THE PATH TO AI RMF 1.0

Oct 19-21, 2021
NIST AI RMF workshop #1

Mar 29-31, 2022
NIST AI RMF workshop #2

Oct 18-19, 2022
NIST AI RMF workshop #3

Jul 29, 2021
RFI seeking input

Dec 13, 2021
AI RMF Concept Paper

Mar 17, 2022
AI RMF 1st Draft

Mar 29-31, 2022
NIST AI RMF workshop #2

Aug 18, 2022
AI RMF 2nd Draft

Jan 26, 2023
AI RMF 1.0
AI RMF Playbook

- Explainable AI paper released Sept 29, 2021
- Comments until Sept 15, 2021
- 106 sets of input
- Comments until Jan 25, 2022
- 59 sets of input
- Listening sessions
- Comments until Apr 29, 2022
- 92 sets of input
- Bias in AI paper released Mar 14, 2022
- Comments on AI RMF and Playbook until Sept 29, 2022
- Call for contributions towards Profiles
WHAT IS THE AI RMF?

Voluntary resource for organizations designing, developing, deploying, or using AI systems to manage AI risks and promote trustworthy and responsible AI.

- Rights-preserving
- Flexibly applied
- Measurable
MANAGING RISKS AND MINIMIZING NEGATIVE IMPACTS

Examples of Potential Harms

Harm to People
- Individual: Harm to a person’s civil liberties, rights, physical or psychological safety, or economic opportunity.
- Group/Community: Harm to a group such as discrimination against a population sub-group.
- Societal: Harm to democratic participation or educational access.

Harm to an Organization
- Harm to an organization’s business operations.
- Harm to an organization from security breaches or monetary loss.
- Harm to an organization’s reputation.

Harm to an Ecosystem
- Harm to interconnected and interdependent elements and resources.
- Harm to the global financial system, supply chain, or interrelated systems.
- Harm to natural resources, the environment, and planet.
AI RISK MANAGEMENT CHALLENGES

- Risk measurement
- Risk tolerance
- Risk prioritization
- Risk integration & management
AI RISKS AND TRUSTWORTHINESS

- Safe
- Secure & Resilient
- Explainable & Interpretable
- Privacy-Enhanced
- Fair - With Harmful Bias Managed
- Accountable & Transparent

Valid & Reliable
AI RMF CORE
Functions, Categories, Subcategories

AI Risk Management Framework

Govern
Map
Manage
Measure
Welcome to the draft NIST AI Risk Management Framework (AI RMF) Playbook – a companion resource for the AI RMF.

The Playbook includes suggested actions, references, and documentation guidance for stakeholders to achieve the outcomes for “Map” and “Govern” – two of the four proposed functions in the AI RMF. Draft material for the other two functions, Measure and Manage, will be released at a later date.

This draft Playbook is being released to allow interested parties the opportunity to comment and contribute to the first complete version, to be released in January 2023 with the AI RMF 1.0. The Playbook is an online resource and will be hosted temporarily on GitHub Pages.

NIST welcomes feedback on this draft Playbook.
Want to Help Others to Use the AI RMF?

Contribute

Suggest guidance about methods, metrics, explanatory papers, successful strategies, etc., for possible inclusion in the NIST Trustworthy and Responsible AI Resource Center.

Engage

Join or launch a collaborative activity – like developing a profile with others in your field or sector

aiframework@nist.gov
Tailored guidance for a specific setting or application based on the requirements, risk tolerance, and resources of the Framework user.
What works….what doesn’t
And how can we improve the AI RMF
What’s Next?

- AI RMF profile(s)
- Work with SDOs on AI standards
- Evaluate AI RMF’s effectiveness
- Crosswalks to other standards, frameworks, etc.
- AI evaluations & test beds
- Trustworthy & Responsible AI Resource Center
- And more …
For More Information…


Alframework@nist.gov