Client Alert

Data, Privacy and Security

AUGUST 12, 2024

For more information, contact:

Kassi Burns
+1 512 457 2073
kburns@kslaw.com

Jesse Snyder
+1 202 383 8925
jsnyder@kslaw.com

King & Spalding

Dallas
2601 Olive St
23rd Floor
Dallas, TX 75201
Tel. +1 404 572 4600

NIST Releases Series of AI Guidelines & Software in Ongoing Response to AI Executive Order

The U.S. Department of Commerce's National Institute of Standards and Technology ("NIST") recently announced the publication of three AI guidelines as well as its release of a software package aimed at helping organizations measure the impact of adversarial attacks on AI system performance. These actions are all in response to President Biden's Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, published on October 23, 2023.

NIST & PRESIDENT BIDEN'S AI EXECUTIVE ORDER

President Biden's AI Executive Order was published with an accompanying Fact Sheet, which included action items for the various departments and agencies falling under the executive branch. Fact Sheet spotlighted, among other things, the need to create new standards for AI safety and security, and it did so specifically in relation to NIST:

- Develop standards, tools, and tests to help ensure that AI systems are safe, secure, and trustworthy;
- Protect Americans from AI-enabled fraud and deception by establishing standards and best practices for detecting AI-generated content and authenticating official content.

Since then, NIST has announced its continued and ongoing efforts to work with private and public stakeholders to fulfill these obligations. As stated by Under Secretary of Commerce for Standards and Technology and NIST Director Laurie Locascio: "We are committed to developing meaningful evaluation guidelines, testing environments, and information resources to help organizations develop, deploy, and use AI technologies that are safe and secure, and that enhance AI trustworthiness."

For example, NIST kicked off these efforts by hosting a workshop in November 2023 to facilitate collaboration efforts, whereby inviting private and public stakeholders to begin the process of identifying working groups for the various deliverables required under the AI Executive Order. These meetings served to path mark the development of NIST's recently published guidelines.

NIST'S NEW AI GUIDELINES

Building on the AI Risk-Management Framework ("AI RMF") published by NIST in July 2024, NIST collaborated with private and public stakeholders, including an open call for comments, to publish final versions of the following AI-related guidelines:

- AI RMF Generative AI Profile (NIST AI 600-1): A companion resource to the AI RMF, this publication provides guidance on issues specific to generative AI;
- Secure Software Development Practices for Generative AI and Dual-Use Foundation Models (NIST Special Publication (SP) 800-218A): A companion resource to NIST's Secure Software Development Framework (SSDF), this guideline is focused on developing well-secured AI systems to (1) reduce the number of vulnerabilities in released software, (2) mitigate the potential impact of the exploitation of undetected or unaddressed vulnerabilities, and (3) address root causes of vulnerabilities to prevent future recurrences;
- A Plan for Global Engagement on AI Standards (NIST AI 100-5): This set of standards was published by NIST to drive the development and implementation of AI-related consensus standards, cooperation and coordination, and information sharing.

In addition to the above final publications, NIST released an initial public draft of Managing Misuse Risk for Dual-Risk Foundation Models, which serves to identify best practices for developers of foundational models to best manage the risks that their models may be deliberately misused to cause harm. NIST has also issued a call for public comments for this draft through September 9, 2024, which will be used to help inform the final version of this document.

NIST'S AI SOFTWARE RELEASE

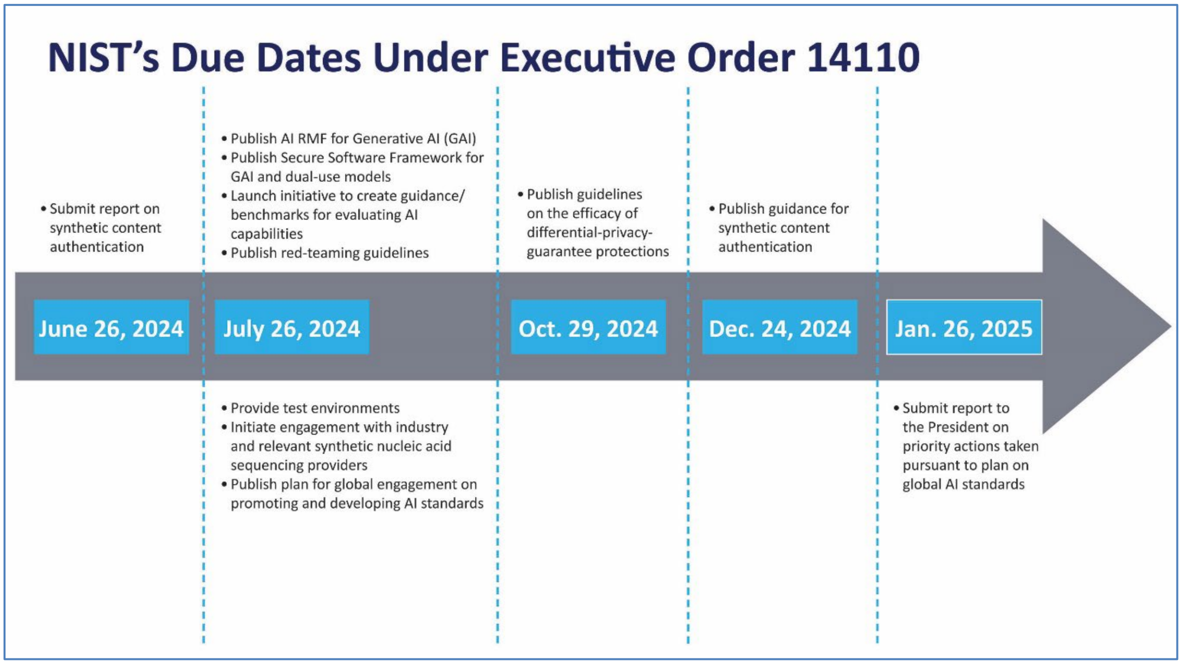
In January 2024, NIST published details about a type of cyberattack unique to AI systems: adversarial machine learning. Threat actors can "corrupt" or "poison" data that might be used by AI systems for training, thereby causing those AI systems to malfunction.

NIST aims to assist organizations through the release of its own open source software tool, Dioptra, which tests the effects of adversarial attacks on AI systems. In doing so, users will be able to select various adversarial tactics that a threat actor might use to make the model perform less effectively and thereby track performance reduction so as to learn how often and under what circumstances the AI system would fail.

> *"For all its potentially transformational benefits, generative AI also brings risks that are significantly different from those we see with traditional software. These guidance documents and testing platform will inform software creators about these unique risks and help them develop ways to mitigate those risks while supporting innovation."*
>
> — *Laurie E. Locascio, Under Secretary of Commerce for Standards and Technology and NIST Director*

UPCOMING NIST AI DELIVERABLES

The diagram below illustrates the effort of NIST as the agency builds out guidelines and standards for the safe, secure, and trustworthy development and use of AI continues in the coming months, with additional key benchmarks set through January 2025.



NIST's Due Dates Under Executive Order 14110

June 26, 2024
- Submit report on synthetic content authentication

July 26, 2024
- Publish AI RMF for Generative AI (GAI)
- Publish Secure Software Framework for GAI and dual-use models
- Launch initiative to create guidance/ benchmarks for evaluating AI capabilities
- Publish red-teaming guidelines
- Provide test environments
- Initiate engagement with industry and relevant synthetic nucleic acid sequencing providers
- Publish plan for global engagement on promoting and developing AI standards

Oct. 29, 2024
- Publish guidelines on the efficacy of differential-privacy-guarantee protections

Dec. 24, 2024
- Publish guidance for synthetic content authentication

Jan. 26, 2025
- Submit report to the President on priority actions taken pursuant to plan on global AI standards

This client alert is part of an ongoing series of client alerts focused on the EU AI Act. King & Spalding will continue to vigilantly monitor developments related to this and other AI-related legislation.

ABOUT KING & SPALDING

Celebrating more than 130 years of service, King & Spalding is an international law firm that represents a broad array of clients, including half of the Fortune Global 100, with 1,300 lawyers in 24 offices in the United States, Europe, the Middle East and Asia. The firm has handled matters in over 160 countries on six continents and is consistently recognized for the results it obtains, uncompromising commitment to quality, and dedication to understanding the business and culture of its clients.

This alert provides a general summary of recent legal developments. It is not intended to be and should not be relied upon as legal advice. In some jurisdictions, this may be considered "Attorney Advertising." View our Privacy Notice.

ABU DHABI	CHARLOTTE	DUBAI	LONDON	NORTHERN VIRGINIA	SILICON VALLEY
ATLANTA	CHICAGO	FRANKFURT	LOS ANGELES	PARIS	SINGAPORE
AUSTIN	DALLAS	GENEVA	MIAMI	RIYADH	TOKYO
BRUSSELS	DENVER	HOUSTON	NEW YORK	SAN FRANCISCO	WASHINGTON, D.C.