AASG
Accountancy AML Supervisors' Group

The use of artificial intelligence to bypass 'customer due diligence' checks

Extract from NCA Amber ALERT of October 2024 (Reference 0752-NECC)

Overview

This JMLIT+ Amber Alert is issued by JMLIT+ Public Private Cryptoassets Forum (PPCF).

The purpose of the alert is to help AML-regulated firms performing 'customer-due-diligence' (CDD) checks understand the threat posed by artificial intelligence (AI) to bypass CDD checks. When a customer, or client, uses AI to complete CDD checks, it is likely that the client is connected to criminality including money laundering or terrorist financing. Furthermore, inaccurate customer information means law enforcement has fewer lines of enquiry, undermining attempts to tackle economic crime.

PPCF interviewed UK financial institutions, cryptocurrency exchanges, third party verification companies and other interested parties for this Alert and it focuses on their experiences. It is important to remember that the presence of behaviour set out below is not conclusive of illicit activity, and should be considered carefully in the context of the customer's circumstances.

Information report

The first reported use of AI to attempt to bypass CDD checks was seen in 2022, with initial uses being of poor quality and sophistication, making them easily detectable by humans and automated systems alike. AI has seen significant improvement since then, providing outputs which pose increasing challenges to the AML-regulated sector performing CDD checks.

Interviewees agreed that attempts at using AI to bypass CDD checks are increasing, as are their sophistication. Despite this, most financial institutions and cryptocurrency exchanges interviewed for this Alert reported under five identified AI attempts per month on average. One interviewee, performing CDD checks for mainly international partners, reported hundreds of AI enabled attempts per week. The reasons for this disparity are unclear.

There are alternatives to using AI to bypass CDD checks, which can be lower in cost and sophistication, and easier to use. Readymade accounts, available in bulk, are available online, as are fake documents which, combined with camera spoofing tools, are alleged to be able to bypass CDD checks. Such alternatives likely explain why large-scale AI attempts to bypass CDD checks are not being reported by all interviewees. Longer term, as CDD checks incorporate more advanced AI systems, such low- sophistication methods will highly likely become easier to detect. It is a realistic possibility that increased AI use in CDD checks will lead to an increase in more sophisticated AI enabled attempts.

Interviewees stated there are multiple examples of AI being used to create fake ID documents, which have then been used to attempt to bypass CDD checks. Such attempts are conducted online as part of the account opening process.

Case Study 1

The website OnlyFake, which generates fake ID documents, including driving licenses and passports, are allegedly able to bypass the CDD checks of some leading UK based cryptocurrency exchanges and financial institutions. For USD 15 each, OnlyFake produces fake driving licenses and passports for 26 nations, including the UK. Up to 100 fake ID documents can be generated at once using Excel spreadsheet data and users have the option to use their own photo or select one from a library. The fake ID documents appear arranged on various domestic surfaces, such as kitchen counters, bedsheets or desks, mimicking the typical presentation for online verifications. At least one cryptocurrency exchange confirmed OnlyFake generated ID documents had been used to successfully create accounts with them, which were then used for criminality. It is unclear how many other institutions are being impacted by OnlyFake generated ID documents.

Interviewees stated there are also numerous examples of AI being used to create video, either pre- recorded or conducted live, which have been used to attempt to bypass CDD checks. Such AI allows the criminal to manipulate the video, rotating the person's head, mimicking how a person would move a phone to capture the sides of their head as part of the CDD process. Some videos, using such rotating movements, can bypass the automated parts of CDD checks,

despite being obvious to human verifiers that they are fake. This highlights the important role human verifiers can play in CDD checks, with some interviewees commenting that hybrid CDD checks, using automated systems to flag suspicious customer applications to a human reviewer, are the best approach to tackling AI enabled bypass attempts. One interviewee added that AI software finds it difficult to fool state-of-the-art multi-frame detection systems.

More advanced AI is capable of generating a realistic, large, zoomed out image showing the person's neck, shoulders and upper body using only a photo ID picture as the source. These generated images are of sufficient quality to be used to create a video file, showing the person moving and even speaking. Combining these techniques with AI software creating fake ID documents, raises the prospect of completely fabricated AI-created identities being used to bypass CDD checks in the future.

Low-cost solutions, in particular 'FaceSwapping', still use AI and can potentially bypass stages of the CDD check process. Such solutions are often cheaper, quicker and more convenient than using more advanced AI, yet offer criminals with lower technical ability options to bypass CDD checks.

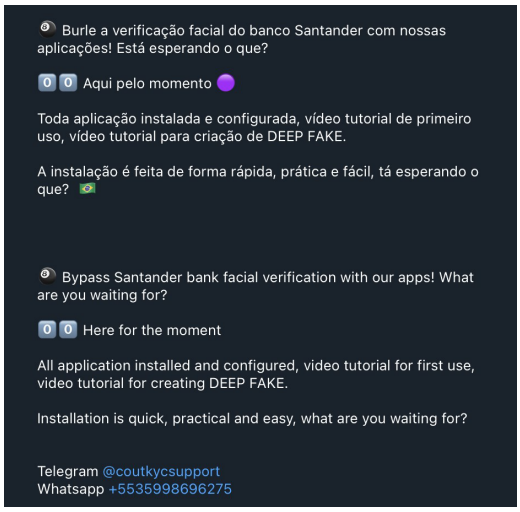
One interviewee noted the use of device emulators as a potential indicator of criminality. Device emulators have been used to trick mobile applications into thinking the criminal is applying using a mobile phone, when they are actually using a desktop computer. The criminal is able to use pre- recorded images and video stored on the desktop, as inputs through the mobile application. The images and videos are moved on the desktop, which then move accordingly on the mobile application. The method can be repeated to generate accounts in bulk.

Some criminals provide services to customers, helping their customers bypass CDD checks using AI. One interviewee has identified examples of criminals offering a service in which their customers pay the criminal a fee and send the criminal an image, in return their customer receives a file which allegedly bypasses CDD checks. Other criminals show software allegedly capable of bypassing UK- based financial institutions, gambling companies, dating sites and cryptocurrency exchanges' CDD checks in YouTube videos. These videos are used as advertising for the criminals' software, guiding those interested to contact the criminal on Telegram. Software is used to move the face of an individual (which may have been deepfaked) or display an ID document on the criminal's computer, deceiving the institution's checks that the individual's face or ID document is that of the individual applying. The videos show the applications being submitted, but it is unclear whether such attempts are successful and some may not work at all or as advertised, in effect being examples of criminals targeting other would-be criminals.

Some of the YouTube videos and Telegram channels use foreign languages, suggesting some of these actors may be operating from overseas or are foreign nationals. One interviewee assessed it is likely that such actors are paid in cryptocurrency, due to the potentially international nature of their customers and ease with which cryptocurrencies can be bought and transferred compared with fiat currency.

Some advertised tools are expensive, including one for 5,000USD. Such tools are unlikely to be used to generate a low amount of images or videos for attempted account creation, as there are much cheaper alternatives. Instead it is more likely they will be used to generate images or videos in bulk, which will be sold to others to recoup the cost of the tool. It is unclear what impact such services are having on the scale of AI use attempts, nor the effectiveness of the outputs.

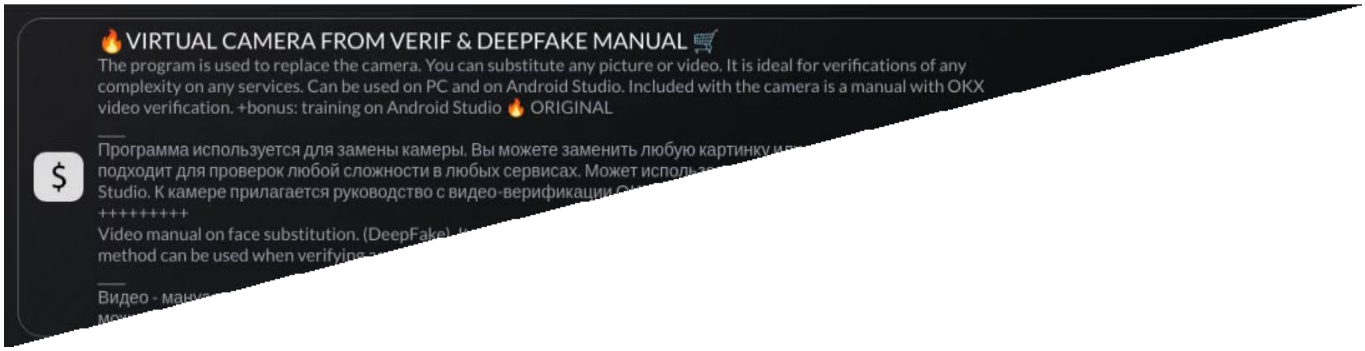
Figure 1 shows a screenshot from a Telegram account offering a tutorial and software allegedly able to bypass CDD checks using AI. The account communicates in both English and Portuguese.



Some dark web platforms advertise CDD bypass methods, which includes guides for customers. Some of these guides suggest the use of different AI software at different stages of the CDD process, indicating that multiple pieces of AI software may be being used to successfully bypass CDD checks.
These suggested AI software packages are widely available, free to use and marketed for other purposes, such as video

gaming. It is unclear whether they can be used in conjunction with one another to successfully bypass CDD checks, but user reviews of the guides suggest they can.

Figure 2 shows a screenshot of a dark web platform advertisement. The advertisement is written in English and Russian.



Multiple interviewees stated that web-based applications were higher risk than mobile ones. Web-based applications are more easily tampered with, especially those allowing applicants to upload pictures of identity documents and pre-recorded videos, providing greater scope for tampering and fewer guarantees that the applicant is who they say they are. Web-based applications require less technical skill and knowledge to bypass CDD checks than are needed when applying by mobile, so are considered higher risk.

Interviewees had different recording practices for AI-enabled attempts to bypass CDD checks. Some were able to analyse the number of attempts that they, or their customers, were experiencing, whilst others were able to provide examples but did not record attempts if they were blocked by their automated systems. Such recording practices made assessing the scale of AI-enabled attempts difficult. Some interviewees did not cluster AI-enabled attempts for further investigation.

Indicators

AI software is constantly being improved and is evolving, with some previously identified indicators of AI use being overcome. One interviewee commented that previously reliable indicators, such as the apparent steadiness of the camera or the individual's apparently flawless skin complexion, are now being amended in more advanced AI software. Such software makes the videos, and individuals in the videos, appear more humanlike, making it more difficult to detect that they are fake.

Interviewees listed a number of indicators of potential AI use to attempt to bypass CDD checks. Such indicators are insufficient to determine criminal or suspicious behaviour and should be combined with other threat indicators when reviewing a CDD application. The determination of suspicion always rests with the reviewer. This list is not exhaustive and is almost certain to change as the technologies advance and evolve:

Data discrepancies or indicators, such as:

- the use of a VPN during the onboarding process;
- the IP address not matching the customer's address or location;
- the metadata not correlating with the image or video being presented; or
- the use of device emulators.

Visual discrepancies or indicators, such as:

- the use of identical backgrounds in multiple images or videos across different users' applications, indicating that the applications may have been submitted in bulk and are malicious;
- human detectors noticing the mouth moving strangely or different to most human behaviour;
- apparent natural lighting conditions being inconsistent with the time of day the image or video was supposedly taken at;
- pixels being out of place or not matching the pixels around them;
- the serial numbers not matching regional numeric conventions of the country it purports to be from (such as a driving license number being in the wrong format); or
- the name not matching regional naming conventions of the country it purports to be from.

Audio discrepancies or indicators, such as:

- the lips not syncing with the sound; or
- the audio having computerised sounding voices.

Other indicators, such as:

- the account being applied for through a web-based browser rather than through a mobile application.

Suspicious Activity Reporting [SARs]

If you know or suspect that there has been money laundering or terrorist financing activity (including as a result of information provided to you by the NCA) and your business falls within the regulated sector, then you are reminded of the obligations to make reports to the NCA under Part 7 Proceeds of Crime Act 2002 and the Terrorism Act 2000. If you decide to make a report in this way you should adopt the usual mechanism for doing so, and it will help our analysis if you would include the reference **0752-NECC** within the text. This reference is specific to the Alerts process; where appropriate, we would ask that this is used in addition to the ongoing use of the Glossary of Terms. Guidance on making suspicious activity reports is available at www.nationalcrimeagency.gov.uk.

The NCA would also welcome any information identified as a result of this alert which does not constitute a SAR. Please email all such information to NECC.PPP@nca.gov.uk. Any information received in this way will be treated in confidence and will be handled in line with the data protection principles.

Data Protection Considerations

Please consider your obligations under the relevant data protection regulations and where necessary remove any related personal data from your systems securely and within a satisfactory timeframe.

Disclaimer

The Accountancy AML Supervisors' Group (AASG) accepts no responsibility for any loss, damage or expense arising in connection with the use of information in this alert. Any use will be taken to signify agreement to these conditions.

Protecting this document

This document uses the United Kingdom's Government Security Classification System (GSCS) and has been graded as OFFICIAL. There are no specific requirements for storage and it can be considered safe for wide distribution within your organisation and for use in staff training or awareness programmes. However, unless otherwise specified, this information is not intended for general public dissemination and should not be included on public facing websites, external mailing lists, social media or other outlets routinely used by you to deliver information to the public without the prior and specific permission of the NCA Alerts team. We therefore request that you risk manage any onward dissemination in a considered way. This document should be disposed of by cross-cut shredder, pulping or incineration.