

# Sora gives deepfakes 'a publicist and a distribution deal.' It could change the internet

[Bobby Allyn](#)

Online safety experts say something else that is happening may be less obvious but more consequential to the future of the internet: OpenAI has essentially rebranded deepfakes as a light-hearted plaything and recommendation engines are loving it.

*OpenAI*

Videos made with OpenAI's Sora app are flooding TikTok, Instagram Reels and other platforms, making people increasingly familiar — and fed up — with nearly unavoidable synthetic footage being pumped out by what amounts to an artificial intelligence slop machine.

Digital safety experts say something else that is happening may be less obvious but more consequential to the future of the internet: OpenAI has essentially rebranded deepfakes as a light-hearted plaything and recommendation engines are loving it. As the videos race across millions of peoples' feeds, perceptions are being quickly reshaped about the truth, and soon, perhaps the basic norms of being online.

"It's as if deepfakes got a publicist and a distribution deal," said Daisy Soderberg-Rivkin, a former trust and safety manager at TikTok. "It's an amplification of something that has been scary for a while, but now it has a whole new platform."

Aaron Rodericks, Bluesky's head of trust and safety, said the public is not ready for a collapse between reality and fakery this drastic.

"In a polarized world, it becomes effortless to create fake evidence targeting identity groups or individuals, or to scam people at scale. What used to be an inflammatory rumour — like a fabricated story about an immigrant or a politician — can now be rendered as believable video proof," Rodericks said. "Most people won't have the media literacy or the tools to tell the difference."

NPR spoke with three former OpenAI employees, who all said they were not surprised the company would launch a social media app showcasing its latest video technology as investor pressure builds for the company to dazzle the world as it did three years ago after the release of ChatGPT.

Like with the chat bot, changes are already coming fast.

OpenAI included numerous guardrails with Sora, including moderation, restrictions on scammy, violent and

pornographic material, watermarks and controls over how one's likeness is used. Some of the safety rails are being gamed by users intent on finding workarounds. And in response, OpenAI scrambles to plug the holes.

One former OpenAI employee who was not authorized to speak publicly said being worried about safety protections dropping over time is a legitimate fear.

"Releasing Sora tells the world where the party is going. AI videos may be the last conquered social media frontier, and OpenAI wants to own it, but as Silicon Valley competes over it, companies will bend the rules to stay competitive, and that could be bad for society," the person said.

Former TikTok manager Soderberg-Rivkin said it is just a matter of time before a developer releases a Sora-esque app with no safety rules, similar to how [Elon Musk built Grok specifically as an "anti-woke"](#) and more unrestrained answer to leading chatbots.

"When there's an unregulated version with no safety rails, it will be used to generate synthetic child sexual abuse material that bypasses current detections," said Rodericks, whose employer, Bluesky has [leaned](#) into customizable content moderation to set it apart from platforms like X, where there are fewer rules. "You'll see state-sponsored actors fabricating realistic news segments and propaganda

to legitimize false narratives."

A spokesperson for OpenAI declined to comment.

## **Experts say might be too late for a 'no AI policy'**

Sora is currently the No. 1 most-downloaded iPhone app, but people can only use the app with an invite code from current users.

Those who have been regularly using it have already noticed how it has become more restrictive. Doing videos of celebrities has become more difficult. Trying to replicate some of the more outlandish videos that have been created like a fake Jeffrey Epstein on a boat heading toward an island, or a replica of Sean "Diddy" Combs' addressing his prison sentence, has gotten trickier. Yet other controversial prompts, like having someone be arrested, or putting someone in a Nazi uniform, are still generating videos.

"We're already at the point where we can't tell what's real and what's not online, and OpenAI and other tech companies will have to solve around that," said a former OpenAI engineer.

*OpenAI*

OpenAI CEO Sam Altman [wrote](#) days after the Sora app was unveiled that rights holders are going to have more control over how their likeness is used, shifting the default approach

from "opt-out" to "opt in." Altman also wrote that Sora will, eventually, share revenue the app makes with rightsholders.

"Please expect a very high rate of change from us; it reminds me of the early days of ChatGPT," he said.

The constant stream of [AI slop](#) filling up everyone's feeds has raised the question of whether users will tire of AI videos as a genre of content.

Most major video platforms now have relatively loose policies around sharing AI videos, but will Sora trigger a backlash? Could that force social media companies to institute a crackdown or ban on AI-generated content? Not likely, said Soderberg-Rivkin, who points out that even if that were to happen, enforcement would run up against just how sophisticated leading AI generators have become.

"If you say no AI use on a social media platform, the fact of the matter is it's getting harder and harder to detect when text, videos and images are AI, which is scary," she said. "A no-AI policy will not stop AI from sneaking in."

## **'The liar's dividend' supercharged like never before**

Another former OpenAI employee who also was not authorized to speak publicly argued that releasing a

deepfake AI social media platform was the right business decision, even if it contributes to the collapse of everyone's shared sense of reality.

"We're already at the point where we can't tell what's real and what's not online, and OpenAI and other tech companies will have to solve around that," said the former OpenAI engineer, using tech lingo for finding a solution to a problem. "But that's not an argument for not trying to dominate this market. You can't stop progress. If OpenAI didn't release Sora, someone else would have."

In fact, Meta is also trying to, with its recent introduction of Vibes, a platform where people can make and share short AI-generated deepfakes. In July, Google introduced Veo 3, an AI video tool. But it wasn't until OpenAI's release of the Sora app that personalized AI slop really took off.

Trust and safety professionals, like Soderberg-Rivkin, say Sora will likely mark a turning point in the history of the internet, the moment when deepfakes went from a mostly one-off phenomenon to the status quo, which may make people disengage with social media more, or at least shatter faith in the integrity of what people are watching online.

Disinformation experts have long warned about a concept [known as](#) "the liar's dividend," when the proliferation of deepfakes allows individuals, especially people in power, to

dismiss real content as fabrications. But now, experts say, that reality is more acute than ever before.

"I'm less worried about a very specific nightmare scenario where deepfakes swing an election, and I'm really more worried about a baseline erosion of trust," Soderberg-Rivkin said. "In a world where everything can be fake, and the fake stuff looks and feels real, people will stop believing everything."