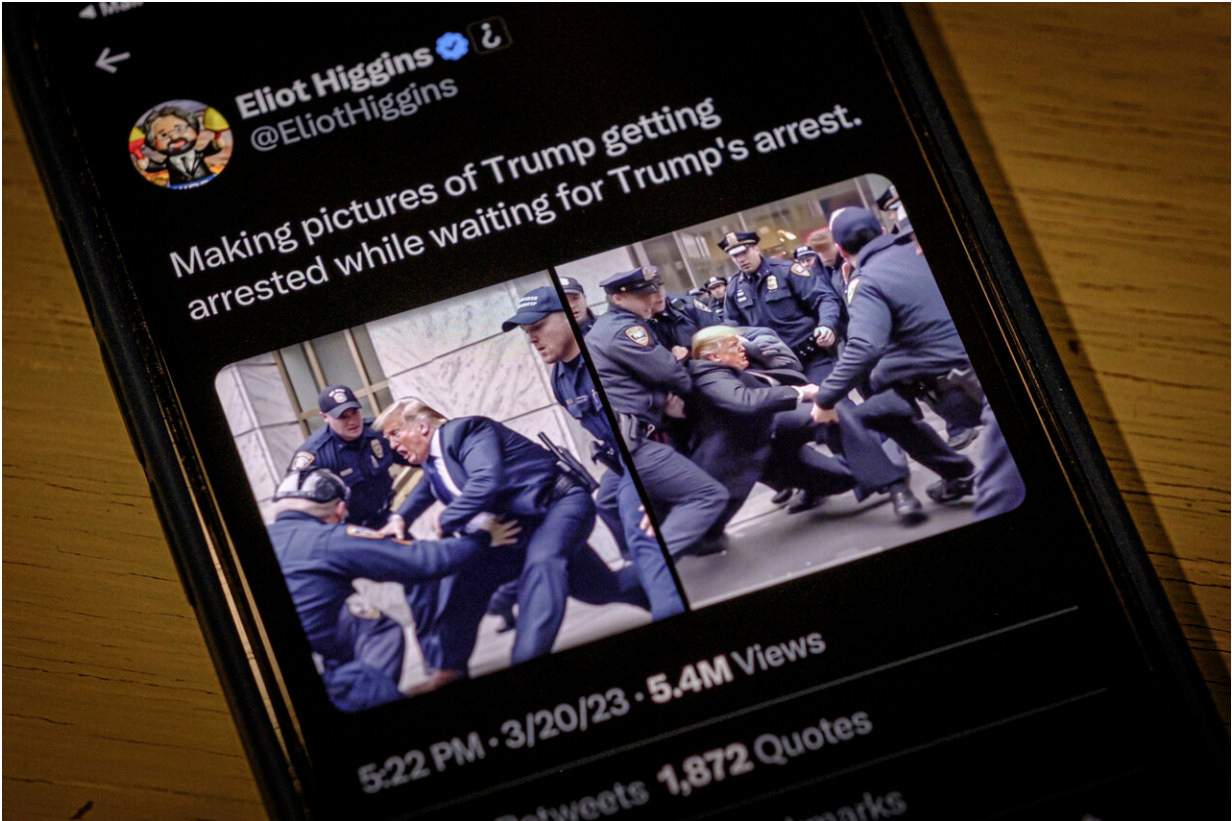# Trump arrested? Putin jailed? Fake AI images spread online

March 24 2023, by ARIJETA LAJKA and PHILIP MARCELO

Images created by Eliot Higgins with the use of artificial intelligence show a fictitious skirmish with Donald Trump and New York City police officers posted on Higgins' Twitter account, as photographed on an iPhone in Arlington, Va., Thursday, March 23, 2023. The highly detailed, sensational images, which are not real, were produced using a sophisticated and widely accessible image generator. Credit: AP Photo/J. David Ake

Former President Donald Trump getting gang-tackled by riot-gear-clad New York City police officers. Russian President Vladimir Putin in prison grays behind the bars of a dimly lit concrete cell.

The highly detailed, sensational images have inundated Twitter and other platforms in recent days, amid news that Trump faces possible criminal charges and the International Criminal Court has issued an arrest warrant for Putin.

But neither visual is remotely real. The images—and scores of variations littering social media—were produced using increasingly sophisticated and widely accessible image generators powered by artificial intelligence.

Misinformation experts warn the images are harbingers of a new reality: waves of fake photos and videos flooding social media after major news events and further muddying fact and fiction at crucial times for society.

"It does add noise during crisis events. It also increases the cynicism level," said Jevin West, a professor at the University of Washington in Seattle who focuses on the spread of misinformation. "You start to lose trust in the system and the information that you are getting."

While the ability to manipulate photos and create fake images isn't new, AI image generator tools by Midjourney, DALL-E and others are easier to use. They can quickly generate realistic images—complete with detailed backgrounds—on a mass scale with little more than a simple text prompt from users.

Some of the recent images have been driven by this month's release of a new version of Midjourney's text-to-image synthesis model, which can, among other things, now produce convincing images mimicking the style of news agency photos.

In one widely-circulating Twitter thread, Eliot Higgins, founder of Bellingcat, a Netherlands-based investigative journalism collective, used the latest version of the tool to conjure up [scores of dramatic images](#) of Trump's fictional arrest.

The visuals, which have been shared and liked tens of thousands of times, showed a crowd of uniformed officers grabbing the Republican billionaire and violently pulling him down onto the pavement.

Higgins, who was also behind a [set of images](#) of Putin being arrested, put on trial and then imprisoned, says he posted the images with no ill intent. He even stated clearly in his Twitter thread that the images were AI-generated.

Still, the images were enough to get him locked out of the Midjourney server, according to Higgins. The San Francisco-based independent research lab didn't respond to emails seeking comment.

"The Trump arrest image was really just casually showing both how good and bad Midjourney was at rendering real scenes," Higgins wrote in an email. "The images started to form a sort of narrative as I plugged in prompts to Midjourney, so I strung them along into a narrative, and decided to finish off the story."

He pointed out the images are far from perfect: in some, Trump is seen, oddly, wearing a police utility belt. In others, faces and hands are clearly distorted.

But it's not enough that users like Higgins clearly state in their posts that the images are AI-generated and solely for entertainment, says Shirin Anlen, media technologist at Witness, a New York-based human rights organization that focuses on visual evidence.

Too often, the visuals are quickly reshared by others without that crucial context, she said. Indeed, an Instagram post sharing some of Higgins' images of Trump as if they were genuine garnered more than 79,000 likes.

"You're just seeing an image, and once you see something, you cannot unsee it," Anlen said.

In another recent example, social media users shared a synthetic image supposedly capturing Putin kneeling and kissing the hand of Chinese leader Xi Jinping. The image, which circulated as the Russian president welcomed Xi to the Kremlin this week, quickly became a crude meme.

It's not clear who created the image or what tool they used, but some clues gave the forgery away. The heads and shoes of the two leaders were slightly distorted, for example, and the room's interior didn't match the room where the actual meeting took place.

With synthetic images becoming increasingly difficult to discern from the real thing, the best way to combat visual misinformation is better public awareness and education, experts say.

"It's just becoming so easy and it's so cheap to make these images that we should do whatever we can to make the public aware of how good this technology has gotten," West said.

Higgins suggests social media companies could focus on developing technology to detect AI-generated images and integrate that into their platforms.

Twitter has a policy banning "synthetic, manipulated, or out-of-context media" with the potential to deceive or harm. Annotations from Community Notes, Twitter's crowd-sourced fact checking project, were

attached to some tweets to include the context that the Trump images were AI-generated.

When reached for comment Thursday, the company emailed back only an automated response.

Meta, the parent company of Facebook and Instagram, declined to comment. Some of the fabricated Trump images were labeled as either "false" or "missing context" through its third-party fact-checking program, of which the AP is a participant.

Arthur Holland Michel, a fellow at the Carnegie Council for Ethics in International Affairs in New York who is focused on emerging technologies, said he worries the world isn't ready for the impending deluge.

He wonders how deepfakes involving ordinary people—harmful fake pictures of an ex-partner or a colleague, for example—will be regulated.

"From a policy perspective, I'm not sure we're prepared to deal with this scale of disinformation at every level of society," Michel wrote in an email. "My sense is that it's going to take an as-yet-unimagined technical breakthrough to definitively put a stop to this."