# **NIST AI Risk Management Policy Framework for Hospitals**

Based on NIST AI RMF Core + NIST AI 600-1 Generative AI Profile

This repository provides a modular, phased AI Risk Management Policy tailored for hospital environments. It aligns with the NIST AI Risk Management Framework (AI RMF Core) and extends compliance with the NIST AI 600-1 Generative AI Profile, addressing risks unique to generative AI technologies such as large language models (LLMs), synthetic data generation, and AI-based summarization in clinical and administrative settings.

Each policy module corresponds to a core function of the NIST AI RMF: Govern, Map, Measure, and Manage, and has been cross-referenced with applicable GenAI-specific guidance per NIST AI 600-1.

# **Framework References**

• NIST AI RMF Core (January 2023)

• NIST AI 600-1 Generative AI Profile (October 2023)

• HIPAA Security Rule (§164.308)

• FDA AI/ML-Based SaMD Action Plan (2024 PCCP Guidance, 2025 Draft Lifecycle Guidance)

• ONC Health IT Certification Criteria

This modular structure enables phased adoption of AI governance practices while scaling to address the emerging risk landscape of generative AI in healthcare.

# **Policy Document Structure/Summaries**

author: jfraze@mycomp.org

## **A. Overview and Governance (GV)**

• A\_GV-1\_policy\_overview.md  
High-level summary of the hospital’s AI risk policy, including trustworthiness goals, GenAI use principles, and compliance alignment.

A\_GV-1\_policy\_overview.md  
High-level summary of the hospital’s AI risk policy, including trustworthiness goals, GenAI use principles, and compliance alignment.

• A\_GV-2\_roles\_responsibilities.md  
Defines governance structure, multi-disciplinary oversight bodies, and accountability roles for AI systems and GenAI deployments.

A\_GV-2\_roles\_responsibilities.md  
Defines governance structure, multi-disciplinary oversight bodies, and accountability roles for AI systems and GenAI deployments.

• A\_GV-3\_ai\_governance\_framework.md  
Hospital-wide governance model including ethical review boards, clinical integration oversight, and AI policy enforcement aligned with AI 600-1. This includes pre-deployment checkpoints, governance review logs, and risk categorization workflows per §2.1 and §2.5.

A\_GV-3\_ai\_governance\_framework.md  
Hospital-wide governance model including ethical review boards, clinical integration oversight, and AI policy enforcement aligned with AI 600-1. This includes pre-deployment checkpoints, governance review logs, and risk categorization workflows per §2.1 and §2.5.

• A\_GV-4\_policy\_enforcement\_phases.md  
Phased implementation approach: Must Do, Should Do, and Recommended guidelines, incorporating GenAI readiness and system maturity.

A\_GV-4\_policy\_enforcement\_phases.md  
Phased implementation approach: Must Do, Should Do, and Recommended guidelines, incorporating GenAI readiness and system maturity.

• A\_GV-5\_human\_oversight\_and\_decision\_accountability.md  
Ensures AI-augmented decisions retain human accountability and oversight in clinical, legal, and administrative workflows.

A\_GV-5\_human\_oversight\_and\_decision\_accountability.md  
Ensures AI-augmented decisions retain human accountability and oversight in clinical, legal, and administrative workflows.

## **B. Map Function (MP)**

• B\_MP-1\_context\_and\_use\_cases.md  
Defines intended AI system and GenAI use cases, including clinical documentation, decision support, triage, scheduling, and patient communication.

B\_MP-1\_context\_and\_use\_cases.md  
Defines intended AI system and GenAI use cases, including clinical documentation, decision support, triage, scheduling, and patient communication.

• B\_MP-2\_data\_origin\_and\_provenance.md  
Details on training data for AI/GenAI models, lineage tracking, synthetic data use, consent, and data quality for sensitive environments.

B\_MP-2\_data\_origin\_and\_provenance.md  
Details on training data for AI/GenAI models, lineage tracking, synthetic data use, consent, and data quality for sensitive environments.

• B\_MP-3\_ai\_system\_inventory.md  
Maintains an inventory of all deployed AI and GenAI systems, noting model types (e.g., LLMs), intended use, ownership, and associated risks.

B\_MP-3\_ai\_system\_inventory.md  
Maintains an inventory of all deployed AI and GenAI systems, noting model types (e.g., LLMs), intended use, ownership, and associated risks.

• B\_MP-4\_responsible\_use\_of\_synthetic\_data.md  
Provides safeguards and documentation protocols for the generation, validation, and use of synthetic datasets in AI development.

B\_MP-4\_responsible\_use\_of\_synthetic\_data.md  
Provides safeguards and documentation protocols for the generation, validation, and use of synthetic datasets in AI development.

## **C. Measure Function (MS)**

• C\_MS-1\_model\_documentation\_and\_traceability.md  
Comprehensive documentation of GenAI systems, including data lineage, training history, fine-tuning metadata, and lifecycle traceability for audit readiness and regulatory disclosure.

C\_MS-1\_model\_documentation\_and\_traceability.md  
Comprehensive documentation of GenAI systems, including data lineage, training history, fine-tuning metadata, and lifecycle traceability for audit readiness and regulatory disclosure.

• C\_MS-2\_explainability\_and\_transparency.md  
Requirements for explainable GenAI outputs, rationale flagging, and transparent decision support in patient-facing or safety-relevant workflows. Includes user-visible model indicators and interpretability safeguards.

C\_MS-2\_explainability\_and\_transparency.md  
Requirements for explainable GenAI outputs, rationale flagging, and transparent decision support in patient-facing or safety-relevant workflows. Includes user-visible model indicators and interpretability safeguards.

• C\_MS-3\_validation\_and\_testing.md  
Validation and Testing Protocols prior to deployment, covering hallucination and truthfulness checks, adversarial prompt testing, clinical accuracy validation, and alignment with GenAI safety thresholds.

C\_MS-3\_validation\_and\_testing.md  
Validation and Testing Protocols prior to deployment, covering hallucination and truthfulness checks, adversarial prompt testing, clinical accuracy validation, and alignment with GenAI safety thresholds.

## **D. Manage Function (MG)**

• D\_MG-1\_risk\_treatment\_and\_controls.md  
Mitigation strategies for generative model misuse, prompt injection, data leakage, PHI re-identification, and other GenAI-related harms.

D\_MG-1\_risk\_treatment\_and\_controls.md  
Mitigation strategies for generative model misuse, prompt injection, data leakage, PHI re-identification, and other GenAI-related harms.

• D\_MG-2\_incident\_response\_and\_model\_escalation.md  
Incident reporting process for GenAI errors, hallucinations, bias events, ethical failures, and regulatory non-compliance.

D\_MG-2\_incident\_response\_and\_model\_escalation.md  
Incident reporting process for GenAI errors, hallucinations, bias events, ethical failures, and regulatory non-compliance.

• D\_MG-3\_model\_monitoring\_and\_event\_logging.md  
Lifecycle governance with checkpoints for generative model fine-tuning, prompt library controls, and continuous monitoring.

D\_MG-3\_model\_monitoring\_and\_event\_logging.md  
Lifecycle governance with checkpoints for generative model fine-tuning, prompt library controls, and continuous monitoring.

• D\_MG-4\_third\_party\_ai\_vendor\_risks.md  
Vendor risk due diligence for hosted GenAI APIs, SaaS models, proprietary LLMs, and integration into EHR or patient-facing tools.

D\_MG-4\_third\_party\_ai\_vendor\_risks.md  
Vendor risk due diligence for hosted GenAI APIs, SaaS models, proprietary LLMs, and integration into EHR or patient-facing tools.

• D\_MG-5\_model\_retirement\_and\_decommissioning.md  
Outlines lifecycle closure procedures for GenAI systems, including audit trails, safe removal, and risk-informed decommissioning.

D\_MG-5\_model\_retirement\_and\_decommissioning.md  
Outlines lifecycle closure procedures for GenAI systems, including audit trails, safe removal, and risk-informed decommissioning.

• D\_MG-6\_ai\_and\_phi\_boundary\_controls.md  
Implements safeguards to prevent PHI leakage and ensure proper access, redaction, and inference boundary protection across GenAI workflows.

D\_MG-6\_ai\_and\_phi\_boundary\_controls.md  
Implements safeguards to prevent PHI leakage and ensure proper access, redaction, and inference boundary protection across GenAI workflows.

## **E. Appendices / Supporting Documents**

• X1\_glossary.md  
Definitions for AI and GenAI terms: hallucination, prompt injection, fine-tuning, synthetic data, zero-shot, etc.

X1\_glossary.md  
Definitions for AI and GenAI terms: hallucination, prompt injection, fine-tuning, synthetic data, zero-shot, etc.

• X2\_policy\_templates.md  
Templates for:  
  
AI/GenAI Use Case Intake  
Bias Impact Assessments  
AI Incident Reports  
Prompt Library Management  
Intake Forms

X2\_policy\_templates.md  
Templates for:

• AI/GenAI Use Case Intake

• Bias Impact Assessments

• AI Incident Reports

• Prompt Library Management

• Intake Forms

• X3\_top-AI-vuln.md  
  
Specific attacks to consider

X3\_top-AI-vuln.md

• Specific attacks to consider

• X4\_output-rules.md  
  
Guidelines for content creation and formatting

X4\_output-rules.md

• Guidelines for content creation and formatting

## **Compliance Mappings**

| Section File | Title | NIST AI RMF Controls |  
|--------------------------------------------------------|------------------------------------------------------|----------------------------------------|  
| AGV-1policyoverview.md | AI Risk Management Policy Overview | GOVERN-1, GOVERN-1.2 |  
| AGV-2rolesresponsibilities.md | AI Roles and Governance Responsibilities | GOVERN-2, GOVERN-4, GOVERN-5 |  
| AGV-3aigovernanceframework.md | Hospital AI Governance Framework | GOVERN-4, GOVERN-5, GOVERN-6 |  
| AGV-4policyenforcementphases.md | Policy Enforcement Phases for AI Risk Controls | GOVERN-1, MANAGE-1.1, MAP-1.5 |  
| AGV-5humanoversightanddecisionaccountability.md | Human Oversight and Decision Accountability | GOVERN-5.2, MEASURE-2 |  
| BMP-1contextandusecases.md | AI System Context and Use Cases | MAP-1, MAP-2 |  
| BMP-2dataoriginandprovenance.md | Data Origin and Provenance for AI Systems | MAP-3, AI 600-1 §2.4.1 |  
| BMP-3aisysteminventory.md | AI System Inventory and Classification | MAP-4 |  
| BMP-4responsibleuseofsyntheticdata.md | Responsible Use of Synthetic Data in GenAI Systems | MAP-2.1, HIPAA §164.514(b), AI 600-1 §2.2.1 |  
| CMS-1modeldocumentationandtraceability.md | Data and Model Documentation for GenAI Systems | MAP-3.3, AI 600-1 §2.2, SP 800-218A §3.2 |  
| CMS-2explainabilityandtransparency.md | Transparency and Explainability in GenAI Use | MEASURE-2.6, AI 600-1 §2.4.1 |  
| CMS-3validationandtesting.md | Validation and Testing for Generative AI Systems | MEASURE-1.3, MEASURE-2.4, AI 600-1 §2.3.1 |  
| DMG-1risktreatmentandcontrols.md | Risk Treatment and Controls for Generative AI | MANAGE-1, MANAGE-2.3, MANAGE-3.1 |  
| DMG-2incidentresponseandmodelescalation.md | Incident Response and Model Escalation Procedures | MANAGE-3.1, MANAGE-4.1, AI 600-1 §2.3.1 |  
| DMG-3modelmonitoringandeventlogging.md | Model Monitoring and Event Logging for GenAI Systems | MANAGE-2.1, MANAGE-4.1, AI 600-1 §2.3.1 |  
| DMG-4thirdpartyaivendorrisks.md | Third-Party Generative AI Vendor Risks | GOVERN-4, MANAGE-2.3, MANAGE-4.2 |  
| DMG-5modelretirementanddecommissioning.md | Model Retirement and Decommissioning Procedures | MANAGE-4.1, MAP-4.2 |  
| DMG-6aiandphiboundary\_controls.md | AI and PHI Boundary Controls | MANAGE-2.2, SP 800-218A §3.1.2 |

# **Primary Reference Document Summaries**

## **AIRMFPlaybook**

– AI RMF Tactics & Implementation Playbook  
An accompanying workbook to AI RMF 1.0 providing tactical actions, metrics, and roles for effective AI risk management. Useful for translating framework outcomes into practical steps

## **NIST.AI.600‑1**

– AI RMF Companion for Generative AI  
A supplement to AI RMF, tailored for generative AI models. It adapts core RMF practices to the unique risks of AI model development and deployment. "Specialized profile on generative AI"

# **Supporting Documents Summaries**

## **NIST.AI.100‑1**

– Artificial Intelligence Risk Management Framework (AI RMF 1.0)  
Provides a structured risk management framework for AI systems, focusing on trustworthy, accountable, transparent design through four core functions: Govern, Map, Measure, Manage

## **NIST.CSWP.29**

– Cybersecurity Framework (CSF) 2.0  
An updated version of the NIST Cybersecurity Framework, introducing a new "Govern" layer on top of core functions (Identify, Protect, Detect, Respond, Recover). Designed as a universal risk management taxonomy

## **NIST.SP.800‑218A**

– Secure Software Development for Generative AI  
A supplement to the SSDF standard (SP 800‑218) focused on secure development practices for generative and dual‑use AI models—covering data sourcing, design, testing, and deployment specifics

## **NIST.SP.800‑61r3**

– Incident Response Guide (Rev. 3)  
A modernized version of NIST SP 800‑61 that aligns incident response practices with CSF 2.0. It's a playbook for preparedness, detection, response, recovery, and risk alignment in cybersecurity incidents

# **Regulatory and Sector-Specific Requirements to Consider**

• HIPAA Privacy and Security Rule

• FDA Good Machine Learning Practice (GMLP) for AI/ML-based Software as a Medical Device (SaMD)

• ONC Certification Criteria for Health IT

• HITECH Act and 21st Century Cures Act for interoperability and patient data access

• State-level medical AI governance (if applicable)