GIBSON DUNN

January 27, 2023

NIST RELEASES FIRST VERSION OF AI RISK MANAGEMENT FRAMEWORK

To Our Clients and Friends:

On Thursday, January 26, 2023, the National Institute for Standards and Technology (NIST) released the first version of its Artificial Intelligence Risk Management Framework (AI RMF 1.0).[1] The framework is intended for voluntary use to help incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products, systems, and services.

AI RMF 1.0 was released after more than 18 months of drafting and workshops, which we have tracked in previous legal updates.^[2] The document reflects about 400 sets of formal comments NIST received from more than 240 different organizations on draft versions of the framework. Speaking at the launch event, Dr. Alondra Nelson, Deputy Assistant to the President and Principal Deputy Director for Science and Society in the White House Office of Science and Technology Policy (OSTP), indicated that OSTP provided "extensive input and insight" into the development of AI RMF 1.0.

As in previous drafts of the AI RMF, the framework is made up of four core "functions":

- Organizations must cultivate a risk management culture, including appropriate structures, policies, and processes. Risk management must be a priority for senior leadership, who can set the tone for organizational culture, and for management who aligns the technical aspects of AI risk management with organizational policies.
- Organizations must understand and weigh the benefits and risks of AI systems they are seeking to deploy as compared to the status quo, including helpful contextual information such as the system's capabilities, risks, benefits, and potential impacts.
- Using quantitative, qualitative, or mixed-method risk assessment methods, as well as the input of independent experts, AI systems should be analyzed for trustworthy characteristics, social impact, and human-AI configurations.
- Identified risks must be managed, prioritizing higher-risk AI systems. Risk monitoring should be applied over time as new and unforeseen contexts, risks, needs, or expectations can emerge.

AI RMF 1.0 also encourages the use of "profiles" to illustrate how risk can be managed through the AI lifecycle or in specific applications using real-life examples. Use-case profiles describe in detail how AI risks for particular applications are being managed in a given industry sector or across sectors (such as large language models, cloud-based services or acquisition) in accordance with RMF core functions. Temporal profiles illustrate current and target outcomes in AI risk management, allowing

GIBSON DUNN

organizations to understand where gaps may exist. And cross-sectoral profiles describe how risks from AI systems may be common when they are deployed in different use cases or sectors.

AI RMF 1.0 is accompanied by:

- AI RMF Playbook—a companion resource that suggests ways to navigate and use the AI RMF across its four core "functions" to incorporate trustworthiness considerations in the design, development, deployment, and use of AI systems.[3]
- AI RMF Roadmap—a list of different initiatives for advancing the AI RMF that NIST hopes organizations will carry out independently or in collaboration with the agency.[4]
- AI RMF Crosswalks—two documents that compare AI RMF 1.0 to 1) an international standard for AI risk management, and 2) the OECD Recommendation on AI, EU AI Act as currently drafted, Executive Order 13960, and the White House's Blueprint for an AI Bill of Rights.[5]
- Various Perspectives—a collection of statements by companies, industry organizations and advocacy organizations in support of AI RMF 1.0.[6]

Comments on AI RMF 1.0 will be accepted until February 27, 2023, with an updated version set to launch in spring 2023.

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

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

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

^[2] Artificial Intelligence and Automated Systems Legal Update (1Q22), available at https://www.gibsondunn.com/artificial-intelligence-and-automated-systems-legal-update-1q22/; Artificial Intelligence and Automated Systems Legal Update (2Q22), available at https://www.gibsondunn.com/artificial-intelligence-and-automated-systems-legal-update(2q22); Artificial Intelligence and Automated Systems Legal Update (3Q22), available at https://www.gibsondunn.com/artificial-intelligence-and-automated-systems-legal-update(2q22);

^[5] Crosswalks to the NIST Artificial Intelligence Risk Management Framework (AI RMF 1.0), available at https://www.nist.gov/itl/ai-risk-management-framework/crosswalks-nist-artificial-intelligence-risk-management-framework.

GIBSON DUNN

[6] Perspectives about the NIST Artificial Intelligence Risk Management Framework, available at https://www.nist.gov/itl/ai-risk-management-framework/perspectives-about-nist-artificial-intelligence-risk-management.



The following Gibson Dunn lawyers prepared this client update: H. Mark Lyon, Frances Waldmann, and Evan Kratzer.

Gibson Dunn's lawyers are available to assist in addressing any questions you may have regarding these developments. If you would like assistance in submitting comments on AI RMF 1.0, please contact the Gibson Dunn lawyer with whom you usually work, or any of the following members of Gibson Dunn's Artificial Intelligence and Automated Systems Group:

Cassandra L. Gaedt-Sheckter – Co-Chair, Palo Alto (+1 650-849-5203, cgaedtsheckter@gibsondunn.com)

H. Mark Lyon – Palo Alto (+1 650-849-5307, mlyon@gibsondunn.com)

Vivek Mohan – Co-Chair, Palo Alto (+1 650-849-5345, vmohan@gibsondunn.com)

Frances A. Waldmann – Los Angeles (+1 213-229-7914, fwaldmann@gibsondunn.com)

© 2023 Gibson, Dunn & Crutcher LLP

Attorney Advertising: The enclosed materials have been prepared for general informational purposes only and are not intended as legal advice. Please note, prior results do not guarantee a similar outcome.