Share:

February 2023

REGULATORY DEVELOPMENTS

AI Risk Management Framework

Igor Gorlach and Nicholas Maietta

On January 26th the National Institute of Standards and Technology (NIST) published the first draft of their Artificial Intelligence (AI) Risk Management Framework (AI RMF) that outlines the process for organizations to assess and address AI-related risks. The AI RMF comes in response to regulatory concerns that "AI tools can be inaccurate, biased, and discriminatory by design and incentivize relying on increasingly invasive forms of commercial surveillance."[1]

Use of AI has been a driving factor in innovation, and its use cases, led by ChatGPT, have recently been dominating technology headlines. AI is increasingly used to scan datasets for patterns, draw inferences from bodies of information, and for communicating with users.

But lawmakers and regulators have concurrently been grappling with ways to manage risks posed by AI, including both privacy and security risks. Studies show little progress in organizations' reported mitigation of AI-related risks,[2] and lawmakers and regulators appear particularly concerned with the use of consumer data and the privacy implications for AI.[3] These concerns are driving the EU's AI Act,[4] as well as the White House's release of an AI bill of rights in October.[5]

NIST AI RMF:

In 2021, Congress directed NIST to develop a voluntary risk management framework for trustworthy AI systems in collaboration with the private and public sectors.[6] In addressing this directive NIST released the AI RMF,[7] along with the AI RMF Playbook,[8] aimed at providing organizations a flexible, structured, and measurable approach to address AI risks and deploy trustworthy AI systems.

NIST frameworks are an important tool for organizations to assess and manage risk, and NIST's Cybersecurity Framework (CSF) plays an important role in risk management of traditional software and information-based systems. The nature of AI systems creates a more complex and novel set of risks, requiring a distinct approach and framework.

The NIST AI RMF and Playbook can be utilized by organizations to show they applied best practices and have taken a comprehensive approach to addressing the risks posed by AI, including privacy, security, bias, inaccuracy, discrimination, and other risks. This is critical given the FTC's focus on AI,[9] as well as recent FTC enforcement actions demonstrating expectations of programmatic and comprehensive privacy programs.[10]

NIST AI RMF – Contents:

The AI RMF is broken into two parts. The first describes and frames foundational concepts, including risks, impacts, and harms associated with AI; the relevant stakeholder groups; and the seven characteristics of trustworthy AI systems: valid and reliable; safe; secure and resilient; accountable and transparent; explainable and interpretable; privacy-enhanced; and fair with harmful bias managed.

The second part describes the 4 core functions for AI risk management which are then broken into 19 categories and 70 subcategories. The 4 AI RMF core function are Govern, Map, Measure, and Manage. The Govern function is about developing a culture of risk management and putting in place the policies, systems, processes, and teams necessary for effective AI risk management. The Map function is focused on information gathering to inform decisions and provide the context to frame risks related to an AI system. The Measure function covers quantitative and qualitative analyses to monitor risks and impacts. And the Managed function is about allocating resources to those mapped and measured risks and benefits associated with an AI system.

The AI RMF Playbook is a web-based document that breaks down the 70 subcategories further, providing more detailed descriptions of the subcategory, suggested actions to address the subcategory, documentation steps and resources on transparency, and references to seek additional information.

In short, the combination of the AI RMF and Playbook provide organizations a systematic and granular way to frame and address AI-related risks.

NIST AI RMF – Next Steps:

The document published by NIST is the first version of the AI RMF. NIST has provided an outline for future development, including alignment with other AI-related standards; standardized methodologies to test and verify AI systems; developing sector-specific AI profiles; and increased guidance related to various elements in the AI RMF, including system tradeoffs, human factors, explainability and interpretability, and methods for developing reasonable risk tolerances.

There are multiple ways for companies to participate in the ongoing development of the AI RMF. NIST is currently seeking industry feedback in developing the revised version of the AI RMF Playbook;[11] broad community involvement in developing tools, benchmarks, and methodologies for evaluating risks in AI;[12] assistance in production of AI RMF profiles and case studies,[13] and suggestions for the future NIST Trustworthy and Responsible AI Center.[14]

---

[1] FTC Report Warns About Using Artificial Intelligence to Combat Online Problems, June 16, 2022, at https://www.ftc.gov/news-events/news/press-releases/2022/06/ftc-report-warns-about-using-artificial-intelligence-combat-online-problems.

[2] *See, e.g.,* McKinsey & Company, The state of AI in 2022—and a half decade in review, December 6, 2022, at https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review.

[3] *See*, *e.g.*, FTC Report Warns About Using Artificial Intelligence to Combat Online Problems, n. 1.

[4] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, at https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/en/pdf.

[5] Blueprint for an AI Bill of Rights, October 4, 2022, at https://www.whitehouse.gov/ostp/ai-bill-of-rights.

[6] 15 U.S.C. § 278h–1(c).

[7] Artificial Intelligence Risk Management Framework (AI RMF 1.0), at: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf.

[8] NIST AI Risk Management Framework Playbook, at https://pages.nist.gov/AIRMF.

[9] *See*, *e.g.*, FTC Report Warns About Using Artificial Intelligence to Combat Online Problems, n. 1; Keep your AI claims in check, February 27, 2023, at https://www.ftc.gov/business-guidance/blog/2023/02/keep-your-ai-claims-check; and Aiming for truth, fairness, and equity in your company's use of AI, April 19, 2021, at https://www.ftc.gov/business-guidance/blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai.

[10] *See*, *e.g.*, https://www.ftc.gov/news-events/news/press-releases/2023/02/ftc-enforcement-action-bar-goodrx-sharing-consumers-sensitive-health-info-advertising.

[11] NIST notes that the revised version of the AI RMF Playbook should be released in the Spring of 2023. *See* NIST AI Risk Management Framework Playbook, at https://pages.nist.gov/AIRMF.

[12] *See* Roadmap for the NIST Artificial Intelligence Risk Management Framework (AI RMF 1.0), at https://www.nist.gov/itl/ai-risk-management-framework/roadmap-nist-artificial-intelligence-risk-management-framework-ai.

[13] *Id*.

[14] *See* AI Risk Management Framework FAQs, at https://www.nist.gov/itl/ai-risk-management-framework/ai-risk-management-framework-faqs; NIST to establish online center for trustworthy and responsible AI resources, September 30, 2022, at https://fedscoop.com/nist-trustworthy-responsible-ai-center.