

National Institute of Standards and Technology releases draft AI Risk Management Framework for feedback

NOV 24, 2022 3 MIN READ

Related Expertise

- [Artificial Intelligence](#)
- [Technology](#)

Authors: [Simon Hodgett](#), [Sam Ip](#), [Sam Dobbin](#)

The U.S. National Institute of Standards and Technology (NIST) has released its [second draft of the NIST AI Risk Management Framework \[PDF\]](#) (AI RMF), as well as its draft of a [companion NIST AI RMF Playbook](#) for feedback on September 29, 2022. Subsequently, NIST held a live online workshop on October 18–19, 2022, for industry-leading AI experts to discuss community feedback on the second draft and next steps for the AI RMF.

AI Risk Management Framework

The AI RMF is being developed to better manage risks to individuals, organizations and society impacted by AI. The AI RMF is voluntary and is intended to improve the incorporation of trustworthiness into the design, development, use and evaluation of AI products, systems and services.

There are common risks in any software or information-based system which include cybersecurity, privacy, etc., but there are also specific risks in AI that are difficult to predict and manage. This is the purpose of the AI RMF: to consider and manage complex AI risks such as the amplification of inequitable outcomes and unintended consequences for individuals and communities. The AI RMF is not a checklist intended to be used in isolated situations; it should instead be broadly incorporated into decision-making and should be used alongside existing regulations and laws, not as a tool to replace them.

The AI RMF organizes AI risk management into functions which govern, map, measure and manage AI risks:

- **Govern:** This function is intended to cultivate AI risk management culture within organizations developing and deploying AI systems. This function ensures that risks and potential impacts are identified, measured and managed effectively and consistently.
- **Map:** This function establishes the context related to an AI system, and risks related to the context are identified. After completing the Map function, users should have sufficient contextual knowledge to inform a go/no-go decision regarding whether to design, develop or deploy an AI system.
- **Measure:** This function measures, analyzes, assesses and tracks AI risk and related impacts through the use of quantitative, qualitative or mixed-method tools and techniques.
- **Manage:** This function allocates risk management resources to mapped and measured

risks. These risks are then prioritized and acted upon based on project impact.

The AI RMF also introduces the concept of profiles, which are used to illustrate and analyze AI risk management processes for specific scenarios. Two noteworthy profiles are mentioned:

- **Use case profiles**, which can be used to offer insight on how to manage risk at different stages of an AI lifecycle in a specific sector, such as hiring or fair housing.
- **Temporal profiles**, which can be used to describe either the current or desired state of specific AI risk management activities within a sector or industry, which can reveal what gaps exist between current risk management processes and those that are being targeted.

Next steps

Feedback on the second draft of the AI RMF, as well as comments solicited at the workshop, will be reviewed and integrated. NIST is planning to release the finalized version of the AI RMF in January 2023. Although the AI RMF is voluntary, once released, we expect it to be widely referenced across various industries in the course of the design, development and deployment of AI systems.

chatgpt

chat gpt

openai

open ai