Morgan Lewis

LAWFLASH

NIST RELEASES NEW AI RISK MANAGEMENT FRAMEWORK

February 08, 2023

AUTHORS

Andrew J. Gray IV

The National Institute for Standards and Technology (NIST) recently released its Artificial Intelligence Risk Management Framework, a flexible set of guidelines that assists artificial intelligence actors, such as organizations and individuals, address the unique risks posed by artificial intelligence.

The framework of the Artificial Intelligence Risk Management Framework (AI RMF 1.0) is broken into two parts. Part 1 addresses foundational information, focusing on how AI actors can frame the potential risks of AI. Part 2 is the core of the AI RMF, centering on four functions: Govern, Map, Measure, and Manage.

PART 1 - FOUNDATIONAL INFORMATION

The AI RMF begins by addressing three categories of harm that AI Actors need to consider when using AI systems:

- Harm to people: violations of individual liberties or threats to physical, psychological, or economic safety and opportunity; discrimination against groups of people; and harm to societal access to democracy and education
- Harm to an organization: interruption of business operations; potential security breaches; and damage to reputation
- Harm to an ecosystem: disruption of global financial or supply chain systems and damage to the environment and natural resources

The framework highlights the fluid nature of risks during the AI lifecycle and highlights seven characteristics that make trustworthy AI systems:

- Valid and reliable
- Safe
- Secure and resilient
- Accountable and transparent
- Explainable and interpretable
- Privacy-enhanced

- Fair with harmful bias managed

Finally, the NIST emphasizes that AI actors should periodically evaluate AI risk management effectiveness in order to improve their abilities to manage AI risks.

PART TWO - CORE AND PROFILES

The AI Risk Management Framework core is composed of four functions: Govern, Map, Measure, and Manage. Each core function is broken down into numerous categories and sub-categories, with each having identified actions that should be performed continuously throughout the AI system lifecycle. NIST has published an accompanying AI RMF Playbook to help organizations navigate the RMF and achieve their outcomes through suggested actions that organizations can apply in their own context.

Govern

The Govern function centers around aspects of compliance and evaluation and is the cornerstone that enables other functions to process. Strong governance requires organizations to have structures, systems, processes, and teams in place to create and implement the policies, accountability structure, and organizational culture necessary for identifying and mitigating AI risks. Some examples of key elements to the Govern function include assigning roles and responsibilities related to the other core functions, developing policies and procedures to ensure legal and regulatory compliance, and establishing systems to collect, consider, prioritize, and integrate feedback from external sources.

Map

The Map function establishes the context to frame risks related to AI systems. AI systems rely on a complex network of interdependent activities, and understanding the interplay of these activities can make it easier for AI actors to anticipate the impacts of AI systems. Outcomes from the Map function provide the basis for the Measure and Manage functions.

Measure

The Measure function employs tools and methodologies to analyze, assess, benchmark, and monitor AI risks and related impacts. AI actors should track metrics for trustworthy characteristics, social impacts, and human interactions. Users should conduct comprehensive software testing and performance assessment and document and report the results. AI actors are encouraged to provide results for independent review to improve the effectiveness of their testing and minimize the potential for internal bias.

Manage

The Manage function concerns the allocation of resources to mapped and measured risks with the goal being to maximize AI benefits and minimize negative impacts. Policies developed in the Govern function will provide guidance as to which risks should be prioritized based on their likelihood and potential impact. Risk treatment under the Manage function should include plans on how to respond to, recover from, and communicate about incidents or events that cause harm.

CONCLUSION

The NIST AI RMF and accompanying playbook provide a thorough set guidelines, actions, and outcomes for individuals and organizations to implement in their development of and interactions with AI systems. While adherence is voluntary, organizations adopting the framework will be better prepared for the unique and often unforeseen risks that AI systems present.

CONTACTS

If you have any questions or would like more information on the issues discussed in this LawFlash, please contact any of the following:

Authors
Andrew J. Gray IV