# Document Title: Generative AI Standard

## Revision: 03

## Owner Department: Information Security

### Table of Contents

## Table of Contents

## 1. Purpose

This document provides detailed standards and guidelines for implementing the Generative AI Policy.

## 2. Scope

This standard applies worldwide to all Western Digital directors, officers, employees, and anyone else that has access to Company Information Systems or Data utilized for GenAI including workers, contractors, and other third parties (hereinafter collectively referred to as "you" or "workforce"). Western Digital or Company should be defined here or elsewhere in the Policy to be Western Digital Corporation and all of its majority-owned subsidiaries.

Applies to all departments using or developing GenAI tools for tasks including but not limited to:

- Text generation (e.g., reporting, documentation).
- Code generation (e.g., script automation).
- Image or video generation (e.g., defect simulations).
- Conversational AI (e.g., internal copilots, chatbots).
- Document summarization or translation.
- Designing and product prototyping.

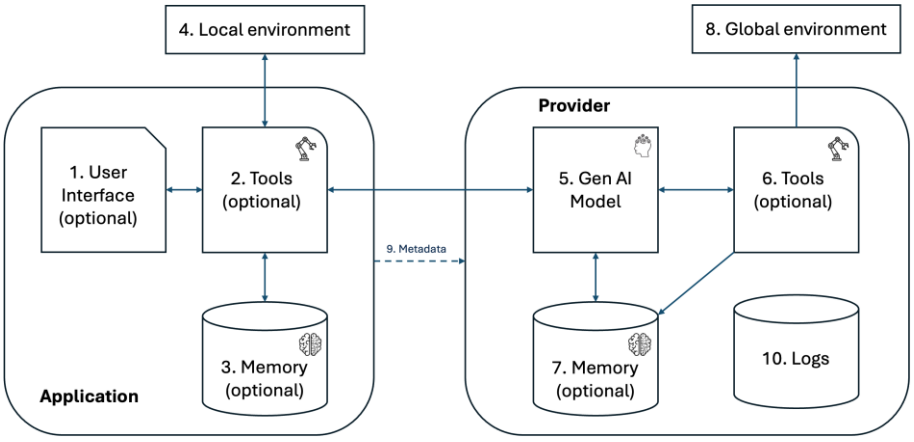
### 2.1 Exclusions

Does not apply to non-GenAI related efforts and development.

## 3. Standards

### 3.1 GenAI Categories

The following is a simplified structure of a GenAI solution to help better illustrate the typical structure and building blocks of GenAI solutions, and the associated risks that can be introduced.




1. Interface (typically user interface); optional if the application is an unattended agent.
2. Tool component that provides integration with a local environment and knowledge; optional if the application relies on tool integrations running on the provider side, or is not used at all.
3. Local knowledge base e.g. user's prompt, in-app vector database, preferences, etc.; optional if the application relies on knowledge stored on the provider side.
4. Local environment accessible to tools, e.g. internal file stores or databases; optional as not always used in a specific implementation.
5. GenAI model, analogous to the brain of the solution.
6. Tools running in provider's environment, e.g. on IaaS platform or a local server.
7. Knowledge base on the provider side; might contain personal, workgroup, and enterprise settings and knowledge objects, e.g. a vector store.
8. Global environment, e.g. other cloud apps, databases, or Internet as a whole.
9. Metadata and telemetry collected by provider; optional if Western Digital is the provider.
10. Metadata and telemetry storage that might be used by the provider for GenAI model training and other purposes.

Commented [JD1]: @Slava Feige requesting to please update visual and annotations to match revised language

### 3.2 GenAI building blocks

The following tables contain further information on GenAI building blocks, common examples, and the corresponding approval requirements. For systems using aspects in different approval categories, all approval requirements must be considered.

#### 3.2.1 GenAI User Interfaces

**User Interfaces** include UI components and agent systems that interact with users and GenAI models. They can collect usage data, alter communications, and execute actions similar to those of a logged-in end user.*

| User Interfaces | Examples | Approval requirements |
|---|---|---|
| Business-to-Consumer (B2C) | ChatGPT, Microsoft Copilot | Enterprise vendor agreement as required per Procurement policy. |
| Business-to-Business (B2B) | ServiceNow, Salesforce | Enterprise vendor agreement as required per Procurement policy, including Third Party Risk Management review. |
| Open-Source | LangFlow, Microsoft Agent Swarm | GenAI review as required per the Corporate GenAI Policy. |
| Custom | Built within systems of record today | GenAI review as required per the Corporate GenAI Policy. |

** See Appendix A for details.*

#### 3.2.2 GenAI Models

**GenAI models** generate data such as text, images, markup data, and commands in response to provide input. A model's behavior is determined by the training conducted by the base model creator (lab) and modifications introduced by model suppliers.

| GenAI Models | Examples | Approval requirements |
|---|---|---|
| Base models from established US-based labs/companies | ChatGPT4.5, ChatGPT-O1, Sonnet, Gemini, Grok, LLama | Per this policy, a GenAI review is required and must include a Legal & Compliance review. |
| Base models from established EU-based labs/companies | Mistral | |
| Base models from the rest of the world-based labs/companies | DeepSeek, QwQ, Kandinski | |
| Open-source models in unmodified form (original weights) published by original producers | Open-source model weights published on Hugging Face by their authors | |
| Open-source models in packaged form that can be validated to original weights | Packages, e.g. GGUF, produced from original models by 3rd parties | Per this policy, a GenAI review is required and must include a Legal & Compliance review and a Supply Chain risk management review. |
| Fine-tuned models from specialized labs governed by a supplier agreement | ServiceNow | Per Procurement policy, an Enterprise vendor agreement and a Third-Party Risk Management review are required. |
| Fine-tuned models from suppliers not governed by a supplier agreement | Fine-tuned models published on Hugging Face by third parties, e.g. Gemma Writer. | Per this policy, a GenAI review is required and must include a Legal & Compliance review. |
| Specialized models from labs governed by a supplier | AlphaFold | Per Procurement policy, an Enterprise vendor agreement and a Third-Party Risk Management review are required. |

#### 3.2.2.1 Model selection

- Where possible, utilize GenAI models already approved by the GenAI enablement program.
- Internal LLMs must comply with security and hosting policies (cloud/on-prem).

#### 3.2.2.2 Prompt Engineering Standards

- Prompts must be repeatable, auditable, and tested for bias or hallucinations.
- Use prompt templates where possible.

#### 3.2.2.3 Output Validation & Human Oversight

- All outputs from GenAI must be reviewed by human experts before being used in production or shared externally.
- Outputs must be clearly labeled as AI-Generated if distributed beyond internal use.

### 3.3 GenAI Providers

**Providers** supply compute resources for operating GenAI models, cloud-hosted applications and agents. In this role, they have access to model inputs and outputs, as well as additional metadata they may retain. They can also impose specific guardrails, constraints, or cost-related factors.

| Providers | Examples | Approval requirements |
|---|---|---|
| Local on device (edge) | On your laptop, phone, camera: (e.g. Whisper, Humo) | Must not violate Acceptable Use and Cyber Security Policy. |
| Local on-premises | Western Digital data center (e.g. Tabline) | |
| Private cloud | Western Digital's Bedrock instance (e.g. Bedrock hosted LLama-3) | |
| Enterprise cloud | OpenAI models in Enterprise mode, Microsoft Copilot with enterprise license | Per Procurement policy, an Enterprise vendor agreement and a Third-Party Risk Management review are required. |
| Public cloud | Microsoft Copilot free, OpenAI monthly subscription | Per this policy, a GenAI review is required and must include a Legal & Compliance review. |

### 3.4 GenAI Memory

**Memory** provides domain-specific knowledge that alters model output. This includes both pre-trained and acquired (operational) knowledge, e.g. files, databases, user's provided input, etc. An application or agent can rely on several memory stores.

| Memory | Examples | Approval requirements |
|---|---|---|
| Local on device (edge) | Local files | Must not violate Acceptable Use and Cyber Security Policy. |
| Local on-premises | Jira, ServiceNow | Must not violate Acceptable Use and Cyber Security Policy. |
| Private sources | Western Digital's data analytics instance | Per this policy, a GenAI review is required and must include a Legal & Compliance review. |
| Public sources | Wikipedia | |
| Enterprise sources | Glint MCP | Per Procurement policy, an enterprise vendor agreement and a Third-Party Risk Management review are required. |

**Tools** allow a GenAI model to interface with an external environment thus extending its functionality. Examples include a connector to ServiceNow, SQL database, etc. Implementation of these tools is outside of the scope of this policy and the application owner must assess and mitigate risks associated with these tools. For tools that connect to other memory stores, all memory section requirements apply.

### 3.5 Data Standards for GenAI Training & Prompting

Does not apply to non-GenAI related efforts and development.

#### 3.5.1 Data Input Quality

- Inputs must be factual, relevant, and free from sensitive data (e.g., IP, PII).
- Use context-rich prompts to reduce hallucination risk.

#### 3.5.2 Training Data Compliance (For Custom Models)

- Only use licensed or proprietary datasets.
- Do not use copyrighted third-party content without rights.
- Follow corporate data governance and privacy policies.

**Commented [AS2]:** "Private Cloud" "Public Cloud" "Enterprise Cloud" to align with Slide 84 of D4 Risk Assessment Framework and renamed to "Private Sources" "Public Sources" and "Enterprise Sources"

**Commented [JD3]:** Please link to policy

**Commented [JD4R3]:** Thank you. Also interesting point you are making; perhaps other teams have not yet updated everything in Agile. Or maybe we are pointing to an old version??

**Commented [AB5R3]:** Done.
FYI - The Data Governance Policy (last reviewed by Yik Ming) is not created using the latest WD format. In fact, it has not WD logo on it either.

### 3.6 Deployment & Integration Standards

#### 3.6.1 Interface Controls

- Deploy GenAI via secure APIs or internal interfaces with access logging.
- Enforce usage limits, rate limiting, and role-based access.

#### 3.6.2 Security Requirements

- Prevent prompt injection, data leakage, and unauthorized model access.
- Use input/output sanitization and monitor for misuse patterns.

#### 3.6.3 Model Hosting

- Follow enterprise hosting standards (cloud isolation, region compliance, encryption).
- For embedded GenAI, only use containerized, sandboxed environments.

### 3.7 Prohibitive Use Cases

GenAI may be **not** used for:

- Autonomous decision-making in safety-critical systems.
- Any content generation shared with customer without human validation (human-in-the-loop).
- Fake data generation for regulatory reporting.
- Bypassing access control, compliance, or security reviews.

Note: This list of prohibited GenAI use cases is not exhaustive. As GenAI technologies and risks continue to evolve, all use cases—whether listed here or not—must adhere to the company's established GenAI review and approval processes. When in doubt, please consult the GenAI enablement team before proceeding further.

## 4. Guidelines

### 4.1 Evaluation frameworks and guidelines

Western Digital uses the following evaluation frameworks and guidelines to review and approve GenAI use case decisions.

- **Tools Rationalization and Duplication**: Where possible, use cases are aligned to existing systems to reduce duplication or redundancy. If existing systems do not meet the minimum viable requirements, new systems or incremental capabilities are added to the asset portfolio only after the GenAI review and approval process described within this policy are completed.
- **Prioritization**: Use cases are evaluated based on value, effort, and strategic alignment. Value considers revenue increases, cost savings, and/or productivity gains. Effort includes estimation of both initial cost and ongoing cost. Strategic Alignment is dependent on Organizational leadership approval and funding approval where required.
- **Buy vs. Build Approach**: The Build vs. Buy approach determines whether to build a new GenAI solution in-house or buy off-the-shelf by prioritizing scalability, compatibility, and long-term cost of ownership. The approach focuses on the minimum viable need to succeed, using existing assets to avoid redundancy where possible. Data security, compliance, and risk will also be assessed using the standard third party risk management process whenever evaluating third party vendors (see TPRM-Policy for more details).
- **Tools Rationalization**: Use case requirements should be aligned to previously approved assets (tools/platforms/data products/use cases) to reduce duplication or redundancy. If existing tools do not meet the minimum viable requirements of the requestor, incremental capabilities may be added to the asset portfolio following a review process.

Commented [JD6]: Please refer to Procurement policy

### 4.2 Risk Management

The potential risks and benefits of each proposed GenAI use case are assessed using a risk evaluation framework. The evaluation framework assigns a GenAI risk rating (low, medium, high, and prohibited) to each use case. The framework considers GenAI risks around privacy, intellectual property, security, third party, ethical, safety, and performance risks, and recommends applicable controls to mitigate those risks.

### 4.3 Procurement of GenAI Tools

GenAI tools that are procured or licensed from third parties must be reviewed in line with the following considerations:

- **Due Diligence**: Before procuring or licensing a GenAI tool from third parties, review whether the GenAI tool aligns with this or related policies, standards, and procedures.
- **Contracting**: Agreements with third parties providing the GenAI tool must include clauses requiring compliance with this or related policies, standards, and procedures.
- **Integration and Monitoring**: All required metadata (e.g., owner, data sources, etc.) for third party GenAI tools must be maintained in an enterprise-approved AI inventory and undergo ongoing monitoring and evaluation to ensure continued compliance with related policies, standards, and procedures. Any non-compliance issues must be promptly addressed through appropriate remedial actions.

### 4.4 Incident Response and Reporting

Please refer to the **Generative AI Policy** for details.

| Possible Incident Types | Examples |
|---|---|
| Data Leakage & Exposure | Input of sensitive IP into AI models. Theft of data and/or models. |
| Compliance Violations | AI processes data in violation of legal or regulatory compliance requirements. |
| Cyber Attacks | Attempting to steal sensitive information or modify behavior of GenAI systems. |
| Malicious AI Use | Malicious use of AI, bypassing security policies. Includes fraud, scams, and similar activities. |
| AI design and implementation failures | Unauthorized introduction of AI models into the design or production process. Increased attack surfaces through poor design. |

## 5. Review

This standard will be reviewed periodically as necessary to reflect changes in regulations, business needs, or technology. These standards will be reviewed and updated periodically to reflect advances in GenAI technology, emerging best practices, and evolving societal expectations.

## 6. Reference Documents

| Document Number | Document Title |
|---|---|
| To be updated | Generative AI Policy |

## 7. Revision History

| Revision | Changes | Originator | Function |
|---|---|---|---|
| 1.0 | Initial Western Digital Release | Ali Baig | Information Security |
| 2.0 | Used latest Western Digital Template | Ali Baig | Information Security |
| 3.0 | Updated Standards & Guideline Sections.<br>Added links to Policies. | Justin Duren<br>Ali Baig<br>Salman Mohamed | Information Security |