

The Promise of Biometric Authentication Versus the Threat of Deepfakes

Modernizing Identity Access Management to improve both user convenience and security





Executive Overview

In recent years, the rise of deepfake technology has introduced a new and concerning threat to online security. Deepfake videos, images, and voices, powered by sophisticated artificial intelligence algorithms, allow cybercriminals to create convincingly realistic impersonations of individuals, including their facial expressions and voices.

Bad enough that these videos can fool humans, but worse, as organizations turn to biometric authentication to replace passwords, deepfakes expose vulnerabilities that many first-generation biometric authentication systems had not anticipated and cannot address.

This whitepaper delves into the intricacies of deepfake attacks, exploring their types, tactics used for identity impersonation, and the role of biometric authentication with liveness detection and injection attack detection in mitigating these threats.

By understanding the threats and risks presented by deepfake attacks and implementing appropriate security measures, Information Technology and Security executives can modernize identity and access management to improve the user experience and better safeguard their organization while protecting their workers' and customers' online accounts from compromise.

Introduction

In 1997 when IBM's Deep Blue beat grandmaster Garry Kasparov in a game of chess, Machine Learning entered the modern lexicon. A decade later and enabled by progressively more powerful chip sets, machine learning evolved into deep learning, including layered, hierarchical logic that equipped computers with the ability to recognize complex patterns buried in large datasets, draw inferences, and self-adjust without manual intervention. Deep learning then found its way into the hearts and minds of consumers embedded in Siri, Alexa, and so many other virtual assistants as they listened and responded dutifully to voice commands.

Fast forward to 2014 and learning algorithms move beyond pattern recognition and predictions. Ian Goodfellow has invented the Generative Adversarial Network (GAN) pitting one deep learning neural network against another in a challenge to produce a fake image from statistical analysis of an original. The two networks work as adversaries in an iterative cycle to produce, grade and refine the image for accuracy. Iterations follow at the speed of light. Generative AI is born. Computers now create digital media without following explicit instructions.



Today, generative AI is in the movies and television commercials we watch and in the computer games we play. It is driving cars on city streets, predicting protein structures and folding (e.g., AlphaFold), accelerating materials design (e.g., decomposable plastics) and helping clinicians analyze CAT scans to name just a few commercial applications. It is no wonder criminal elements are using it to attack the places where we work, family and friends we hold dear, and ultimately our identity, assets, and livelihoods.

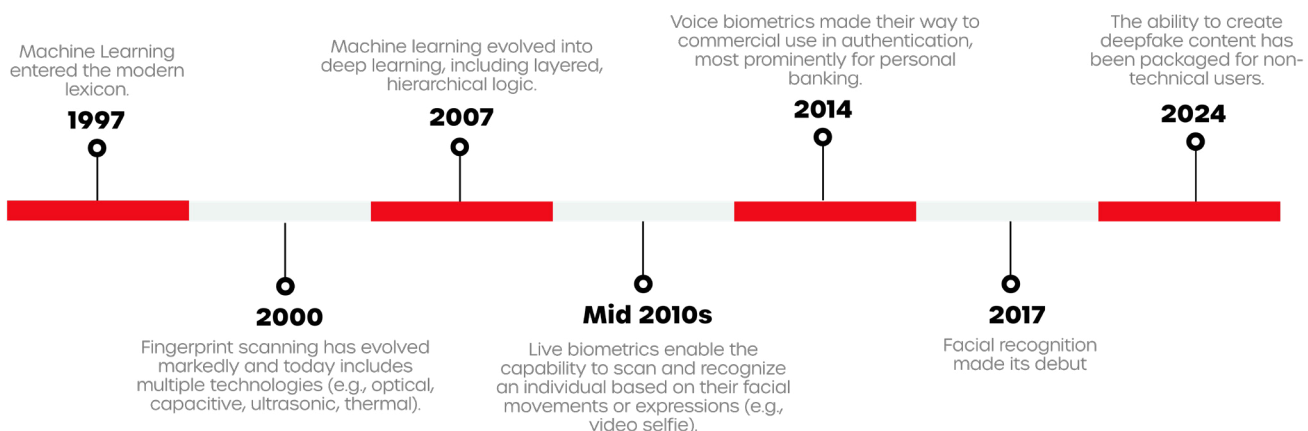
But unlike 2014, today you do not need to be one of the brightest minds in data science to create a deepfake. Less than a decade later, the ability to create deepfake content has been packaged for non-technical users. Inspired to advance learning, creative expression, and entertainment, but repurposed for far less noble intentions by political extremists, social media trolls, cybercriminals, and others, this new generation of synthetic media blurs the lines between fantasy and reality. It is a fool's game to ignore the threats and the risks.

Anything that can be digitally sampled can be deepfaked ... an image, video, audio, even texting to mimic the style and syntax of the sender. Equipped with any one of a half dozen or so widely available tools and a training dataset (e.g., YouTube videos), even an amateur can produce deepfakes of sufficiently quality to coax **one finance worker out of \$25m** or **swindle a few thousand bucks** from an unsuspecting grandparent.

At the crux of it, what makes deepfakes so effective ... a few million years or so of human evolution that has programmed each of us to trust and believe what we see and hear. Moreso than any cyber threat before, deepfakes attack human sensibilities, inspiring fear, affecting our decision making and causing us to act in ways that seem valid, but are inspired by a ruse.

For human decision-making, a measure of awareness, education and street smarts combined with governance and oversight in the corporate domain can help mitigate the threats. But, what about deepfakes turned against biometric authentication?

It is no coincidence that deepfakes are surging at a time when organizations are moving in mass to passwordless authentication and away from antiquated, message-based multi factor authentication (MFA). The "who you are" or inherence factor that biometrics represent is replacing the "what you know" in the form of passwords, one-time codes, or personal factoids, such as the make and model of your first car.



Biometric authentication fundamentally alters the traditional tradeoff between security and convenience by improving both, but they usher into the forefront many inescapable questions about efficacy and security. Are they safer than passwords? What new risks do they introduce? What about privacy? And, which methods are more trustworthy than others?

Let's look at the various types of biometrics, their benefits, limitations, and the attack vectors exploited by deepfakes. But first, just a bit of context on the types of deepfake attacks.

Types of Attacks

Deepfake attacks come in two varieties:

1

Presentation Attacks

2

Injection Attacks



Presentation Attacks

Presentation attacks involve presenting a fake image, rendering or video to a camera or sensor for authentication. Examples include:

Print attacks:

- 2D image
- 2D paper mask with eyes cut out
- Photo displayed on smart phone
- 3D layered mask
- Replay attack of a captured video of the legitimate user

Deepfake attacks:

- Face swapping
- Lip syncing
- Voice cloning
- Gesture / expression transfer
- Text-to-speech





Injection Attacks

Injection attacks involve manipulating the data stream or communication channel between the camera or scanner and the authentication system. Such is common with a virtual device or man-in-the-middle (MITM) attack. Using automated software intended for application testing, a cybercriminal with access to an open device can inject a passing fingerprint or face ID into the authentication process, bypassing security measures and gaining unauthorized access to online services.

Examples Include:

- Uploading synthetic media
- Streaming media through a virtual device (e.g., cameras)
- Manipulating data between a web browser and server
(i.e., man in the middle)

Types of Biometric Authentication

Biometric markers for use in authentication comprise a virtually endless list of possibilities including retina, iris, hand/palm, ear, vein, heartbeat, signature, gait, DNA and even odor. For our purposes, we will focus on the few that comprise the broadest use and widest commercial adoption:

1

Device Biometrics

2

Voice Biometrics

3

Live Biometrics





Device Biometrics (e.g., TouchID, FaceID)

Fingerprint scanning has evolved markedly from its first consumer introduction in 2000 and today includes multiple technologies (e.g., optical, capacitive, ultrasonic, thermal). As an example, Apple's "Touch ID" uses a combination of capacitive and ultrasonic imagery to provide a match probability of 1 in 50,000.

Facial recognition made its debut about a decade later and in 2017 increased match probability to 1 in 1,000,000 by combining machine learning with depth sensing cameras that can capture multiple 3D infrared images from various angles to model the face.

Most leading device-level systems do not store biometric images but rely on mathematical representations. The data supporting those models is encrypted in communication with the OS and while in storage in the Secure Enclave of the device. This makes compromising the device biometric for presentation attacks difficult at best.

Fingerprints do not change with age and offer the convenience of logging in with a simple touch. Faces do change over time, but machine learning can discern typical changes making an even more convenient user experience.

While injection vulnerabilities on open devices have been reported in various Android devices, injection attacks targeting device biometrics are rare and require physical access to a device.

Device biometrics address user convenience, but they fall short of high identity assurance security requirements, such as accessing a corporate VPN, privileged systems, or high value consumer purchases. This is because with device-level biometrics enrollment is informal.

Specifically, anybody who gains administrative access to a device can register their biometric. This could include family, friends, pranksters as well as those with criminal intent. So, while a biometric scan can match a biometric previously registered on the device, the identity associated with that biometric is not established with high assurance. Whether device-level biometrics offer sufficient security is conditioned on the risk associated with granting systems access.



Voice Biometrics

After slow baking over a 40-year timespan, voice recognition began widespread adoption in the early 2000s following advancements in machine learning. A little over 15 years later, voice biometrics made their way to commercial use in authentication, most prominently for personal banking.

Like device-level biometrics, authentication via voice biometrics requires a real-time match against a reference sample captured during registration. Similarly, the typical voice biometrics system does not capture a voice recording, but rather a “voiceprint” comprised of a mathematical model, encrypted, and securely stored. The modeling is sufficiently sophisticated to match the way an individual pronounces words, and includes various other unique factors including tonality, intonations, cadence, and pitch.


Match probability for voice depends on several factors including voice quality, but it is highly accurate though slightly less user friendly than device biometrics because the user may have to repeat a specific phrase captured at registration.

Registration of voice biometrics shares some of the same shortcomings as device biometrics. Any person able to recite accurate personal details can register and then use voice biometrics. This again addresses user convenience but offers little security beyond knowledge-based factors.

For presentation attacks, voice conversion and text-to-speech both use generative AI trained on a voice sample, for example, from an online video or voice recording. The synthetic media can then be presented to the audio mic in the device. Technically, injection attacks are also possible, but it is relatively easy to detect a virtual device.

To the human ear, deepfake audio can be exceedingly difficult to detect. Sophisticated defenses using generative AI can detect slight variations in pitch, speech patterns, and choice of words. Statistical analysis of audio spectral features and even computer vision modeling of vocal pitch, timbre, the dynamic range of voice wavelengths and other inconsistencies have also been refined. And, in an approach reminiscent of antivirus solutions, AI defenses improve when fed modelling data from new fakes as they appear, which makes continuous monitoring and sampling essential for AI defenses to resemble actual learning.





For voice authentication, an “ensemble” approach to security can be highly effective. This combines multiple detection technologies and comparison of a real-time voice sample to a reference sample of the individual captured at registration. But, again, binding of the registered biometric to an identity relegates voice authentication to self-assertion, which inevitably falls on the weak side of user authentication.



Live Biometrics

Live biometrics enable the capability to scan and recognize an individual based on their facial movements or expressions (e.g., video selfie). It has been commercially available since the mid-2010s. Enabled by AI and the widespread availability of front-facing, depth sensing cameras on mobile devices, adoption for authentication with live biometrics has accelerated rapidly over the past few years.


Live biometrics -- and in a lesser (static) form the facial scan -- fundamentally alters the historical tradeoff between security and convenience by significantly improving both. First, unlike the FaceID technology used in device biometrics, both a facial scan and video selfie can be matched to a pre-verified offline credential during registration. This can be a government issued document such as driver’s license, state ID, or passport, or other machine-readable directory such as an HR (Human Resources) system containing employee photos. For this to happen accurately and securely, several activities need to happen in concert.

During the registration process, AI can perform facial matching in real-time by comparing the biometric to the picture on a valid credential. AI can compensate for changes to the face, for example, from aging when matching a person to a valid, but dated credential. At the same time, the credential needs to be verified digitally with the issuing authority. The data between multiple credentials compared or triangulated and photo inspected for manipulation.

Successful registration significantly upgrades identity assurance from “self-asserted” (e.g., NIST Identity Assurance Level 1 or IAL1) to “verified identity” (e.g., IAL2). Verified identity to IAL2 lends itself to supporting several use cases for Know Your Customer (KYC), Strong Customer Authentication (SCA) and Employment Eligibility Reporting (e.g., I9).

Certification of the identity verification (IDV) system to test methodologies such as ISO 30107-3 Presentation Attack Detection (PAD) can validate defenses against basic presentation attacks such as 2D cut out masks, photos, 3D layered masks and replay attacks.





But, with the growing availability of tools for non-technical users to create deepfake videos, a combination of both passive and active liveness detection can be used to harden defenses. Passive liveness detection runs in the background and uses AI to inspect the biometric for signs of image manipulation such as dull shadows around the eyes, unrealistic facial hair, overly smooth or wrinkled skin, fictitious moles, unnatural lip color or lighting, and more.

Active liveness detection asks users to perform facial expressions or movements, again using AI to inspect for strange reflections, compression artifacts and other telltale signs when evaluating a match between the biometric and the one previously registered.

Deepfake defenses are also needed to monitor for an injection attack where virtual camera software can be used to feed pre-recorded deepfake videos into the authentication process, bypassing liveness detection checks. The countermeasure here is to detect when a virtual device is being used and to disallow access.

The continuous improvements in facial recognition algorithms, coupled with the integration of liveness detection mechanisms and monitoring for injection attacks provide compelling defenses against deepfake video attacks. As a result, many organizations and service providers have begun incorporating live selfie video authentication as part of their identity verification processes to enhance security and user experience.

Combating Biometric Spoofing with 1Kosmos

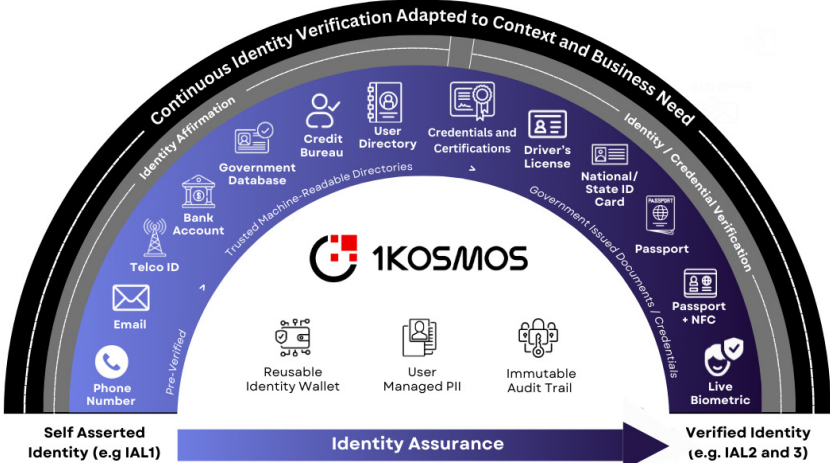
1Kosmos innovative solutions are designed to modernize and simplify the IT infrastructure for IAM while combatting the new threats posed to biometric authentication by deepfakes.

Our state-of-the-art platform is the first to support both remote identity verification and passwordless MFA and to do so in a way that puts users in control of their own PII (Personal Identifiable Information). This approach has several benefits as it:

- Supports straight-through user onboarding with minimal manual oversight
- Secures new account origination, login, and transactions from identity spoofing
- Creates a reusable digital wallet for high assurance login to digital services
- Provides instant validation of user qualifications, competencies, authorities, etc.
- Simplifies security by eliminating centralized honeypots of user data

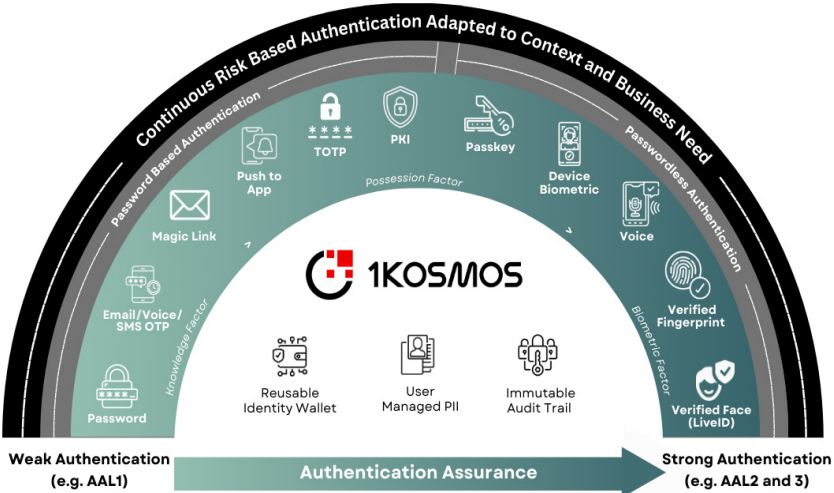


Just as important, 1Kosmos enables tailorable onboarding journeys and flexible deployment options that tune identity verification and authentication controls to the risk profile of different types of systems, transactions, or classes of users. This allows IT and Security administrators, for example, to tighten security controls for accessing privileged systems while adopting more modest controls where transaction risk is low or where user friction needs to be at a minimum.



1Kosmos identity verification supports tailorable browser-and mobile-based user onboarding journeys that span a wide range of use cases from Self-Attested to Verified Identity.

To defeat deepfakes, 1Kosmos LiveID can perform both “active” liveness (requiring the user to perform randomized expressions) and “passive” liveness, one without the user’s involvement. Additionally, 1Kosmos utilizes true-depth camera functionality to prevent presentation attacks and offers an SDK (Software Development Kit) to protect against camera manipulation to prevent an injection attack.



1Kosmos authentication supports multiple authentication use cases spanning business-to-worker, business-to-customer, and government-to-citizen.



Alongside these advances, 1Kosmos also offers the following features:

- **Anti-Spoofing Algorithms:** 1Kosmos anti-spoofing algorithms detect and differentiate between genuine biometric data and spoofed data. Our algorithms analyze factors like texture, temperature, color, and movement to determine the authenticity of the biometric sample, catching virtual/hardware camera and JavaScript injections and ensuring the validity of the transmitted identity.
- **Data Encryption:** 1Kosmos ensures that biometric data is encrypted both during transmission and storage to prevent unauthorized access. Implementing strict access controls and encryption protocols prevents man-in-the-middle and protocol injections, ensuring the validity of the transmitted identity.
- **Adaptive Authentication:** 1Kosmos uses additional signals to verify the user identity based on factors such as networks, devices, applications, and end user context to appropriately present authentication or re-authentication methods to users and to support the context of the utilized method.

Finally, 1Kosmos performs rigorous technical audits to ensure quality coding and to identify and address any vulnerabilities. These include:

- **Regular Audits and Penetration Testing:** 1Kosmos conducts regular audits and penetration testing to identify and address vulnerabilities, including access to a user's biometric data. This helps ensure that security measures are effective and up to date.
- **Regulatory Compliance:** 1Kosmos has certified compliance with regulations and standards related to biometric authentication, data protection, and security. These include the National Institute of Standards and Technology (NIST 800-63-3), UK DIATF, ISO PAD, FIDO2, and SOC II.
- **Human "Failover":** 1Kosmos offers 24x7 staffed call centers to assist when an attack is detected or if a user has trouble completing a verification process.





Conclusion

The technology to create compelling deepfake videos has been consumerized and is now widely available to non-technical users. As the prevalence of deepfake technology continues to rise, so does the threat deepfakes pose to online security. Cybercriminals leverage a variety of tools and tactics to create bogus accounts and execute deepfake account takeover attacks, exploiting vulnerabilities in authentication systems and preying on unsuspecting users.

To combat this threat effectively, IT (Information Technology) and Security executives must remain vigilant, staying abreast of the latest developments in deepfake technology and implementing appropriate security measures to protect against account takeover and identity fraud.

For attacks targeting people, organizations should undergo cybersecurity training and implement appropriate policies and procedures to guard against well intentioned workers who have simply fallen for the spoof.

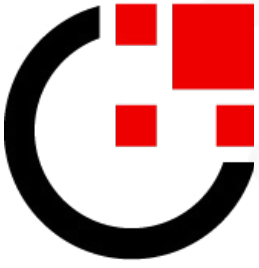
For organizations looking to modernize identity and access management to eliminate the identity gap with passwords and other knowledge-based approaches, biometric authentication offers an opportunity to improve both the user experience and security, but this is not a one-size fits all proposition. Different forms of biometrics offer varying levels of identity assurance.

By leveraging advanced technologies and taking an “ensemble” approach utilizing various counter defenses in concert including performing passive and active liveness detection, detecting and defeating injection attacks, and verifying identity credentials in real time at user registration, live biometrics can offer the highest level of identity assurance available for granting access to digital systems.

Ample standards exist to define the capabilities of these identity verification and authentication systems. Certification to these standards requires rigorous and ongoing testing to verify efficacy. These include the National Institute of Standards and Technology (NIST 800-63-3), UK DIATF, ISO PAD, FIDO2 and SOC II.

By deploying systems certified to these standards, IT and Security executives can efficiently mitigate the risks posed by deepfake attacks to deliver high assurance trust for digital interactions with workers and customers. This can prevent the creation of accounts based on synthetic or stolen identities, secure legitimate accounts from takeover, and prevent identity-related fraud.





About 1Kosmos

1Kosmos enables remote identity verification and passwordless multi-factor authentication for workers, customers and residents to securely transact with digital services. By unifying identity proofing, credential verification and strong authentication, the 1Kosmos platform prevents identity impersonation, account takeover and fraud while delivering frictionless user experiences and preserving the privacy of users' personal information. 1Kosmos performs millions of authentications daily for government agencies and some of the largest banks, telecommunications, higher education and healthcare organizations in the world. The company is funded by Forgepoint Capital and Gula Tech Adventures with headquarters in East Brunswick, New Jersey.

For more information, visit www.1kosmos.com and follow us on [X](#) and [LinkedIn](#).