Top experts warn against 'malicious use' of AI

February 21 2018, by Marlowe Hood

Last year, robotics and AI entrepreneurs—including SpaceX CEO Elon Musk and British astrophysicist Stephen Hawking—petitioned the United Nations to ban autonomous killer robots, warning that the digital-age weapons could be used against civilians

Artificial intelligence could be deployed by dictators, criminals and

terrorists to manipulate elections and use drones in terrorist attacks, more than two dozen experts said Wednesday as they sounded the alarm over misuse of the technology.

In a 100-page analysis, they outlined a rapid growth in cybercrime and the use of "bots" to interfere with news gathering and penetrate social media among a host of plausible scenarios in the next five to 10 years.

"Our report focuses on ways in which people could do deliberate harm with AI," said Sean O hEigeartaigh, Executive Director of the Cambridge Centre for the Study of Existential Risk.

"AI may pose new threats, or change the nature of existing threats, across cyber-, physical, and political security," he told AFP.

The common practice, for example, of "phishing"—sending emails seeded with malware or designed to finagle valuable personal data—could become far more dangerous, the report detailed.

Currently, attempts at phishing are either generic but transparent—such as scammers asking for bank details to deposit an unexpected windfall—or personalised but labour intensive—gleaning personal data to gain someone's confidence, known as "spear phishing".

"Using AI, it might become possible to do spear phishing at scale by automating a lot of the process" and making it harder to spot, O hEigeartaigh noted.

In the political sphere, unscrupulous or autocratic leaders can already use advanced technology to sift through mountains of data collected from omnipresent surveillance networks to spy on their own people.

"Dictators could more quickly identify people who might be planning to

subvert a regime, locate them, and put them in prison before they act," the report said.

Likewise, targeted propaganda along with cheap, highly believable fake videos have become powerful tools for manipulating public opinion "on previously unimaginable scales".

An indictment handed down by US special prosecutor Robert Mueller last week detailed a vast operation to sow social division in the United States and influence the 2016 presidential election in which so-called "troll farms" manipulated thousands of social network bots, especially on Facebook and Twitter.

Another danger zone on the horizon is the proliferation of drones and robots that could be repurposed to crash autonomous vehicles, deliver missiles, or threaten critical infrastructure to gain ransom.

Autonomous weapons

"Personally, I am particularly worried about autonomous drones being used for terror and automated cyberattacks by both criminals and state groups," said co-author Miles Brundage, a researcher at Oxford University's Future of Humanity Institute.

The report details a plausible scenario in which an office-cleaning SweepBot fitted with a bomb infiltrates the German finance ministry by blending in with other machines of the same make.

The intruding robot behaves normally—sweeping, cleaning, clearing litter—until its hidden facial recognition software spots the minister and closes in.

"A hidden explosive device was triggered by proximity, killing the

minister and wounding nearby staff," according to the sci-fi storyline.

"This report has imagined what the world could look like in the next five to 10 years," O hEigeartaigh said.

"We live in a world fraught with day-to-day hazards from the misuse of AI, and we need to take ownership of the problems."

The authors called on policy makers and companies to make robot-operating software unhackable, to impose security restrictions on some research, and to consider expanding laws and regulations governing AI development.

Giant high-tech companies—leaders in AI—"have lots of incentives to make sure that AI is safe and beneficial," the report said.

Another area of concern is the expanded use of automated lethal weapons.

Last year, more than 100 robotics and AI entrepreneurs—including Tesla and SpaceX CEO Elon Musk, and British astrophysicist Stephen Hawking—petitioned the United Nations to ban autonomous killer robots, warning that the digital-age weapons could be used by terrorists against civilians.

"Lethal autonomous weapons threaten to become the third revolution in warfare," after the invention of machine guns and the atomic bomb, they warned in a joint statement, also signed by Google DeepMind co-founder Mustafa Suleyman.

"We do not have long to act. Once this Pandora's box is opened, it will be hard to close."

Contributors to the new report—entitled "The Malicious Use of AI: Forecasting, Prevention, and Mitigation"—also include experts from the Electronic Frontier Foundation, the Center for a New American Security, and OpenAI, a leading non-profit research company.

"Whether AI is, all things considered, helpful or harmful in the long run is largely a product of what humans choose to do, not the technology itself," said Brundage.