PRESCIENT SECURITY

NIST AI RMF 1.0:
Understanding Trustworthy AI Governance & Attestation

NIST AI 100-1 (Artificial Intelligence Risk Management Framework) provides organizations with a structured framework to govern, design, build, and deploy trustworthy AI systems. The goal is to operationalize AI by integrating risk management across the entire AI lifecycle and to move beyond traditional IT controls to address safety, bias, as well as model robustness.

The framework establishes four core functions: GOVERN, MAP, MEASURE, and MANAGE. The functions are designed to help AI developers and deployers balance innovation with proactive risk mitigation and audit-ready accountability.

Here is what has changed:

Attestation replaces nonexistent "certification": NIST AI RMF 1.0 is explicitly voluntary and non-certifiable by NIST. Third-party assessments which result in an Independent Attestation Letter to verify framework alignment.

Continuous lifecycle risk over static control checklists: Risk management spans system design, training data curation, model validation, deployment, and decommissioning instead of only point-in-time audits.

Trustworthy AI Characteristics become core metrics: Risk is measured against seven defined trustworthy AI dimensions: valid & reliable, safe, secure & resilient, accountable & transparent, explainable & interpretable, privacy-enhanced, and fair with harmful bias managed.

Core functions at a glance

Function	Primary focus	Best fit	Scope & categories	Submission / output path
GOVERN (GV)	Culture, organizational risk processes, legal compliance & AI accountability	Enterprise-wide leadership, AI ethics boards, and legal/ compliance teams	6 Categories (19 subcategories)	AI Governance Charter, Responsible AI Policy & Accountability Matrix
MAP (MP)	Context framing, categorization, AI actor mapping & impact/ hazard identification	Product managers, data scientists, and system architects	5 Categories (18 subcategories)	AI System Impact Assessment & Context-of-Use Baseline
MEASURE (MS)	Quantitative & qualitative evaluation, red-teaming, bias & drift testing	ML engineers, QA teams, independent AI evaluators	4 Categories (18 subcategories)	Model Evaluation Reports, Red-Teaming Results & Performance Benchmarks
MANAGE (MN)	Risk prioritization, mitigation controls, fallback mechanisms & monitoring	MLOps teams, incident responders, security engineers	4 Categories (13 subcategories)	AI Risk Register, Fallback Operating Procedures & Continuous Monitoring Plan

The core functions explained

GOVERN (GV)

The operational baseline that anchors organizational accountability and AI risk culture.

- Establishes leadership oversight, legal compliance, and workforce AI competency.
- Governs inventory and third-party AI/LLM vendor risk management.
- Defines formal procedures for model decommissioning and risk escalation.
- Serves as the primary operational bridge to ISO/IEC 42001 certification.

MAP (MP)

The contextual discovery phase that identifies potential harms before deployment.

- Characterizes the specific business context, intended domain, and end-user populations.
- Maps dependencies across foundation models, proprietary data, and external APIs.
- Uncovers potential societal, organizational, and individual impacts and failure modes.
- Informs which risk metrics must be tracked during the Measure phase.

MEASURE (MS)

The empirical verification phase that validates system performance and safety.

- Implements testing, evaluation, verification, and validation (TEVV) workflows.
- Assesses model accuracy, bias, drift, hallucination rates, and adversarial robustness.
- Utilizes independent red-teaming and prompt-injection testing.
- Provides repeatable metrics to substantiate third-party attestation findings.

MANAGE (MN)

The operational execution phase that treats risks and monitors models in production.

- Allocates risk treatment strategies (mitigate, transfer, avoid, accept).
- Deploys automated guardrails, output filters, and human-in-the-loop oversight.
- Establishes fallback, shutdown, and incident response procedures for rogue behavior.
- Feeds operational insights back into the Govern and Map functions.

What makes the functions different from traditional security frameworks

Requirement/ dimension	GOVERN	MAP	MEASURE	MANAGE
Primary artifacts	Policies, RACI, oversight charter	System profiles, data provenance	TEVV reports, red-team logs	Incident response, drift monitors
Evaluation method	Organizational & process review	Architecture & data-flow analysis	Empirical testing & model validation	Operational monitoring & controls
Applicability focus	Entire organization / all AI systems	Per system / use-case context	Per model / pipeline version	Live production deployments
Independent validation	Governance readiness audit	Scoping & boundary review	Model testing & benchmark validation	Control operating effectiveness
ISO 42001 alignment	Clauses 4–10 (Management System)	Annex A.5 (Assessing impacts of AI systems)	Annex A.6 (AI system life cycle)	Annex A.9 (Use of AI systems)

Why customers choose Prescient Security

- We test your actual pipelines, and models,to validate resilience against adversarial prompt injection, data poisoning, and harmful bias.
- Your engagement results in a a credible, auditor-signed Attestation Letter that satisfies stakeholders
- With us you can map NIST AI RMF findings to ISO/IEC 42001 requirements, to create a direct path to accredited certification without duplicate effort.
- We evaluate live ML systems using automated tooling, OSCAL/API integration, and developer-friendly workflows.

Technical Guidance Callout: NIST AI RMF 1.0 does not issue compliance certificates. Organizations seeking formal certification should leverage their NIST AI RMF maturity evidence toward ISO/IEC 42001 (Artificial Intelligence Management System). A dual-framework assessment reduces audit overhead while satisfying both US enterprise buyers and global regulatory demands.