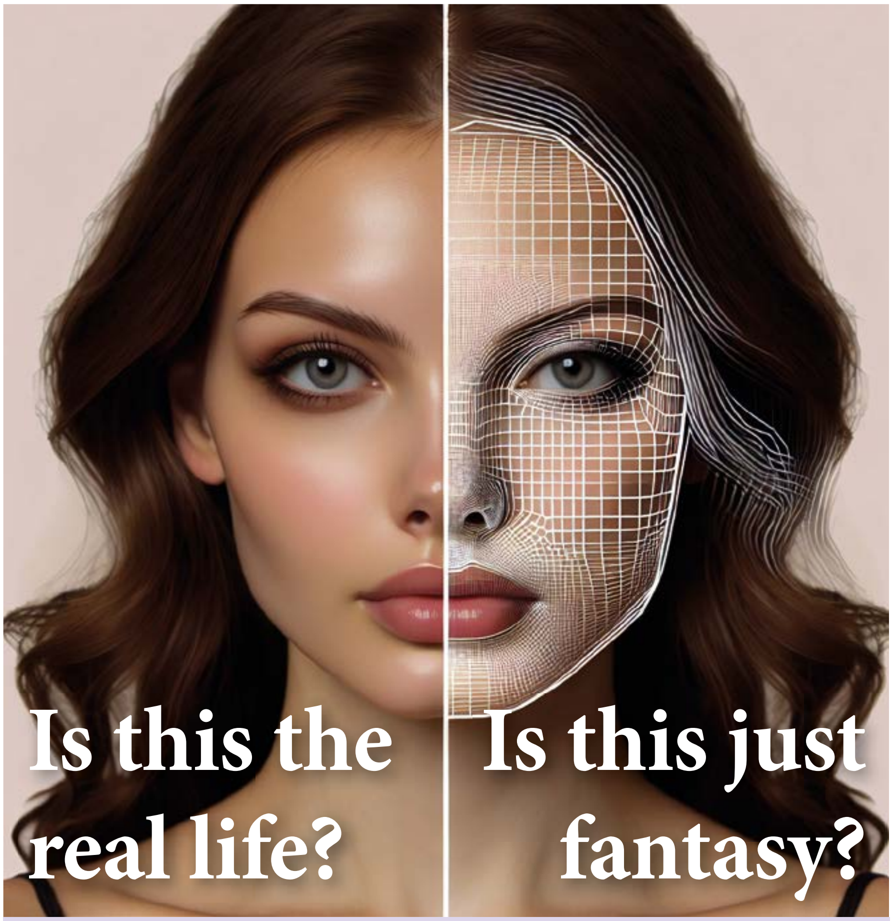

The challenge for the pensions industry to spot AI-generated 'deepfakes', and the role they may play in identity fraud

Imagine the horror of accidentally authorising £25 million of your company's money to be sent to fraudsters.

In January 2024, this happened to one unfortunate Hong Kong employee of British engineering company Arup, who was tricked into sending HK$200 million (approx £20 million) to criminals that had used AI to create a 'deepfake' group video call pretending to be the company's senior officers 'authorising' the transfer.

While this may not have occurred in the pensions sector "it's fair to say that AI-powered ID fraud represents a genuine and growing risk across all corners of financial services," Trafalgar House client director, Daniel Taylor, warns.

"While we haven't yet seen a headline pensions case involving deepfakes or voice cloning, we'd be naïve to think we're immune."

Using AI for ID fraud

AI can be used in a number of ways to create 'deepfakes', also known as 'synthetic profiles' or 'generative AI', for the purpose of impersonation fraud.

For instance, to pass ID authentication checks, AI can be used to create fake images or clone the voice of an individual – it is estimated that only 20-30 seconds of audio is needed to make a convincing clone. These can both be used to make a deepfake video in as little as 45 minutes using freely-available software.

AI can also generate convincing fake identification documents, such as passports and driver's licences, using personal information and images found online.

Biometric spoofing is also a

Deepfakes

Deepfakes are AI-generated forgeries of a person's face and/or voice. They can be replicated as images, video or audio and also used 'live' to deceive biometric authentication systems. In the financial world, they can be used by fraudsters to bypass security systems to steal assets such as data or money.

AI can also be used to fake official documents, such as passports and driving licences, to commit identity fraud. Entire 'synthetic' identities can also be created for the purpose of committing fraud.

Deepfakes statistics

- Only 20-30 seconds of audio is estimated to be needed to create a convincing voice clone.
- Deepfake videos can be created in as little as 45 minutes using freely available AI software.
- Deepfakes online increased from 500,000 in 2023 to eight million in 2025.
- In 2023, the average cost of deepfake financial fraud was $500,000 (approx £372,250) per incident.
- 42.4% of all fraud attempts are now AI-driven.
- Deepfake fraud rates increased by 2,137% between 2021-24.
- 40% of deepfake attacks target the financial sector.
- Only 24.5% of deepfakes are successfully detected by humans.

possibility, where AI-generated fingerprints and facial recognition can bypass authentication systems.

This can also all be used to create an entire synthetic identity creation – complete fake personas with supporting documentation.

"AI has moved identity fraud into a whole new league. We're no longer talking about poorly worded phishing emails – we're now in the age of deepfakes, cloned voices, and synthetic identities," Taylor says.

"With freely available AI tools, bad actors can convincingly mimic someone's voice or face and they've even fabricated 'live' video calls that fool biometric systems. In short: Fraudsters no longer need to steal your ID – they can manufacture one."

In September, cybersecurity firm Deepstrike reported that the number of deepfake files surged from 500,000 in 2023 to eight million in 2025, and that fraud attempts using deepfakes spiked 3,000 per cent in 2023, costing an average of $500,000 (approx £372,250) per incident.

According to digital identity solution provider, Signicat's 2024 report on AI-driven identity fraud, 42.4 per cent of all fraud attempts are now AI driven, with deepfake fraud rates having surged by 2,137 per cent over three years. It also revealed that 40 per cent of all deepfake attacks were directed at the financial sector, second only to cryptocurrency.

Signicat also found that 'spoofing' – aka the use of fake biometric credentials to defeat identity checks – rose from 7.58 per cent in 2021 to a projected 12.83 per cent in 2024. Meanwhile, 'injection attacks', which involve the insertion of synthetic or deepfake videos into authentication streams, bypassing conventional fraud detection layers, saw an increase from 1.51 per cent to 6.27 per

Notable examples of deepfake scams

- ***Arup (January 2024)***

An employee of British engineering firm Arup in Hong Kong was tricked into transferring HK$200 million (approximately £20 million) to fraudsters. The criminals used AI to create a deepfake video call impersonating senior executives of the company, seemingly authorising the payment.

- ***Elon Musk (August 2024)***

An 82-year-old retiree in the United States was defrauded of $690,000 (approx £513,780) after being targeted by a deepfake video featuring an AI-generated version of Elon Musk. The video promoted a fake 'radical investment opportunity', which the retiree believed to be legitimate.

- ***WPP CEO (May 2024)***

Fraudsters impersonated Mark Read, the CEO of WPP, using an AI-generated voice and a fake video call. They set up a WhatsApp account using his public image, staged a Microsoft Teams meeting that appeared to be with senior executives, and attempted to persuade an agency leader to form a new business and divulge funds and personal details.

- ***Joe Biden (2024)***

AI was used to generate a synthetic voice of then-President Joe Biden in a robocall targeting voters in the New Hampshire Democratic primary. The call falsely urged Democrats not to vote, claiming their votes would not be counted.

- ***Volodymyr Zelenskyy (2022)***

A deepfake video circulated that showed Ukrainian President Volodymyr Zelenskyy apparently ordering the armed forces to surrender. The video was widely shared on social media and even appeared briefly on a hacked Ukrainian news site, before being debunked.

- ***Brad Pitt (2024)***

A 53-year-old French woman was scammed out of approximately €830,000 (around £724,630) after being tricked into believing she was in a romantic relationship with Hollywood actor Brad Pitt. Over 18 months, scammers used deepfake images and videos, including fake hospital footage, to build trust and manipulate her emotions. Claiming Pitt was in financial trouble due to frozen bank accounts and medical issues, they convinced the woman to send money multiple times under false pretences, including customs fees and treatment costs.

cent over the same period.

Aviva workplace pension policy manager, Dale Critchley, notes the use of AI ID fraud within the insurance sphere.

"Scammers are often quick to take advantage of any technology that helps them commit fraud. Given the similarities between insurance and pensions in terms of data and money involved, the risk is likely to be transferable," he warns.

The banking sector has also begun to see AI ID fraud, Isio head of pension administration operations, Neil Brady, says, "and the pensions industry tends to follow the path of the banking sector".

Pensions industry vulnerability

The pensions industry is arguably more vulnerable than the banking sector to this type of fraud.

"Pensions are a perfect storm for fraudsters: Large pots of money, inconsistent ID verification processes, and a predominantly disengaged member base. Unlike banks, which often have real-time, high-frequency interaction with customers, pensions operate in low-touch, high-value environments – ideal for impersonation and synthetic ID fraud," Taylor explains.

"One convincing deepfake, one cloned voice, one breached biometric could be all it takes for a fraudster to redirect funds or access member data. Add to that the fact that many schemes are still catching up digitally, and the defences aren't always as sophisticated as the scams."

However, Aon partner, Paul McGlone, notes: "The controls around payment of benefits tend to be more rigorous than, for example, making a bank payment, and for most members the amount that sits in their pension account probably isn't sufficient to warrant the effort that goes into creating a convincing deepfake video or voice clone. Of course, that could change."

Lumera head of technical research, Rebecca Morgan, believes so, as the pensions sector is not transferring money around often, but when it does "it is really large amounts, which does make it a target, and the infrequent nature of fund movements makes it harder to spot anomalies".

The opportunities for this type of fraud to occur within pensions are threefold.

"AI-powered ID fraud represents a genuine and growing risk across all corners of financial services"

PASA Identity Management Working Group chair, and LexisNexis Risk Solutions head of identity strategy, UK&I, Lorraine Salmond, firstly highlights how there is the particularly lucrative opportunity to dupe a trustee or employee within the pensions supply chain by using AI to authorise a colleague's permission to transfer large volumes of money or data to the criminals.

Schemes need to think about protecting trustees against this risk, McGlone states.

"Many trustee boards routinely have their voices and images recorded from online trustee meetings, without knowing where those records are stored and for how long. With so much recorded content it wouldn't be difficult for someone to create deepfake a trustee, so processes need to be in place that would prevent unauthorised actions being approved. Third-party providers also have a role to play in challenging unusual requests from trustees where they suspect potential ID theft," he explains.

Salmond also notes how AI could be used to impersonate an individual pension scheme member, to convince those managing the pension scheme to transfer retirement savings to the criminal's account. That can happen in real time through face and voice synthesis to pass selfie and voice checks within pension portals or call centres, she explains.

And finally, the pension scheme member themselves can be fooled by a deepfake into handing over their pension pots to fraudsters.

Last month, The Pensions Regulator (TPR) sent out an industry alert regarding impersonation fraud generally, noting that members between the ages of 50 and 69 are at the greatest risk of these impersonation techniques, with 55 per cent of the reported victims in this age range.

In August 2024, it was reported that an 82-year-old retiree in the USA fell victim to a deepfake of Elon Musk 'recommending' a 'radical investment opportunity'. Over the course of several weeks the retiree drained his retirement fund to invest $690,000 (approx £513,780) in the scam, with the money lost to the fraudsters.

As well as the risks of financial loss or data breaches, schemes falling foul of such scams would face operational inconvenience, reputational damage and potential regulatory or legal issues.

As DLA Piper partner, Matthew Swynnerton, notes, "in the absence of general statutory regulation, the UK continues to rely on the General Data Protection Regulation (GDPR) framework to regulate AI usage. This places a significant compliance burden on scheme administrators, particularly where digital verification tools are integrated into member portals".

"Due to the rapid nature of AI developments, schemes must evolve to ensure their application of the GDPR accurately addresses potential new security risks," he adds.

Minimising the risk

The increasing sophistication of these deepfake scams means the 'uncanny valley' effect – the sense of unease or revulsion a person experiences when subconsciously noting that a computergenerated human figure (or humanoid robot) is not really human – can no longer be relied upon. According to Deepstrike, there is just a 24.5 per cent successful human detection rate of deepfakes.

But just as AI can generate ID fraud, so too can it be used to tackle the problem.

"We really need to make sure that pension companies are using AI to combat AI fraud, because it is not good enough to rely on humans spotting something that is so clever. You can't tell if it's a voice or photo fake as deepfakes fool the human eye all the time. But AI can be programmed to spot some of the things that don't quite hang together, such as the pixelation not being 100 per cent correct," Pension Scams Industry Group chair, Margaret Snowdon, explains.

"AI software can be low cost and can analyse information such as emails, letters or telephone calls coming in. AI can look through these and analyse the language and content. It can look at whether a document is original or whether it's been doctored or monitoring odd responses from customers," she adds.

Salmond gives the example of literal masks being used to con ID video checks, but that AI software can determine if it is fake by monitoring if blood is flowing in the *[masked]* images.

"AI tools can also be used to determine if the member's device being used is in the expected country or being held in a standard way. It also checks whether the email address or bank account is newly set up, which can be a sign of fraud," she adds.

According to Lumera AI product owner and data scientist, Alessandro Alviani, relying on what is unique to individuals and therefore useful for security– such as fingerprints, the iris or face – can now be cloned, and so multi-factor authentication provides extra layers of security to minimise fraud.

"We've got to focus on multiple angles, preventing fraud by using tools such as multi-factor authentication, continuous activity monitoring and protecting data from cyber-attacks, as stolen data becomes the fuel for fraud attempts downstream," Aptia chief growth officer, Andy Seed, says.

LexisNexis Risk Solutions 2021 *Digital Dilemma Pension Report* found that people would be receptive to that, as 77 per cent of respondents supported the idea that pension apps should have the same log-in security levels as online banking apps.

Despite the use of multiple tools to minimise the risk of AI ID fraud, "it's still the human that is getting fooled. These scam tricks that have always existed – it's the same confidence trick as the scam artist knocking on the door, but it's just the technology that has changed", Bravura proposition lead EMEA, Jonathan Hawkins, says.

Therefore, training staff on this issue is essential.

"There are some practical tips pension schemes could consider, to help protect themselves and their members from AI identity fraud. This might include training people to spot deepfakes and fake documents and verifying sources. However, arguably the most important aspect is to stay informed – stay vigilant and keep up to date with

What to do if subject to an AI ID fraud

The pension scheme should:

1. Alert their fraud prevention office or another nominated individual immediately of the incident who can provide immediate advice.
2. Establish a small and confidential working group with necessary counter fraud skills, expertise and experience to contribute to an investigation, ensuring compliance with related legal requirements.
3. Refer for specialist investigation as soon as there is a suspicion of fraud that is substantiated by at least one reliable piece of information.
4. Leave all response and recovery options open. An investigation may reveal the fraud is much bigger than initially suspected.

The pension scheme should not:

1. Panic or succumb to initial reactions as this risks any investigation being compromised
2. 'Tip off' the suspected person, so to advise the team(s) to use delay tactics on any activity occurring on the case
3. Take any action without seeking external professional advice.

PASA Cybercrime & Fraud Working Group member, Gillian Baker

emerging fraud tactics," Critchley says.

This training may enable the pensions industry, in turn, to educate savers on the risk of AI ID fraud. Seed expects scheme managers and trustees to rely on their advisers, especially administrators, to help protect their members.

"Schemes and trustees can play their part, but members do have a role to play," McGlone says.

"Historically the pensions industry has warned members about pension scams, and increasingly schemes are warning their members about issues such as phishing attempts and how to ensure they don't fall victim to ID fraud. Some schemes are now reminding members how they will engage with them, in a similar way that the banks tell customers 'we will never call and ask for your password' or explaining how to report anything suspicious," he explains.

Growing awareness

But before schemes can educate their members of the risk of deepfakes, awareness within the pensions industry itself arguably needs to increase.

According to Taylor, "awareness is patchy at best".

"Most trustees understand the general cyber risk landscape, but few are actively talking about AI-powered ID fraud. Many still view AI as a productivity tool, not a threat. There's a blind spot here and the clock is ticking. The pace of AI innovation is staggering, but the pensions industry doesn't move as quickly. That creates a dangerous mismatch between threat evolution and defence readiness," he warns.

TPR's industry alert on impersonation scams last month helps to raise awareness of the issue generally, with its revelation that between October 2024 and March 2025 almost a third of Action Fraud reports referenced attempts to bypass pension scheme defences and exploit security vulnerabilities to gain unauthorised access to members' accounts.

Swynnerton also highlights how TPR and the Pension Scams Action Group have recently joined forces "to detect and take down fraudulent websites targeting individuals accessing their pension accounts – all powered by machine learning technology".

> **"AI has moved identity fraud into a whole new league... fraudsters no longer need to steal your ID – they can manufacture one"**

Looking ahead

According to AI education company Sidecar, AI is on a 'sprint', with its computational power doubling approximately every six months.

"Perhaps more worryingly, AI has the intelligence to 'learn' or be taught failed techniques as it goes, which makes *[AI ID fraud]* an increasing risk as techniques improve," Seed says.

According to LexisNexis Risk Solutions' 2023 *Digital Pensions Fraud* whitepaper, "as of today, the digital threat in the pensions space is still relatively low as the industry is behind the curve from a digital perspective". *[It notes that 29 per cent of schemes do not utilise electronic identity verification for UK-based members].*

However, as the power of, and uses for, AI develops, so too will the way it is used by the pensions industry.

"The digital transformation of the UK pensions sector, through the upcoming pensions dashboards and member expectations regarding a digital service akin to other financial sectors, could make the pensions industry more vulnerable to AI ID fraud," Salmond warns.

In particular, "dashboards are going to open up more opportunities for ID fraud, by revealing untraced funds to the general public, and also to fraudsters, so we have got to think about how we protect ourselves", Brady states.

Also, "banks and other financial institutions are increasingly using biometric checks such as videos and voice checks; those may start to be more prevalent in pensions over time," McGlone states.

Swynnerton notes that holding biometric data is subject to heightened legal scrutiny under Article 4(14) of the UK GDPR.

"Biometric systems have previously been viewed as a safer method of identity verification; however, the growing sophistication of AI technology means they are increasingly vulnerable to manipulation," he explains.

"In order to process this special category data, pension schemes employing facial recognition or similar technologies must obtain explicit consent under Article 9, as well as satisfy a lawful basis requirement under Article 6. They must also adhere to strict obligations as data controllers, by implementing technical and organisational measures to mitigate the risk of identity fraud. This includes safeguarding against unauthorised and fraudulent access, for which the emergence of deepfakes and AI-produced documentation presents a new challenge."

So, should the pensions industry be concerned about the challenges AI may bring with regards to identification fraud?

According to Taylor, "concerned, yes; panicked, no".

"AI ID fraud is a real and evolving threat but it's also manageable with the right safeguards, mindset, and oversight," he explains.

"The key is proactivity. Waiting until something goes wrong is no longer an option. If we want members to trust digital pensions, we need to make sure that the industry isn't trusting digital illusions."

Written by Laura Blows