Governing AI in Health and Human Services

Applying the principles of data governance to Copilot and other generative AI in HHS environments

In Time Tec · April 2026

Abstract. As AI enters health and human services, technology leaders are being asked to govern something they cannot fully see. Prompts, retrievals, and AI-generated outputs cross data boundaries drawn before generative AI existed. This paper applies the established principles of data governance to the AI context — the challenges that generative AI introduces, the best practices that address them, and the AI-specific tooling in the Microsoft ecosystem that technology leaders can turn to today.

Why Governance Comes First

For technology leaders in Health and Human Services, the governance question is a consequential one. Agency data are sensitive under HIPAA, 42 CFR Part 2, Title IV-E, the Child Abuse Prevention and Treatment Act, CJIS, and state-level statutes. The populations served include children, elderly residents, benefit recipients, and people in crisis. The cost of a governance failure is measured in lives affected, not just audit findings. And yet the benefits of applying AI to this data are too significant to forgo: faster eligibility determinations, reduced case backlogs, and earlier identification of risk and opportunity.

This paper is structured around six principles of data governance, adapted for the realities introduced by generative AI. The scaffold draws from the NIST AI Risk Management Framework 1.0 and its Generative AI Profile (NIST AI 600-1), the HIPAA Security Rule's administrative, physical, and technical safeguards, and FedRAMP High baselines as implemented in government community clouds.

Six Principles for Governing AI in HHS

1. Data Classification and Sensitivity Labeling

You cannot govern what you have not classified. Every downstream policy depends on knowing, in machine-readable form, what a given document or record is. Generative AI amplifies the cost of getting this wrong: an unclassified file is merely risky in a shared folder, but the same file indexed by semantic search becomes a risk that surfaces in answers to prompts the original owner never anticipated. AI does not invent new access; it reveals access that was already there.

- Establish a classification taxonomy before AI deployment
- Label automatically where possible
- Require labels for new content

Within the Microsoft ecosystem, sensitivity labels from Purview Information Protection travel with content into Microsoft 365 Copilot, Copilot Studio, and Dataverse-backed Copilot experiences. Purview Data Security Posture Management for AI (DSPM for AI) provides a tenant-wide view of where AI is processing which data at which sensitivity. Purview DLP, through its Copilot policy location (generally available since March 2026), can block Copilot from processing prompts or items that match specified sensitive information types or label combinations.

2. Identity, Access, and Least Privilege

The oldest rule of data governance is amplified in an AI context: a user should see what their role permits, and no more. The AI-specific challenge is that semantic search collapses soft boundaries. A shared folder that a user technically had access to but would never have browsed on their own is now one prompt away. Industry surveys consistently identify over-permissioning as a leading cause of Copilot rollout delays and post-deployment data exposure incidents.

- Remediate access sprawl before AI rollout, not after
- Run a permissions audit
- Turn on sharing restrictions
- Apply conditional access
- Adopt just-in-time privilege elevation for administrators

The investment in identity hygiene pays back on every subsequent AI use case, because AI inherits whatever access model it runs on top of.

3. Privacy and Protection of PII, PHI, and AI-Generated Content

The most common privacy misconception about AI is that enterprise services behave like consumer chatbots. A public chatbot that learns from user conversations, has no tenant boundary, and offers no contractual data protection is fundamentally different from an enterprise AI service running under a Business Associate Agreement within a dedicated compliance boundary.

In the enterprise case (i.e., Microsoft 365 Copilot with Enterprise Data Protection, Azure OpenAI Service, and equivalent offerings from other cloud vendors), three protections apply that consumer tools lack:

1. Prompts and responses are logically isolated to the tenant and are not visible to other customers.
2. The vendor contractually commits that Customer Data, prompts, and completions are not used to train or improve foundation models.
3. The service operates within the vendor's own compliance boundary under a data protection addendum.

These are the mechanisms by which AI protects stored personal information from future sharing — not vague assurances, but isolation, contract, and boundary.

For HHS workloads specifically, Microsoft 365 Copilot for Enterprise is covered under Microsoft's HIPAA Business Associate Agreement. That coverage does not eliminate an agency's own obligations to DLP, sensitivity labels, minimum-necessary access, audit logging, breach notification, and workforce training, but it does establish the contractual foundation for using PHI in a Copilot workflow. Governance must extend to the AI's outputs as well: a generated summary, risk indicator, or reunification score carries the same sensitivity as the source records that produced it, and must be labeled, logged, and access-controlled accordingly.

The practical rule is simple: if a tool will ever touch PHI, confirm the BAA covers that specific product, verify the data residency boundary, and document the configuration that enforces it.

4. Tenant Isolation and Training-Data Exclusion

Agencies reasonably want written assurance that an AI vendor will not use their data to train its models. For Microsoft 365 Copilot and Azure OpenAI Service, that commitment is set out in the Microsoft Online Services Terms and the Data Protection Addendum. The isolation extends technically as well: Azure OpenAI runs model weights hosted in Microsoft Azure rather than forwarding prompts to OpenAI's infrastructure, and Copilot's tenant data does not leave the Microsoft 365 service boundary except as the customer configures. The same assurance is not automatic for every AI tool an employee might reach for; consumer tiers operate under different terms. A governance program must distinguish—in policy, procurement, and network controls—between the enterprise services that carry these commitments and the consumer services that do not.

5. Auditability and Human Validation of AI Outputs

Generative AI outputs are probabilistic. They can be grounded, cited, and constrained, but they are not deterministic. They can be wrong in ways that look right. The question leaders reasonably ask is how to trust findings that no human could feasibly verify one by one. The answer is twofold:

1. Make outputs auditable
2. Design human-in-the-loop checkpoints to examine critical failure modes

On auditability, Microsoft 365 Copilot logs every interaction, including prompt, response, and grounded source, to the Purview Audit log, and Azure OpenAI emits equivalent telemetry. Purview Communication Compliance can scan Copilot interactions against policy, and Microsoft Security Copilot provides investigation tooling for AI-related incidents. These records are what an agency produces when a determination is challenged, an investigation opens, or a public records request arrives.

On validation, a sound AI deployment does not ask the system to make decisions. It asks the system to surface evidence that is grounded in the record, cited to sources, and ranked by confidence. Retrieval-augmented generation patterns that preserve document, page, and passage references make this practical. A human decision-maker can evaluate this type of evidence in minutes rather than hours, making it economical. A caseworker reviewing several hundred AI-surfaced citations in a day can cover what would otherwise take a year of manual review, while judgment stays where it belongs: with the human. Test against labeled evaluation sets before going live, sample outputs continuously, measure agreement with human reviewers, and never let AI output flow into a constituent-affecting decision without a reviewer in the loop.

6. Deployment Boundary and Data Residency

For HHS data in the United States, the governance answer is almost always cloud, but a specific kind of cloud. Microsoft 365 Copilot is now generally available in GCC High, which meets FedRAMP High, DFARS, ITAR, and CJIS requirements. Wave 2 capabilities, including GPT-5 access and research agents, are shipping to GCC High through the first half of 2026. Data in GCC High remains in U.S. data centers managed by screened U.S. personnel, with web grounding off by default to prevent sensitive content from crossing the commercial boundary. Azure Government provides comparable guarantees for custom Azure OpenAI deployments and for workloads requiring CJIS adjudication. On-premises remains viable for narrow cases, but the governance overhead is substantially higher and should be chosen when a compliance driver requires it, not by default.

Regarding token economics, Microsoft 365 Copilot pairs a flat per-user license for baseline productivity experiences with consumption-based charges for custom agents and higher-volume usage, while Azure OpenAI

Service workloads are metered per token against the agency's own subscription, with a Provisioned Throughput option for predictable capacity and cost. Consumption models demand the same discipline agencies already apply to compute: quotas on tokens or messages per minute, rate limiting through Azure API Management, managed-identity authentication, and Cost Management budgets with alert thresholds well below the ceiling. For HHS budgeting, the relevant metric is not units consumed but case-hours displaced, expressed as cost per case, per decision, or per document processed.

Putting It Into Practice

Plan for AI readiness ahead of any deployment. Start with the data estate: inventory where sensitive information lives, apply sensitivity labels, and run a permissions remediation pass on SharePoint, OneDrive, Teams, and Dataverse to clear the oversharing that typically accumulates in long-lived tenants. Pick a narrow pilot population with clean content rather than an agency-wide rollout, and turn on Purview DSPM for AI, DLP coverage for the Copilot location, and the Purview audit log before the first user touches a prompt.

Pair the technology work with a standing cross-functional governance group that has authority over approvals, incident response, and review cadence. Baseline every pilot against measurable outcomes such as hours per case, days to decision, or error rate, so displaced effort can be weighed against consumption cost. Schedule a quarterly review of AI workloads the way you already schedule access reviews, and retire what is no longer earning its footprint.

AI readiness can be started this quarter with the tools and teams the agency already has.

References and Further Reading

- NIST AI Risk Management Framework 1.0 — the foundational governance framework referenced throughout.
- NIST AI 600-1: Generative AI Profile (July 2024) — cross-sectoral profile of AI RMF 1.0 with over 400 mitigation actions.
- HIPAA & HITECH Microsoft Compliance Offering — coverage of Microsoft services under HIPAA, including BAA terms.
- Microsoft Purview for Microsoft 365 Copilot — end-to-end data security and compliance configuration for Copilot.
- Purview DSPM for AI — Deployment Considerations — tenant-wide posture management for AI usage.
- Azure OpenAI Data Privacy and Customer Data Commitments — contractual specifics on training exclusion and tenant isolation.
- Microsoft 365 Copilot in GCC High — GA in GCC High and Wave 2 roadmap for 2026.
- Microsoft 365 Copilot rollouts slowed by data security, ROI concerns (Computerworld, June 2025) — summary of Gartner’s June 2025 survey of 132 IT leaders, which found 40% of organizations delayed M365 Copilot rollouts by three months or more due to data-oversharing concerns.