'Deepfakes' ranked as most serious AI crime threat

August 4 2020

Credit: Pixabay/CC0 Public Domain

Fake audio or video content has been ranked by experts as the most worrying use of artificial intelligence in terms of its potential applications for crime or terrorism, according to a new UCL report.

The study, published in *Crime Science* and funded by the Dawes Centre for Future Crime at UCL (and available as a policy briefing), identified 20 ways AI could be used to facilitate crime over the next 15 years. These were ranked in order of concern—based on the harm they could cause, the potential for criminal profit or gain, how easy they would be to carry out and how difficult they would be to stop.

Authors said fake content would be difficult to detect and stop, and that it could have a variety of aims—from discrediting a public figure to extracting funds by impersonating a couple's son or daughter in a video call. Such content, they said, may lead to a widespread distrust of audio and visual evidence, which itself would be a societal harm.

Aside from fake content, five other AI-enabled crimes were judged to be of high concern. These were using driverless vehicles as weapons, helping to craft more tailored phishing messages (spear phishing), disrupting AI-controlled systems, harvesting online information for the purposes of large-scale blackmail, and AI-authored fake news.

Senior author Professor Lewis Griffin (UCL Computer Science) said: "As the capabilities of AI-based technologies expand, so too has their potential for criminal exploitation. To adequately prepare for possible AI threats, we need to identify what these threats might be, and how they may impact our lives."

Researchers compiled the 20 AI-enabled crimes from academic papers, news and current affairs reports, and fiction and popular culture. They then gathered 31 people with an expertise in AI for two days of discussions to rank the severity of the potential crimes. The participants were drawn from academia, the private sector, the police, the government and state security agencies.

Crimes that were of medium concern included the sale of items and services fraudulently labelled as "AI", such as security screening and targeted advertising. These would be easy to achieve, with potentially large profits.

Crimes of low concern included burglar bots—small robots used to gain entry into properties through access points such as letterboxes or cat flaps—which were judged to be easy to defeat, for instance through

letterbox cages, and AI-assisted stalking, which, although extremely damaging to individuals, could not operate at scale.

First author Dr. Matthew Caldwell (UCL Computer Science) said: "People now conduct large parts of their lives online and their online activity can make and break reputations. Such an online environment, where data is property and information power, is ideally suited for exploitation by AI-based criminal activity.

"Unlike many traditional crimes, crimes in the digital realm can be easily shared, repeated, and even sold, allowing criminal techniques to be marketed and for crime to be provided as a service. This means criminals may be able to outsource the more challenging aspects of their AI-based crime."

Professor Shane Johnson, Director of the Dawes Centre for Future Crimes at UCL, which funded the study, said: "We live in an ever changing world which creates new opportunities—good and bad. As such, it is imperative that we anticipate future crime threats so that policy makers and other stakeholders with the competency to act can do so before new 'crime harvests' occur. This report is the first in a series that will identify the future crime threats associated with new and emerging technologies and what we might do about them."

More information: *Crime Science*, DOI: 10.1186/s40163-020-00123-8

Provided by University College London