



Accenture Cybersecurity

AI Security Evolution

Security becomes an increasing factor for AI

AKAMAI – Cyber Security Meetup Munich
6. Oktober 2026

Agenda



01

**Evolving AI results
in evolving risk**

02

**New threats require
new answers**

03

Solution concepts

04

**Discussion &
Appendix**

AI agent adoption is a complex challenge for organizations.

Agentic architecture is reshaping enterprise workflows — and **expanding the attack surface**.

Organizations that don't **embed security from the start** risk amplified exposure.

Nearly 9 in 10

organizations identify **AI-related vulnerabilities as the fastest-growing cyber risk**.*

52%

of organizations have already been breached **via non-human identities**** — the fastest-growing attack vector — and we're already at an **82:1 machine-to-human ratio *****

Frontier AI models

Over 99% of the vulnerabilities discovered by Mythos Preview **have not yet been patched** ²

Organizations must embed security and guardrails early in their transformation journey

— **while proactively preparing for an evolving AI compliance landscape.**



How risk evolves – from assistants to agents to agentic AI



AI Assistants

AI Assistants help end users with tasks & queries in mostly reactive response mode

- Built for user interaction
- Reactive in response to user inquiries
- Limited autonomy



AI Agents

AI Agents proactively take action on a particular task.

- Built for specific tasks or decisions
- Proactive and adaptive
- Higher level of autonomy



Agentic AI

Agentic AI is a dynamic, multi-agent autonomous systems that collaborate and adapt across enterprise environments

- Built for autonomous collaboration
- Dynamic, multi-agent orchestration
- Adapts across enterprise environments

AI Assistants Compartmentalized risks

The challenge

Balancing innovation with security guardrails while enabling controlled AI development.

- **Isolated Environment Risks**
Even in sandboxes, AI agents can extract sensitive data, introduce prompt injection vectors, and generate misleading outputs
- **Security Debt Accumulation**
Early focus on capability over controls creates technical debt as authentication and access patterns remain underdeveloped
- **Shadow AI Proliferation**
Inconsistent governance across teams leads to untracked agent deployment and unmonitored data handling
- **Control-innovation Tension**
Risk of either over-constraining innovation or under-protecting critical systems during experimental phase

Agentic AI Dynamic expanded risk boundaries

The challenge

Establish scalable security architecture for autonomous AI systems while enabling business value.

- **Access Boundary Breakdown**
AI Agent autonomy challenges traditional (static) security access controls creating potential for privilege escalation risk.
- **Development-to-Production Inheritance**
Production deployments carry over excessive permissions and weak controls from testing environments, expanding exploitable attack surfaces
- **Agent Interaction Blind Spots**
Multiple autonomous agents operating across systems create unpredictable behaviors and security gaps invisible to traditional monitoring
- **Opaque Decision Lineage**
Agent actions become increasingly difficult to trace and explain, complicating incident response and regulatory compliance

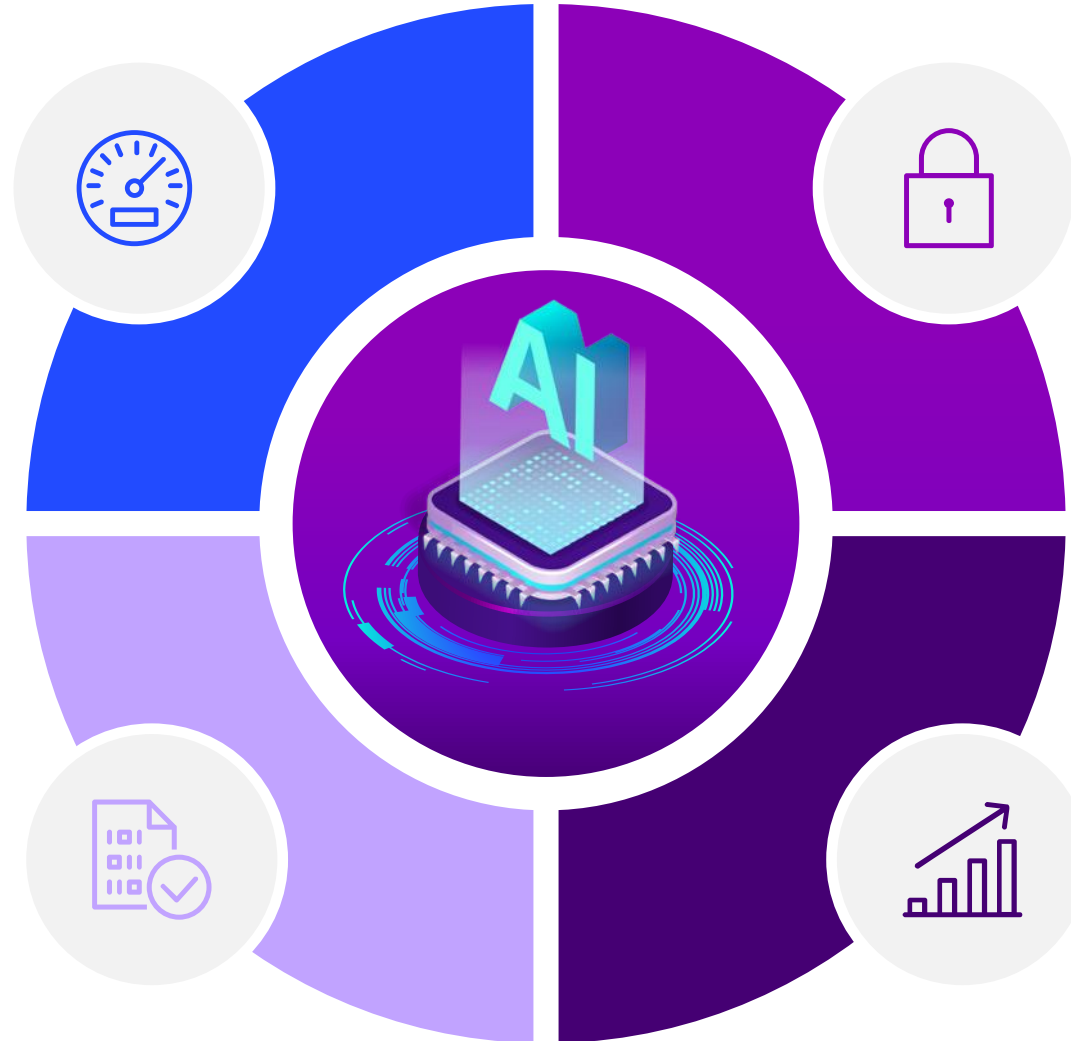
The Big Picture: What would it mean if your organization could securely scale critical Agentic AI uses cases with speed & confidence?

Faster deployments of AI agents – using your existing tools – without compromising trust

Stakeholder confidence and streamlined regulatory compliance

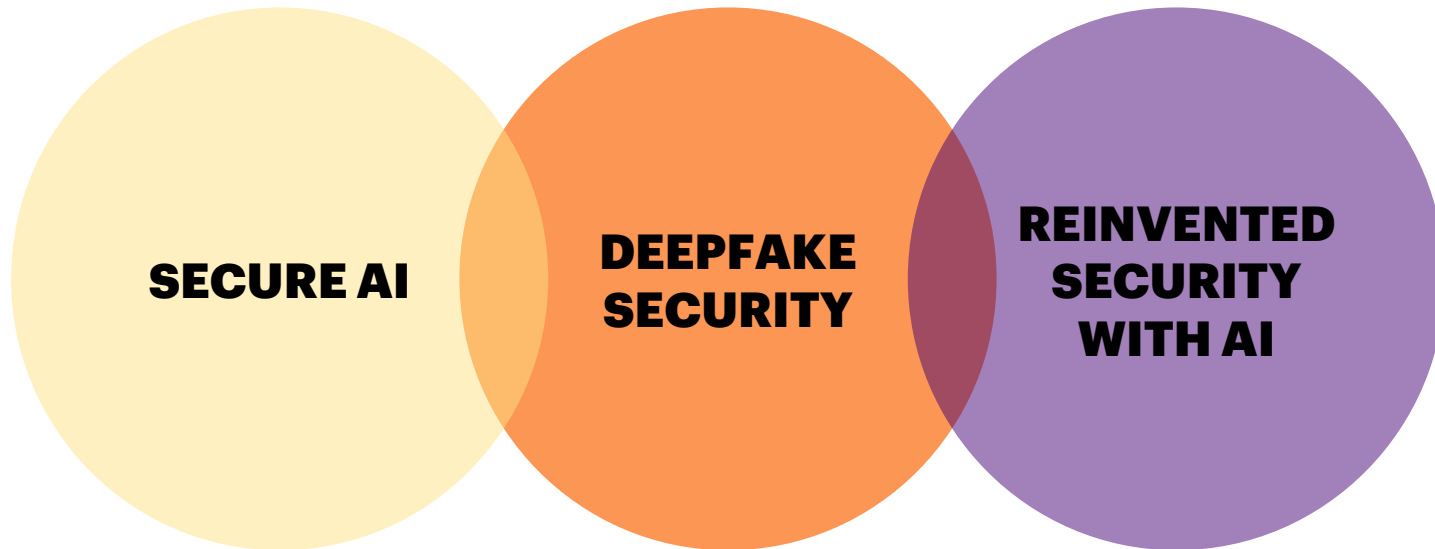
Operational resilience against emerging AI-driven threats

Business growth, as secure AI adoption correlates with higher productivity and revenue gains



AI Security is a multi dimensional approach

AI requires various categories of risks and threat mitigation concepts



Definition

Secure AI refers to the development and deployment of AI systems that prioritize the protection of data, privacy, and integrity

Reinvented Security with AI involves leveraging AI to enhance security measures and streamline processes

Deepfake Security addresses the threats posed by deepfake technology

02

**A new Threat Landscape
requires new answers**



Deepfake technology is a highly-sophisticated tool in a Threat Actor's social engineering toolbox

Social Engineering



Deepfake



Baiting



Phishing



Spear Phishing



SMiShing



Tailgating



Pretexting



Scareware



DNS Spoofing



Business Email Compromise

Threat Actor Objectives

IP Theft, Financial, and
Reputational Harm

Unauthorized Access
to Critical Systems

Business Disruption
and Loss of Trust



The Mythos Moment: A structural shift in the cyber threat landscape

Anthropic's Claude Mythos is a frontier AI model that **autonomously discovers, chains, and exploits zero-day vulnerabilities** across every major operating system and browser — including flaws that survived decades of human review.

It was not built for offense. These capabilities emerged as a byproduct of advances in code reasoning

Three details every enterprise leader needs to understand:

1. Non-experts can weaponize it overnight
2. It synthesizes across systems, turning low-risk findings into end-to-end attacks
3. You cannot prevent Cyber-criminals from getting access to this technology - Mythos is the first example of capabilities; cyber-attackers will build to weaponize themselves

Research snapshot

Impacts observed



Thousands of zero-day vulnerabilities identified across every major OS and browser

Key insights

83%

first-attempt exploit success rate vs. near-zero prior

~\$0

marginal cost for non-experts to produce working exploits

20h

Mean Time-to-Exploit in 2026
(from 2.3 years in 2018)

We anticipated this shift and invested heavily in re-inventing our cybersecurity offerings, powered by Anthropic.



A Structural Shift in the Cyber Security Landscape

Claude Mythos is not an isolated incident — it is an early signal of a fundamental shift. AI is changing the threat landscape at its core:



AI as an attacker's tool

Threat actors can now discover, exploit, and weaponize vulnerabilities at machine speed, dramatically compressing the window between disclosure and exploitation



Current controls are no longer sufficient

Traditional patch cycles, rule-based detection, and manual triage were built for a different era — they cannot keep pace with AI-augmented attacks



Vulnerability management must evolve

The discipline needs to move from periodic, reactive processes to continuous, intelligence-driven, and increasingly automated operations



The question is no longer *if* AI changes security — it is *how fast* organizations can adapt.



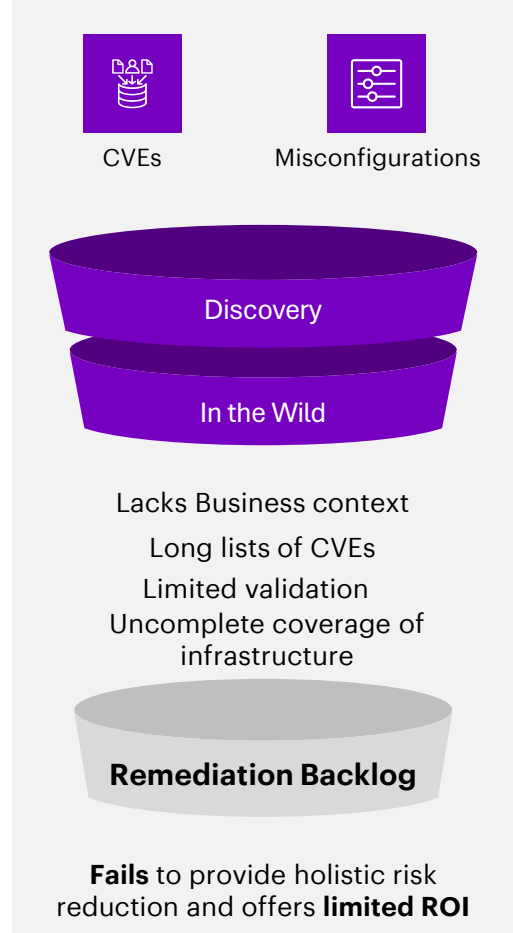
**Mythos is not the
problem.**

It is the moment the problem
became impossible to ignore.

A single pane of glass is required for defining activities that result in significant security posture improvements

Difference between VM and CTEM

Vulnerability Management



14%

of breaches involved exploitation of **vulnerabilities** for initial access, a **3x** year-over-year increase*

The Vulnerability / Breach Disconnect

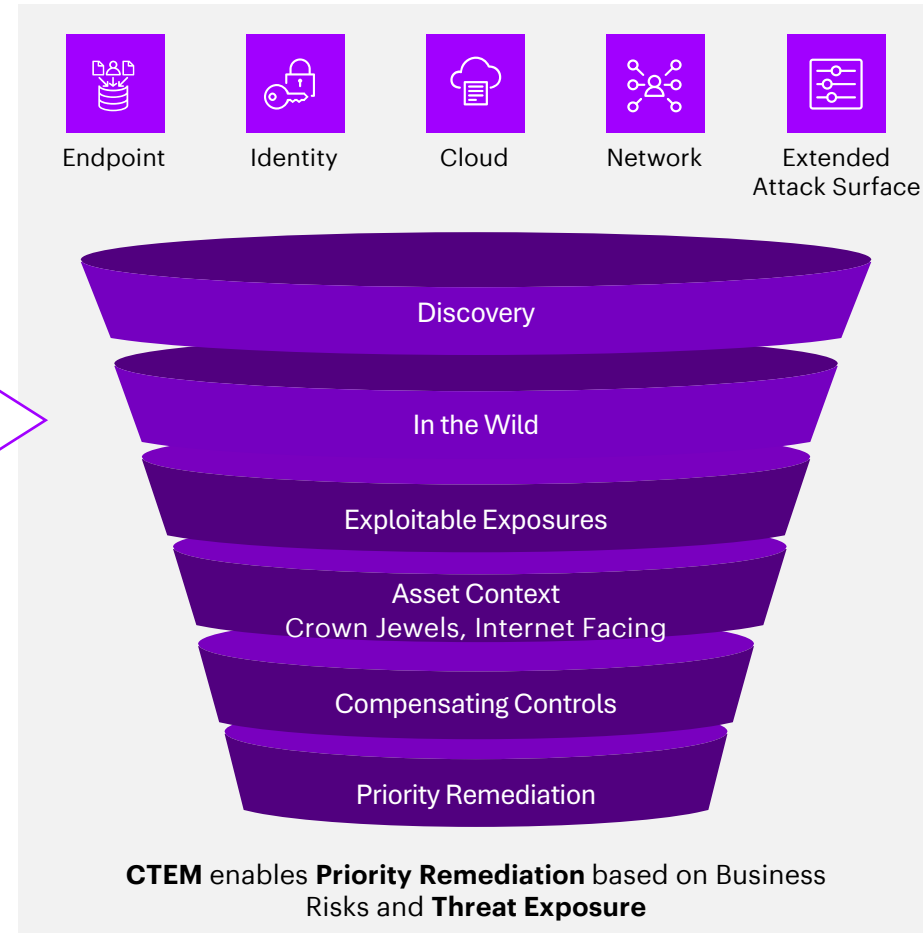
85%

of successful cyber intrusions involve compromised credentials or exposed identities as a means to propagate an attack**

*Verizon 2024 Data Breach Investigations Report (DBIR)

**Accenture Cyber Response Intrusion-derived intelligence

Continuous Threat Exposure Management



03

So What?



1. Fixing the basics to increase protection level & the effort for the attacker

1 Privileged Access Management The keyring. Control who holds the keys to everything.

2 Active Directory Security AD compromise is organization-wide compromise.

3 Multifactor Authentication Non-negotiable. Remove the password as the single point of failure.

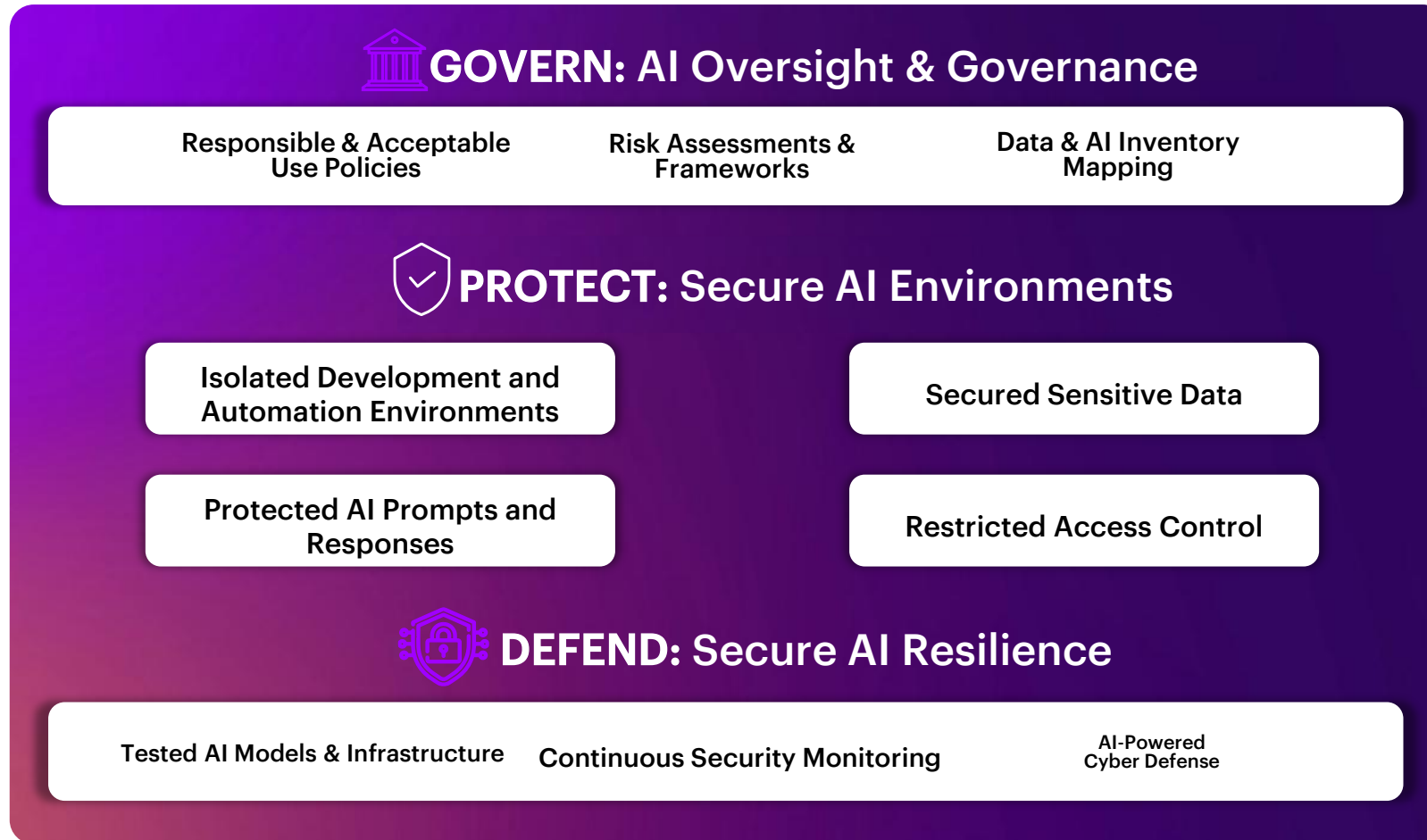
4 Endpoint Detection & Response Containment and isolation at AI-driven speed — not just alerting.

5 Identity-Based Segmentation Contain blast radius architecturally before the incident starts.

6 Monitoring and Response Continuous visibility with the operational capability to act.

Apply tightly around systems of mass destruction — AD, virtualisation layer, privileged identity infrastructure.

2. Establish a trustworthy set of AI capabilities to mitigating risk



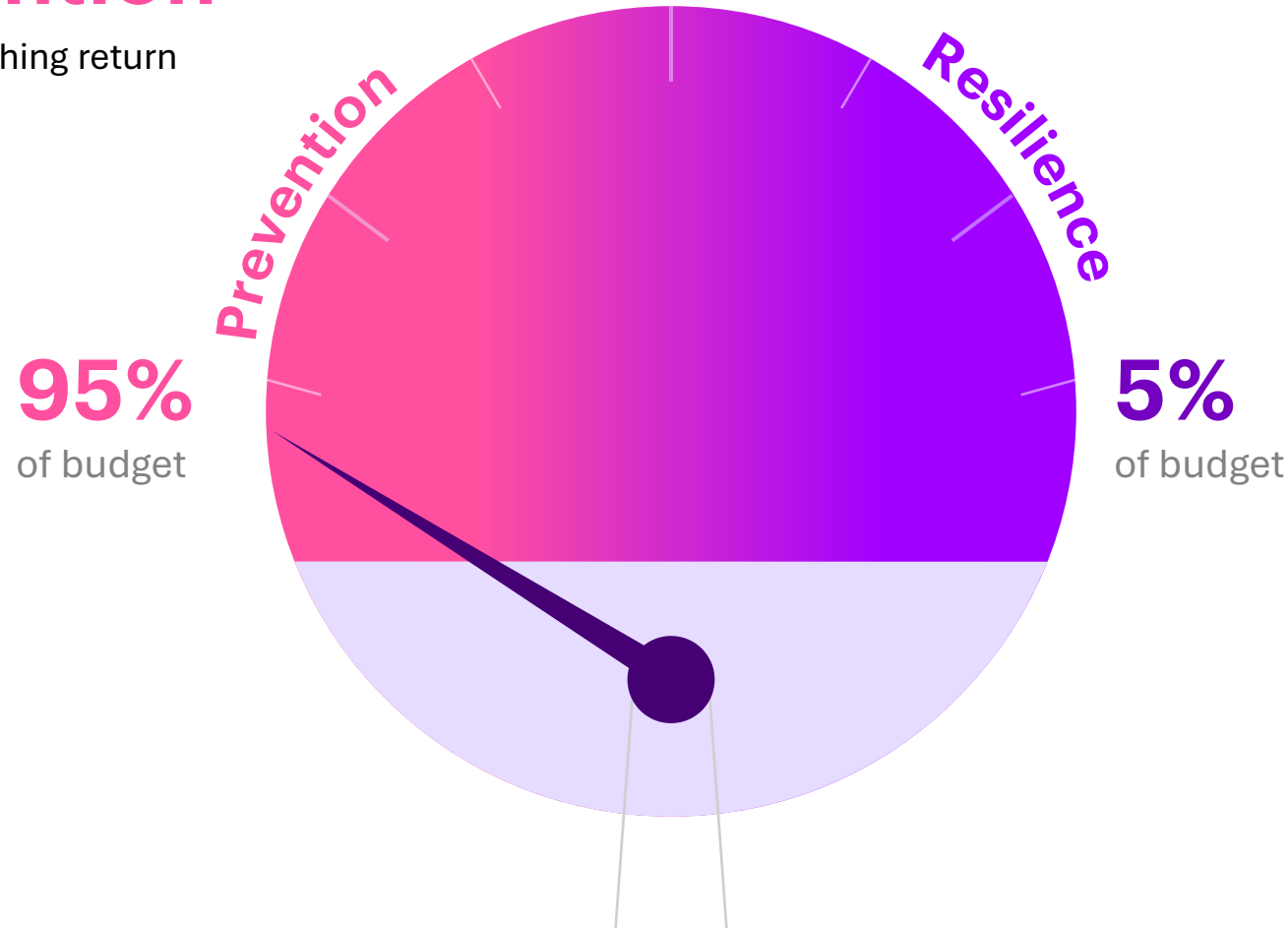
3. Paradigm shift from Prevention to Resilience

Broad prevention

Sprawling · expensive · diminishing return

Resilience

Underfunded · untested



Broad scanning,
compliance tooling,
perimeter controls,
and more

Recovery

Concentrate prevention where it pays. Fund resilience with the rest.



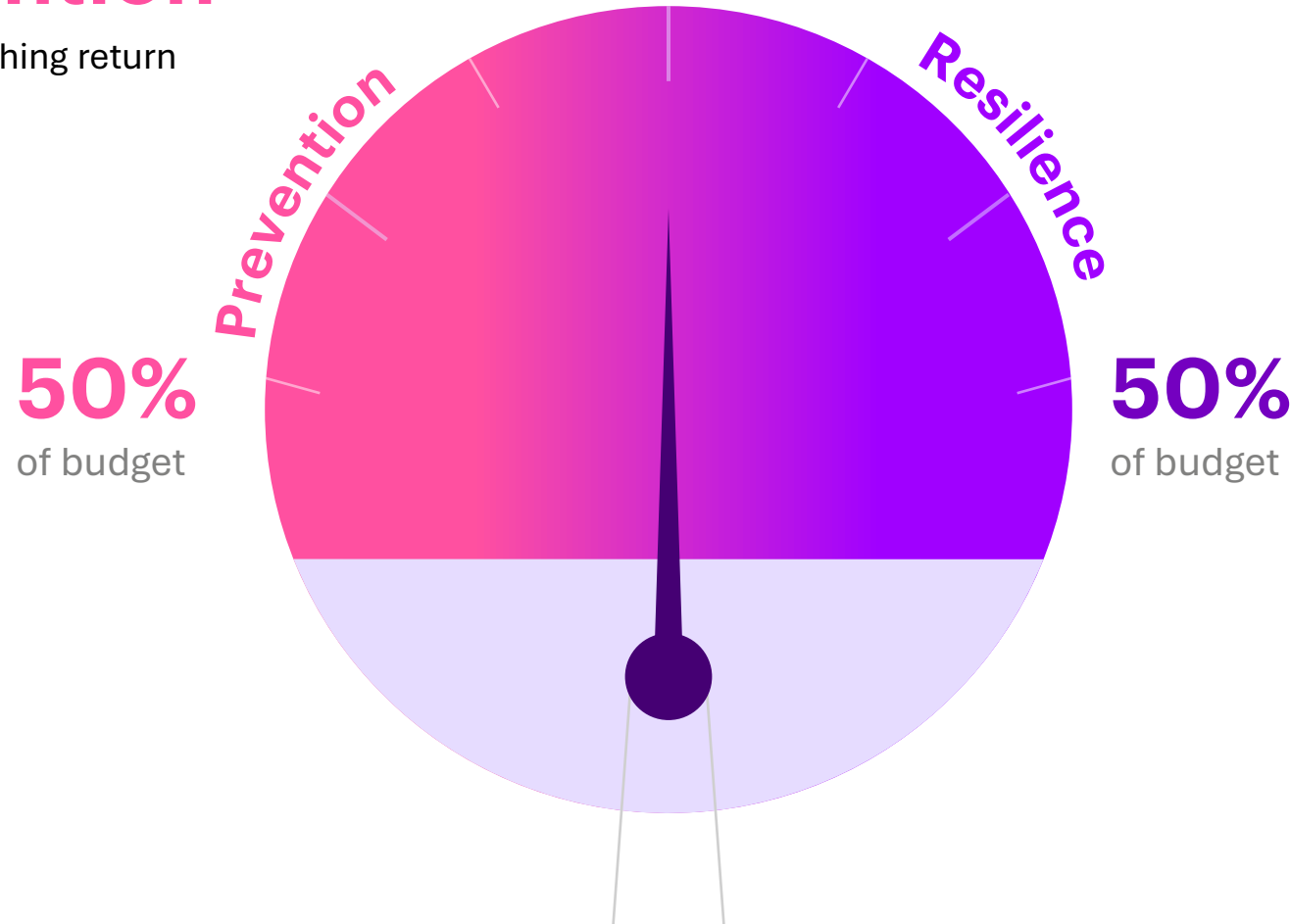
3. Paradigm shift from Prevention to Resilience

Broad prevention

Sprawling · expensive · diminishing return

Resilience

Underfunded · untested



Broad scanning,
compliance tooling,
perimeter controls,
and more

Recovery

Concentrate prevention where it pays. Fund resilience with the rest.



4. Follow AI Governance frameworks and regulations

Overview



ISO 42001

- AI management system standard
- Structured way to manage risks and opportunities associated with AI



BSI STANDARDS

- BSI AIC4: AI Security in Cloud
- BSI LLM Security
- BSI Vulnerability Report Prompt Injections



EU AI ACT

- Risk-based approach
- List of high-risk applications & requirements
- Conformity assessment for AI systems



NIST RMF & CONCRETIZATIONS

- NIST AI 600-1 Artificial Intelligence Risk Management Framework
- NIST SP 800-218A (ipd) Secure Software Development Practices for Generative AI and Dual-Use Foundation Models
- NIST AI 100-4 2 (ipd) Reducing Risks Posed by Synthetic Content



MITRE ATLAS

- Indicates an adaption from ATT&CK
- Shows the progression of tactics used in attacks with ML techniques



OWASP LLM SECURITY

- LLM AI Cybersecurity & Governance Checklist
- OWASP Top 10 for LLM

6. Follow a programmatic approach for Agentic AI Security

Each step of the journey is tailored to your specific needs

Do Now



Establish Agent Oversight & Risk Management

Outcomes

- Comprehensive visibility into AI agent risks and security posture
- Prioritized action plan addressing immediate gaps and risks
- Baseline governance structure for responsible AI agent deployment

Do Next



Secure Environments for Agent Deployments

Outcomes

- Secure, compliant environments for agent development and deployment
- Reduced unauthorized access risks through effective IAM governance
- Increased visibility and response capabilities through proactive security monitoring

Do Later



Advanced Monitoring & Continuous Improvement

Outcomes

- Full operational security monitoring and threat detection for agentic systems
- Continuous security validation and improvement processes
- Advanced runtime protection and anomaly response capabilities

Thank You

Raimund Wilhelmer
raimund.wilhelmer@accenture.com

