

Seven Leading AI Development Firms Agree to Conduct Principles That May Be Regarded as Initial Industry 'Best Practices'

Article By:

Gary De Waal

Daniel J. Davis

Seven leading developers of artificial intelligence technology voluntarily agreed to comply with eight specific commitments ([Commitments](#)) to help mitigate the risks of AI, as announced in a ceremony at the White House on July 21, 2023. These eight commitments emanate from three core principles that the White House said “must be fundamental to the future of AI: safety, security and trust.”

These commitments included agreements to:

1. test all major public releases of new AI models, relying on both internal and, at least in part, independent experts. Testing will consider, among other matters, national security implications; cyber security; and societal risks (e.g., bias and discrimination);
2. share information across companies, governments, civil society and academia on managing AI risks, including joining forums or cooperate in other ways to “... develop, advance and adopt shared standards and best practices for ... AI safety;
3. treat AI model weights as “core intellectual property,” and maintain adequate cybersecurity protection and protection from threats from insiders appropriate for the most sacrosanct intellectual property and trade secrets. According to the Commitments, “[t]hese model weights are the most essential part of an AI system, and the companies agree that it is vital that the model weights be released only when intended and when security risks are considered.” (Generally, AI model weight reflects how much weight any input to an AI system will have on its output.);
4. encourage third-party discovery and reporting of system issues and vulnerabilities to help identify issues after an AI system is disseminated. These could include instituting bounty systems or contests to encourage reporting;
5. develop and institute methods so users will know when content is AI generated (e.g., watermarking);
6. help ensure that users understand the capabilities and limitations of all new and material public releases of an AI system;
7. prioritize research to help ensure that neither detrimental biases nor discrimination is propagated by AI systems, and that privacy is protected; and

8. develop and deploy AI systems to help “address society’s greatest challenges,” including preventing cancer and detecting it early as well as mitigating cyber threats.

These commitments, however, may be important for all firms – not just the firms that agreed to them – including financial services firms, to consider as part of any governance implemented for developing and/or deploying AI for their own purposes (to the extent relevant). This is because these commitments could be regarded as among the first “best practices” in AI and could take on added significance in the absence of relevant legislation.

In May 2023, the White House released a blueprint for an “[AI Bill of Rights](#),” that also included principles to guide the “design, use and deployment automated systems to protect the American public in the age artificial intelligence.”

For additional information on the Commitments, click [here](#).

"Companies that are developing [AI] technologies have a responsibility to ensure their products are safe." -- The White House



[www.whitehouse.gov/...](https://www.whitehouse.gov/)

©2025 Katten Muchin Rosenman LLP

National Law Review, Volume XIII, Number 229

Source URL: <https://natlawreview.com/article/seven-leading-ai-development-firms-agree-to-conduct-principles-may-be-regarded>