

TIER

THE EIGHT TIERS · PART 6 · RISK MANAGEMENT FRAMEWORKS · 8 MIN READ

6

NIST AI RMF Explained: Four Functions, 72 Subcategories

Tier 6 of eight: a voluntary framework that became load-bearing, and a statutory safe harbour it never asked for

4

functions

19

categories

72

subcategories

7

trustworthiness
characteristics

0

legal obligation to follow it



The NIST AI RMF — formally the NIST AI Risk Management Framework, AI 100-1 — organises AI risk management into four functions: Govern, Map, Measure and Manage. Those span 19 categories and 72 subcategories, plus seven trustworthiness characteristics. It published in January 2023 and it is entirely voluntary.

It is also, by some distance, the most load-bearing document in American AI compliance. Which is a strange thing to be able to say about something nobody is obliged to follow, and the story of how it happened is worth five minutes.

HOW IT WORKS

The four functions

The framework is technology-agnostic and sector-neutral. It tells you how to think about AI risk rather than what to build — which is its strength, and also why teams find it harder to operationalise than a control catalogue. You can't hand someone AI 100-1 and say "implement that."

CROSS-CUTTING GOVERN GOVERN is the cross-cutting one, and the one almost everybody underinvests in. Risk culture, policies and procedures, accountability structures, workforce competence, engagement with affected communities, third-party and supply-chain risk. Everything in the other three functions rests on it. If Govern is thin, the rest is theatre.		
MAP MAP establishes context. Categorise the system. Understand its capabilities, intended use, goals, expected benefits against costs. Identify the risks and who they land on. Map is where most programmes discover their inventory isn't good enough — which is uncomfortable and also the most useful thing the framework does for you early.	MEASURE MEASURE applies metrics. Select the methods, evaluate against the trustworthiness characteristics, set up monitoring, gather feedback on whether any of it is working. This is where governance stops being a policy exercise and becomes evaluation engineering.	MANAGE MANAGE allocates resources to prioritised risks, plans responses and recovery, handles third-party risk, documents and communicates.

Running through all of it: seven trustworthiness characteristics.

Valid and reliable	Safe	Secure and resilient	Accountable and transparent
Explainable and interpretable	Privacy-enhanced	Fair, with harmful bias managed	

They map closely onto the requirements in Chapter III of the EU AI Act, which is not a coincidence and is very convenient.

COMPANION DOCUMENTS

The companion you'll actually use

NIST AI 600-1, the Generative AI Profile, published July 2024. If you only read one companion document, read this one.

It identifies twelve risk categories specific to generative AI — and adds more than two hundred suggested actions mapped back to the four functions.

AI 600-1 · twelve generative AI risk categories		200+ suggested actions
01 Confabulation	02 Dangerous or violent recommendations	03 Data privacy
04 Environmental impacts	05 Harmful bias	06 Human-AI configuration
07 Information integrity	08 Information security	09 Intellectual property
10 Obscene and degrading content	11 Value chain and component integration	12 CBRN information

Two hundred actions is a lot, and you won't do all of them. But as a checklist against which to ask "have we even thought about this?", it's the best free thing available.

The rest of the ecosystem, briefly:

AI 100-2	covers adversarial machine learning.
SP 800-218A	covers secure software development for generative AI and dual-use foundation models.
IR 8596	in preliminary draft since December 2025, bridges the AI RMF and Cybersecurity Framework 2.0.
In development	Control overlays for SP 800-53 and a critical infrastructure profile are in development — NIST released a concept note for the latter in April 2026.

FOUR REASONS

Why a voluntary framework ended up mattering this much

Four reasons, and together they're the thing that makes this document different from the dozens of other well-meaning frameworks published since 2023.

01

It became a statutory safe harbour.

Texas TRAIGA makes substantial compliance with the AI RMF an enforcement safe harbour. No other voluntary framework has that status anywhere in US law. A legislature looked at a NIST document and wrote it into statute as a defence.

02

It's the federal baseline.

OMB memoranda governing federal agency AI use and procurement reference it directly.

03

It's the procurement default.

This is the one that'll reach you first. It is the most common baseline in American enterprise vendor questionnaires, which means your customers will ask about it whether or not a regulator ever does.

I've spent a lot of my career on the receiving end of enterprise security questionnaires, and I'd make a prediction: for most companies reading this, the first time AI governance becomes a real deadline won't be a regulator. **It'll be a 180-question document from a customer's procurement team with "NIST AI RMF" somewhere in it and a deal attached to the answer.**

04

It crosswalks well.

The function-and-subcategory structure maps cleanly onto ISO 42001's Annex A controls, onto EU AI Act articles, onto state statutory requirements. Build your control library against the AI RMF and 42001 together and you've covered most of what everything else asks for.

**NIST AI RMF + ISO
42001**

→ ISO 42001
Annex A

→ EU AI Act
articles

→ State statutes

That last point is the practical advice. If you're going to pick two things, pick those two.

CAVEAT

The caveat I'd want you to know about

Colorado's original AI Act offered an affirmative defence for organisations using a recognised risk management framework, and it named the NIST AI RMF explicitly. That was genuinely useful. It said: adopt something credible, get legal protection.

That defence did not survive into the 2026 replacement law.

I'm not telling you that to put you off the framework. I'm telling you because it changes why you should adopt it. Adopt the AI RMF because it produces the evidence a regulator or an auditor or a customer will ask you for. Don't adopt it because a statute currently rewards you — legislatures can withdraw that, and in Colorado's case one already has.

One more thing to watch. The framework is under revision as part of the 2025 AI Action Plan. As of now, AI RMF 1.0 is still the current version and the revision is still in progress — but if a version 2.0 lands, a great deal of published guidance becomes wrong rather than merely dated.

FAQ

Things people ask me

Is the NIST AI RMF mandatory?

No, it's voluntary. It functions as an enforcement safe harbour under Texas TRAIGA and as a baseline for US federal agencies, which is a kind of mandatory that doesn't look like one.

What are the four functions of the NIST AI RMF?

Govern, Map, Measure and Manage, across 19 categories and 72 subcategories.

How does the NIST AI RMF relate to ISO 42001?

They're complementary and I'd use both. The AI RMF gives you the risk methodology; ISO/IEC 42001 gives you the certifiable management system and the control catalogue. They crosswalk cleanly, and between them they cover most of what other frameworks ask for.

NEXT IN THE SERIES · PART 7

OWASP LLM Top 10

Where governance stops being paperwork and turns into test cases.

The Eight Tiers series — Part 0: The eight kinds · Part 1: Comprehensive statutes · Part 2: The treaty nobody is watching · Part 3: Where the obligations actually are · Part 4: The regulators who were already there · Part 5: ISO 42001 · [Part 6: NIST AI RMF](#) · Part 7: Security taxonomies · Part 8: Principles · Part 9: Putting it together