

# AI firms should face prison over creation of fake humans, says Yuval Noah Harari

Publication Date: 2023-07-06

Author: Hannah Devlin

Section: Technology

Tags: Artificial intelligence (AI), Yuval Noah Harari, Computing, Consciousness, Neuroscience, Social media, ChatGPT, news

Article URL: <https://www.theguardian.com/technology/2023/jul/06/ai-firms-face-prison-creation-fake-humans-yuval-noah-harari>



The creators of AI bots that masquerade as people should face harsh criminal sentences comparable to those who trade in counterfeit currency, the Israeli historian and author Yuval Noah Harari has said. He also called for sanctions, including prison sentences, to apply to tech company executives who fail to guard against fake profiles on their social media platforms. Addressing the UN's AI for Good global summit in Geneva, the author of *Sapiens* and *Home Deus* said the proliferation of fake humans could lead to a collapse in public trust and democracy. "Now it is possible, for the first time in history, to create fake people – billions of fake people," he said. "If this is allowed to happen it will do to society what fake money threatened to do to the financial system. If you can't know who is a real human, trust will collapse. "Maybe relationships will be able to manage somehow, but not democracy," Harari added. The advent of ChatGPT and other large language models means AI bots can not only amplify human content, but also artificially generate their own content at scale. "What happens if you have a social media platform where ... millions of bots can create content that is in many ways superior to what humans can create – more convincing, more appealing," he said. "If we allow this to happen, then humans have completely lost control of the public conversation. Democracy will become completely unworkable." Part of the solution, he argued, should be the introduction of severe criminal sentences for those who create bots and swamp public forums with artificially generated content. "It is very easy to create fake money – whether it is coins or bank notes," he said. "In order to protect the financial system, governments have very strict rules against counterfeit money." If tech companies were required to take countermeasures, or face 20-year jail sentences, he predicted they would "quickly find ways to prevent the platforms from becoming overwhelmed with fake people". Harari also called for greater investment into AI safety, suggesting companies should be legally required to commit 20% of investment spending to researching risks and how to manage them. Harari's comments followed an announcement by OpenAI, the maker of ChatGPT, on Thursday that it would dedicate 20% of its computing power to AI safety, including creating a research team dedicated to understanding how to control and steer AI and stop it from going rogue. "We need to invest a lot more in safety and regulation, said Harari.