The Office of Science and Technology Policy (OSTP), a part of the Executive Office of the President, recently published a white paper titled "The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People" (Blueprint). This Blueprint offers a nonbinding framework for the responsible development of policies and practices around automated systems, including artificial intelligence (AI) and machine learning (ML) technologies.

Background

The Blueprint comes on the heels of bipartisan executive orders seeking to balance the benefits and potential risks of AI and ML technologies, and direct executive agencies and departments to develop policies around the responsible use of AI and ML. In one of his final executive orders, former President Trump required federal agencies to adhere to a set of principles when deploying AI technologies with the intention of fostering public trust in AI. And in one of his first executive orders, President Biden directed executive departments and agencies to address systemic inequities and embed fairness in decision-making processes. Some U.S. government agencies, including the Government Accountability Office (GAO) and the U.S. Department of Energy (DOE), have already developed frameworks for identifying and mitigating risks in the use of AI technology.

Scope and Purpose

The Blueprint is a nonbinding framework designed "to support the development of policies and practices that protect civil rights and promote democratic values" with automated systems. The OSTP emphasizes that the Blueprint is not a binding regulation. Rather, it outlines five guiding principles that the OSTP asserts should be applied to any automated system that could meaningfully affect civil rights, equal opportunity, or access to critical resources and services. The Blueprint also includes a handbook that provides detailed guidance on how to implement these principles in practice.

The Blueprint does not specifically define or limit itself to "artificial intelligence." Instead, it places any "automated system" under its scope, which is broadly defined as "any system, software, or process that uses computation as whole or part of a system to determine outcomes, make or aid decisions, inform policy implementation, collect data or observations, or otherwise interact with individuals and/or communities." The OSTP notes the definition explicitly includes, but is not limited to, AI, ML, and other data processing techniques. By its terms, the definition likely includes a much broader set of technologies that are not traditionally considered AI or ML.

The Five Principles

The OSTP Blueprint identifies five principles that should guide the "design, use, and deployment" of automated systems, which are summarized as follows:

1. **Safe and Effective Systems.** Automated systems should be developed in consultation with diverse stakeholders to help identify risks. They should be designed to protect from foreseeable harms and undergo both pre-deployment testing and ongoing monitoring to mitigate potential harmful outcomes.
2. **Algorithmic Discrimination Protections.** Developers and deployers should take efforts to protect the public from algorithmic discrimination from their automated systems. These efforts should include proactive equity assessments, use of representative data, and ensuring accessibility in the design process. After the system is deployed, there should be ongoing disparity testing and mitigation.
3. **Data Privacy.** Automated systems should have built-in data privacy protections that give users agency over how their data is used. Only data that is strictly necessary for the specific context should be collected and any data collection should conform to users' reasonable expectations. Developers and deployers

should seek consent regarding collection and use of data using consent requests that are brief and easily understandable. Sensitive data, including health, work, education, criminal, and children's data, should receive enhanced protections. Surveillance technologies should be subject to heightened oversight and continuous surveillance, and monitoring should not be used in education, work, housing, or in other contexts where its use is likely to limit rights, opportunities, or access.

4. **Notice and Explanation.** The public should be informed of where and how an automated system is being used and how it affects outcomes. Automated systems should come with publicly accessible, plain language documentation that provides notice that an automated system is being used and describes the function of the system, the purpose of the automation, the entity responsible for the system, and the outcomes.
5. **Human Alternatives, Consideration, and Fallback.** Where appropriate, users should be able to opt out from automated systems in favor of a human alternative. Automated systems should be connected to a fallback and escalation process with human consideration in the event that an automated system produces an error, fails, or an affected party otherwise wants to contest an automated decision.

For all five principles, the Blueprint emphasizes the use of independent evaluations and public reporting wherever possible to confirm adherence to the principles.

The OSTP has also published guidance on how to apply the Blueprint, as well as a 41-page "Technical Companion" which explains, for each of the five principles, (1) why the principle is important, (2) what should be expected of automated systems, and (3) how these principles can move into practice.

Foreshadowing the Future of AI Regulation

The Blueprint is the latest in a series of guidelines and frameworks recently published by various government entities and international organizations concerning the safe use of AI technologies, including the European Union's *Ethics Guidelines for Trustworthy AI*, the Organisation for Economic Co-operation and Development's *AI Principles*, and the GAO's *AI Accountability Framework*.

While these efforts have not yet yielded binding regulations or obligations, the growing focus on mitigating the potential harms of AI by both government entities and nongovernmental organizations suggests that future regulation is likely, particularly if an industry is unable to self-regulate against the potential harms. These guidelines and frameworks also portend the types of regulations that the future may bring. For example, they suggest that new laws, agency guidance, and industry policies could all be used to effectuate the goals of the Blueprint. The Artificial Intelligence, Machine Learning & Robotics industry group at Perkins Coie will continue to monitor changes to the AI regulatory landscape to better help clients navigate potential legal and regulatory issues during the development, testing, and launch of AI and ML products and services.

Authors

Marc S. Martin

Partner
MMartin@perkinscoie.com 202.654.6351

Nicola Menaldo

Partner
NMenaldo@perkinscoie.com 206.359.8000

David St. John-Larkin

Partner
DLarkin@perkinscoie.com 303.291.2365

Tyler D. Robbins

Associate
TRobbins@perkinscoie.com 202.654.3313

Explore more in

Technology Transactions & Privacy Law Privacy & Security Communications Advertising, Marketing & Promotions

Related insights

Update

FERC Meeting Agenda Summaries for October 2024

Update

New White House Requirements for Government Procurement of AI Technologies: Key Considerations for Contractors