INSIDER INTELLIGENCE | eMarketer

JUN 03
Gadjo Sevilla

Google takes a stand on deepfake AI, bans training models in lab

Article

The news: **Google** is banning the training of AI systems that can be used to create deepfakes on its platforms like **Google Colaboratory**, per TechCrunch.

Disallowing deepfakes: Deepfakes, which superimpose a person's face on top of another to create realistic videos, have become increasingly realistic thanks to artificial intelligence. AI can match body movements, microexpressions, and skin tones more accurately than CGI animation.

BleepingComputer and **Unite.ai** spotted updated terms of use, including deepfake-related work, in the disallowed projects list on Google's **Colab**, a service allowing coders to write and execute arbitrary computer code through web browsers.

- Colab has been a key platform for running demos within the AI research community. Google has taken a laissez-faire attitude on what code it allows on Colab, potentially attracting nefarious users.
- Users of the open-source deepfake generator **DeepFaceLab** recently received error messages after trying to run DeepFaceLab in Colab.
- The warning read: "You may be executing code that is disallowed, and this may restrict your ability to use Colab in the future. Please note the prohibited actions specified in our FAQ."
- Some deepfake code will still run without errors or warning. In context, **FaceSwap**, a deepfake app, still runs without issue.

The bigger picture: "Deterring abuse is an ever-evolving game, and we cannot disclose specific methods as counterparties can take advantage of the knowledge to evade detection systems," a Google spokesperson told TechCrunch. "In general, we have automated systems that detect and prohibit many types of abuse."

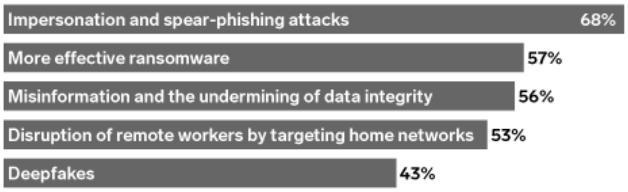
- Previous Colab restrictions voiding terms of service include running denial-of-service attacks, password cracking, and downloading torrents, which are potentially illegal activities.
- Deepfakes are being used more often by hackers to spread disinformation and fraud. **Deepfakes online increased from around 14,000 in 2019 to 145,000 in 2021.**

What's next? Google and other companies looking to regulate deepfake AI will need to enforce more comprehensive controls to effectively clamp down on its use.

- Lopsided deepfake regulation could result in a rise of disinformation as well as a proliferation of extortion and fraud schemes.
- In context, deepfake scams are increasing, per SHRM. Forrester Research estimated that these scams cost $250 million in 2020.

How AI May Be Used Against Organizations in Possible Cyberattacks According to Business Leaders Worldwide, Jan 2021

% of respondents



Note: respondents were asked to choose all that apply
Source: MIT Technology Review and Darktrace, "Preparing for AI-Enabled Cyberattacks," April 8, 2021

265577 InsiderIntelligence.com