EXECUTIVE BRIEFING 05

# Securing Generative AI in the Enterprise

Protecting sensitive information, models and workflows as generative AI becomes embedded in everyday work.

| DATA | ACCESS | OUTPUT | MONITOR |
|---|---|---|---|
| Prevent leakage | Control use | Validate results | Detect misuse |

## EXECUTIVE PREMISE

**Generative AI is becoming a general-purpose interface to enterprise information and work. Employees use it to draft, analyze, code and search; applications embed it into customer and operational workflows. Security must therefore move beyond blocking public tools and create a controlled path for productive adoption.**

## Executive Summary

Generative AI is becoming a general-purpose interface to enterprise information and work. Employees use it to draft, analyze, code and search; applications embed it into customer and operational workflows. Security must therefore move beyond blocking public tools and create a controlled path for productive adoption.

For leadership, the objective is to translate emerging AI capability into controlled enterprise value. That requires visibility, accountable ownership, proportionate safeguards and evidence that controls continue to work after deployment.

## Why This Matters Now

- Users may disclose confidential information to services with inappropriate data handling terms.
- Prompt injection and malicious content can manipulate AI-enabled applications or connected tools.
- Generated code, analysis and factual claims may be wrong while appearing highly credible.
- Retrieval systems can expose information that a user was never authorized to see.
- Plugins, connectors and model supply chains expand the attack surface beyond the model itself.

## Executive Takeaways

| | |
|---|---|
| **1. Establish visibility** | Know where material AI is used and who owns it. |
| **2. Govern by consequence** | Apply stronger controls where decisions or exposure matter most. |
| **3. Secure by design** | Build safeguards into data, identity, architecture and workflow. |
| **4. Preserve accountability** | Automation never removes executive or business ownership. |
| **5. Assure continuously** | Monitor outcomes, incidents, suppliers and changing conditions. |

# The Yakamichi Executive Framework

**1. Approved Use** Define sanctioned platforms, permitted data classes and acceptable business use.

**2. Identity & Access** Use enterprise identity, least privilege and role-aware access to AI applications and connected data.

**3. Data Protection** Prevent sensitive prompts, files and retrieval sources from crossing unauthorized boundaries.

**4. Application Security** Threat-model prompt injection, insecure output handling, plugins, APIs, secrets and agent/tool integrations.

**5. Output Assurance** Require validation for consequential decisions, code, legal/financial content and external publication.

**6. Monitoring & Response** Log material use, detect policy violations, investigate incidents and adapt controls as models change.

# Maturity Model

| Level | Characteristics | Executive Meaning |
|---|---|---|
| **1 — Ad Hoc** | Unmanaged experimentation; fragmented ownership. | Risk is poorly understood. |
| **2 — Emerging** | Initial policy, reviews and accountable owners. | Controls exist but are inconsistent. |
| **3 — Managed** | Enterprise process, risk-tiering and reporting. | Decisions become repeatable and auditable. |
| **4 — Adaptive** | Continuous monitoring, testing and improvement. | Governance supports confident scale. |

# Operating Principle

The control environment should be proportionate to consequence. Low-risk experimentation should remain usable, while high-impact systems receive deeper assessment, stronger approval, tighter access and more continuous assurance.

## Board and Executive Responsibilities

- Which generative-AI services are employees actually using today?
- What information is prohibited from entering external or unapproved models?
- Can AI-enabled applications retrieve data beyond the user's underlying authorization?
- Who validates high-impact generated outputs before they become actions or published information?
- How are prompt injection, data leakage and third-party model changes monitored?

## Common Failure Patterns

| | |
|---|---|
| **Policy without ownership** | Rules exist but no executive or business owner is accountable for outcomes. |
| **Inventory gaps** | Leadership cannot govern systems it does not know exist. |
| **One-time review** | Approval is treated as permanent even as models, data and suppliers change. |
| **Technology-only governance** | Legal, operational, ethical and business consequences are separated from technical risk. |
| **Unclear escalation** | Teams lack thresholds for pausing use, accepting risk or escalating material issues. |

## Metrics That Matter

Executive reporting should emphasize material systems inventoried, high-risk assessments completed, overdue remediation, policy exceptions, incidents, supplier exposure, monitoring coverage, human-oversight effectiveness and unresolved risks accepted above defined thresholds.

## 180-Day Executive Action Plan

**0–30 DAYS | Discover usage**

Identify approved and shadow generative-AI use, data flows and high-risk workflows.

**31–60 DAYS | Establish safe channels**

Publish acceptable-use rules and provide approved enterprise tools with appropriate data protections.

**61–90 DAYS | Secure integrations**

Review retrieval, plugins, APIs, identities, secrets and authorization boundaries.

**91–120 DAYS | Test applications**

Perform security and misuse testing for prompt injection, leakage and unsafe output handling.

**121–180 DAYS | Operationalize assurance**

Introduce monitoring, incident response, user guidance and periodic reassessment of models and suppliers.

## Executive Checklist

- ☐ Approved generative-AI platforms defined
- ☐ Acceptable-use and prohibited-data rules published
- ☐ Enterprise identity and access controls enabled
- ☐ Sensitive-data protections applied
- ☐ Retrieval authorization tested
- ☐ Plugins, connectors and APIs reviewed
- ☐ Prompt-injection testing performed
- ☐ High-impact outputs require validation
- ☐ Usage and security events monitored
- ☐ Model and supplier changes periodically reassessed

## Yakamichi Perspective

**Generative AI security works best when the enterprise offers a safe road instead of merely erecting roadblocks. Governance, secure architecture and practical user guidance can preserve both productivity and information protection.**

The aim is disciplined adoption: enough governance to protect intelligence, secure operations and manage enterprise risk without suffocating legitimate innovation.

### REQUEST A CONFIDENTIAL CONSULTATION

United States: +1 202 413 3763 | Maryland, USA
Nigeria: +234 811 095 0651 | Abuja, Nigeria

YAKAMICHI • Artificial Intelligence • Cybersecurity • Counter-Surveillance • Corporate Risk Intelligence