# Document Metadata

author: jfraze@mycomp.org

title: Data and Model Documentation for GenAI Systems

nist\_function: Map

priority\_phase: Must

last\_reviewed: 2025-06-24

status: draft

## **Purpose**

To ensure transparent, complete, and lifecycle-aligned documentation of datasets, model architectures, training parameters, and evaluation criteria used in hospital-deployed generative AI systems.

## **Scope**

Applies to all GenAI systems developed or deployed across clinical, administrative, and operational domains, including LLMs, NLP assistants, structured summarizers, and generative chat interfaces.

## **Policy Statement**

All GenAI systems must maintain detailed and accessible documentation encompassing data lineage, model structure, intended uses, known limitations, and prior evaluations. Documentation should support reproducibility, audit readiness, and regulatory disclosure.

## **Roles and Responsibilities**

• Model Developers: Maintain version-controlled records of datasets, model types, and fine-tuning configurations.

• Compliance Officers: Validate that documentation supports HIPAA audit, ONC certification, and FDA GMLP standards.

• AI Risk Officers: Review and update documentation in response to system changes or lifecycle transitions.

## **Implementation Phases**

### **Must Do**

• Maintain a "Model Factsheet" including training data origin, base model, and risk classification (AI RMF MAP-3.3, AI 600-1 §2.2).

• Document dataset composition, including any synthetic data and PHI filtering mechanisms.

• Version all fine-tuning steps, with associated datasets and performance metrics (SP 800-218A §3.2).

### **Should Do**

• Use model cards and datasheets for standardized disclosure.

• Annotate known limitations, bias risks, and hallucination rates.

• Integrate documentation into MLOps or software lifecycle tools.

### **Recommended**

• Share non-sensitive portions of documentation with third-party safety bodies.

• Include documentation of fallback behavior and safety interlocks.