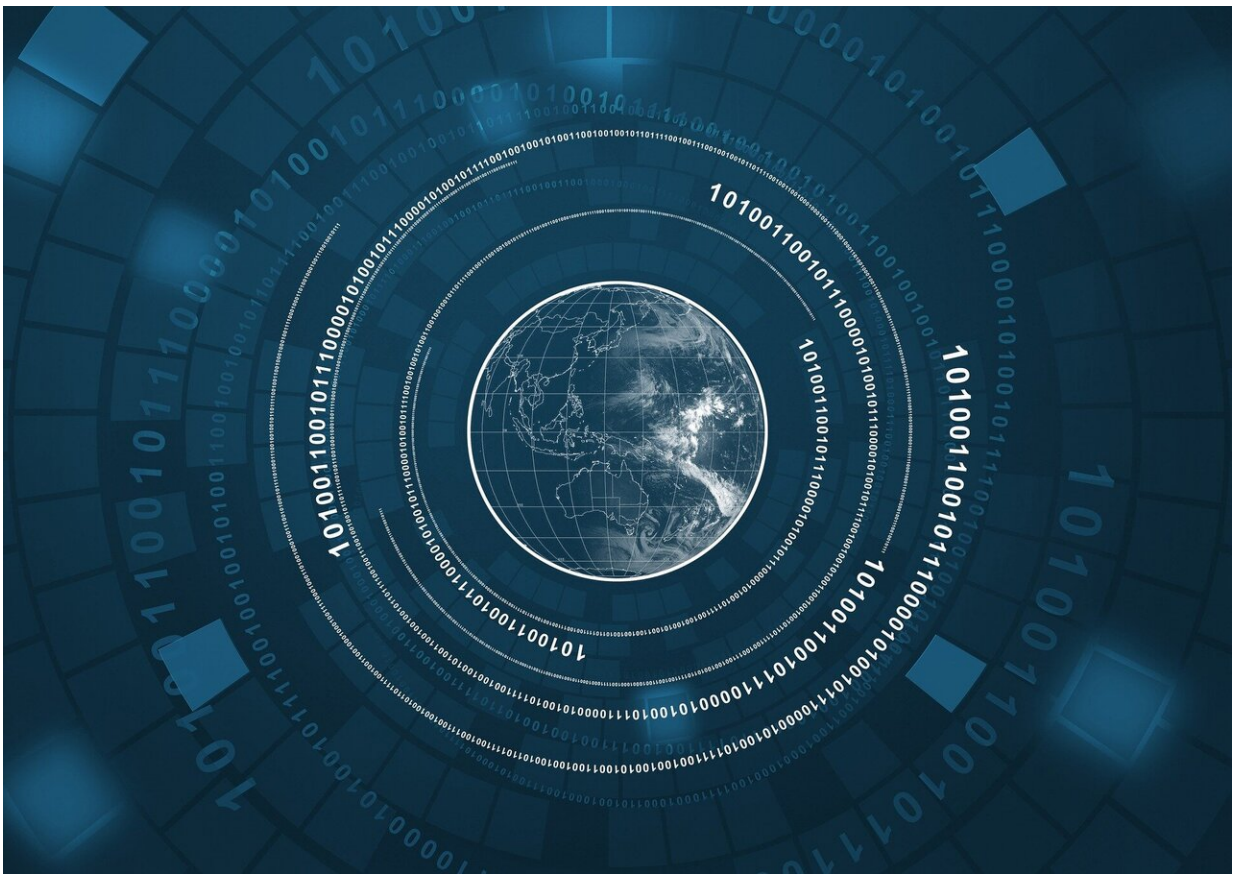
Researcher: Biden administration executive order tackles AI risks, but lack of privacy laws limits reach

November 3 2023, by Anjana Susarla



Credit: CC0 Public Domain

The comprehensive, even sweeping, set of guidelines for artificial

intelligence that the White House unveiled in an executive order on Oct. 30, 2023, show that the U.S. government is attempting to address the risks posed by AI.

As a researcher of information systems and responsible AI, I believe the executive order represents an important step in building responsible and trustworthy AI.

The order is only a step, however, and it leaves unresolved the issue of comprehensive data privacy legislation. Without such laws, people are at greater risk of AI systems revealing sensitive or confidential information.

Understanding AI risks

Technology is typically evaluated for performance, cost and quality, but often not equity, fairness and transparency. In response, researchers and practitioners of responsible AI have been advocating for:

- algorithm auditing
- standard reports on AI models
- credentials for otherwise opaque AI systems
- comprehensive risk mitigation practices
- AIs that are transparent to the public
- a recognition of the harms caused by AIs that make predictions about people

The National Institute of Standards and Technology (NIST) issued a comprehensive AI risk management framework in January 2023 that aims to address many of these issues. The framework serves as the foundation for much of the Biden administration's executive order. The executive order also empowers the Department of Commerce, NIST's home in the federal government, to play a key role in implementing the

proposed directives.

Researchers of AI ethics have long cautioned that stronger auditing of AI systems is needed to avoid giving the appearance of scrutiny without genuine accountability. As it stands, a recent study looking at public disclosures from companies found that claims of AI ethics practices outpace actual AI ethics initiatives. The executive order could help by specifying avenues for enforcing accountability.

Another important initiative outlined in the executive order is probing for vulnerabilities of very large-scale general-purpose AI models trained on massive amounts of data, such as the models that power OpenAI's ChatGPT or DALL-E. The order requires companies that build large AI systems with the potential to affect national security, public health or the economy to perform red teaming and report the results to the government. Red teaming is using manual or automated methods to attempt to force an AI model to produce harmful output—for example, make offensive or dangerous statements like advice on how to sell drugs.

Reporting to the government is important given that a recent study found most of the companies that make these large-scale AI systems lacking when it comes to transparency.

Similarly, the public is at risk of being fooled by AI-generated content. To address this, the executive order directs the Department of Commerce to develop guidance for labeling AI-generated content. Federal agencies will be required to use AI watermarking—technology that marks content as AI-generated to reduce fraud and misinformation—though it's not required for the private sector.

The executive order also recognizes that AI systems can pose unacceptable risks of harm to civil and human rights and the well-being of individuals: "Artificial Intelligence systems deployed irresponsibly

have reproduced and intensified existing inequities, caused new types of harmful discrimination, and exacerbated online and physical harms."

What the executive order doesn't do

A key challenge for AI regulation is the absence of comprehensive federal data protection and privacy legislation. The executive order only calls on Congress to adopt privacy legislation, but it does not provide a legislative framework. It remains to be seen how the courts will interpret the executive order's directives in light of existing consumer privacy and data rights statutes.

Without strong data privacy laws in the U.S. as other countries have, the executive order could have minimal effect on getting AI companies to boost data privacy. In general, it's difficult to measure the impact that decision-making AI systems have on data privacy and freedoms.

It's also worth noting that algorithmic transparency is not a panacea. For example, the European Union's General Data Protection Regulation legislation mandates "meaningful information about the logic involved" in automated decisions. This suggests a right to an explanation of the criteria that algorithms use in their decision-making. The mandate treats the process of algorithmic decision-making as something akin to a recipe book, meaning it assumes that if people understand how algorithmic decision-making works, they can understand how the system affects them. But knowing how an AI system works doesn't necessarily tell you why it made a particular decision.

With algorithmic decision-making becoming pervasive, the White House executive order and the international summit on AI safety highlight that lawmakers are beginning to understand the importance of AI regulation, even if comprehensive legislation is lacking.

Provided by The Conversation