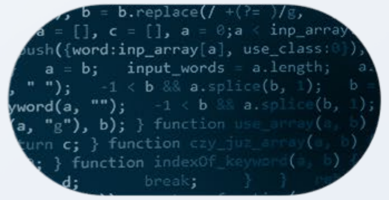
iProov

Identity Crisis in the Digital Age:

Using Science-Based Biometrics to Combat Malicious Generative AI

<not all is what it seems
sure who you can tru.st>
2053020394
<alert> <!>
<not all is what it
<are you/ sure who you can
2053020394 030494020 29094-95-n29394-n0090
<alert> <!>
falso 1939-492
<not all is what it seems> <n.ow> ms
seems> <n.ow>
tru.st> <?>
29094-95-n29394-n0090
<alert> <!> <!> <!> 01293
350112309493
<codigo rojo>
<not all is what it seems> <n.ow>
<are you/ sure who you can tru.st> <?>

All images showing people within this report have been created using generative AI tools, and are intended to demonstrate the visual capabilities of generative AI



Contents

Introduction

The potential applications of artificial intelligence (AI) are immense. AI aids us in everything from early cancer diagnoses to alleviating public administrative bureaucracy to making our working lives more productive.[1,2]

Generative AI is a form of machine learning that is capable, amongst other things, of generating media, such as imagery, audio, and text. While it has many positive uses, the technology is also employed for illicit purposes, such as creating content to spread disinformation online. Some experts believe this trajectory could lead to an “identity crisis”: a world where the public cannot trust the media or their elected officials, organizations cannot trust their users, and individuals cannot trust one another in remote settings.[3]

Identity fraud can be the inflection point for a domino effect of other crimes committed. Once a threat actor infiltrates a system using a spoofed or synthetic identity, funds gained can be laundered and funneled into other illicit activities, such as narcotics trafficking, terrorist financing, and even human trafficking. Likewise, personally identifiable information (PII) obtained through identity fraud-related data breaches can be sold over the dark web and used to support new identity fraud schemes.

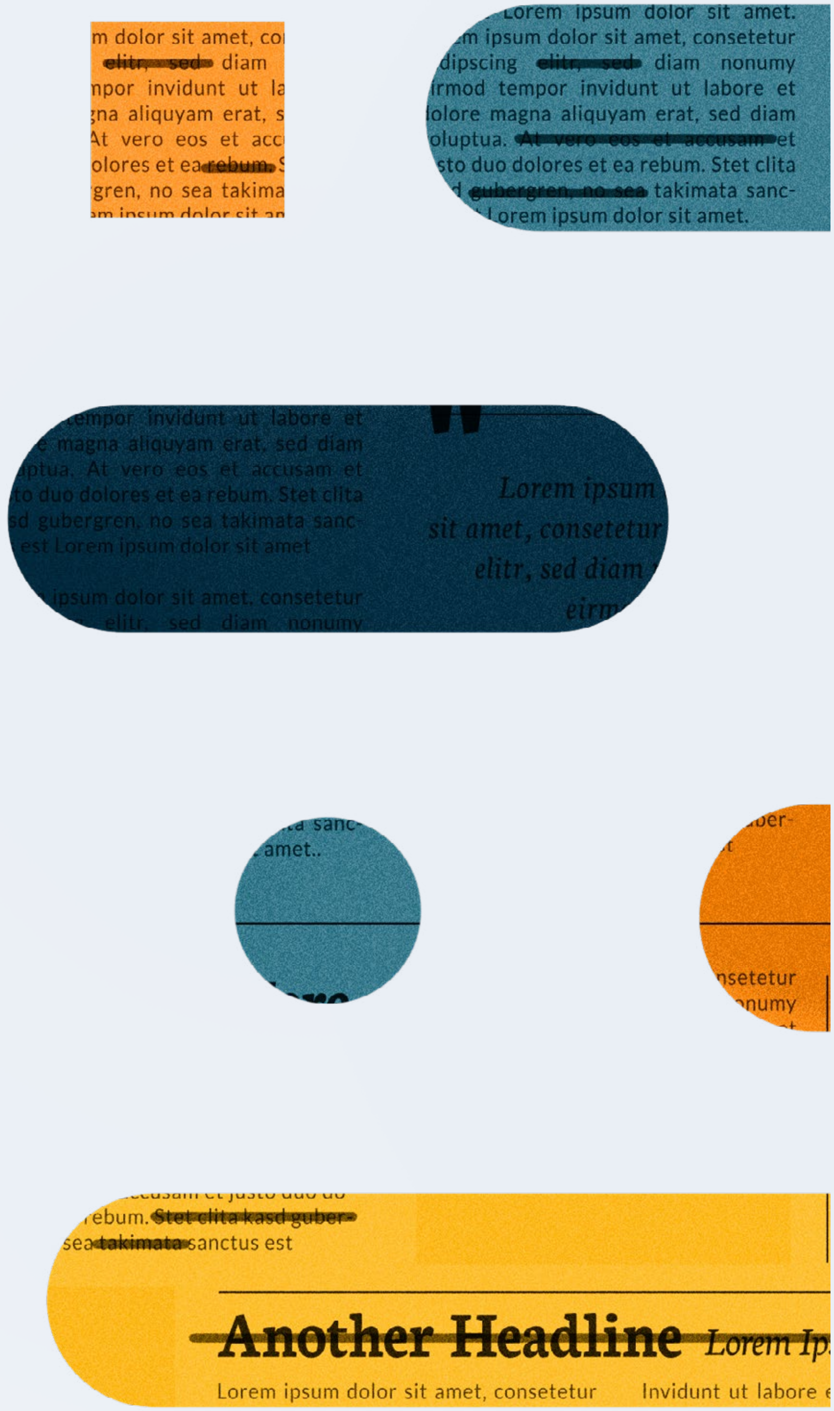
Regulators are racing to curb the dangerous elements of generative AI while simultaneously harnessing its potential. The EU is at the forefront of this agenda. The AI Act, a critical part of the EU’s Digital Services Act, is taking a risk-based approach, even prohibiting technologies deemed to be high-risk.[4]

In an increasingly digital world, resilient identity verification is essential to assure remote individuals and entities are who they claim to be, not AI-generated synthetic media.

This report focuses on the criminal side of generative AI and advises on the technology and processes needed to combat it.

[1] *No longer science fiction, AI and robotics are transforming healthcare, PWC, 2022*
[2] *Can AI help governments clean out bureaucratic “Sludge”?, LinkedIn, 2023*
[3] *Lack of AI Regulation Amidst Deep Fake Identity Crisis: Crypto and Blockchain Can Help, LinkedIn, 2023*
[4] *The EU AI Act’s Risk-Based Approach: High-Risk Systems and What They Mean for Users, European Commission, 2023*

Generative AI for Bad: Fake News and Disinformation



Trust that a public figure is who they claim to be or that consequential pieces of information are genuine is essential to the functioning of society. Due to the proliferation of AI-generated fake news, the UN Secretary-General recently warned of a “trust deficit” that threatens to undermine nations’ abilities to reach sustainable development goals.[5] Likewise, identity compromise in the digital age can lead to fraud on a massive scale, regulatory fines, and, in some cases, irreversible reputational damage, or worse.

The utilization of generative AI by nation-state actors to spread disinformation online is becoming a great concern for governments, particularly as the US and Europe head into election season.

AI “will almost certainly be used to generate fabricated content, AI-created hyperrealistic bots will make the spread of disinformation easier, and the manipulation of media for use in deepfake campaigns will likely become more advanced.”[6]

- UK National Cyb er Security Centre

The rapid advancement in the sophistication of synthetic media today makes it extremely difficult to detect malicious use, putting organizations into situations where they need to minimize the negative impact caused by these attacks. A more effective approach is for organizations to use advanced technologies and processes to prevent AI-generated attacks before they can harm individuals and entities.

[5] *Secretary-General Highlights ‘Trust Deficit’ amid Rising Global Turbulence, in Remarks to Security Council Debate on ‘Upholding United Nations Charter’, United Nations, 2020*
[6] *Deepfakes could disrupt next British election, GCHQ warns, The Times, 2023*

How Can Organizations Assure the Identity of Remote Individuals?

As synthetic media becomes more pervasive, organizations that operate digitally must implement tools to enable individuals to prove they are the rightful owners of their genuine identity, content, and credentials. These technologies must be reliable, inclusive, and usable to establish trust and drive adoption.

Assuring that a remote person is who they claim to be is essential for many high-risk use cases, including enrolling a new banking customer, onboarding a new employee, or even verifying the authenticity of a court witness or journalist.

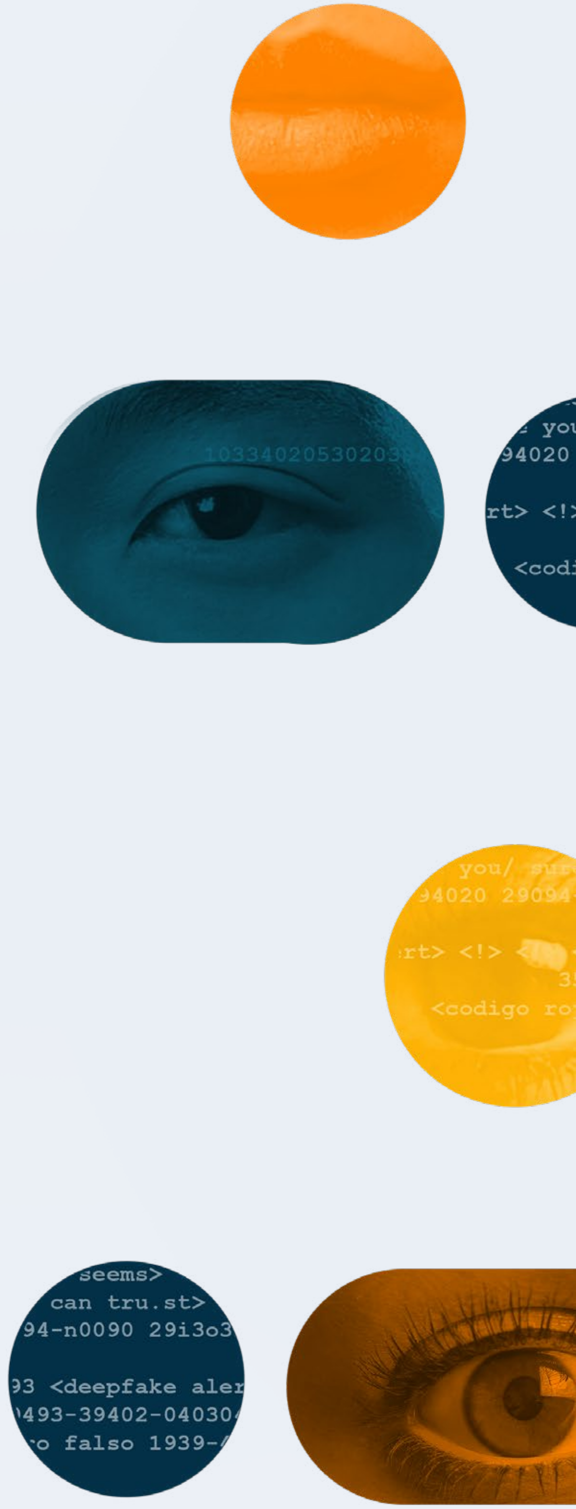
The organization must bind the individual's digital identity to their real-world, government-issued ID to establish their identity. The digital identity should be reusable so individuals can authenticate with the same credentials throughout their user lifecycle.

Traditional identity verification methods typically relied on in-person checks, which are no longer feasible in today's digital landscape. Organizations must verify individuals' identities remotely to meet user demand for convenient digital experiences. One such method is video call verification, in which the individual is asked to participate in a live two-way conversation with a trained operator over a video conferencing tool while presenting their government-issued ID on camera. The operator then verifies the individual's identity by matching the trusted document to the user's face.

One drawback of video call verification is that it relies on humans' ability to detect whether the individual's document and face are genuine, which is increasingly problematic given the sophistication of AI-generated media.[7] Studies consistently show that even trained humans have a low success rate at detecting synthetic media. In 2023, the *Journal of Cyber Security* published a research paper that assessed individuals' ability to differentiate between deepfake images and those of real people, along with their self-reported confidence level in their answers. Of nearly 300 participants, the success rate of spotting a deepfake was only 50%.

In a high-profile case, ethical hackers at the German-based Chaos Computer Club were able to circumvent video call verification technology and a human operator using AI-generated imagery and a forged ID.[8]

Organizations must, therefore, leverage the positive aspects of artificial intelligence to assist them in ensuring a remote individual is who they claim to be. Biometric face verification has proven to be a reliable tool for achieving this goal. By binding an individual's digital identity to their biometric data, organizations can verify their identity with a high degree of accuracy.



[7] Testing human ability to detect 'deepfake' images of human faces, Oxford Academic, 2023
[8] Chaos Computer Club hacks Video-Ident, Chaos Computer Club, 2022

AI for Good:

Biometric Face Verification

When used with biometric face verification, AI enhances the accuracy, security, and speed of the identity verification process. Users can remotely verify their identity by scanning an individual's trusted document and their face. Deep learning models such as convolutional neural networks (CNNs) detect and match the images. At the same time, liveness detection uses computer vision to ensure that the imagery is of a real person and not a non-living spoof, such as a deepfake, mask, or other synthetic media.

Face Verification Has Many Advantages

Identity Verification

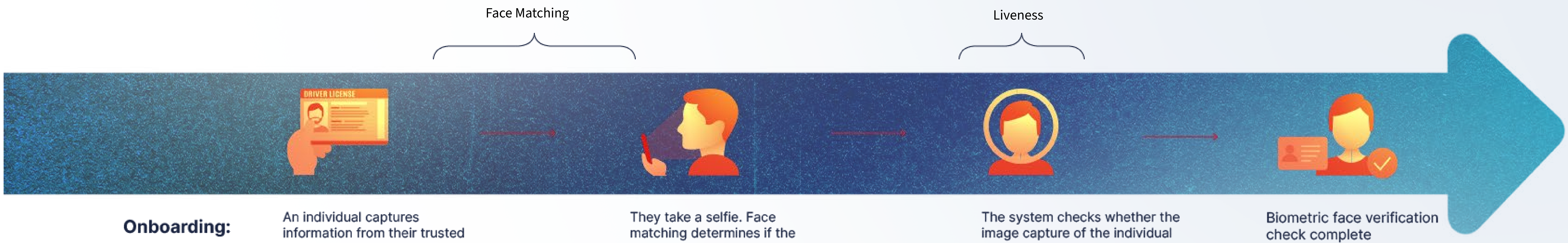
Aims to assure that a remote individual is who they claim to be by binding their digital identity to their trusted government-issued ID. This increases the security of the remote onboarding process.

Reusable Digital Identity

Data is encrypted and stored as a biometric template once the individual has verified their identity. They can use their face as a credential to authenticate throughout the user lifecycle.

User Convenience

Facial biometrics are inherent to an individual and cannot be lost, forgotten, or compromised, unlike knowledge or possession factors. Since people always have their faces with them, they can verify or authenticate from anywhere.



Face Matching

Liveness

Onboarding:

An individual captures information from their trusted government-issued ID

They take a selfie. Face matching determines if the two images are the same

The system checks whether the image capture of the individual is 'live' and not a spoof

Biometric face verification check complete

Biometric Face Verification vs. Facial Recognition:

A Critical Difference

The EU's AI Act classifies face recognition technologies as "high-risk," meaning that the technology is permissible for a blanket ban.[9] Human rights campaigners, such as Amnesty International, have welcomed the decision, stating that there is "no human rights-compliant way to use [face recognition]."[10]

It is important to distinguish between face recognition and biometric face verification. With face verification, the user opts into the process knowingly to verify themselves. Alternatively, facial recognition captures the user's imagery without their knowledge or consent.

Face Verification

Individuals prove who they are to access services

Consent-based

One-to-one

Face Recognition

Used for surveillance

Operates without the individual's consent

One-to-many

[9] Regulating facial recognition in the EU, European Parliament, 2021
[10] EU: European Parliament adopts ban on facial recognition but leaves migrants, refugees and asylum seekers at risk, Amnesty International, 2023

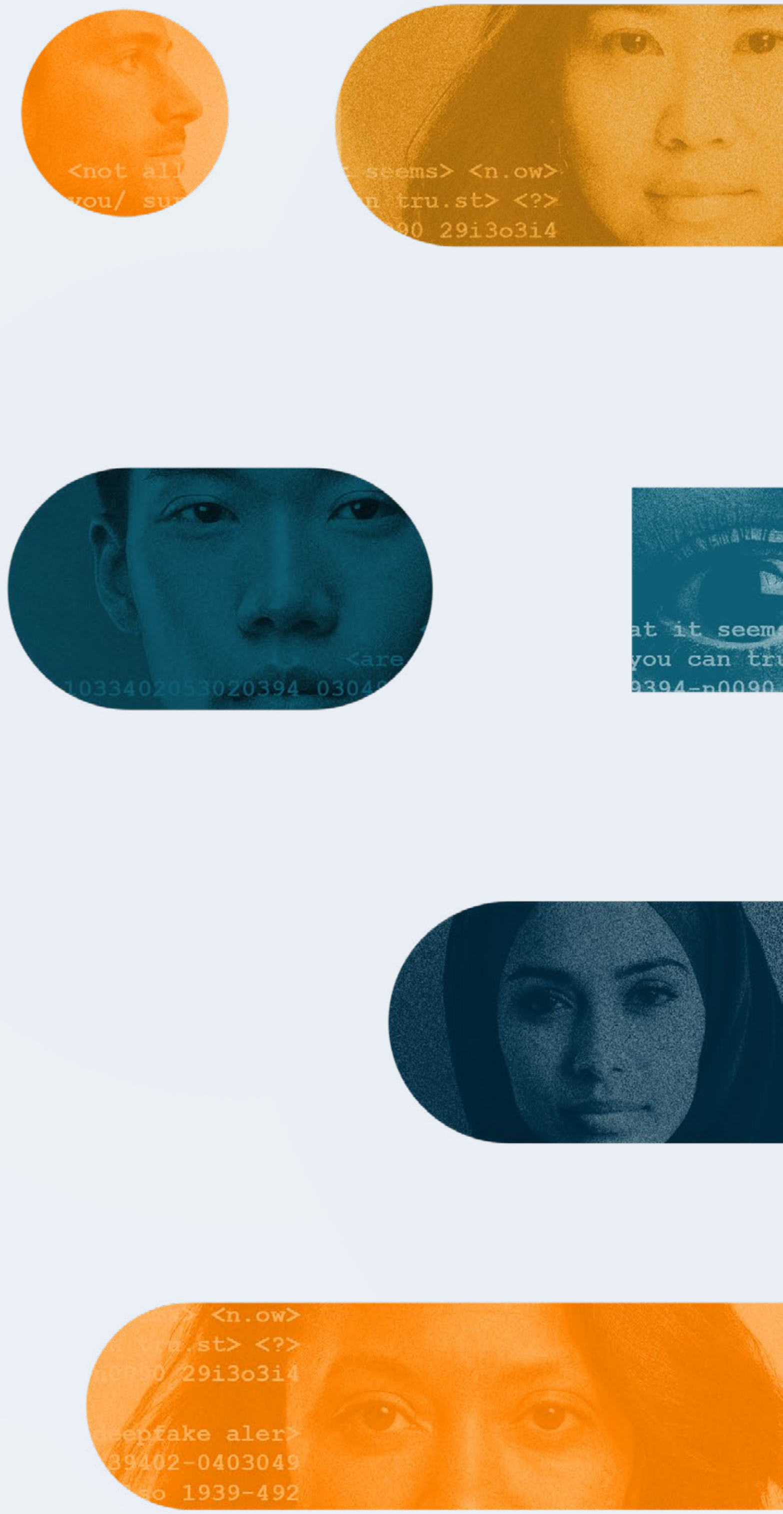
Identity Assurance with Science-based Face Verification

Biometric face verification is one of the most reliable and convenient methods to verify identity remotely. However, not all facial biometric technologies are created equal. To do this, science-based computer-vision technology that leverages artificial intelligence is needed to confidently detect that an individual is a ‘live’ person (not a gen AI spoof) and that they are genuinely present and verifying in real-time.

To achieve this, face verification technologies must utilize a one-time biometric. A one-time biometric is a challenge-response mechanism that is never repeated in a user’s lifetime.

The random nature of the technology makes it unpredictable, impervious to replay attacks – when threat actors inject previous authentication attempts to bypass the system – and incredibly challenging to reverse engineer. It is the only way to mitigate injection and generative AI attacks effectively by detecting genuine presence with high assurance.

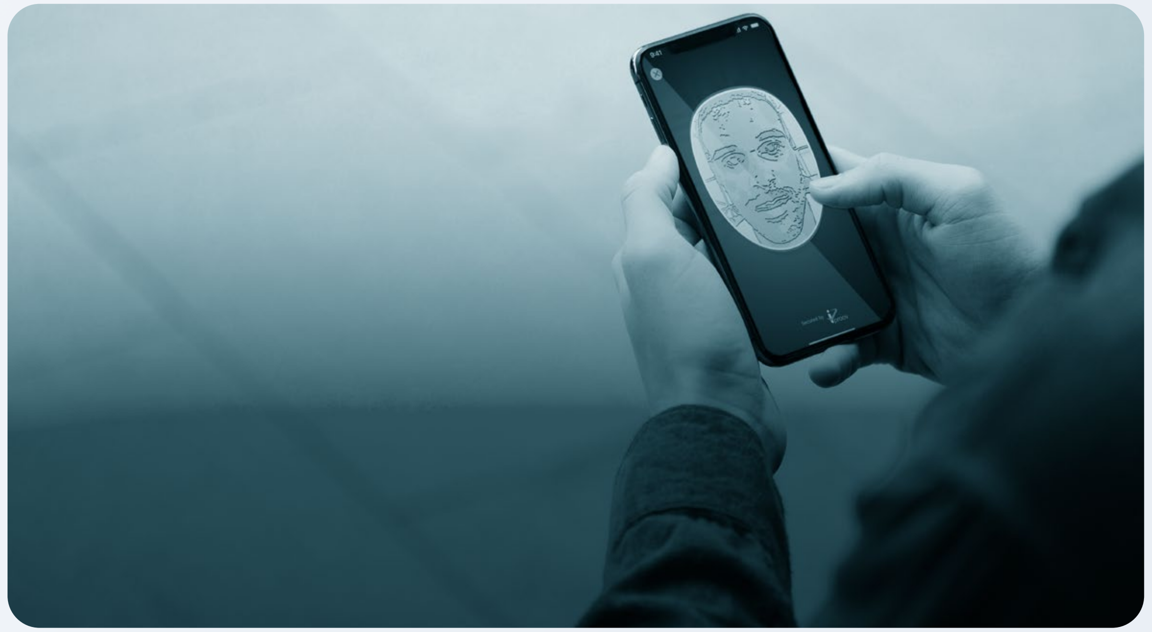
However, applying a one-time biometric is not the only factor contributing to the security of face verification technologies. How the technology is deployed – on-device, on-premise, or cloud-based – also plays a key role.



On-device face biometrics, like Apple's Face ID®, keeps the biometric data, face matching, and liveness detection on the device and can only authenticate the individual's face that is registered, which can easily be set and is not tied to government-issued identity documents such as a passport or driving license.

On-premise technologies bind biometric data to a trusted identity document but can be reverse-engineered and susceptible to spoofing. As the defenses are static, the solution quickly becomes outdated and vulnerable.

Cloud-based face verification is hosted on the vendor's servers, is difficult to reverse engineer, and can bind a digital identity with a trusted document. The vendor has full control over algorithm updates and can adapt defenses accordingly.



As biometric threats evolve rapidly, simply reacting to potential attacks is insufficient to ensure adequate defenses. Instead, biometric vendors must adopt a science-based, multimodal approach that allows for reactive measures and builds resilience against current and future threats.

Threat Monitoring and Intelligence Sharing

Biometric face verification should not be perceived as an out-of-the-box product but rather as a service that the vendor delivers to the organization continuously. This must include ongoing scientific research and active monitoring of the biometric threat landscape, combining good artificial intelligence with human expertise and processes.

Active threat monitoring is essential to combat generative AI for several reasons:

Responding to Novel Threats

Face verification vendors need real-time insight into threat actor methodologies to adapt defenses quickly.

Future-Proofing SDKs (Software Development Kits)

Through exposure and insight into attempted attacks across multiple geographies, platforms, and devices, vendors can continually fortify their SDKs, not just against the attacks of today but also the future.

Responding to Zero-Day Vulnerabilities

In mission-critical use cases, vendors must proactively monitor and block unknown and novel attacks, closing any vulnerability gaps within hours, not weeks or months.

Intelligence Sharing

Face verification vendors with biometric threat intelligence must share insights with their customers.

Securing Face Verification:

Why Presentation Attack Detection is No Longer Enough

Organizations that value security, such as governments and banks, rely on biometric face verification to safeguard their assets. However, it's predictable that financially motivated bad actors will attempt to bypass this technology using the methods described below.

While faces cannot be stolen, they can be copied or synthetically rendered. However, simply creating a mask, face swap, or deepfake on its own is not a threat. The two principal types of biometric attacks are presentation attacks and digital injection attacks.

Presentation Attack

The threat actor presents an object, such as an image, mask, or device, displaying an image or synthetic media to the camera.

Digital Injection Attack

The threat actor injects real or synthetic imagery, such as generative AI-driven deepfakes, into the data stream.

While presentation attacks remain rife, the much greater threat lies in digital injection attacks, which are infinitely scalable and more difficult to detect.

iProov's research into the biometric threat landscape shows that during the second half of 2022, digital injection attacks outnumbered persistent presentation attacks 10 to 1.[11]

Presentation attacks are well-known, and many vendors have been accredited for presentation attack detection (PAD), such as the most recent NIST Face Analysis Technology Evaluation (FATE) report.[12] However, synthetic image injection attacks are evolving fast and perpetuated by advancements in generative AI and are less understood by the wider industry. Currently, no testing bodies exist for digital injection or AI-generated attack detection.

How Generative AI Is Used Against Face Verification

The use of generative AI has gained widespread acceptance and ceased to be thought of as a futuristic concept. Crime-as-a-Service marketplaces have enabled custom-built attacks to be purchased at modest prices[13] and pre-packed attack tools with 'how-to' guides are readily available from code depositories, helping low-skilled threat actors to launch sophisticated attacks. Consequently, threat actors utilize generative AI to exploit digital identity verification in numerous ways.

2053020394 0
<are you/ sur
2053020394 030494020 29094
<alert> <!> <!>
<not all is what it seems> <n.ow> ms
<not all is what it seems> <n.ow>
<are you/ sure who you can tru.st>
94-n0090
<alert> <!> <!> <!> 01293 <deepfake alert>
350112309493-39402-0403049
<codigo rojo> rostro falso 1939-492
<not all is what it seems> <n.ow>
<are you/ sure who you can tru.st> <?>

[11] *iProov Biometric Threat Intelligence Report, iProov, 2023*
[12] *Face Technology Evaluations - FRTE/FATE, NIST, 2023*
[13] *Cybersecurity and Authentication, Raconteur, 2023*

Synthetic identity fraud (SIF) is a rapidly growing financial crime in the US that exploits the onboarding processes of organizations. SIF fraudsters combine invented and stolen personally identifiable information (PII) to create an identity that does not exist in the real world. This allows threat actors to create new accounts with synthetic identities, default on credit and loans, launder money, and, in some cases, fund terrorist activities. The Deloitte Center for Financial Services expects SIF to generate at least a $23bn loss by 2030.[14]

Generative adversarial networks (GANs), threat actors leverage GANs, a generative AI technology, to create synthetic imagery that makes synthetic identities more plausible. Synthetic imagery of a face depicting a non-existent person is generated using this technology and matched with forged IDs to circumvent facial biometrics during onboarding and throughout the user lifecycle. To learn more about the relationship between synthetic identity fraud and generative AI, refer to the iProov report: Stolen to Synthetic: The Evolution of Identity Fraud and the Need for Resilient Identity Verification.

Face swaps: Another approach involves using generative AI to create synthetic imagery to take over the account of genuine users during authentication. Face swaps, for example, are synthetic imagery where an attacker superimposes the biometric template of an authorized user over their face. By combining the traits of one face with the appearance of another, the threat actor can attempt to circumvent the biometric face verification technology and gain unauthorized access to the account.

To become resilient to the evolving threat landscape, face verification technologies must deliver beyond presentation attack detection and deploy methods to thwart the greater threat of gen AI and digital injection attacks.

[14] How with biometrics shape the future of banking?, TheBanker, 2023

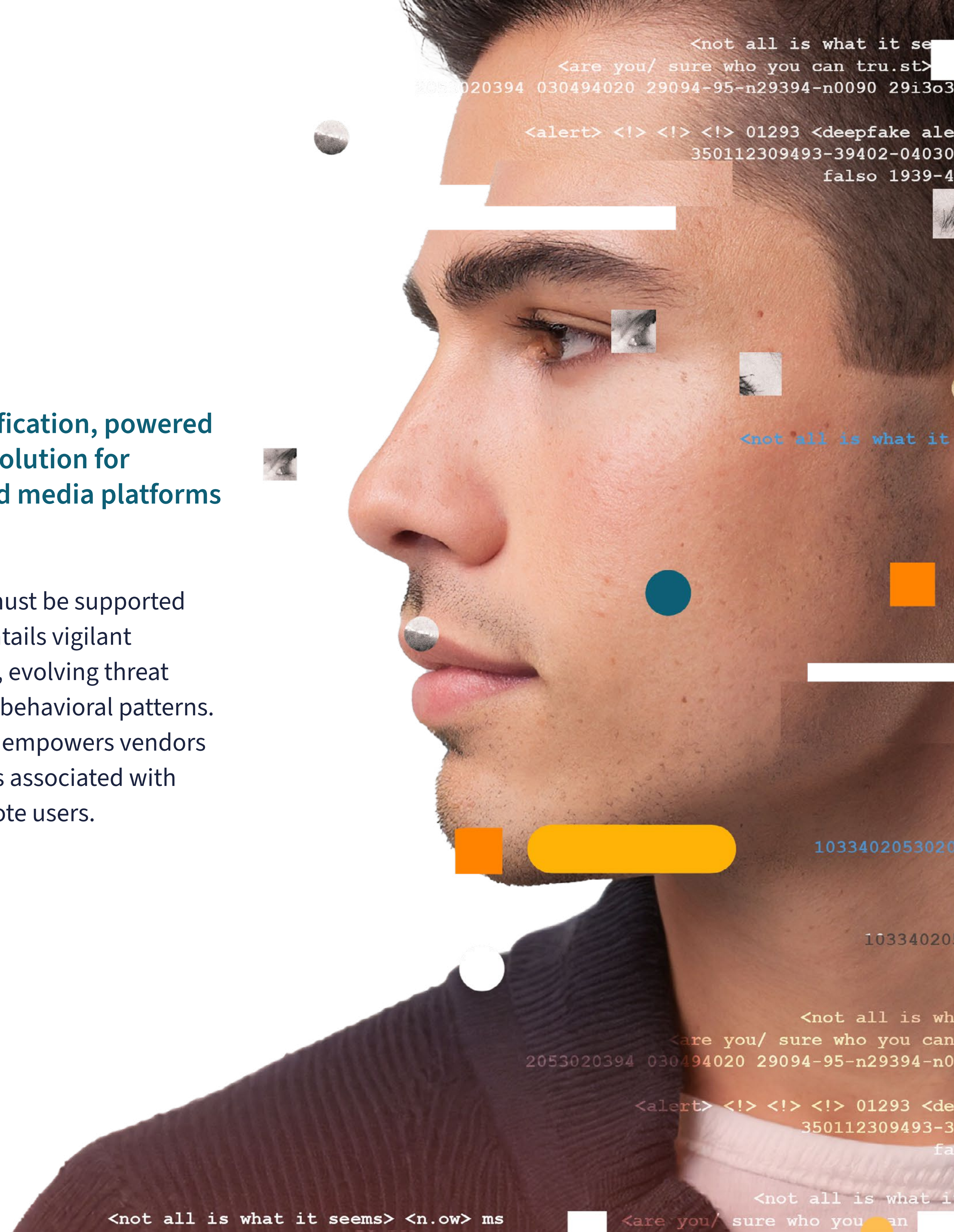
Summary

Generative AI's dangerous applications are becoming increasingly evident to policymakers, organizations, and the public. In the US, for example, a federal judge has limited the Biden administration's communication with social media firms, citing the companies' inability to tackle harmful content and AI-generated disinformation as the reason.[15] Meanwhile, the threat of AI-generated attacks has driven two-thirds of fintechs to boost fraud response budgets.[16]

Establishing trust in individual or entity identities is crucial to combat the negative aspects of generative AI. Organizations must be confident that remote users are who they claim to be.

Cloud-based biometric face verification, powered by advanced AI, offers the best solution for governments, organizations, and media platforms to verify identities.

In addition, the technology utilized must be supported by ongoing scientific research that entails vigilant observation of attack methodologies, evolving threat vectors, persistent threat actors, and behavioral patterns. Accessing this intelligence effectively empowers vendors and organizations to combat the risks associated with generative AI and instills trust in remote users.


<not all is what it se
<are you/ sure who you can tru.st>
020394 030494020 29094-95-n29394-n0090 29i3o3
<alert> <!> <!> <!> 01293 <deepfake ale
350112309493-39402-04030
falso 1939-4
<not all is what it
1033402053020
10334020
<not all is wh
<are you/ sure who you can
2053020394 030494020 29094-95-n29394-n0
<alert> <!> <!> <!> 01293 <de
350112309493-3
<not all is what i
<not all is what it seems> <n.ow> ms
<are you/ sure who you an

[15] Biden officials must limit contact with social media firms, BBC, 2023
[16] Threat of Deepfakes Drives Two-Thirds of FinTechs to Boost Fraud Budgets, PYMNTS, 2023

Verify Your Online Users' Identities Securely and Conveniently

Right Person, Real Person, Right Now

Get an iProov Demo Today >

Find out why so many governments, banks, identity providers, and other organizations trust iProov to verify and authenticate user identities online.

Experience a demo of our facial biometric technology – used by the US Department of Homeland Security, the UK Home Office, the UK National Health Service (NHS), the Australian government, the Singaporean government, Eurostar, Rabobank, ING, and many more.

Protect Your Organization From Cybercriminals and Fraud

- Deliver high customer completion rates
- Deliver outstanding usability
- Deliver inclusive accessibility
- Deliver maximum privacy for your customers