

Microsoft releases AI tool for photorealistic copying of faces and voices

Publication Date: 2023-11-17

Author: Kari Paul

Section: Technology

Tags: Artificial intelligence (AI), Microsoft, Deepfake, Computing, news

Article URL: <https://www.theguardian.com/technology/2023/nov/17/microsoft-azure-ai-video-deepfakes>



Microsoft announced its latest contribution to the artificial intelligence race at its developer conference this week: software that can generate new avatars and voices or replicate the existing appearance and speech of a user – raising concerns that it could supercharge the creation of deepfakes, AI-made videos of events that didn't happen. Announced at Microsoft Ignite 2023, Azure AI Speech is trained with human images and allows users to input a script that can then be “read” aloud by a photorealistic avatar created with artificial intelligence. Users can either choose a preloaded Microsoft avatar or upload footage of a person whose voice and likeness they want to replicate. Microsoft said in a blog post published on Wednesday that the tool could be used to build “conversational agents, virtual assistants, chatbots and more”. The post reads: “Customers can choose either a prebuilt or a custom neural voice for their avatar. If the same person's voice and likeness are used for both the custom neural voice and the custom text to speech avatar, the avatar will closely resemble that person.” The company said the new text-to-speech software is being released with a variety of limits and safeguards to prevent misuse. “As part of Microsoft's commitment to responsible AI, text to speech avatar is designed with the intention of protecting the rights of individuals and society, fostering transparent human-computer interaction, and counteracting the proliferation of harmful deepfakes and misleading content,” the company said. “Customers can upload their own video recording of avatar talent, which the feature uses to train a synthetic video of the custom avatar speaking,” the blog post reads. “Avatar talent” is a human posing for the AI's proverbial camera. The announcement quickly elicited criticism that Microsoft had launched a “deepfakes creator” – which would more easily allow a person's likeness to be replicated and made to say and do things the person has not said or done. Microsoft's own president said in May that deepfakes are his “biggest concern” when it comes to the rise of artificial intelligence. In a statement, the company pushed back on the criticism, saying the customized avatars are now a “limited access” tool for which customers have to apply and be approved for by Microsoft. Users will also be required to disclose when AI was used to create a synthetic voice or avatar. “With these safeguards in place, we help limit potential risks and empower customers to infuse advanced voice and speech capabilities into their AI applications in a transparent and safe manner,” Sarah Bird of Microsoft's responsible AI engineering division said in a statement. The text-to-speech avatar maker is the latest tool as major tech firms have raced to capitalize on the artificial intelligence boom of recent years. After the runaway popularity of ChatGPT – launched by Microsoft-backed firm OpenAI – companies like Meta and Google have pushed their own artificial intelligence tools to the market. With AI's rise has come growing concerns about the capabilities of the technology, with the OpenAI CEO, Sam Altman, warning Congress that it could be used for election interference and safeguards must be implemented. Deepfakes pose particular danger when it comes to

election interference, experts say. Microsoft launched a tool earlier this month to allow politicians and campaigns to authenticate and watermark their videos to verify their legitimacy and prevent the spread of deepfakes. Meta announced a policy this week requiring the disclosure of the use of AI in political ads and banning campaigns from using Meta's own generative AI tools for ads.