Work From Clone:

How to Prevent Your Organization From Hiring Deepfakes

Introduction

A recent FBI alert warned organizations of a new threat to businesses and governments around the world: cybercriminals are using deepfakes to create fake people and apply for remote job vacancies.

Why? It's not just the pay check. If a criminal can succeed in getting hired to a role that has access to personal or other data, they can steal that data or hold it for ransom for financial gain.

The FBI alert brings a number of things into question:

- What can happen to an organization if it hires a deepfake?
- Can we trust the human eye and video conferencing to detect deepfakes?
- How aware are individuals within organizations of the deepfake threat?
- What can organizations do to ensure they don't hire people that don't exist? And ensure only the right people have access to sensitive data and systems?

The report seeks to answer these questions.

The Problem: FBI Warns of a New Threat in Job Applications

In June 2022, the FBI issued a public service announcement alerting organizations to a new cybersecurity risk: deepfake employees.

According to the alert, there has been a rise in cybercriminals using 'deepfakes and stolen Personally Identifiable Information (PII) to apply for a variety of remote work and work-at-home positions.'

The FBI reported that the attackers used deepfakes to apply for roles in 'information technology and computer programming, database, and software-related job functions.'

The Bureau became privy to these fraudulent ongoings when they received a number of IC3 complaints (Internet Crime Complaint Center) from the organizations carrying out the interviews. Complainants reported that they detected spoofing and deepfake videos when speaking with candidates.

Strange things happened during the calls, such as the mouths and lips not aligning with the words being said on screen.

What Are Deepfakes?

Deepfakes are videos or images created using AI-powered software to show people saying and doing things that they didn't say or do.

Think of the thing that you are least likely to ever say. Now imagine your friends, family, or employer being shown a convincing video of you saying it. It is easy to see the potential for malicious misuse.

Deepfakes are created using artificial neural networks, which means that they can be produced increasingly easily to look authentic and convincing.

Not all deepfakes are malicious or dangerous. Many have been used for social sharing and entertainment. But they have also been employed in hoaxes, revenge porn, and increasingly, fraud and impersonation.

How Are Deepfakes Used for Cybercrime?

Deepfakes are increasingly being used to commit cybercrime, usually for financial gain. There are several methods for criminals to do this. Two examples:

Synthetic Identity Fraud:

This type of criminal activity involves a criminal collating separate elements from real identities to create a persona that doesn't actually exist. Using this synthetic identity, they then apply for credit cards or government support programs, or open accounts for the purposes of money laundering. Deepfakes can be used to make the fake person look more credible during the enrollment process.

Identity Theft:

Using deepfake tools, attackers can convincingly pose as another, real person. By impersonating this individual, they can cause reputational damage, authorize transactions (such as payments), or convince others to take action, such as divulging sensitive information.

How Are Deepfakes Used For Job Applications?

Gina urgently needs to hire a new support team member for her IT department at a local government. The role is remote to support inclusion and give the organization the best chance of finding a capable candidate.

One of the candidates has an impressive resume and has been responsive to emails. In the video interview, Gina notices some sync issues between the audio and video but she puts it down to the connection. She offers the role and the person is hired.

But there wasn't a bad connection. In actuality, the individual Gina hired wasn't a real person, but a synthetic identity cobbled together from separate pieces of personal data. When Gina conducted the interview, she was actually speaking to a deepfake.

Now employed within the local government, the attacker behind the deepfake has access to a lot of PII. They can then sell this data, hold the local government to ransom for it, or use it for further cybercrimes.

.

What Can Happen to Your Organization if You Hire a Deepfake?

The FBI reported that criminals are using deepfakes to try and secure remote employment in technology roles. There are several motivations for this, each damaging to an organization and designed for financial gain.

A Monthly Pay Packet

Deepfakes are scalable. This means it's easy for the criminal to repeat the same attack over and over again. Gaining employment in a number of different organizations could lead to a tidy sum being transferred to the criminal's bank account every month.

However, this is not the biggest danger of hiring a deepfake.

Further Cyberattacks

Hypothetically, once the attacker has succeeded in acquiring a tech role, they then potentially have access to the organization's computer systems, as well as its data.

This provides an opportunity to carry out further cyberattacks, such as injecting malware.

Data Theft

Once successfully employed within the organization, the criminal has access to 'customer PII, financial data, corporate IT databases and/or proprietary information', according to the FBI alert.

There is no end to the damage the criminal can cause with this data. It's especially dangerous for organizations that hold lots of sensitive PII, like governments. The criminal can sell this data to bad actors or use it to hold organizations to ransom.

The Scalability Issue

Moonlighting – the act of being employed in several organizations at once and clocking in the same hours – and employee fraud are nothing new. For years, corrupt staff members have committed criminal acts within the workplace.

But before the rise of remote working and the development of deepfake technology, these criminal acts were limited. Realistically, how many jobs could you hold down if you had to be present for interviews and show up at least occasionally at the office? Likewise, cases of employee fraud usually involved one individual scamming one organization.

These days, many organizations offer entirely remote employment, meaning you no longer need to be even in the same country as your employee. Meanwhile, the tools used to create deepfakes have become more affordable and readily available.

Combine remote working with the relative ease of developing deepfakes and you have a real, scalable issue. In theory, attackers will be able to apply for dozens, even hundreds of remote roles.

Can Humans Spot a Deepfake During a Video Interview?

From the moment we are born, we are equipped with an innate ability to recognize faces. You would think that no matter how convincing deepfakes get, surely we will still be able to tell fact from fiction?

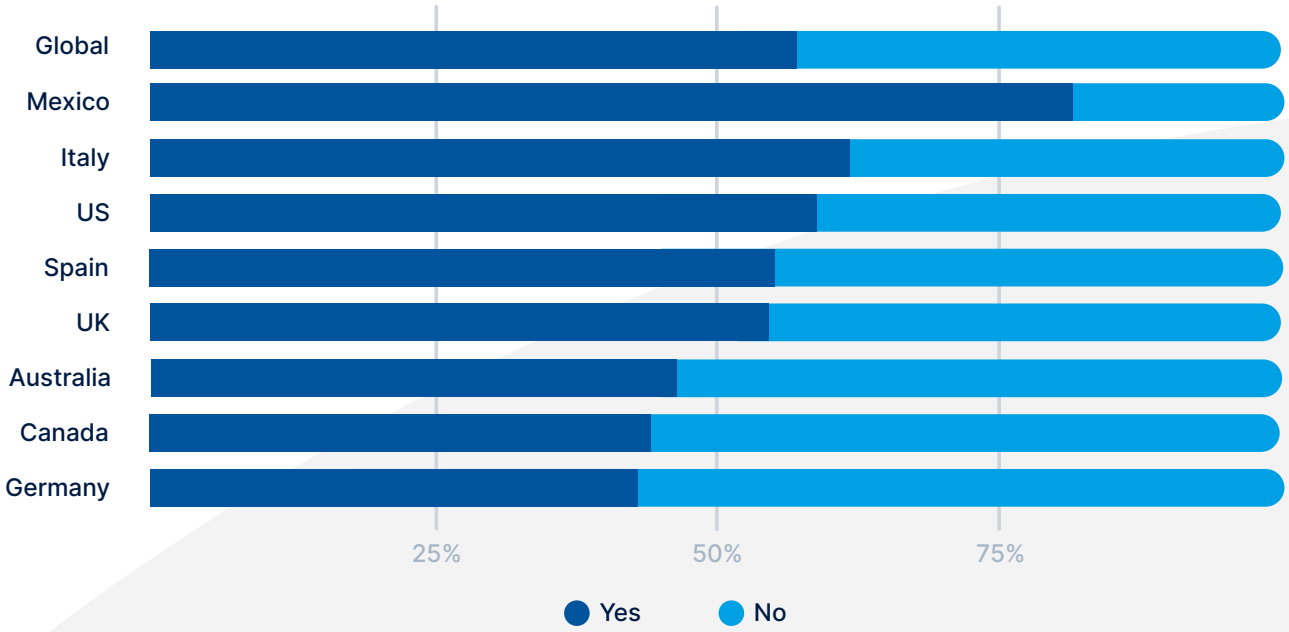
Not the case. As humans, we are not adept at successfully detecting deepfakes, especially as they become more convincing. The human eye can be easily spoofed. Therefore, technology is required to help us spot these types of attacks.

But what's more worrying is that most people erroneously believe that they could spot a deepfake.

iProov recently carried out a survey of 16,000 consumers around the world. **57%** were confident that they could tell the difference between a real video and a deepfake. And that confidence is growing: in 2019, this figure was only **37%**.

Mexicans were the most confident in their own abilities with **82%** of those surveyed saying that they could spot a deepfake.

Do You Think You Would Be Able To Tell the Difference Between a Real Video and a Deepfake?



But our overconfidence is misguided. Studies have shown that we, as humans, are overwhelmingly inept at distinguishing real faces from deepfakes.

Take, for example, the work of Pavel Korshunov and Sebastien Marcel from the Idiap Research Institute, a facility built to examine artificial and cognitive intelligence. In 2020, Korshunov and Marcel conducted a study into human ability in detecting deepfakes and compared it with AI technology.

The study was simple. Subjects were shown progressively more convincing deepfakes, interspersed with real videos, and asked 'is the face of the person in the video real or fake?'

The results showed that the subjects were easily spoofed by the 'well-made' deepfakes. Only **24%** of the subjects detected that these deepfake videos were not real. What's more, only **71%** of subjects correctly identified deepfakes in the easy category!

There's a caveat here. Korshunov and Marcel admitted that the experiment conditions need to be taken into account. The subjects were expecting to view deepfakes. As such, the researchers believe that if a deepfake were to be delivered to an unsuspecting audience, the number of people successfully detecting them to be fake would be 'significantly lower.'

Despite our self-assurance, humans are poorly equipped to detect deepfakes. Even more so when we're not expecting them.

Are People Within Organizations Aware of the Risk of Deepfakes?

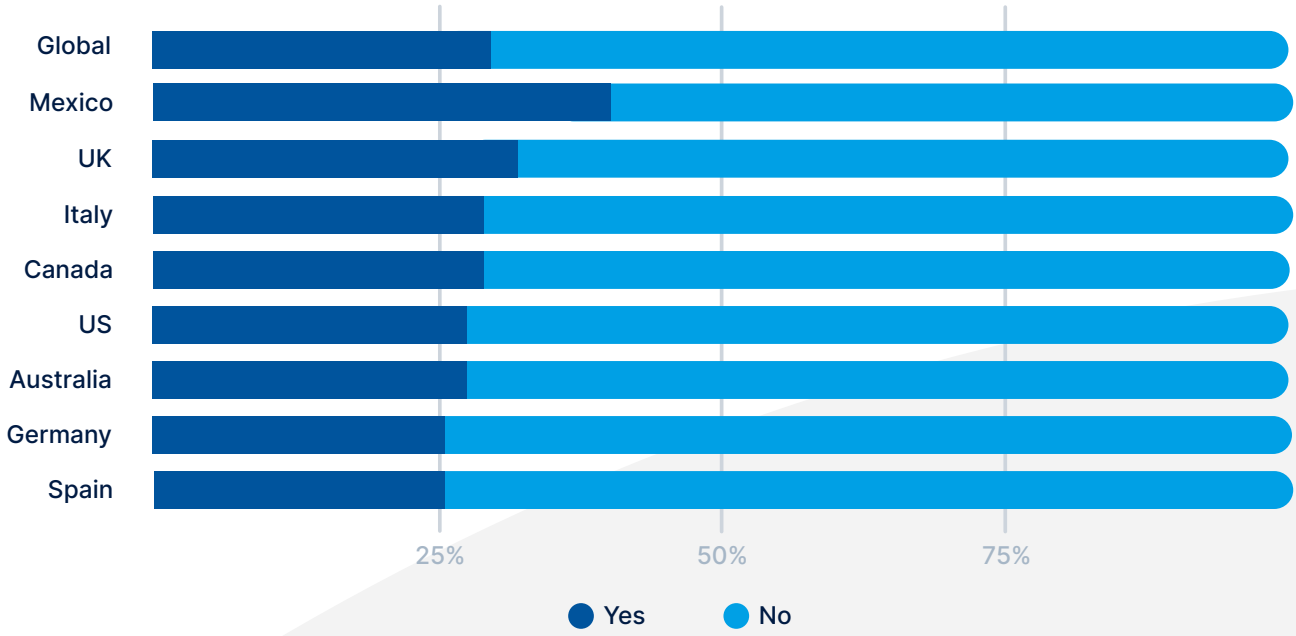
The good news is that the dangers of the deepfake threat are not lost on the general public.

For one, public awareness has evolved in the last few years. In iProov's survey back in 2019, only **13%** of respondents knew what a deepfake was. In 2022, that number had jumped to **29%**.

But that still leaves **71%** of people who don't even know what a deepfake is.

It seems that as the deepfake threat develops, so will public awareness. It's even made its way into the cultural zeitgeist. Popular British soap drama Coronation Street recently included a plot line whereby a major character's face was imposed onto the body of a naked woman using deepfake technology.

Do You Know What a Deepfake Is?



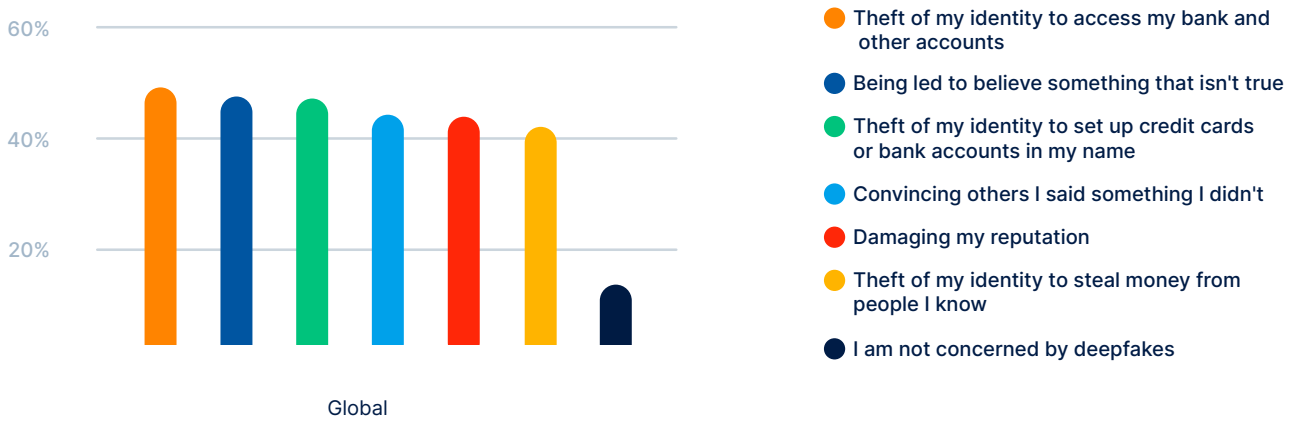
Although not all deepfakes are created for nefarious purposes, the overriding feeling of the global population is one of apprehension.

75% of respondents in iProov's global survey agreed with the statement, 'deepfakes will make it hard for people to trust what they see online'. At the same time, **62%** said they were 'dangerous' and **58%** said they were a 'growing concern.' Only **9%** of those asked said they thought deepfakes were 'harmless'.

People are right to be worried. As we've seen, deepfake technology, when used by the wrong people, poses a real, scalable threat to organizations and individuals.

If organizations are to maintain public trust, they must show a commitment to combatting these new online dangers.

Which of the Following Worries You Most About How Deepfakes Could Be Used Against You?



Video Calls Are No Longer an Effective Security Measure

As we've seen, the human eye has a poor track record of identifying deepfakes. This brings into question the efficacy of video calls as a measure to verify a person's identity.

The FBI warned of the rising threat of criminals using deepfakes to apply for jobs due to complaints from organizations saying they suspected a number of their interviewees of being digitally synthesized. What's not clear from the alert, however, is how many deepfake attempts succeeded in spoofing the interviews.

The technology used to develop deepfakes has become so sophisticated that it can now also bypass the camera. This is what is known as a 'digitally injected attack.'

And it's not hard to produce. With a simple plug-in, attackers can create what's called a 'real-time deepfake'. This involves superimposing images to distort a video. This video can then be streamed into video conferencing communication channels.

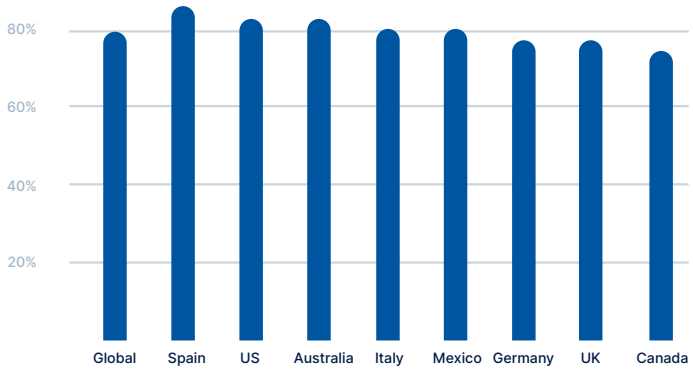
A lack of skill as humans, combined with ever-more refined deepfake technology, means that we can no longer trust video calls when interviewing candidates. As such, relying on them as a vetting technique or for identity verification creates a security risk.

Organizations Must Take Steps To Tackle the Deepfake Threat

If organizations are to safeguard against deepfakes, then they must equip themselves with the right tools to verify an individual's identity. These tools need to be used across the board, from onboarding customers to interviewing remote candidates.

The public agrees. According to iProov's 2022 data, **80%** globally said they'd be more likely to use an online service that prevented deepfakes. The demand for further protection is growing. The demand for deepfake detection has grown since 2019.

Would You Be More Likely To Use an Online Service That Had Measures in Place To Prevent Deepfakes Being Used?



How Can You Ensure You're Interviewing a Real Person?

Organizations that hire remote workers should be concerned about the deepfake threat. But there is a solution.

To combat deepfakes, organizations must ascertain whether the interviewee is the right person, a real person and authenticating right now.

Facial biometric technology can be used to verify that someone is who they say when carrying out an activity online, such as opening a bank account or applying for a job. But there are different levels of facial verification. We'll discuss them here:

Liveness Detection

Deepfakes can be used to spoof verification processes in a number of ways. The simplest way is that someone creates a deepfake photo or video and presents it to a camera. This is what's known as a presentation attack. Other presentation attacks involve presenting a mask or image to the camera.

Liveness, such as iProov's Liveness Assurance™, is a form of facial verification that can distinguish whether what's been shown to the camera is an actual person or a presentation attack. Therefore, liveness can prevent a criminal from acquiring a role within a company if they present a deepfake to the camera.

But as we've seen, technology has progressed to the point where attackers can bypass the camera and stream the deepfake directly into the video stream. If an attack is digitally injected, liveness on its own cannot provide a safeguard. It has no way of telling whether the individual is both a live person and genuinely there right now.

Genuine Presence Assurance®

Genuine Presence Assurance (GPA) is the only way to check that a remote individual, who is asserting their identity, is the right person, a real person and that they are authenticating right now.

GPA provides liveness detection, meaning it has the ability to identify whether a user is a real human being and not a presentation attack. But it also detects digitally injected attacks – those that often use deepfakes to bypass the device sensors and spoof the system.

Whereas liveness detection alone can usually protect against known threats – mostly presentation threats (physical and digital artifacts shown to a screen) – GPA delivers defences against known, new and evolving synthetic digital injection attacks.

It does this using iProov's Flashmark™ technology, which illuminates the remote user's face with a unique, randomized sequence of colors. The analysis of the returning color sequence assures the genuine presence of the user. This mitigates the risk of replays or synthetic manipulation, preventing spoofing.

How Does It Work?

Stefan is applying for an IT support role. He completes the application process and submits his resume. He is then asked to verify his identity by scanning his ID document – such as a driver's license – and then his face using iProov's Genuine Presence Assurance.

This verifies that he is the right person, a real person and that he is authenticating right now. When it comes to the interview, he can authenticate himself to prove that he is the same person. The hiring organization can then be confident that Stefan is not a deepfake.

The Deepfake Threat: An Arms Race

As deepfakes become more sophisticated, so too should the biometric security that is used to combat them.

At iProov, we use machine learning technology, people, and processes to detect and block cyber attacks, including deepfakes. In doing this, we are constantly learning from attacks. This helps prevent fraud, theft, money laundering, and other serious online crime today and tomorrow.

Next Steps

For a demo: to see how remote biometric face verification can detect deepfakes, please e-mail contact@iproov.com and we will contact you to arrange a demo.

For more information: visit our website at www.iproov.com to find out more about how we use face verification to help organizations to complete secure processes online.

Methodology

This report is based on research carried out by an independent agency on behalf of iProov in April-May 2022. Eight countries were included in the research (the US, Canada, Mexico, Germany, Italy, Spain, the UK and Australia) with 2,000 consumers surveyed in each country.

contact@iproov.com

iproov.com