The new cyber battleground

Building resilient, secure AI capabilities across the enterprise

KPMG. Make the Difference.

January 2026

What's inside

Foreword

The last 18 months have seen a prolific increase in activity around cyber analytics, automation, and AI security services in the Middle East region, and indeed globally. Many organizations are struggling to keep up with the pace at which this domain is moving, amid uncertainty about trust, time to value, return on investment, workforce upskilling, and a potential global AI bubble waiting to burst.

Do you have the skills and knowledge to identify critical use cases and implement it with a Trusted AI governance framework across your business? Do you really need AI – or can you solve your challenges with analytics and automation to prove value? These questions are common and necessary to solve "problem zero": where and how does one start?

In addition to this, cyber adversaries are rapidly advancing their use of AI tools and development techniques to accelerate wide-scale attacks. In this publication we explore how organizations can realize value and return quickly around AI security.

Current challenges

Today, organizations are facing challenges with trust and security of AI, as well as deriving value from the AI investment.

As we saw recently, Anthropic was used to launch a fully unmanned cyber-attack across multiple sectors and organizations. This introduces a new threat paradigm and threat actors, mandating organizations to quickly understand and combat the very same AI models and solutions we are using on a day-to-day basis

This raises several fundamental questions about the trustworthiness and security of Large Language Models and solutions freely available in the market.

These solutions are already being integrated into organizations, and have rapidly become trusted solutions, without full consideration of the risks involved.

This raises a wider question around the trustworthiness of AI security, and how organizations can ensure that use cases and AI solutions being implemented do not compromise the risk posture of your organization. Many organizations are considering the establishment of Trusted AI frameworks and processes; however, early adoption is slow. This is a critical first step to reducing the risk of internal AI compromise, but this work impedes the race to create and implement AI security use cases.

The intelligent age has arrived

66% of people use AI regularly.

83% of people believe the use of AI will result in a wide range of benefits.

Trust remains a critical challenge

46% of people are globally are willing to trust AI systems.

AI regulations

70% believe there is a need for national and international AI regulation.

AI at work

66% rely on AI output without evaluating accuracy.

56% are making mistakes in their work due to AI.

KPMG's trusted AI framework

Equity
AI solutions must be designed to reduce or eliminate bias against individuals, communities and groups.

Transperancy
They must be transparent and ensure a clear understanding of what happens in each solution throughout the AI lifestyle.

Explainability
They must be developed and offered by addressing the questions of how and why a solution has reached a given colcusion.

Accountability
Human oversight must be integrated into the full AI lifecycle to manage risks and comply with applicable regulations.

Cybersecurity
AI solutions must be safeguarded against cybercriminals, disinformation and other adverse events.

Privacy
AI solutions must be designed to comply with applicable privacy and data protection laws and regulations.

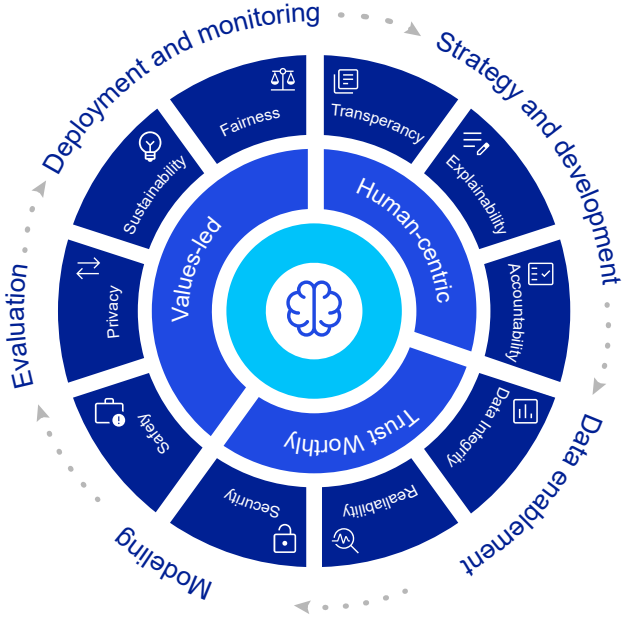
Sustainability
They must be designed to be energy efficient, minimize carbon emissions, and promote a cleaner enviroment.

Data integrity
The data used must be obtained in compliance with applicable regulations and assessed for accuracy, integrity, and quality.

Reliability
AI solutions must function systematically according to their intended purpose, scope and desired level of accuracy.

Safety
AI solutions must be designed and implemented to prevent harm to individuals, companies and assets.

Many organizations over the past 12 months have raced straight to technical AI solutions and use cases, without having the required governance in place to use AI safely and securely. This supports the view from many AI experts that over 95% of AI use cases and solutions are failing to realize any value, having proceeded straight to AI agents as the solution, before clearly understanding the challenges they are trying to solve.



Critical to the success of your AI security strategy and objectives is identifying what problems need to be solved, and this can be achieved in several ways—with analytics, automation, AI, or a combination. AI is not always the solution and should not be the starting point as it is still emerging and not fully trusted - yet. Quick Time to Value (TTV) can be realized through process automation in the first instance, provided you understand your end-to-end internal processes across Information and Cyber Security Operations teams. These are also critical building blocks for implementing AI use cases, which require process maturity, good quality data, experts to train models, and supporting infrastructure to run AI agents. This needs to be a balanced approach with a clear business case and Return on Investment (RoI) metrics.

The current failure rate of 95% suggests there is still a significant gap between the accelerated (and perceived) need to use AI for security and the cost benefit analysis and investment justification.

Where to start: securing AI

Securing AI begins with strengthening its foundations, using known and tested AI governance and security frameworks. Organizations must treat data as a protected asset by enforcing lineage, provenance, rigorous cleaning, bias mitigation, and continuous monitoring throughout the training pipeline. Model protection then becomes the next layer: applying guardrail frameworks, adversarial robustness testing, watermarking for IP protection, and controlled model registries to ensure models behave predictably and resist manipulation. Infrastructure must also be hardened through secure pipelines, trusted execution environments, immutable deployments, and policy-driven configurations that prevent drift, contamination, and supply chain compromise.

True AI resilience emerges when these capabilities operate as a unified security fabric. By embedding security into MLOps (Machine Learning Operations) pipelines and Musicos (Model Use, Integration, and Control Operating Systems), organizations ensure that every model update, retraining cycle, and API integration is authenticated, monitored, version-controlled, and auditable. Automated drift detection, secure retraining workflows, AI posture management, and continuous red-teaming transform AI from a static model into a living system that defends and improves itself.

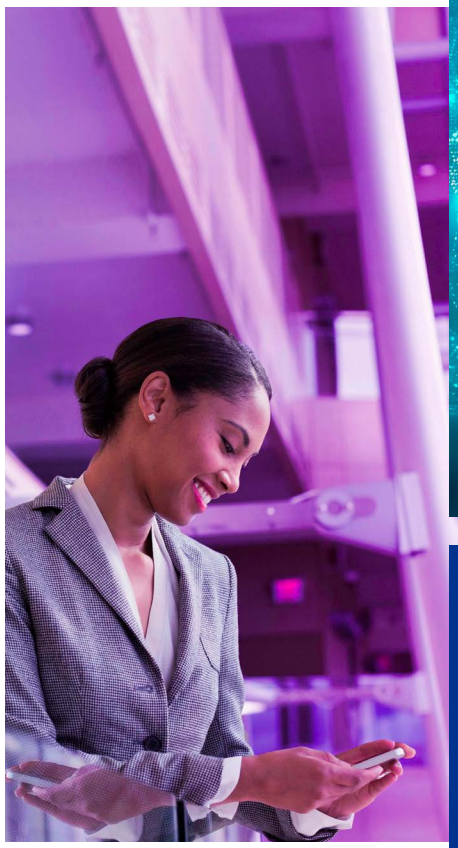
> “Securing AI is no longer about bolting controls; it is about engineering trust into the lifecycle - data, models, infrastructure, and operations, so AI can scale safely, confidently, and sustainably.”

AI for enhancing intelligent protection

AI transforms cybersecurity from reactive defense to intelligent, adaptive protection. By shifting static rules to semantic and probabilistic reasoning, AI helps security teams detect what truly matters much earlier and with greater accuracy. It learns behavioral baselines, understands context, prioritizes alerts by business impact, and turns fragmented evidence into coherent narratives.

AI is elevating cybersecurity by strengthening how organizations understand, manage, and protect their digital environments. Rather than treating identity, data, cloud, and operations as separate silos, AI connects signals across them, learning patterns of normal behavior, identifying subtle deviations, and surfacing risks before they escalate. It enhances visibility across the estate: how identities are being used, how sensitive data flows, how cloud configurations drift, and how development pipelines evolve. The starting point is simple: introduce AI where fragmentation and manual effort are highest (correlation, enrichment, misconfiguration detection, and behavioral insights) and let it amplify the effectiveness of existing controls.

As organizations progress, AI becomes a strategic enabler across the entire cyber lifecycle. It streamlines secure deployment and management of the security controls, strengthens governance through continuous evidence gathering and policy checks, enhances threat intelligence by contextualizing global signals, and improves resilience by automating routine tasks while keeping humans firmly in control.



Ultimately, AI doesn't replace core cybersecurity professionals and capabilities; it binds them together, making them more adaptive, consistent, and scalable. When AI is woven into detection, response, engineering, governance and cloud security as a common intelligence layer, organizations shift from reactive protection to proactive resilience. The journey begins with augmenting what teams already do today and evolves into a security model where AI continuously reinforces the organization's defenses end-to-end.

A practical roadmap for trusted AI adoption

Understanding how to safely and securely use AI to defend your organization in the future is non-negotiable. Adversaries are active in this space, and attacks will again become faster and more sophisticated with the use of AI agents and tools. If you are starting out, identify easier problems to solve while keeping things simple, in addition to acquiring the skills and knowledge to be successful and realize quick time to value and return on investment.

Below, we outline five steps we believe organizations

- **Establish trusted AI security governance .**
- **Understand the problem you want to solve, then choose the right option (Analytics, Automation, AI, Agentic AI etc.) to show value.**
- **Upskill employees to both secure and use AI for security.**
- **Use the tools you have today to create iterative value, then scale (e.g. Microsoft Co-pilot Studio).**
- **Focus on addressing easier challenges first to realize quick time to value and return on investment.**

Sources

https://www.anthropic.com/news/disrupting-AI-espionage

https://kpmg.com/xx/en/our-insights/ai-and-technology/trust-attitudes-and-use-of-ai.html

https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf

Gillespie, N., Lockey, S., Ward, T., Macdade, A. & Hassed, G. (2025). Trust, attitudes and use of artificial intelligence: A global study 2025. The University of Melbourne and KPMG. DOI 10.26188/28822919. The research cited in this report was led by Professor Nicole Gillespie and Dr. Steve Lockey at the University of Melbourne.

Contact us

Trevor Niblock
Partner, Digital Trust
KPMG Middle East
tniblock@kpmg.com

Shirish Jangid
Director, Digital Trust
KPMG Middle East
sjangid1@kpmg.com

www.kpmg.com