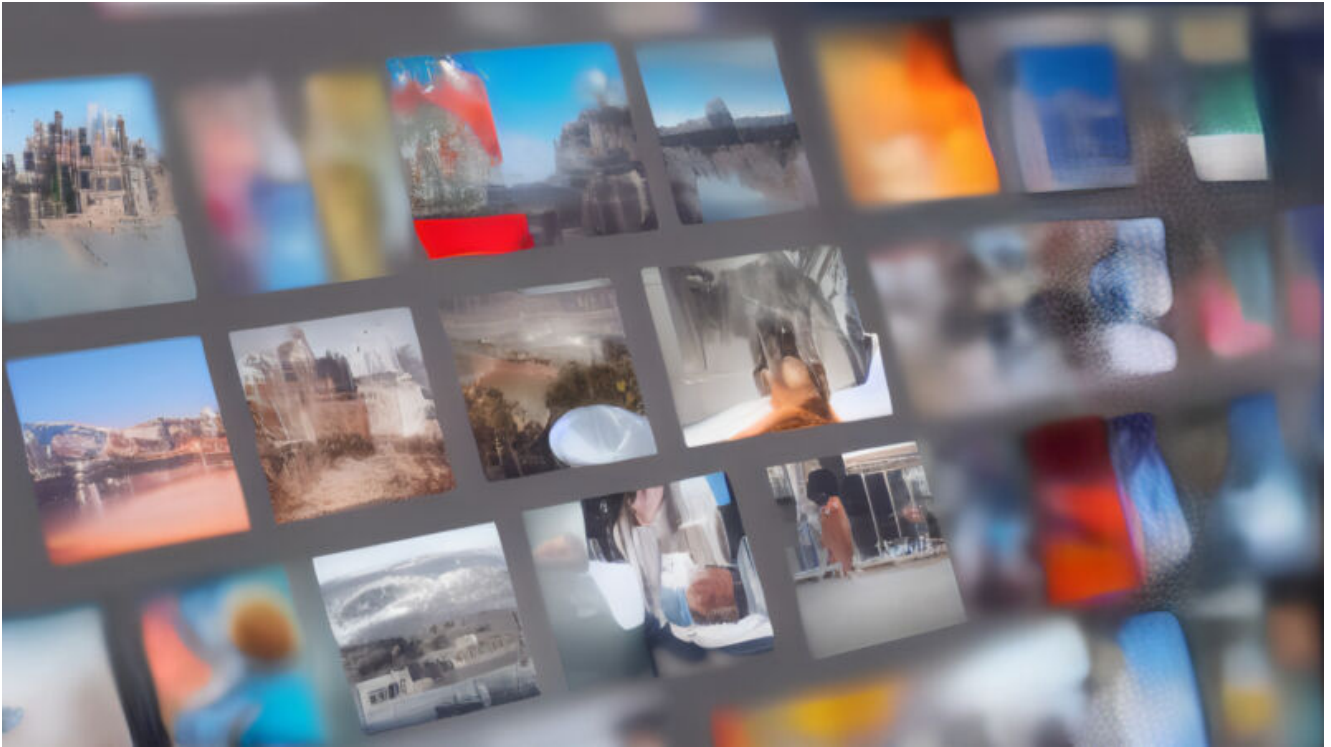


Among AI dangers, deepfakes worry Microsoft president most



Enlarge / An AI-generated image of a “wall of fake images.”
Stable Diffusion

reader comments

28 with

On Thursday, Microsoft President Brad Smith announced that his biggest apprehension about AI revolves around the growing concern for deepfakes and synthetic media designed to deceive, Reuters reports.

Smith made his remarks while revealing his “blueprint for public governance of AI” in a speech at Planet World, a language arts museum in Washington, DC. His concerns come when talk of AI regulations is increasingly common, sparked largely by the popularity of OpenAI’s ChatGPT and a political tour by

OpenAI CEO Sam Altman.

Smith expressed his desire for urgency in formulating ways to differentiate between genuine photos or videos and those created by AI when they might be used for illicit purposes, especially in enabling society-destabilizing disinformation.

“We’re going have to address the issues around deepfakes. We’re going to have to address in particular what we worry about most foreign cyber influence operations, the kinds of activities that are already taking place by the Russian government, the Chinese, the Iranians,” Smith said, according to Reuters. “We need to take steps to protect against the alteration of legitimate content with an intent to deceive or defraud people through the use of AI.”

Smith also pushed for the introduction of licensing for critical forms of AI, arguing that these licenses should carry obligations to protect against threats to security, whether physical, cybersecurity, or national. “We will need a new generation of export controls, at least the evolution of the export controls we have, to ensure that these models are not stolen or not used in ways that would violate the country’s export control requirements,” he said.

Advertisement

Last week, Altman appeared at the US Senate and voiced his concerns about AI, saying that the nascent industry needs to be regulated. Altman, whose company OpenAI is backed by Microsoft, argued for global cooperation on AI and incentives for safety compliance.

In his speech Thursday, Smith echoed these sentiments and insisted that people must be held accountable for the problems caused by AI. He urged for safety measures to be put on AI systems controlling critical infrastructure, like the electric grid and water supply, to ensure human oversight.

In an effort to maintain transparency around AI technologies,

Smith urged that developers should develop a “know your customer”-style system to keep a close eye on how AI technologies are used and inform the public about content created by AI, making it easier to identify fabricated content. Along these lines, companies such as Adobe, Google, and Microsoft are all working on ways to watermark or otherwise label AI-generated content.

Deepfakes have been a subject of research at Microsoft for years now. In September, Microsoft’s Chief Scientific Officer Eric Horvitz penned a research paper about the dangers of both interactive deepfakes and the creation of synthetic histories, subjects also covered in a 2020 article in FastCompany by this author, which also mentioned earlier efforts from Microsoft at detecting deepfakes.

Meanwhile, Microsoft is simultaneously pushing to include text- and image-based generative AI technology into its products, including Office and Windows. Its rough launch of an unconditioned and undertested Bing chatbot (based on a version of GPT-4) in February spurred deeply emotional reactions from its users. It also reignited latent fears that world-dominating superintelligence may be just around the corner, a reaction that some critics claim is part of a conscious marketing campaign from AI vendors.

So the question remains: What does it mean when companies like Microsoft are selling the very product that they are warning us about?

<https://arstechnica.com/?p=1942347>