WHITE HOUSE ADDS ITS VOICE TO GROWING CONSENSUS ON AI ETHICAL REQUIREMENTS

Eliza Davis

Martin Tully

On October 4, 2022, the White House Office of Science and Technology Policy ("OSTP") released the Blueprint for an AI Bill of Rights, laying out core principles for the responsible use of AI (artificial intelligence) systems. While the Blueprint is not binding and does not create any right or duty, it reinforces recent state legislation and federal agency regulatory guidance around transparency, testing, and auditing of AI system behavior and related data privacy issues.

Regulators such as the Federal Trade Commission have announced their intent to "crackdown" on corporate surveillance by any means, including AI tools. Further, AI tools that rely upon personal data collected from consumers will likely need to navigate both state and federal data privacy laws restricting the use and retention of such information. To meet these requirements, corporations will increasingly need leading edge guidance in data law and AI system behavior.

The Regulatory Landscape

The Blueprint continues a theme seen in recent AI initiatives at several levels of government requiring transparency regarding the use and behavior of AI systems. Laws relating to AI have been passed by Illinois (Artificial Intelligence Video Interview Act, 2020), Maryland (HB 1202 re Use of Facial Recognition Services, 2020), and New York City (Int. No. 1894-A requiring bias audits for AI employment use, 2021). More proposed AI legislation is pending in several other states.

Federal agencies have also focused on transparency, bias, and privacy issues. The U.S. Department of Defense recently published a Responsible AI Implementation Pathway in June 2022, following its 2020 Ethical Principles for Artificial Intelligence. The Federal Trade Commission ("FTC") warned last year that it might take enforcement actions against violations involving AI that are unfair practices under FTC Act Section 5, or involve discrimination under the Fair Credit Reporting Act ("FCRA") or the Equal Credit Opportunity Act ("ECOA"). Similarly, the Consumer Financial Protection Bureau ("CFPB") has outlined options to ensure AI-generated home valuations are "accurate and fair." The Department of Justice ("DOJ") and Equal Opportunity Commission ("EEOC") each released technical assistance documents relating to AI use and protection from discrimination against disabled persons.

AI regulation is, if anything, moving even faster in Europe. In April 2021, the European Commission presented its proposal for a European Union regulatory framework on AI. The proposed Artificial Intelligence Act ("AI Act") is the first proposed law to focus on

regulating AI. The AI Act has many similarities with the EU General Data Protection Regulation ("GDPR")– including potentially significant consequences for organizations that develop, sell, or use AI systems, as well having an extraterritorial reach.

The EU's AI Act proposes to ban "unacceptable risk" AI uses, such as social scoring. In addition, it creates a "high-risk application" category that would be subject to specific legal requirements. Finally, there would be "minimal risk" applications that would still be subject to the GDPR but not additional legal obligations. Notably, the AI Act includes an enforcement mechanism for each Member State to designate at least one national authority to supervise the AI Act application and compliance.

The AI Bill of Rights in Detail

The White House OSTP's Blueprint for an AI Bill of Rights results from a year-long process that involved panel discussions, meetings, and comments received from the public. It contains five principles:

1. Safe and Effective Systems: Systems should reflect input from diverse communities, stakeholders, and domain experts to help ensure they are safe and effective while identifying potential risks and concerns. This principle supports AI system testing, vetting of the data used, and continuous auditing for accuracy and bias. Developers and users of AI should provide regular reports regarding AI systems, including what they do, how they work, performance expectations, descriptions of training data, and potential concerns.

2. Algorithmic Discrimination Protections: Systems should be designed and implemented in an equitable way, without discrimination against protected classes. Designers, developers, and deployers of AI systems must proactively and continuously take steps to protect against the use of data that will serve as a proxy for protected demographic features. This principle requires training to include representative datasets. Independent "algorithm impact assessments" should be performed whenever possible.

3. Data Privacy: Individuals should have control over their personal data and should be protected from pervasive surveillance. The collection of data for AI systems should be limited to only what is "strictly necessary for the specific context." Consent to use of personal data should be sought in a way that it can be given by data subjects in an appropriate and meaningful way. AI systems should incorporate privacy by design, and sensitive data should be subject to very strict controls. In addition, continuous surveillance should not be used where it may limit rights, opportunities, or access.

4. Notice and Explanation: Individuals should know when AI is impacting a decision. Accessible, plain language should be used to provide notice and explain algorithmic systems so that individuals will understand the systems and their impacts, and updates should be provided. Reporting should be made public when possible.

5. Human Alternatives, Considerations, and Fallback: Individuals should be able to opt out of AI systems when possible, and should be able to access a person to address problems. There is a right for meaningful recourse when errors occur, including a fallback and escalation processes, and an appeal processes.

For each principle, the OSTP Blueprint includes an explanation of why it is important, including examples of problematic AI uses that have contradicted that principle. For example, it describes AI tools that deployed police to neighborhoods that did not actually have the highest crime rates; discriminatory AI hiring tools; AI employee surveillance used to track and intervene in discussions about union activity; denials of benefits to individuals after an AI system's criteria changed invisibly; and a fraud warning system that withheld wages without giving people a chance to explain themselves or receive human review.

The Blueprint also provides expectations for AI systems for each principle. The expectations generally involve disclosing and explaining the AI use, engaging with affected communities or individuals, diligent testing and monitoring for accuracy and equitable impact, other audits, and reporting. Finally, each principle is also supported by possibilities on how the principle can be implemented. It refers to initiatives underway by various agencies, state legislatures, industry groups, and individual employers.

Takeaways

The OSTP Blueprint for an AI Bill of Rights adds the White House's voice to existing guidance regarding the importance of the ethical use of AI. Federal and state regulation of AI will almost certainly continue to evolve quickly, alongside related regulation of personal data use and privacy. Organizations can best prepare for the future by implementing ethical AI practices when possible, including protecting private data, anticipating and measuring potential discriminatory impacts, and creating compliance programs subject to regular review and updates.

For assistance with or additional information on this topic, please contact Martin Tully at mtully@redgravellp.com or Eliza Davis at edavis@redgravellp.com.