection and response. In addition to reducing time to detect and respond, using AI for cyber resilience also helps bend the cost curve.



A strategic blueprint for modern cyber resilience

A more cyber resilient approach involves a coordinated strategy across six reinforcing pillars:

1. **Strengthen foundational controls across cloud, identity, data, applications, infrastructure and endpoints.**

   Consistency, hygiene, segmentation and strong identity governance create a secure baseline that reduces exploitable gaps. This includes renewed emphasis on zero-trust-aligned identity controls, such as eliminating long shelf-life credentials, enforcing phishing resistant multi-factor authentication (MFA), using short-lived tokens and restricting lateral movement pathways. This is basic cyber hygiene which will never go away.

2. **Integrate security and broader risk management principles into engineering and design from the outset.**

   Shifting security upstream into architecture, data management, development and testing reduces systemic weaknesses before they reach production. Proactive design patterns must also extend to software development pipelines, which hold elevated privileges and present common vectors for privilege escalation.

3. **Adopt AI-assisted defensive capabilities that match the speed of modern threats.**

   Automated detection, correlation and response functions improve the ability to identify and contain threats that operate at machine speed. Emerging AI-driven triage and investigation agents are beginning to reduce alert fatigue and accelerate investigative workflows, which improves containment.

4. **Govern AI with responsible practices that reinforce resilience.**

   Effective AI governance establishes clear ownership, lifecycle controls, guardrails, integrity monitoring and secure MLOps. Responsible AI frameworks reduce unintended behavior and strengthen trust, which in turn supports cyber resilience. Given AI's dependence on training data, ongoing data quality verification and integrity monitoring help the models behave as intended.

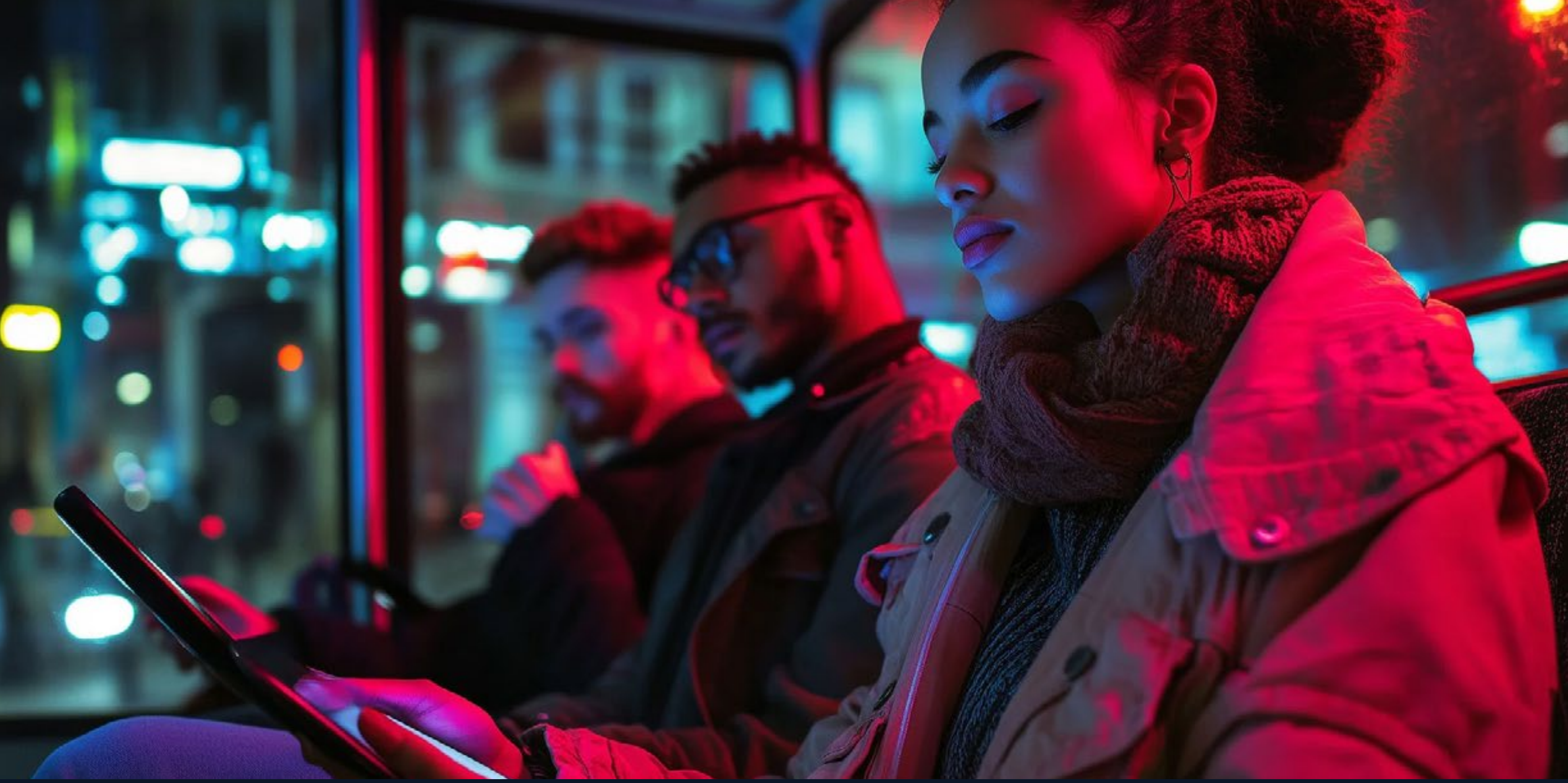
5. **Embed cybersecurity within the broader enterprise risk and resilience framework.**

   Cyber must be unified across the broader Enterprise and Technology Risk and Resilience program. Business continuity and disaster recovery plans are not enough to deal with this risk. True proactive resilience and risk management is required to effectively protect an organization. This extends across the full attack surface – on-premise, across cloud providers, through software as a service (SaaS) vendors and via other critical third parties – so disruptions can be absorbed without material impact.

6. **Hack yourself first.**

   Learn from attacker techniques and treat this information as an opportunity to better understand your attack surface, proactively identify issues and leverage new techniques, like data scanning at scale, to drive internal improvements and reduce exposure.

These pillars create a cyber-resilient strategy that aligns prevention, detection, response, governance and recovery in a cohesive model.



Conclusion

These disclosures offer a preview of how AI will shape the next decade of cyber risk. Attacks are becoming more automated, adaptive and intertwined with the technologies that organizations depend on. These developments reflect the NAVI characteristics outlined in the 2025 EY Global Risk Transformation Study and reinforce the need for a modern approach to resilience.

Cyber resilience, responsible AI governance and secure engineering are no longer separate domains. They are interconnected components of a forward-looking strategy. Organizations that integrate these capabilities will be better positioned to navigate a world where AI influences both the threats they face and the defenses they must deploy.

Summary

Cyber resilience is increasingly critical as cybersecurity faces rapid AI-driven transformation. AI is automating attacks, making them more adaptive and challenging traditional defenses. Organizations must strengthen their foundational security measures, such as zero trust and proactive risk governance, to combat these evolving threats. The integration of AI into both offensive and defensive strategies highlights the need for a holistic approach to cyber resilience. By embedding security within the broader enterprise risk framework and adopting AI-assisted capabilities, organizations can better navigate the complexities of modern cyber threats and enhance their overall resilience against future attacks.

About this article

Authors

Ayan Roy
EY Americas Cybersecurity Competency Leader

Kapish Vanvaria
EY Global and Americas Risk Consulting Leader

Scott McCowan
EY Global Consulting Risk Markets Leader

Contributors

Traci Gusher
EY Americas AI and Data Leader

George Haggar
EY Americas Digital and Technology Risk Management Solution Leader

Dan Mellen
EY Global and US Cyber Chief Technology Officer

EY | Building a better working world

EY is building a better working world by creating new value for clients, people, society and the planet, while building trust in capital markets.

Enabled by data, AI and advanced technology, EY teams help clients shape the future with confidence and develop answers for the most pressing issues of today and tomorrow.

EY teams work across a full spectrum of services in assurance, consulting, tax, strategy and transactions. Fueled by sector insights, a globally connected, multidisciplinary network and diverse ecosystem partners, EY teams can provide services in more than 150 countries and territories.

All in to shape the future with confidence.

EY refers to the global organization, and may refer to one or more, of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. Information about how EY collects and uses personal data and a description of the rights individuals have under data protection legislation are available via ey.com/privacy. EY member firms do not practice law where prohibited by local laws. For more information about our organization, please visit ey.com.

Ernst & Young LLP is a client-serving member firm of Ernst & Young Global Limited operating in the US.

US SCORE no. 29979-261US
2602-10442-CS
ED None

This material has been prepared for general information purposes only and is not intended to be relied upon as accounting, tax, legal or other professional advice. Please refer to your advisors for specific advice.

ey.com