

Apple agrees to abide by White House AI safeguards

July 26 2024



Credit: Armand Valendez from Pexels

The White House on Friday said that Apple has joined more than a dozen tech firms committed to following safeguards intended to curb the risks of artificial intelligence.

Amazon, Google, Microsoft and OpenAI were among the AI sector rivals who joined US officials in unveiling the voluntary pact a year ago.

President Joe Biden's administration said at the time that it had secured commitments from the companies "to help move toward safe, secure, and transparent development of AI technology."

The safeguards include tech companies simulating attacks on AI models in a testing technique referred to as "red teaming" to expose flaws or vulnerabilities.

Testing of AI models or systems is to include societal risks and national security concerns such as cyber assaults and developing [biological weapons](#), according to the White House.

Companies that signed on to the commitment are to work on sharing information with each other and the government about AI dangers and attempts to circumvent defenses.

Apple in June unveiled "Apple Intelligence," its suite of AI features for its coveted devices as it looks to reassure users that it is not falling behind on the AI frenzy.

Apple's announcement included a partnership with OpenAI that would make ChatGPT available to iPhone users on request.

Biden late last year signed an executive order setting new safety standards for AI systems and requiring developers to share results of safety tests with the US government.

Biden's executive order was touted by the White House as "the most sweeping actions ever taken to protect Americans from the potential risks of AI systems."

Shortly after it was issued, Vice President Kamala Harris gave a major AI policy speech at a gathering of politicians, tech industry figures and academics.

The event focused on growing fears about the implications of advanced AI models that have prompted concerns around everything from job losses and cyber-attacks to humankind losing control of the very systems it created.

© 2024 AFP

Citation: Apple agrees to abide by White House AI safeguards (2024, July 26) retrieved 27 July 2024 from <https://techxplore.com/news/2024-07-apple-white-house-ai-safeguards.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.