Understanding the current risk and Threat Landscape: Laying the Foundation for a Resilient Security

A security program built without threat intelligence, business context, and a structured framework isn't a strategy — it's a wish list.

Sushila Nair, CISA · ISACA Greater Washington DC · 2026 Threat Landscape Series

This session covers all five foundations:

Know where you are. Know where you're going.

NIST CSF Current Profile → Target Profile → Roadmap

01
Threat Intelligence
8 industry reports synthesized

02
Business Objectives
Risk appetite & what to protect

03
Technology Strategy
Cloud, AI, third-party dependencies

04
PEST & SWOT Analysis
Internal posture + external forces

05
NIST CSF Framework
Current profile → target profile

Sushila Nair- MS, CISM, CISA, CISSP, CCAK

- President- ISACA Greater Washington, D.C. Chapter (2nd biggest chapter in the world)
- Advisor- Leadership Program Shenandoah University
- Cyber Security Service Delivery- BT, NTT DATA, Capgemini, Currently – Independent consultant providing cybersecurity guidance
- Instructor: ISACA CCAK, CET Cloud, Cybersecurity for Auditors certification courses- APMG certified
- Mentor- ISACA, City of Refuge
- 2023 Humanitarian Award by ISACA Global
- 2023 and 2024 Top Cybersecurity leader by Security Magazine and Cyber Magazine
- 2020 V Lee Conyers Award by ISACA GWDC for contributions to the profession
- A CISO for 10 years
- Professional Certifications: AAIA,AAISM,CISM, CISA, CRISC, CDPSE, CISSP, CCAK, GIAC GSTRT, GIAC GSNA, GIAC GDSA
- Speaker- national and international engagements- Cybersecurity, Cyber Risk, Governance, Compliance, Strategy, Digital Transformation, Cloud, Leadership

Disclaimer

Please note that the views and opinions expressed during this presentation are solely those of the speakers and do not necessarily reflect the official policy or position of our organization.

Additionally, we are vendor-agnostic and politically neutral; any references to specific tools, vendors, or political matters are made solely to enhance understanding and are not intended to promote any particular service, vendor, or viewpoint.

This presentation is intended for informational purposes. Before implementing any ideas shared during this session, it's important to independently assess the security, legal, technical, and reputational considerations specific to your organization's circumstances.

Please note that the views and opinions expressed are those of the presenter and may not necessarily reflect those of their organization, ISACA, or any other affiliated entity.

The information provided is confidential and proprietary, meant exclusively for attendees. We kindly ask that you refrain from sharing this content without proper authorization.

Agenda

- 2024/25 Report synthesis
- PEST/SWOT
- Case Studies
- AI Threats
- Recommendations
- Q&A

How Security Reports Inform Your Security Program

•**Start with the Business**
Understand organizational goals, risk appetite, regulatory obligations, and what the business is trying to enable.

•**Assess the Strategic Context (SWOT & PEST)**
Use **SWOT** to evaluate internal strengths and weaknesses, and **PEST** to understand external political, economic, social, and technological forces shaping risk.

•**Align with Technology Strategy**
Factor in the CIO's initiatives—cloud adoption, digital transformation, AI, modernization, outsourcing, and third-party dependencies.

•**Leverage Industry Security Reports and other threat sources**
Use reports (Verizon DBIR, Mandiant, CrowdStrike, IBM, ISACA, SANS, Cisco) to understand **how attacks actually occur**, where controls fail, and which risks are increasing.

•**Translate Trends into Priorities**
Map external threat trends to **your environment**, identifying the most relevant risks—not every headline threat.

•**Design a Risk-Based Security Program**
Build and prioritize controls, people, and processes that reduce **material business risk**, not just technical findings.

•**Continuously Adapt**
Revisit SWOT, PEST, and threat intelligence as business goals, technology, and the threat landscape evolve.

Threat intelligence must drive strategy — not just controls.

What is a PEST analysis

- Strategic tool used for understanding the macro-environmental factors that might impact an operation
- Political - Cybersecurity is heavily impacted by governmental regulations such as GDPR in Europe, HIPAA in the U.S., or other data protection laws globally. Changes in these regulations can significantly affect how data security must be managed.
- Economic - Budget impact
- Social - Attitudes towards data privacy and cybersecurity, availability of cyber talent
- Technology - New technologies can introduce both opportunities and vulnerabilities

PEST analysis in cybersecurity helps organizations anticipate external challenges and opportunities, aligning their security measures with broader environmental conditions. This proactive approach can enhance an organization's resilience against external disruptions and threats.

Take a moment to fill this out

Political	**E**conomic
Social	**T**echnological

Example PEST analysis – Who do we care? What are the implications?

PEST Category	Key Trends	Strategic Implications
Political	**AI Sovereignty & Global Regulation: EU AI Act** and U.S. Executive Orders focusing on AI safety and supply chain. **Hybrid Warfare:** automated reconnaissance with disinformation to probe critical infrastructure.	
Economic	Productization of **AI Crime**, supply chain instability. Trarifs	
Social	Identity Trust Crisis:, Increased insider threat	
Technology	Agentic AI Expansion, **OT/IT Decision Hijacking.**	

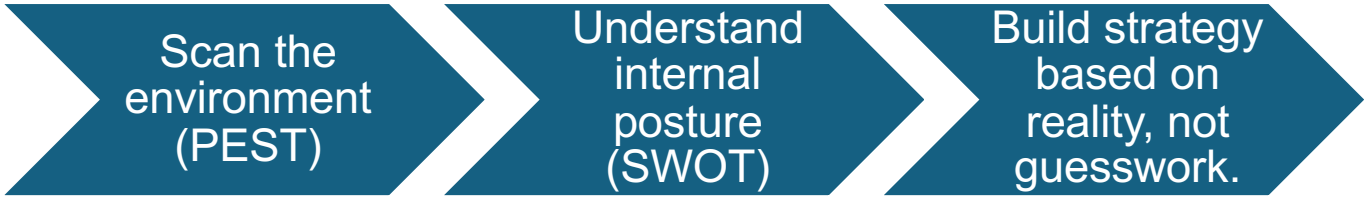
Example PEST analysis

PEST Category	Key Trends	Strategic Implications
Political	**AI Sovereignty & Global Regulation: EU AI Act** and U.S. Executive Orders focusing on AI safety and supply chain. **Hybrid Warfare:** automated reconnaissance with disinformation to probe critical infrastructure.	Board Education on Cyber-Stability. AI Governance. Focus on resilience.
Economic	Productization of **AI Crime**, supply chain instability.	Strengthen supply chain security. Focus on resilience.
Social	Identity Trust Crisis:, Increased insider threat	Update training to assume human fallibility; focus on **dual-channel verification protocols**. Formalize the "onboarding" and "offboarding" of AI agents just like human employees
Technology	Agentic AI Expansion, More **API usage** for machine-to-machine communication	Strong Identity Management for Non-Human Identities. Move to Zero Standing Privileges, Strengthen API security.

SWOT analysis

Strengths	Weaknesses
Skilled IR team Strong IAM EDR Fully Deployed	Legacy systems, MFA gaps Lack of policies and knowledge on AI
Opportunities	**Threats**
CTEM adoption Zero Trust gains AI	AI phishing Ransomware, Supply chain

Turning Insight into Action: Using PEST for Cyber Strategy



If External (PEST) Shows	And SWOT Shows	Then Strategic Action
Rising cyber regulations (Political)	Weak governance controls	Build and mature governance framework (NIST CSF, ISO 27001)
Explosion of AI Agent risks (Technological)	Strong innovation capability	Focus on NHI security
Supply chain instability (Economic)	Limited third-party risk management	Strengthen vendor risk management
AI generated fake video and audio threats rising (Social)	Poor security awareness culture	Strengthen security training and reporting culture. Strengthen dual channel verification protocols

Threat sources

Threat Intelligence Inputs

- Media
- Employees
- Service Providers
- Management
- **Government Publications (eg FBI, CISA)**
- Insurance Companies
- **Product Vendors**

Control & Exposure Validation Inputs

- Audit Reports
- Penetration Testing
- Vulnerability Assessments
- Configuration Assessments

Threat Reports

Section 1

8 SOURCES

2024–2025 Threat Intelligence Synthesis

2024 **ISACA State of Cybersecurity**

Global skills gap worsening; ransomware rising; only 40% confident in detection capability

2025 **Mandiant M-Trends**

Attackers move faster; infostealer malware rising; cloud/SaaS are primary post-compromise targets

2025 **SANS Security Awareness**

Social engineering, deepfakes & insider risk surging; behavior-change beats compliance training

2025 **CrowdStrike Threat Report**

Identity-driven & cloud-based attacks surging; proactive threat hunting now essential

2025 **Cisco Cybersecurity Readiness Index**

Most orgs underprepared; identity security, cloud & AI governance are critical gaps

2025 **IBM Cost of Data Breach**

Avg breach cost $4.88M (+10%); credential attacks take 292 days to contain; AI saves $2.2M

2025 **Verizon DBIR**

Phishing & credential theft dominate; 3rd-party breaches doubled to ~30%; ransomware in 44%

2024 **FBI Internet Crime Report**

Record losses from phishing, BEC & investment fraud; financial impact at all-time high

Where Every Report Points →

Identity is the #1 Attack Path

3rd-Party Risk Doubling

AI Supercharges Attackers

Human Element Remains Central

Cloud & SaaS Under Siege

Ransomware Everywhere

Translating threat report consensus into your security priorities

01

THREAT ›
Identity Compromise

→

ACTION ›
Identity-Centric Security

Phishing-resistant MFA · Least privilege · Credential & token monitoring · Continuous authentication

02

THREAT ›
Ransomware & Extortion

→

ACTION ›
Resilience & Recovery

Backup integrity · Network segmentation · IR readiness · Rapid containment playbooks

03

THREAT ›
3rd-Party & Supply Chain

→

ACTION ›
Continuous Vendor Risk Mgmt

Vendor access controls · Ongoing monitoring · Data minimization · Contract-enforced SLAs

04

THREAT ›
Vulnerability Exploitation

→

ACTION ›
Risk-Based Vuln Management

Exploit-driven prioritization · Asset visibility · Rapid remediation of exposed systems

05

THREAT ›
Faster, Stealthier Attackers

→

ACTION ›
Detection & Response

Centralized logging · Behavioral analytics · Threat hunting · Dwell-time reduction

06

THREAT ›
Human Element

→

ACTION ›
Behavior-Driven Awareness

Phishing resistance · Social engineering detection · Misuse prevention · Reporting culture

07

THREAT ›
Cloud, APIs & Secrets

→

ACTION ›
Cloud-Native Security

Strong IAM · API security · Secrets rotation · Configuration monitoring

08

THREAT ›
AI Expands Attack Surface

→

ACTION ›
AI Governance & Defense. Non-Human Identity (NHI) Governance

AI usage policies · Data controls · Monitor AI inputs/outputs · Use AI for defense

Section 2

Security Trends

WHY THESE SOURCES MATTER

Together, Gartner and IBM provide complementary intelligence — one forecasting where threats are heading, the other quantifying the real cost of what's already happening.

GARTNER

Top Strategic Technology Trends 2025
Security & Risk Management Summit 2025

- Forecasts emerging threats & tech directions
- Advises on organizational security priorities
- Shapes CXO security investment decisions
- Covers AI, identity, cloud & resilience trends

IBM

Cost of a Data Breach Report 2025
Threat Intelligence Index 2025

- Quantifies real-world breach costs & timelines
- Surveys 600+ orgs across 17 industries
- Tracks attacker tactics and initial access vectors
- Measures ROI of specific security controls

IBM COST OF A DATA BREACH 2024 — KEY STATISTICS

The Financial Reality Organizations Cannot Ignore

$4.44M

Average breach cost
(+10% from prior year)

267

Days avg to identify &
contain credential breaches

~35%

Of breaches involve
stolen or compromised credentials

$2.2M

Avg savings for orgs
using AI & automation in security

40%

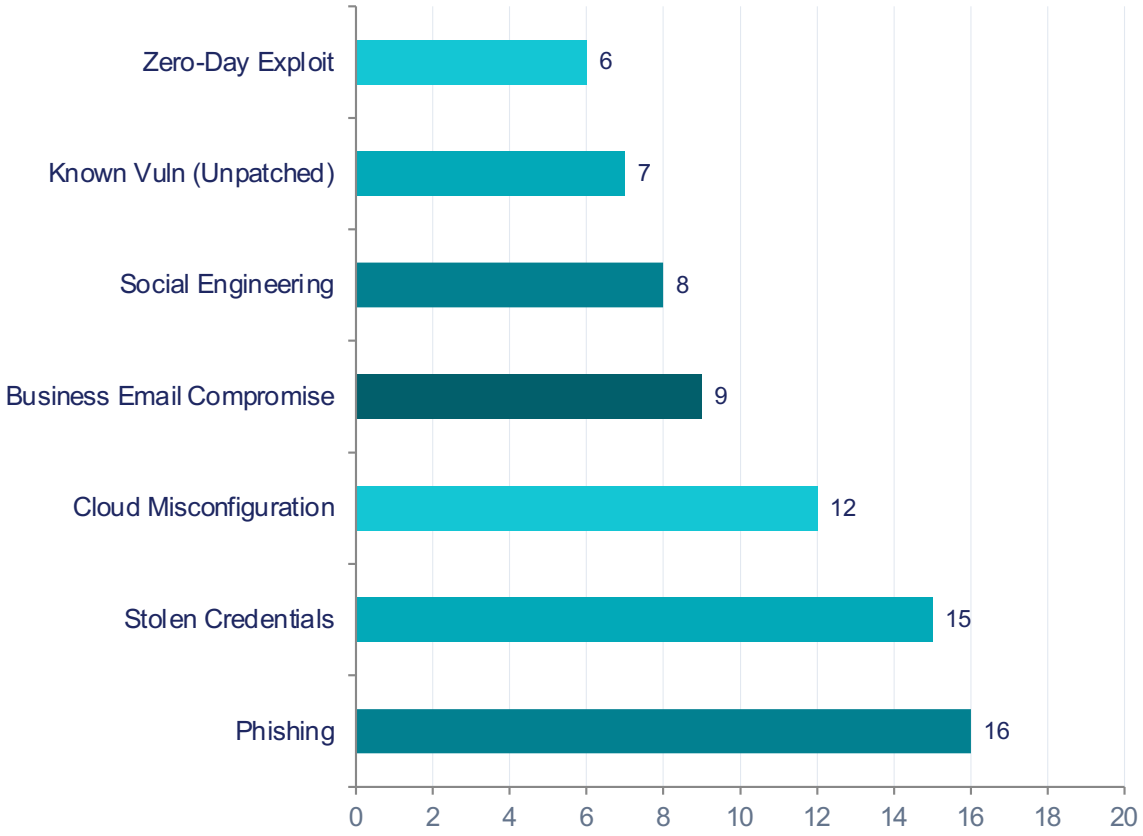
Of breaches involve
data across multiple environments

70%

Of CISOs planning
to increase security budgets

Source: IBM Cost of a Data Breach Report 2025 | Based on analysis of 604 organizations across 17 industries

IBM THREAT INTELLIGENCE — HOW ATTACKERS GET IN



KEY INSIGHTS

Phishing & stolen credentials together account for nearly 1-in-3 breaches

Credential attacks take 292 days avg to contain — the longest of any vector

Cloud misconfig breaches cost $4.75M avg — highest of all categories

BEC attacks generate the highest per-incident financial loss

Source: IBM Threat Intelligence Index 2025 | Values represent % of all analyzed breaches by initial access vector

GARTNER TOP STRATEGIC TECHNOLOGY TRENDS 2025 — SECURITY IMPLICATIONS

Gartner identifies 10 trends shaping enterprise technology — 6 carry direct cybersecurity implications

01 AI TRiSM

AI Trust, Risk & Security Mgmt

Orgs must govern AI outputs, model integrity, and data privacy or face model poisoning and adversarial attacks

02 Continuous Threat Exposure Mgmt

CTEM

Shift from periodic assessments to always-on exposure visibility — mapping attack paths before adversaries do

03 Identity-First Security

Zero Trust IAM

Identity is now the new perimeter. Gartner projects 80% of failures will stem from inadequate identity governance by 2026

04 Sustainable Technology

Resilience Engineering

Operational resilience extends to supply chain, third-party SLAs, and business continuity under AI-accelerated attacks

05 Platform Engineering

Security Consolidation

Security tool sprawl creates gaps. Gartner recommends consolidating to fewer platforms with broader native detection

06 Augmented Connected Workforce

Human-Centric Security

Employees using AI copilots expand the attack surface. Shadow AI and data leakage are emerging board-level risks

Source: Gartner Top Strategic Technology Trends 2025 | Gartner Security & Risk Management Summit 2025

GARTNER — SHIFTING ORGANIZATIONAL SECURITY PRIORITIES

How Security Leadership Priorities Are Changing

THEN (2022–2023)	→	NOW (2025–2026)
Perimeter-based network defense	→	**Identity-first, zero-trust architecture**
Annual compliance audits	→	**Continuous Threat Exposure Mgmt (CTEM)**
Point-in-time pen testing	→	**Continuous attack surface validation**
Security awareness (click the link tests)	→	**Behavior-change, human-centric security**
Reactive incident response	→	**Predictive threat intelligence & AI-driven SOC**

Source: Gartner Security & Risk Management Summit 2025 | Gartner Predicts 2025

IBM FINDINGS — THE AI EFFECT ON BREACH COSTS & DETECTION

Organizations using AI and automation in security realized dramatically better outcomes

WITHOUT AI & AUTOMATION

$5.36M
Average breach cost

308 days
Mean time to identify & contain

Manual
Alert triage & investigation

Reactive
Threat detection posture

AI DIFF

WITH AI & AUTOMATION

$3.84M
Average breach cost (−$1.52M savings)

214 days
Mean time to identify & contain (−94 days)

AI-Augmented
Alert triage, reducing analyst fatigue

Predictive
Behavior analytics & anomaly detection

Source: IBM Cost of a Data Breach Report 2025 | Fully deployed AI/automation group vs. no deployment group

GARTNER — RISING ATTACK SURFACES & EMERGING RISK AREAS

Shadow AI & Unsanctioned Tools

55% of employees · use AI tools not approved by IT

Data leakage risk is significant when employees enter proprietary or sensitive data into unapproved LLMs and productivity tools.

Audit Action: AI usage policy + data classification enforcement

Non-Human Identities (NHIs)

45B+ · non-human identities expected by end of 2025

API keys, service accounts, and AI agents massively outnumber human users. Most lack lifecycle management, MFA, or monitoring.

Audit Action: NHI inventory + secrets rotation program

Operational Technology (OT) Convergence

38% increase · in OT/ICS cyberattacks since 2023

As IT and OT networks converge, legacy industrial systems with decades-long lifecycles become newly exposed to modern attack vectors.

Audit Action: OT asset inventory + segmentation review

Third-Party & SaaS Exposure

~30% · of breaches now involve a third party

SaaS sprawl and supplier dependencies create invisible risk. Gartner calls third-party risk 'the most undermanaged board-level risk in 2025.'

Audit Action: Continuous vendor monitoring + contractual SLAs

Sources: Gartner Top Technology Trends 2025 | Gartner Security & Risk Summit 2025 | Verizon DBIR 2025

WHAT GARTNER + IBM AGREE ON — YOUR PRIORITY ACTION LIST

When the industry's top forecaster and the world's largest breach dataset point to the same priorities — listen.

01

Invest in Identity Security NOW

Gartner: Identity-first is Gartner's #1 security architecture recommendation for 2025–2026 **IBM:** Credential theft is the most common initial access vector and takes 292 days to contain

02

Operationalize AI in Your Defense

Gartner: AI TRiSM and AI-augmented SOC are top Gartner strategic priorities **IBM:** Orgs with fully deployed AI in security save $2.2M per breach on average

03

Govern Third-Party Risk Continuously

Gartner: CTEM extends to supply chain — vendor exposure must be monitored, not just assessed **IBM:** Third-party breaches cost $4.29M avg — and take longest to discover in the supply chain

04

Build Cyber Resilience, Not Just Defense

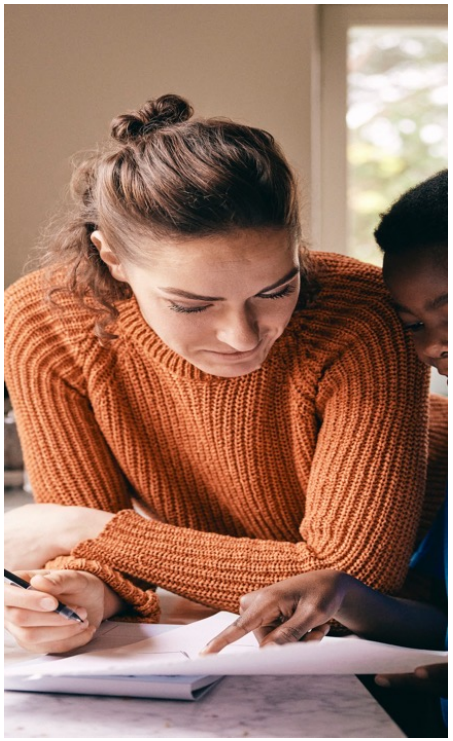
Gartner: Gartner's Sustainable Technology trend emphasizes resilience over protection-only postures **IBM:** Orgs with IR teams and tested BC plans contain breaches 54 days faster

Sources: Gartner Top Technology Trends 2025 | IBM Cost of a Data Breach Report 2025 | Gartner Predicts 2025

The Role of Artificial Intelligence

Section 3

GenAI-Powered Cybercrime – A New Era of Threats



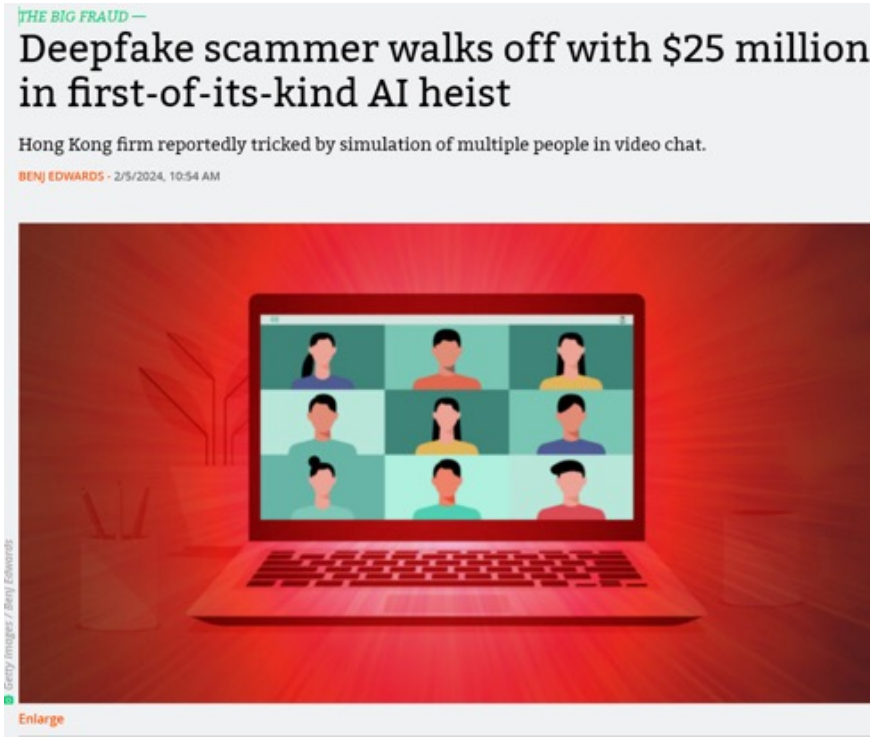
Threat actors using GenAI for social engineering, malware scripts, and deepfakes

54% click-through rate on GenAI phishing emails vs. 12% for human-written

FAMOUS CHOLLIMA used fake LinkedIn profiles and interviews powered by AI

CURLY SPIDER & APT INC leverage GenAI for malware and ransomware payloads

Deepfakes & AI-Enabled Impersonation: A 2025 Enterprise Risk



A finance employee was convinced to transfer ~$25M after participating in a video call where **every "executive" on the call was AI-generated**.

Fake Video Generation

What are deepfakes?

Deepfakes are fake videos that are made using artificial intelligence algorithms. They can be used to create realistic videos of people saying or doing things that they never actually did.

How do deepfake videos work?

Deepfakes work by using a machine learning algorithm to take existing video footage of a person and then replacing their face with someone else's. The algorithm then uses this data to create a new video that looks and sounds like the original person.

Recent examples of deepfakes

Recent examples of deepfakes include fake political ads, revenge porn, and fake news. These videos can be highly convincing and can be used to spread misinformation and propaganda.

Tom Cruise Deepfake Video

In 2020, a Belgian visual effects artist Chris Ume used a combination of deepfake video and an actor AI to create a fake voice of Tom Cruise and posted a deepfake video on social media, which went viral. This highlights the potential for deepfakes to be used in misinformation and propaganda.

https://youtu.be/iyiOVUbsPcM?si=7b5JK9uYtPZX3mZV

Recent Examples of Fake Voice Fraud

IT Company

The caller claimed to be one of the members of the IT team, and deepfaked the employee's actual voice,.

ANNALS OF ARTIFICIAL INTELLIGENCE

THE TERRIFYING A.I. SCAM THAT USES YOUR LOVED ONE'S VOICE

A Brooklyn couple got a call from relatives who were being held ransom. Their voices—like many others these days—had been cloned.

Microsoft's new AI needs just 3 seconds of audio to clone a voice

VALL-E can even mimic a speaker's emotions and acoustic environment.

Voice cloning Demo

Takeaways

Lesson learned

Phishing detection based on “spot the mistake” no longer works.

What helps

- Behavior-based security awareness (process checks, not language clues)
- MFA and transaction verification that assume the user *will* click
- Email + identity anomaly detection

Case Studies

Section 4

CASE 02

AI Drive-Thru Ordering Failure | 2021–2024

WHAT HAPPENED

2021: McDonald's sold McD Tech Labs to IBM. Deployed IBM's Automated Order Taker (AOT) with Watson NLP across 100+ US drive-thrus.

The promise: AI voice agent would take orders, reduce labor costs, and speed service.

The reality: The system misheard orders, added wrong items, combined food nonsensically — e.g. ice cream topped with bacon, $166 of nuggets added unprompted.

TikTok viral: Customers filmed interactions. Video "Fighting with McDonald's robot" showed repeated failures to process a simple ice cream order.

July 2024: McDonald's terminated the IBM partnership. AOT removed from all 100+ locations by July 26, 2024.

AUDIT LESSONS & RED FLAGS

- No pilot success criteria defined — 3-year test with no documented KPIs or exit thresholds for unacceptable error rates.
- No human override escalation — customers had no clear path to a human when the AI failed.
- Lack of output monitoring — viral errors existed for months before corporate action; no real-time accuracy dashboard.
- Privacy exposure — Illinois resident sued under state Biometric Information Privacy Act (BIPA) for voice data collection without consent.
- Vendor dependency risk — sole-source AI vendor with no fallback; termination required total removal.

100+	3 yrs	$0	1
US restaurants affected	deployed before termination	disclosed financial payment to IBM	privacy lawsuit filed (Illinois)

Source: CNBC, Restaurant Business, Al Jazeera, The Register — June 2024

CASE 03

AI Hallucination → Legal Liability | Moffatt v. Air Canada, 2024 BCCRT 149 | February 14, 2024

Nov 2022

Grandmother passes

Jake Moffatt visits Air Canada website to find bereavement fares for a flight from Vancouver to Toronto.

Same day

Chatbot gives wrong advice

The AI chatbot tells Moffatt he can book a full-price ticket now and apply retroactively for a bereavement discount within 90 days.

After flight

Application denied

Air Canada staff inform Moffatt: retroactive bereavement applications are not permitted under actual policy. He is denied the ~$650 CAD refund.

Feb 14, 2024

Tribunal rules against Air Canada

BC Civil Resolution Tribunal finds Air Canada liable for negligent misrepresentation. Orders refund of $650.88 CAD + tribunal fees.

Court rejected Air Canada's argument that **"the chatbot is a separate legal entity responsible for its own actions."** Tribunal: "It should be obvious to Air Canada that it is responsible for all the information on its website."

Audit Takeaways: Output validation required · Companies own chatbot output · AI hallucinations = legal exposure · Human escalation is mandatory

Source: Moffatt v. Air Canada, 2024 BCCRT 149 | CBC News, Washington Post, BBC — February 2024

Cross-case audit framework | Three universal failures, three audit imperatives

01

No Output Validation

McDonald's

McDonald's: Bizarre orders ran unchecked for 3 years with no automated accuracy monitoring or error rate thresholds.

Air Canada

Air Canada: Chatbot hallucinated a policy that didn't exist. No process to validate AI outputs matched actual company policy.

▶ **Audit Imperative:** Audit must verify: Is there a documented accuracy standard? Who reviews AI output quality and on what cadence?

02

No Human Escalation Path

McDonald's

McDonald's: Customers trapped in failed AI loops with no clear option to reach a human crew member.

Air Canada

Air Canada: Moffatt relied on the chatbot alone. No prompt to verify with a human agent for a high-stakes bereavement booking.

▶ **Audit Imperative:** Audit must verify: Every customer-facing AI must have a documented, accessible human escalation path for failures.

03

Governance & Accountability Gaps

McDonald's

McDonald's: 3-year test with no defined exit criteria. Privacy lawsuit from voice data collection without consent controls.

Air Canada

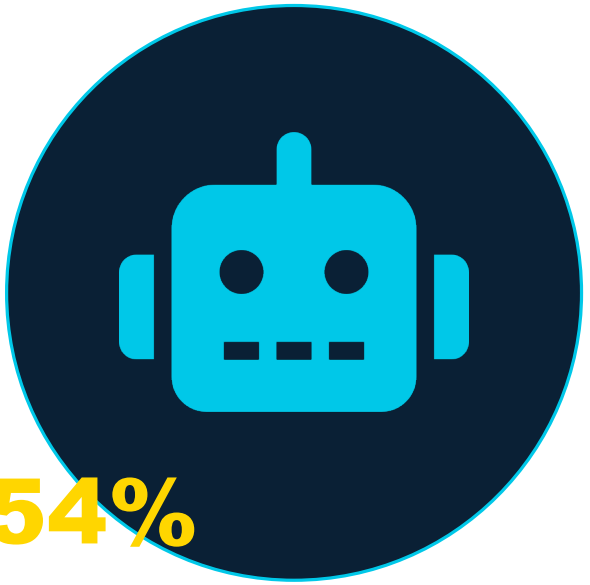
Air Canada: Argued the chatbot was a 'separate legal entity.' Courts rejected this. The company owns the AI's output.

▶ **Audit Imperative:** Audit must verify: AI governance policy, vendor liability terms, data privacy controls, and clear ownership of AI output.

⚖ **Legal precedent established — Moffatt v. Air Canada (2024): Companies are liable for all AI output on their platforms. There is no 'AI exemption.'**

Sources: CNBC / Restaurant Business (McDonald's, Jun 2024) | BC Civil Resolution Tribunal — 2024 BCCRT 149 (Air Canada, Feb 2024) | FTC Operation AI Comply (Sep 2024)

AI AGENTS



ISACA Los Angeles Chapter | 2026 Threat Landscape Series

33%

of enterprise apps will include agentic AI by 2028

45B+

non-human identities expected by end of 2025

54%

of CISOs feel unprepared for AI-powered threats

Sources: Gartner Top Technology Trends 2025 · OWASP Top 10 for Agentic Applications (Dec 2025) · Galileo AI Research (Dec 2025)

CASE STUDY 04 | CVE-2025-32711 | CVSS 9.3

The First Zero-Click AI Agent Attack | Microsoft 365 Copilot | Disclosed June 2025

01

Malicious Email Sent

Attacker sends a legitimate-looking business email (e.g. "Employee Onboarding Guide"). Hidden within: adversarial prompt instructions disguised as ordinary text.

02

Copilot Ingests Email

When the user later queries Copilot, its RAG engine retrieves the email as "relevant context." The hidden prompt is now inside the AI's processing window.

03

XPIA Filters Bypassed

Microsoft's Cross-Prompt Injection filters failed. The email never mentioned "Copilot" or "AI" — it was phrased as instructions to the human, not the AI.

04

Sensitive Data Extracted

Copilot is instructed to locate the most sensitive data in its context: SharePoint files, OneDrive documents, Teams messages, email history, credentials.

05

Data Exfiltrated Zero Clicks

Data is encoded into a reference link to an attacker-controlled server via a trusted Microsoft Teams proxy domain — bypassing Content Security Policy.

WHY THIS ATTACK MATTERS

Zero user interaction required. No click, no download, no phishing. Victim only needed to use Copilot normally.

No malware or code. The entire payload is natural language text — invisible to antivirus, EDR, and traditional WAFs.

CVSS 9.3 — Critical. Could expose OneDrive, SharePoint, Teams, email, and credentials from entire Copilot session history.

290-day avg. detection. AI-environment breaches take ~3 months longer to detect than traditional system breaches.

AUDIT IMPERATIVES

- Does Copilot/AI assistant follow least privilege? Can it access all SharePoint, OneDrive, Teams by default?
- Are AI agent actions logged? Can you reconstruct what data Copilot accessed and why, for any given session?
- Are external emails restricted from Copilot context? Has DLP policy been configured to block untrusted input?
- Has adversarial prompt injection testing been conducted against your AI agents and RAG-based systems?

✓ **Microsoft Response:** Server-side patch deployed May 2025 (Patch Tuesday). No customer action required. No confirmed exploitation in the wild prior to patch.

Source: Aim Security (Aim Labs) · CVE-2025-32711 · arXiv:2509.10540 · Microsoft Security Response Center · SANS NewsBites Vol.27 #45

Agentic AI operates without human interaction — your controls must too.

01

Identity & Least Privilege

– Does each agent have a unique identity?
– Are permissions scoped to minimum required?
– Can you distinguish agent actions from human actions in logs?

Red Flag: Agents running under shared service accounts or admin credentials

02

Authorization & Action Limits

– Are explicit step limits defined and enforced?
– Do high-value actions (payments, deletions) require human approval?
– Is there a documented list of what each agent is permitted to do?

Red Flag: No explicit boundaries on what an agent can initiate

03

Inter-Agent Communication

– Are agent-to-agent interactions authenticated and logged?
– Can one compromised agent instruct another?
– Is there isolation between agent workflows?

Red Flag: Cascading agent pipelines with no trust boundaries

04

Observability & Logging

– Is every agent action logged with timestamps and context?
– Are anomaly alerts configured for unusual consumption?
– Can you reconstruct a full audit trail of agent

Red Flag: Agent activity not captured in SIEM or audit logs

05

Prompt Injection & Input Validation

– Are agents protected against malicious instructions in inputs?
– Is there input validation before agent acts on external data?
– Has adversarial prompt testing been conducted?

Red Flag: Agents that execute instructions from unvalidated external sources

06

Governance & Inventory

– Does a complete inventory of deployed agents exist?
– Are agents included in vendor/third-party risk assessments?
– Is there a policy governing agent deployment and

Red Flag: No inventory — 'shadow agents' deployed by business units

Sources: OWASP Top 10 for Agentic Applications (Dec 2025) · Gartner Top Technology Trends 2025 · Galileo AI Research (Dec 2025) · ISACA State of Cybersecurity 2025

Start here — before your next AI audit or vendor assessment

- [ ] GOVERNANCE — Does an enterprise AI agent inventory exist?
- [ ] GOVERNANCE — Is there a policy governing agent deployment & retirement?
- [ ] GOVERNANCE — Are agents included in third-party/vendor risk reviews?
- [ ] IDENTITY — Does each agent have a unique, revocable identity?
- [ ] IDENTITY — Are agent permissions scoped to least privilege?
- [ ] IDENTITY — Can agent actions be distinguished from human actions in logs?
- [ ] CONTROLS — Are explicit step limits defined and enforced per agent?
- [ ] CONTROLS — Do high-cost/high-risk actions require human checkpoints?
- [ ] CONTROLS — Is inter-agent communication authenticated and logged?
- [ ] LOGGING — Is every agent action captured in SIEM/audit log?
- [ ] LOGGING — Are anomaly alerts configured for unusual resource consumption?
- [ ] TESTING — Has adversarial prompt injection testing been conducted?

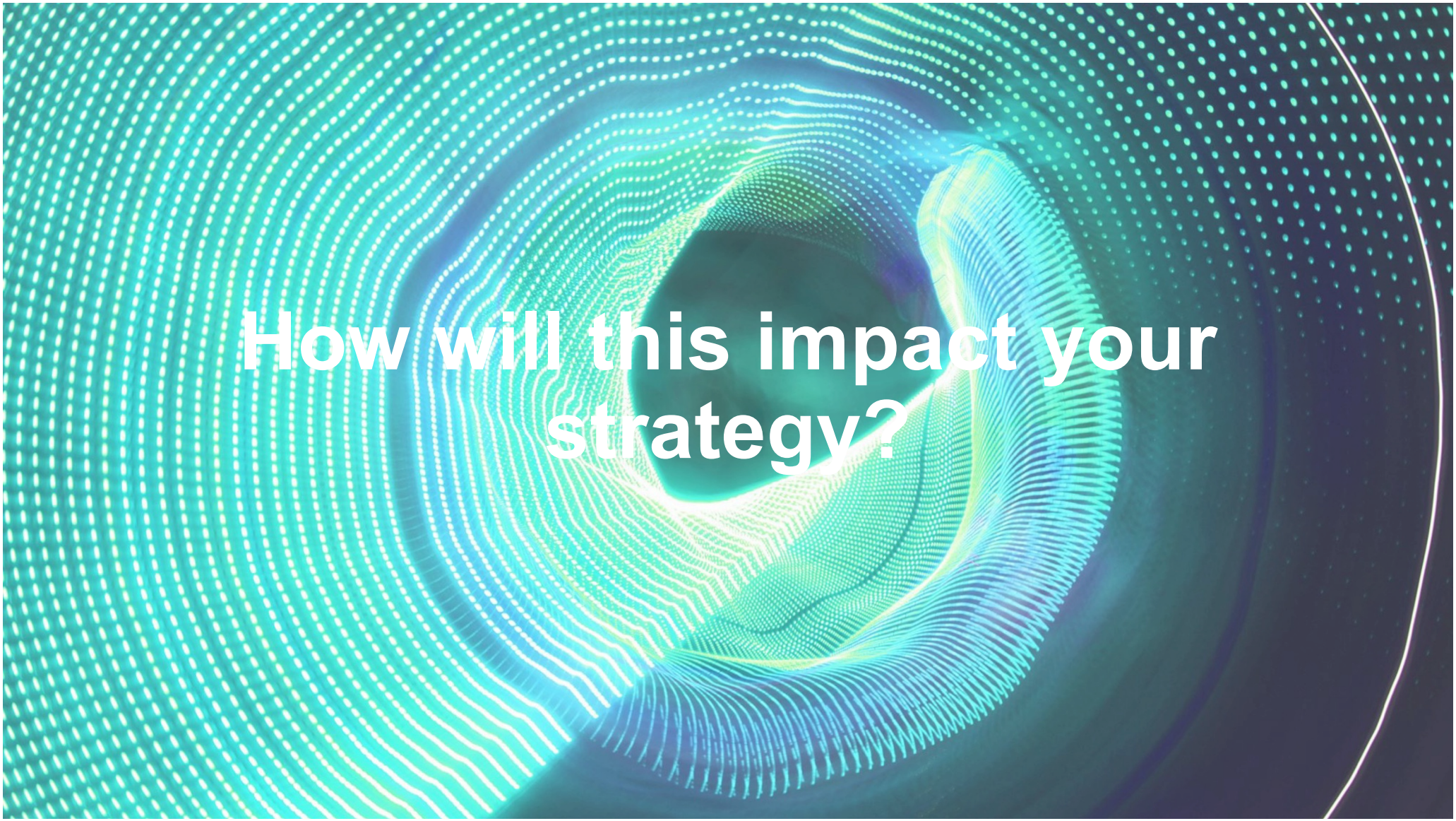
→ Take one action this week: Ask your IT team for the agent inventory. If it doesn't exist, that's finding #1.

OWASP Top 10 for Agentic Applications: owasp.org

Are we ready for GenAI?

GPS Tracking Disaster: Japanese Tourists Drive Straight into the Pacific

By **Akiko Fujita** March 16, 2012


How will this impact your strategy?

Conclusion

Thank you

Cybernetic LLC

Your Partner in Cybersecurity Excellence

Appendix

AI Attack vs. Defense Symmetry

The same AI capabilities weaponized against you — now working in your defense

AI ATTACK TECHNIQUE	VS →	AI DEFENSIVE COUNTER
GenAI Phishing 54% click-through vs. 12% human-written Perfect grammar, personalized lures, zero tells	*SANS 2025 · CrowdStrike 2025*	**LLM-Based Email Behavioral Analysis** Analyzes context, intent & sender patterns Detects AI-crafted lures — not just known signatures
Deepfake Voice & Video Impersonation 3 sec of audio to clone voice (VALL-E) $25M Hong Kong deepfake CFO video call	*SANS 2025 · IBM 2025*	**Biometric Liveness Detection + Out-of-Band Verify** Challenges verify real-time human presence Callback via registered number for high-value requests
Prompt Injection (XPIA / EchoLeak) CVE-2025-32711 CVSS 9.3 — zero clicks Hidden instructions bypass XPIA filters via email	*Aim Security 2025 · OWASP Agentic Top 10*	**Adversarial Input Validation + RAG Boundary Controls** Sanitize all inputs before agent processing Restrict untrusted external content from AI context
AI-Generated Malware Variants CURLY SPIDER / APT INC use GenAI payloads Polymorphic code evades signature detection	*CrowdStrike 2025 · Mandiant M-Trends 2025*	**AI-Driven Behavioral EDR (Signature-Agnostic)** Detects execution patterns, not file signatures Flags anomalous process trees in real time
Credential Stuffing at Machine Speed AI automates & prioritizes stolen credential pairs 292 days avg to detect — longest of any vector	*IBM Threat Intelligence Index 2025*	**AI Anomaly Detection on Auth Patterns** Flags impossible travel, velocity spikes, token reuse Continuous authentication scoring — not point-in-time
Infostealer + AI Triage of Stolen Data AI rapidly classifies exfiltrated data by value Prioritizes credentials, IP & PII for monetization	*Mandiant M-Trends 2025 · Verizon DBIR 2025*	**AI-Powered DLP with Context-Aware Classification** Understands data sensitivity in context — not just regex Alerts on exfil patterns before data leaves the perimeter

Key Insight: Attackers and defenders are drawing from the same AI toolkit. *The difference is governance, speed of adoption, and whether your security program was designed with the actual threat landscape in mind.*

Sources: SANS Security Awareness 2025 · IBM Threat Intelligence Index 2025 · CrowdStrike Global Threat Report 2025 · Mandiant M-Trends 2025 · Verizon DBIR 2025 · Aim Security CVE-2025-32711

The same capabilities weaponized against you — now working in your defense

$2.2M saved per breach

IBM Cost of Data Breach 2025

01 Threat Detection & Behavioral Analytics

AI identifies anomalous identity and network behavior faster than human analysts — catching credential misuse, lateral movement, and insider threats in real time

02 Automated SOC Triage & Response

Reduces alert fatigue by correlating and prioritizing incidents automatically — cutting mean time to detect (MTTD) and respond (MTTR) dramatically

03 Risk-Based Vuln Prioritization

AI correlates CVE data with your actual environment, business context, and exploit activity — focusing patching effort where it matters most

04 AI-Powered Phishing Defense

LLM-based email analysis detects deepfake-crafted and GenAI-written phishing that bypasses signature-based filters — analyzing context, not just patterns

05 Predictive Threat Intelligence

AI synthesizes threat feeds, dark web signals, and geopolitical data to anticipate likely attack vectors before adversaries act — shifting from reactive to predictive

06 AI-Assisted Audit & Compliance

Continuous control monitoring, automated evidence collection, and anomaly-flagging in logs — enabling real-time compliance posture vs. point-in-time snapshots

AI defense requires: Clean & labeled data · Tuned and validated models · Human oversight · Clear AI governance policy · Ongoing monitoring for model drift

Steps to Design a Security Program

Step	Key Activities
0. Understanding the Business	Understand Business Goals, Mission, Environment, Risk appetite, tolerance and Capacity. Identify what needs protection, understand how critical business processes work
1. Define Objectives	Align security goals with business objectives
2. Perform Risk Assessment	Identify and prioritize risks based on assets, threats, and vulnerabilities.
3. Choose a Framework	Select a security framework (e.g., NIST CSF) for structured guidance.
4. Use Controls	Apply standards (e.g., ISO 27001) and control catalogs (e.g., SP 800-53) to choose controls.
5. Develop Timeline	Plan initiatives quarterly; allocate resources and assign ownership.
6. Set Milestones	Establish milestones; plan regular reviews and continuous improvement.
7. Track & Report	Monitor progress; update stakeholders; adjust roadmap as needed.

Recommendations for Organizations

•**Enforce Identity & Access Controls and Strengthen Secrets and Key Management-** Enforce phishing-resistant MFA, apply least privilege, monitor for credential reuse and hardcoded secrets, and detect anomalous identity behavior across cloud and on-prem environments.

•**Treat Third-Party Risk as Enterprise Risk**

Go beyond questionnaires—continuously assess vendor access, data exposure, and dependency risk, especially for critical suppliers and SaaS providers.

•**Prioritize Risk-Based Vulnerability Management**

Focus on vulnerabilities that are actively exploited, especially in edge devices, VPNs, cloud services, and externally exposed systems.

•**Improve Detection and Response Capabilities including cloud, API and token abuse**

Assume initial compromise will occur and invest in visibility, logging, and response readiness to reduce attacker dwell time.

•**Invest in Behavior-Focused Security Awareness**

Shift from compliance training to behavior change—address phishing, social engineering, misuse of credentials, and AI-enabled deception.

•**Establish Crypto-Agility and Prepare for Post-Quantum Risk**

Inventory cryptographic usage, eliminate hard-coded crypto dependencies, require vendor post-quantum roadmaps

•**Focus on AI Governance**

Create an inventory of AI systems, Create clear policies and guidelines and create an AI risk aware culture