Artists use tech weapons against AI copycats

December 24, 2023

Online Desk: Artists under siege by artificial intelligence (AI) that studies their work, then replicates their styles, have teamed with university researchers to stymy such copycat activity.

US illustrator Paloma McClain went into defense mode after learning that several AI models had been "trained" using her art, with no credit or compensation sent her way.

"It bothered me," McClain told AFP.

"I believe truly meaningful technological advancement is done ethically and elevates all people instead of functioning at the expense of others."

The artist turned to free software called Glaze created by researchers at the University of Chicago.

Glaze essentially outthinks AI models when it comes to how they train, tweaking pixels in ways indiscernible by human viewers but which make a digitized piece of art appear dramatically different to AI.

"We're basically providing technical tools to help protect human creators against invasive and abusive AI models," said professor of computer science Ben Zhao of the Glaze team.

Created in just four months, Glaze spun off technology used to disrupt facial recognition systems.

"We were working at super-fast speed because we knew the problem was serious," Zhao said of rushing to defend artists from software imitators.

"A lot of people were in pain."

Generative AI giants have agreements to use data for training in some cases, but the majority if digital images, audio, and text used to shape the way supersmart software thinks has been scraped from the internet without explicit consent.

Since its release in March of 2023, Glaze has been downloaded more than 1.6 million times, according to Zhao.

Zhao's team is working on a Glaze enhancement called Nightshade that notches up defenses by confusing AI, say by getting it to interpret a dog as a cat.

"I believe Nightshade will have a noticeable effect if enough artists use it and put enough poisoned images into the wild," McClain said, meaning easily available online.

"According to Nightshade's research, it wouldn't take as many poisoned images as one might think."

Zhao's team has been approached by several companies that want to use Nightshade, according to the Chicago academic.

"The goal is for people to be able to protect their content, whether it's individual artists or companies with a lot of intellectual property," said Zhao.

– Viva Voce –

Startup Spawning has developed Kudurru software that detects attempts to harvest large numbers of images from an online venue.

An artist can then block access or send images that don't match what is being requested, tainting the pool of data being used to teach AI what is what, according to Spawning cofounder Jordan Meyer.

More than a thousand websites have already been integrated into the Kudurru network.

Spawning has also launched haveibeentrained.com, a website that features an online tool for finding out whether digitized works have been fed into an AI model and allow artists to opt out of such use in the future.

As defenses ramp up for images, researchers at Washington University in Missouri have developed AntiFake software to thwart AI copying voices.

AntiFake enriches digital recordings of people speaking, adding noises inaudible to people but which make it "impossible to synthesize a human voice," said Zhiyuan Yu, the PhD student behind the project.

The program aims to go beyond just stopping unauthorized training of AI to preventing creation of "deepfakes" — bogus soundtracks or videos of celebrities, politicians, relatives, or others showing them doing or saying something they didn't.

A popular podcast recently reached out to the AntiFake team for help stopping its productions from being hijacked, according to Zhiyuan Yu.

The freely available software has so far been used for recordings of people speaking, but could also be applied to songs, the researcher said.

"The best solution would be a world in which all data used for AI is subject to consent and payment," Meyer contended.

"We hope to push developers in this direction."