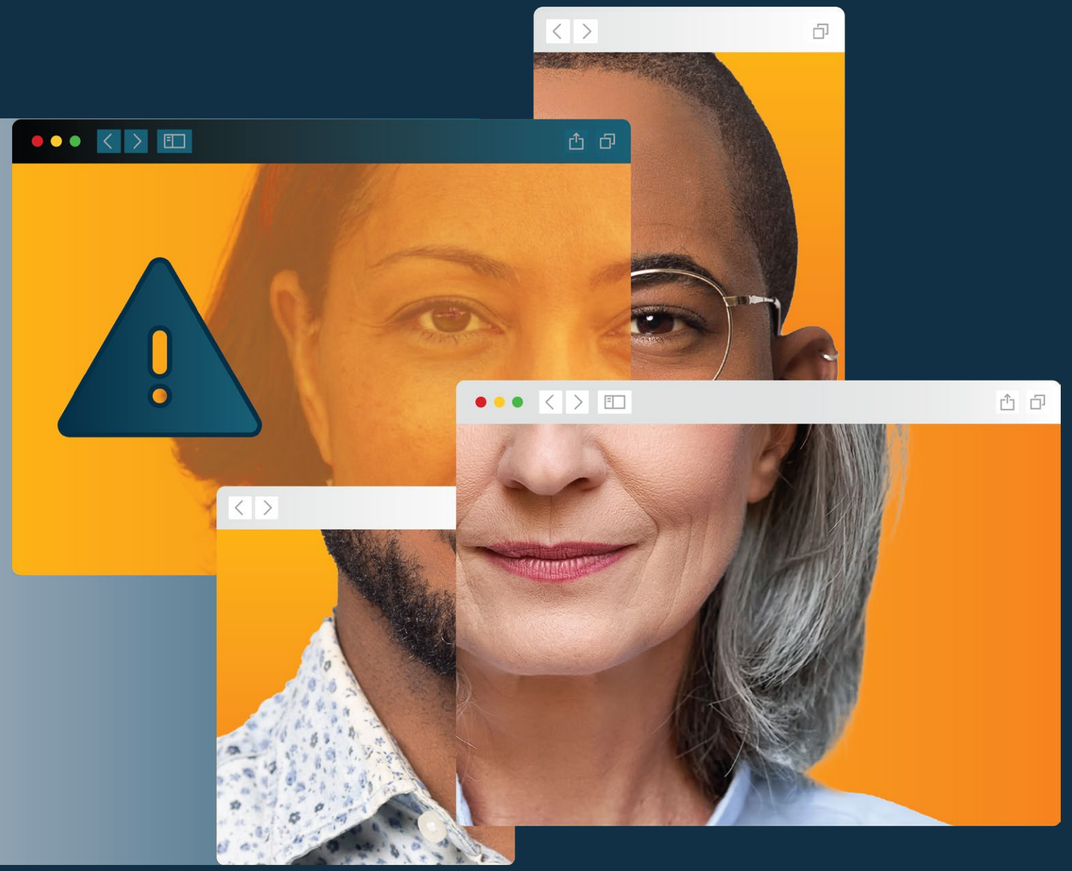
iProov

Threat Intelligence Report 2024

The Impact of Generative AI on Remote Identity Verification



Contents

Foreword

Digital ecosystems continue to grow and expand at record levels as organizations and governments seek to provide remote access and services to meet consumer and workforce demand. However, this growth's unintended side effect is an ever-expanding attack surface that, coupled with the availability of easily accessible and criminally weaponized generative artificial intelligence (AI) tools, has increased the need for highly secure remote identity verification.

While AI technology has the potential to streamline and automate processes for beneficial outcomes, it also comes with an equal number of risks to data protection, cybersecurity, and other ethical concerns. Moreover, innovative bad actors are using advanced AI tools, such as convincing face swaps in tandem with emulators and other metadata manipulation methodologies (traditional cyber attack tools), to create new and widely unmapped threat vectors.

This combination has dramatically increased the risk and diversity of identity-related fraud with cyber-enabled financial crimes, such as money laundering, leading the pack as the greatest threat now and into the future by global law enforcement.[1] For these reasons, verifying the authenticity of remote individuals has become more critical than ever before, particularly for high-risk use cases, for example, in banking, employment, and legal proceedings. When done correctly, remote identity verification can confirm that an individual is the right person, a real person, and authenticating in real-time.

Fortunately, "liveness" capabilities in facial biometrics offer a highly usable and effective solution against presentation, digital, and synthetic attacks. Ensuring that the user is the correct and genuine person at the time of authentication is crucial for mission-critical use cases. Therefore, it is essential to ensure the technology's resilience against evolving threats.

This report aims to uncover the remote identity verification threat landscape, providing first-hand insights into the anatomy of an attack and exposing bad actor methodologies, threat trends, and impacts.

The iProov Threat Intelligence Report 2024: The Impact of Generative AI on Remote Identity Verification - reveals the methods and frequency of attacks deployed by threat actors. Focusing mainly on the tools and techniques employed to launch digital injection attacks, which are the most scalable threats due to both the ease with which they can be automated and the rise in access to malware tools.

1. INTERPOL: Global Crime Trend report

Executive Summary

Navigating the Threat Landscape

In the last 24 months, the threat landscape has undergone significant changes. Organizations considering incorporating facial biometrics into their remote identity platforms need to understand the benefits and drawbacks of the various technologies available and the pros and cons of different deployment methods.

Vendors or organizations already leveraging facial biometrics must constantly analyze observed threats to gain a deeper understanding of attackers' methodologies. Operating in the cloud, the iProov Security Operations Centre (iSOC) closely monitors threat actor patterns and techniques, providing unique, data-driven insights. Our Threat Intelligence Team follows and investigates the most prolific bad actors, learning from and adapting to novel threats before they evolve into serious threats.

Gathering Threat Intelligence

Our threat intelligence has revealed the alarming pace at which threat actors deploy and adapt their attacks. This is the second year we have published our Threat Intelligence Report, which aims to uncover threat actor behaviors, popular tools, and the risk synthetic media, particularly those driven by AI-powered tools, pose to remote identity verification.

2023: Biometric Threat Trends

There are two primary attack types observed by the iSOC: presentation attacks and digital injection attacks. Presentation attacks, such as masks or printed imagery held up to a camera, are not scalable and typically do not require complex or technical expertise. In contrast, injection attacks are scalable and involve a more complex multi-step process. To successfully execute an attack, the threat actor must simultaneously compromise various parts of the service. Specifically, they must establish an entry point (injection) and meticulously manipulate the payload (video) to enroll or impersonate an individual successfully.

1. Face swaps are now firmly established as the deepfake of choice among persistent threat actors.
   - We observed an increase in face swap injection attacks of **704%** H2 over H1 2023.

2. Injection attacks are rapidly evolving with significant new threats to mobile platforms.
   - We saw an increase of **255%** in injection attacks against mobile web H2 over H1 2023.
   - Use of emulators continued to grow rapidly with an increase of **353%** H2 over H1 2023.

3. We observed a significant increase in the persistence of threat actors.
   - Among the top threat actors, attack sequences typically lasted more than 60 days, with multiple threat actors engaged in attack sequences over six months.

4. There has been a significant increase in the number of threat groups engaged in exchanging information related to attacks against biometric and video identification systems.
   - Of the groups identified by our analysts, **47%** had been created in 2023.

Global, Indiscriminate Attacks at Scale

In 2022, iProov observed a general trend of 100-200 attacks from the same attacker and location, launched **three times per week** across multiple geographical clusters.

This general trend seems to be consistent across 2023 as well. However, we did observe a considerable **increase in the number of actors and an improvement in the sophistication of the tools used.**

Across both 2022 and 2023, indiscriminate attacks each month ranged from 50,000 to 100,000.

Request a copy of the 2023 iProov Threat Intelligence Report here.

iProov is widely regarded as a pre-eminent provider of facial biometric technology solutions, trusted by prominent organizations across the globe, including the Australian Taxation Office (ATO), GovTech Singapore, UBS, ING, Rabobank, the UK Home Office, the UK National Health Service (NHS), the U.S. Department of Homeland Security (DHS), and many others.

Introduction

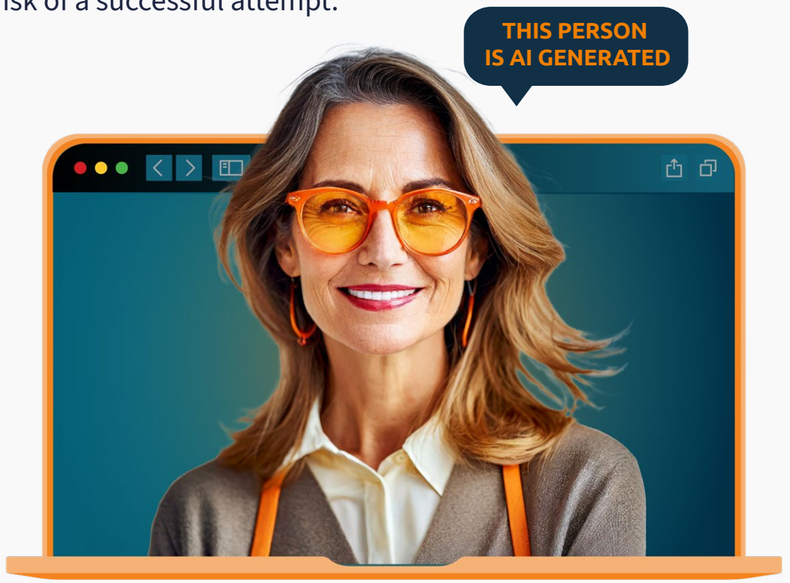
At financial institutions, it's estimated that 95% of synthetic identities are not detected during the onboarding process.[2] The process of checking an individual's identity consists of five parts:

1. Getting evidence of the claimed identity (name, date of birth, and address)
2. Verifying its authenticity
3. Checking its validity over time
4. Assessing the risk of identity fraud
5. Confirming that the identity asserted belongs to the genuine individual.

If the identity document appears genuine, the last line of defense comes down to the capabilities of the remote identity verification process to verify the individual, which is often a manual process. Without the aid of resilient technology designed to detect novel threats, organizations fall at the last verification hurdle.

US Passport forging application.

The previous example showcases a forging tool marketed as functional in the US, allowing users to upload and insert real and synthetic images to bypass both document checks and the remote verification process (human or automated). This is one facet of a more extensive attack methodology that must be addressed to reduce the overall risk of a successful attempt.



2. *Trends in synthetic identity fraud*

Synthetic Media Types

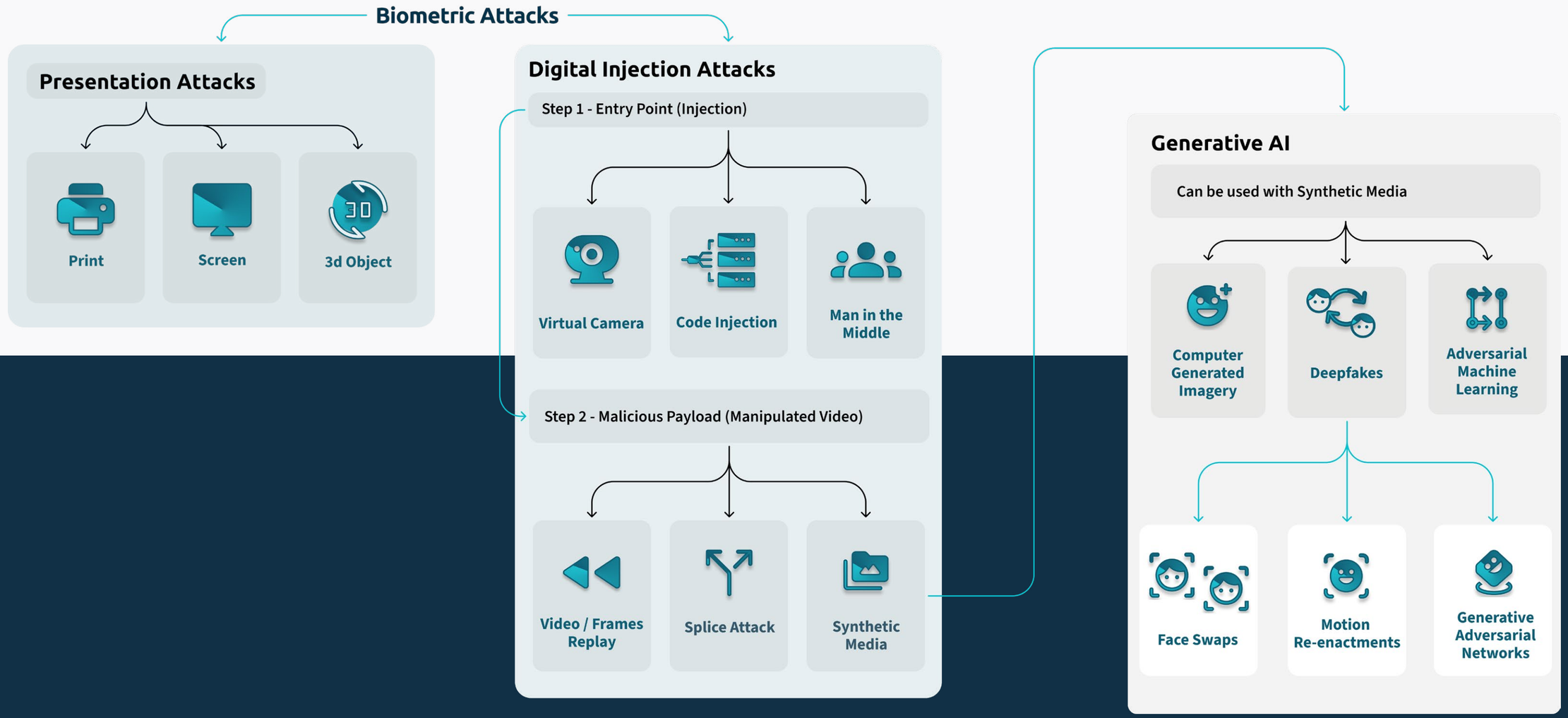
Generative AI and Deepfakes Distinction

The term deepfake has traditionally been used to refer to synthetic imagery created using deep neural networks. In practice, generative AI and deepfakes are now used interchangeably, although a deepfake more likely refers to an image, or occasionally a piece of audio, utilized for malicious purposes. In contrast, generative AI encompasses output in any media (including text) for any purpose. i.e., deepfakes are a subset of generative AI.

The growing trend of face swaps, often used in conjunction with traditional cyber tools, such as emulators, is a cause for concern. It is now easier than ever for anyone to launch an attack within minutes due to the rapid pace of technological advancements in this area. Experts predict that these advancements will continue to progress over the next year[3] according to findings published in the journal 'Vision Research'.

3. Are you for real? Decoding realistic AI-generated faces from neural activity

Taxonomy of Threats



Threat Intelligence Research

iProov's research has discovered more than 60 groups dedicated to creating synthetic imagery worldwide, with varying group sizes ranging from 100 to over 100,000 members. One group that focuses on deepfakes boasts a staggering 114,000 members globally. While some groups are interested in technology, others have malicious intent.

In cases where manual verification backup is in place, discussions in these closed groups claim that intentionally failing the automated biometric process is the easiest way to bypass government identity verification procedures. Steps are shared on how to intentionally fail biometric verification so they can be put through to a human operator, where their members see high success rates.

According to the closed group members, synthetic media fed into the camera stream has a higher rate of success when the process is escalated to a human operator. This has been supported by independent research that found human-operated video identification systems are easily overcome with basic everyday approaches and very little skill.[4]

Example of an Attempted KYC Bypass

Threat actors attempt to simulate motion by manipulating a computer-generated (CG) image.

The technique takes a single still image and animates it with movements - Puppet re-enactment/deepfake

4. Chaos Computer Club Hacks Video-Ident

Technical Innovation Driven by Scientific Expertise

iProov is committed to technical innovation fueled by scientific expertise. Our Science Team comprises specialists in biometrics, computational neuroscience, computer vision, artificial intelligence, and 3-D rendered synthetic imagery. Additionally, our Threat Intelligence and Red Teams are cybersecurity and computer forensics professionals.

Our collaborative, interdisciplinary approach has achieved technological advancements far ahead of the market. Our team of proven experts, with its extensive knowledge and experience, has monitored numerous threat actors and distilled their typical attack patterns into eight common steps:

Identifying and testing the techniques and technologies used by adversaries is crucial. iProov's internal Red Team, supported by iSOC, is unmatched in this area.

Typical Attack Pattern

01 **Identify the Target Company.** These can be government sites, financial institutions, cryptocurrency exchanges, gambling sites, or even dating sites.

02 **Identify the Type of Verification.** Many target companies list the verification providers they use. If it is a liveness solution, threat actors tend to aim for low-cost, easily spoofed providers.

03 **Research Attack Types.** A galaxy of approaches exists if someone aims to exploit or bypass verification. Less secure liveness technologies require significantly less effort and can often be bypassed using cheap or even free tools.

04 **Download a Face Swapping App like DeepFakesWeb.** The application used is often not enough to succeed. Some practice is usually needed to apply a false face to the actor's image.

05 **Document Forgery.** Verification often involves multiple stages, including the presentation of a government-issued ID. Threat Actors will take their desired photo and process it in an illegal document template. In the case of account hijacks, a threat actor may collect imagery of their targeted victim and use that imagery as the basis for a face swap as well as a forged document.

06 **Device Emulator like Android Studio's Emulator.** The inclusion of an emulator allows for mobile verification via a laptop, making the imagery sent to the targeted application easier to manipulate and adapt.

07 **Download a Virtual Camera like OBS Studio.** A virtual camera can simulate the real camera of a mobile device, potentially bypassing a verification company's security.

08 **Configuration.** A threat actor must configure their 'rig' to increase their likelihood of passing. This can involve adjusting settings to perfect their attacks.

iProov's Proactive Threat Response

A threat is a possible security risk that might exploit the vulnerability of a system. An attack is the actual act of exploiting a system's vulnerabilities

Utilizing continuous threat monitoring and mitigation, iProov releases new security updates in parallel as new threats are detected. This work is done by the iProov Security Operations Center, External Threat Intelligence Team, and Red Team, along with the Product and Science Teams.

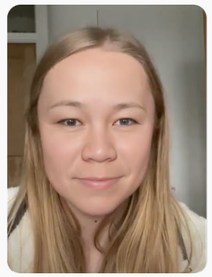
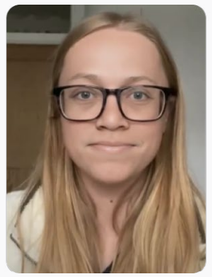
2023 New Trends Discovered by the Threat Intelligence Team

Throughout 2023, iProov's internal Science Team remains vigilant for novel threats. In this section, we highlight the most frequent patterns and methodologies witnessed by our team.

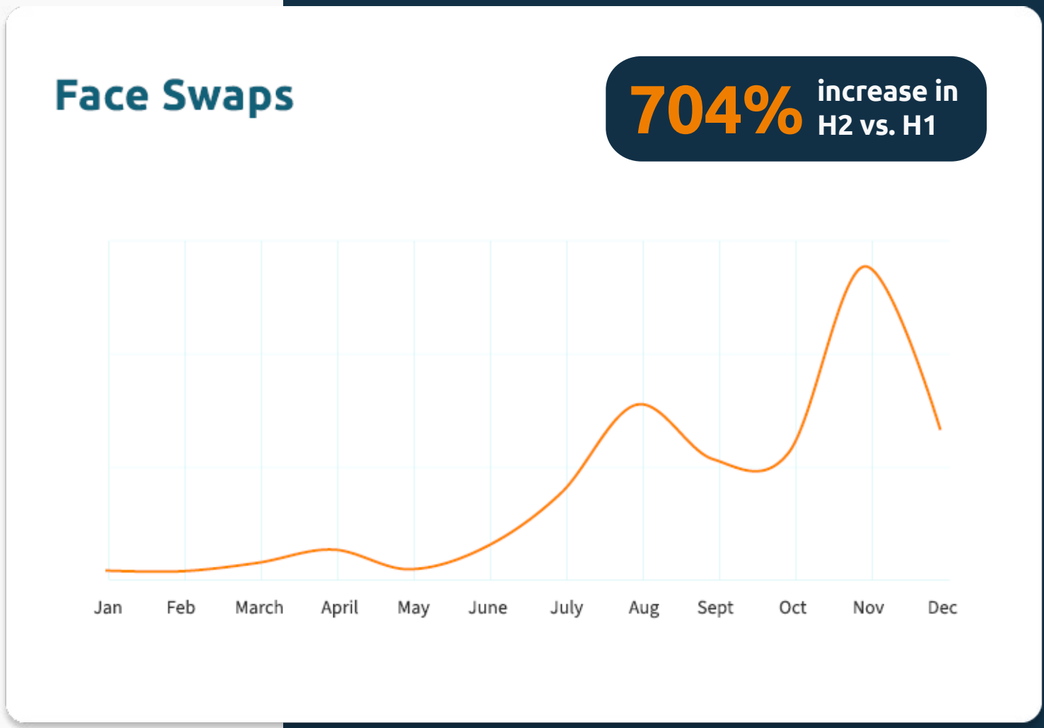
01 Massive Growth in Face Swaps Continues in 2023

Generative AI has revolutionized the field of synthesized media, enabling the creation of highly realistic face swaps and other similar content. However, face swaps derived from generative AI are the primary concern due to their ability to manipulate key traits of images or videos.

We observed an increase in face swap injection attacks of **704%** H2 over H1 2023.


See in motion



Face Swaps Used with Emulators

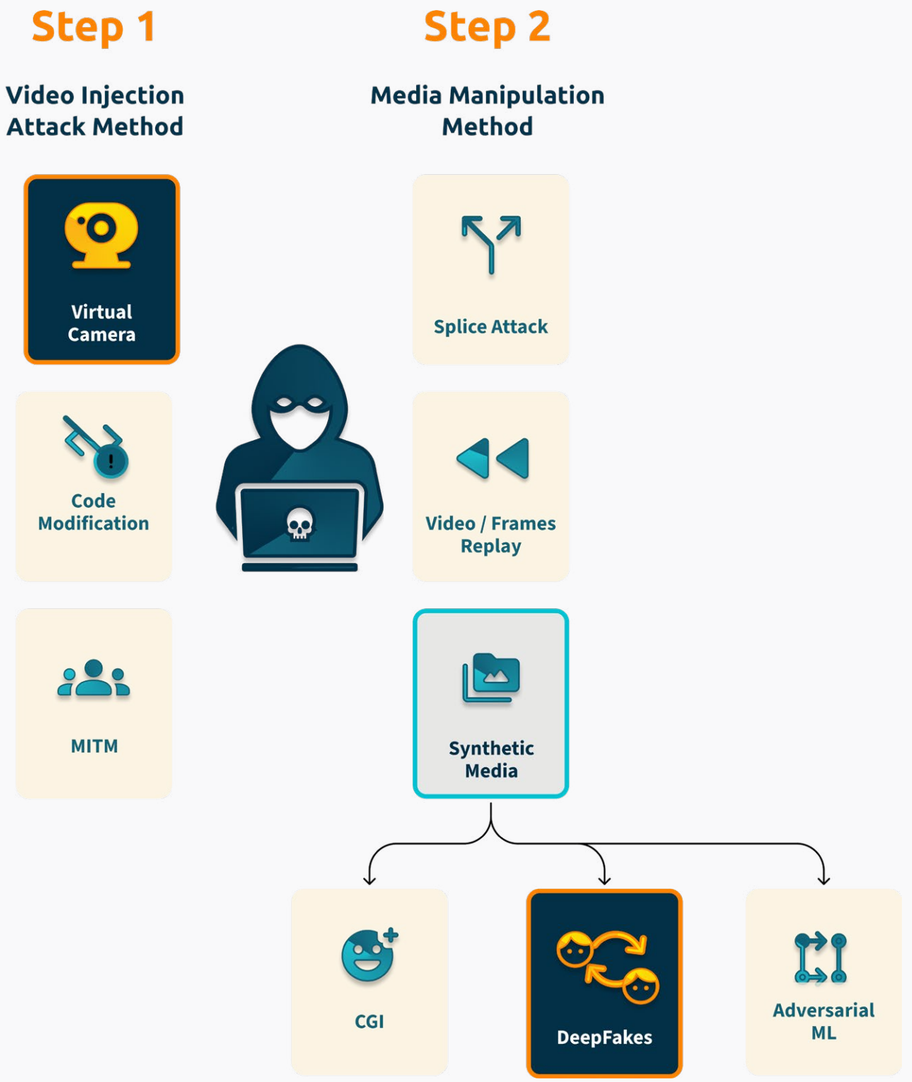
This media, which can be easily generated by off-the-shelf video face-swapping engines, is harnessed by feeding the manipulated or synthetic output to a virtual camera, which many biometric vendors are set up to detect. However, malicious actors in 2023 exploited a loophole by using cyber tools such as emulators to conceal the existence of virtual cameras, making it harder for biometric solution providers to detect.

This irresistible combination has made face swaps and emulators the preferred tools for attackers seeking to perpetrate identity fraud. **Face swaps are now firmly established as the deepfake of choice among persistent threat actors.** Tools such as DeepFaceLive, Swapface, Deepswap, and Swapstream.ai are easily accessible to anyone, including attackers. The fact that most of these tools offer a free tier for users to experiment with makes it even easier for attackers to abuse them without spending any money. This further emphasizes the need for stronger security measures to protect remote verification systems against such attacks.

While presentation and digital injection attacks may have different levels of impact, they can pose a significant threat when combined with traditional cyber attack tools like metadata manipulation.

Persistent and sophisticated threat actors are relentless in their efforts over extended periods, suggesting that they occasionally succeed on other platforms.

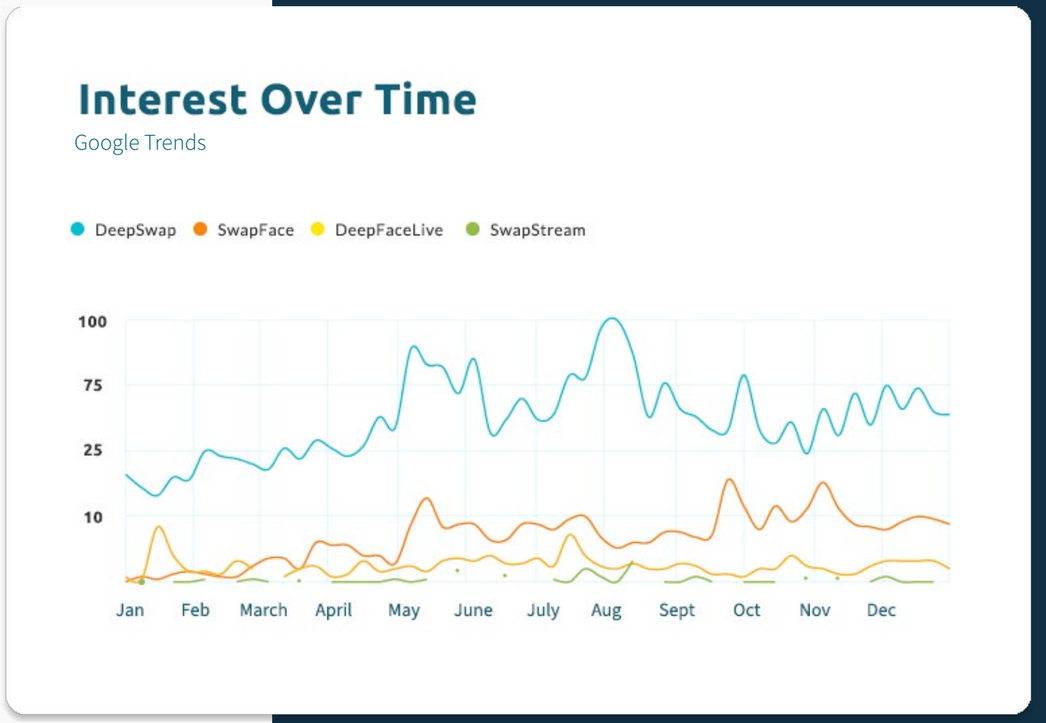
Our analysts continue to track over **110 different face swap tools and repositories.**



Most Popular Tools

The most common face swap tools being used offensively are currently **SwapFace, followed by DeepFaceLive and Swapstream**.

Threat tools available on the market in order of most searched on Google by country/ region - Source: iProov Threat Intelligence Team



Most searched face swap tools 2023

02 Significant Increase in Mobile Web Injections

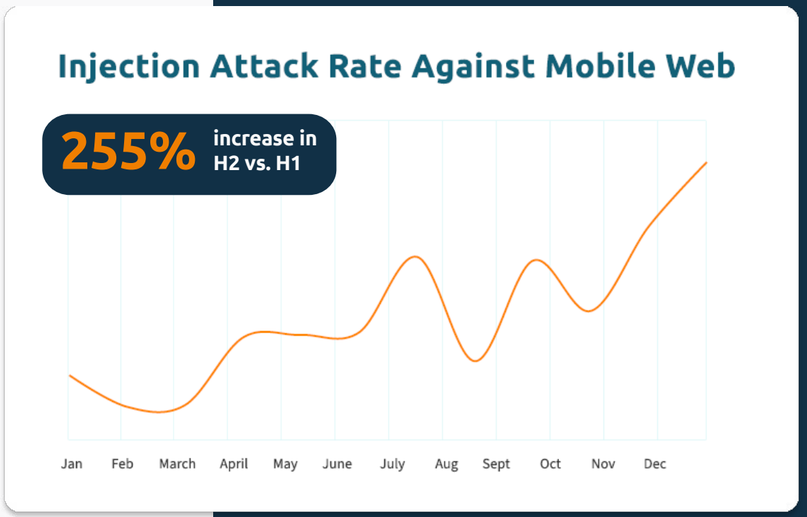
Threat actors continue to launch digital injection attacks across different platforms, using emulators and metadata spoofing. This trend emerged during 2022 and continued to increase throughout 2023.

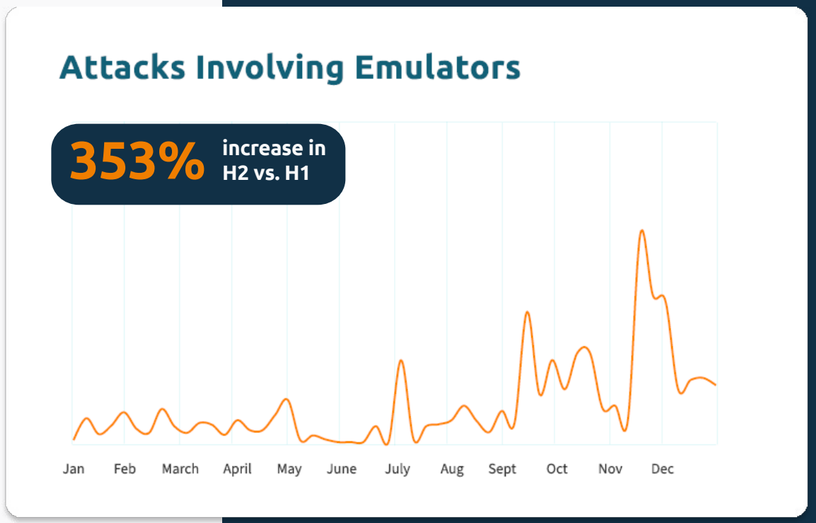
An emulator is a software tool used to mimic a user's device, such as a mobile phone. Over the last 24 months, the iSOC witnessed a considerable rise in threat actors using emulators to attack mobile web platforms as well as native Android and iOS.

Injection attacks are rapidly evolving with significant new threats to mobile platforms.

We saw an increase of **255%** in injection attacks against mobile web H2 over H1 2023.

Use of emulators continued to grow rapidly with an increase of **353%** H2 over H1 2023.

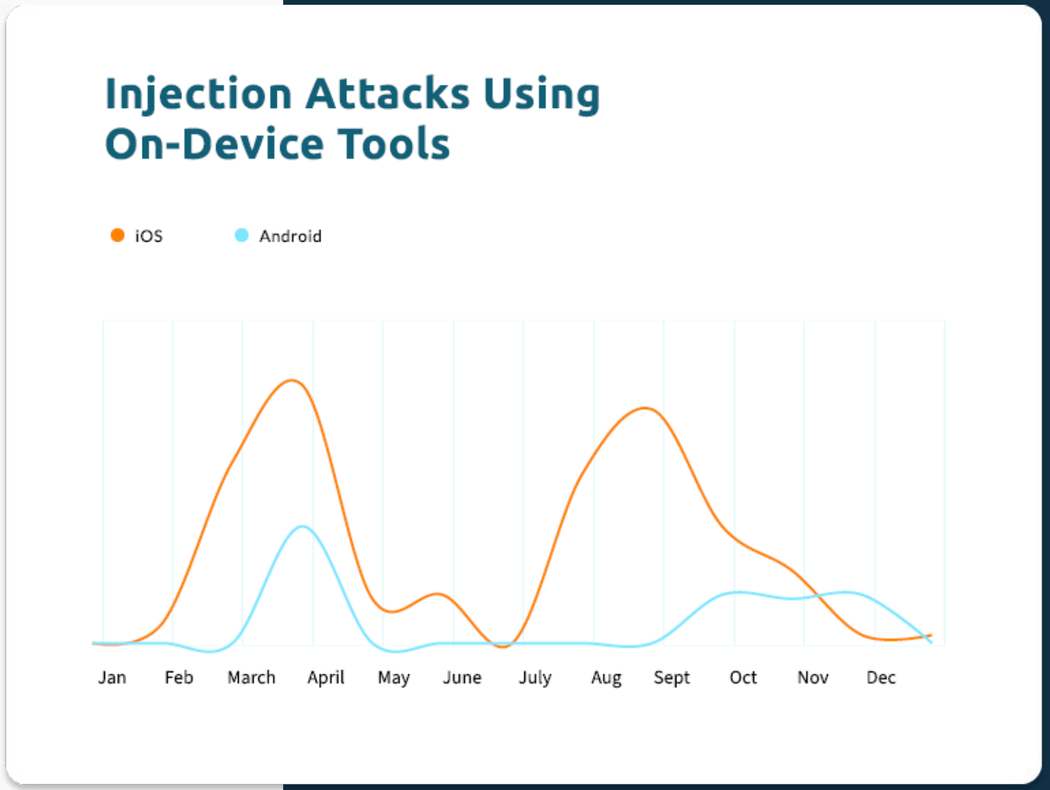




Rise in Native Virtual Cameras

2023 saw a significant advance in native on-device attack tools, including native virtual cameras. We observed many probing attacks from these tools throughout the year, a substantial proportion on iOS.

There is a continued proliferation of conventional virtual webcams, with over 80 being tracked for offensive use.



03 The Persistence of Threat Actor Behavior

We observed a significant increase in the persistence of threat actors.

Among the top threat actors, attack sequences typically lasted over 60 days, with multiple threat actors engaged in attack sequences with a duration in excess of six months.

The distribution of attacks typically revealed periods of high activity interspersed with significant periods of subdued activity.

Threat Actor Personas

Our Threat Intelligence Team follows and investigates the most prolific bad actors, learning from and adapting to novel threats before they evolve into serious threats.

In the field of cybersecurity, threat actors can be classified into three main categories: opportunistic, commercial, and nation-state.

Opportunistic actors are mainly driven by financial gain and typically use basic tools to carry out their attacks. They are often located in countries with more relaxed cybercrime laws. Their tactics often involve phishing, social engineering, and identity theft to obtain as much data as possible.

Commercial actors, on the other hand, are well-funded and highly advanced. They are typically based in countries with strong cybersecurity capabilities and utilize sophisticated techniques such as deepfakes and metadata spoofing. Their attacks are usually more targeted and aimed at specific individuals or organizations.

Nation-state actors are the most powerful and resourceful of the three categories. They are state-sponsored and have vast resources at their disposal. They are highly skilled and use advanced techniques to gain access to sensitive information. Their primary targets include critical infrastructure, government agencies, and military organizations.

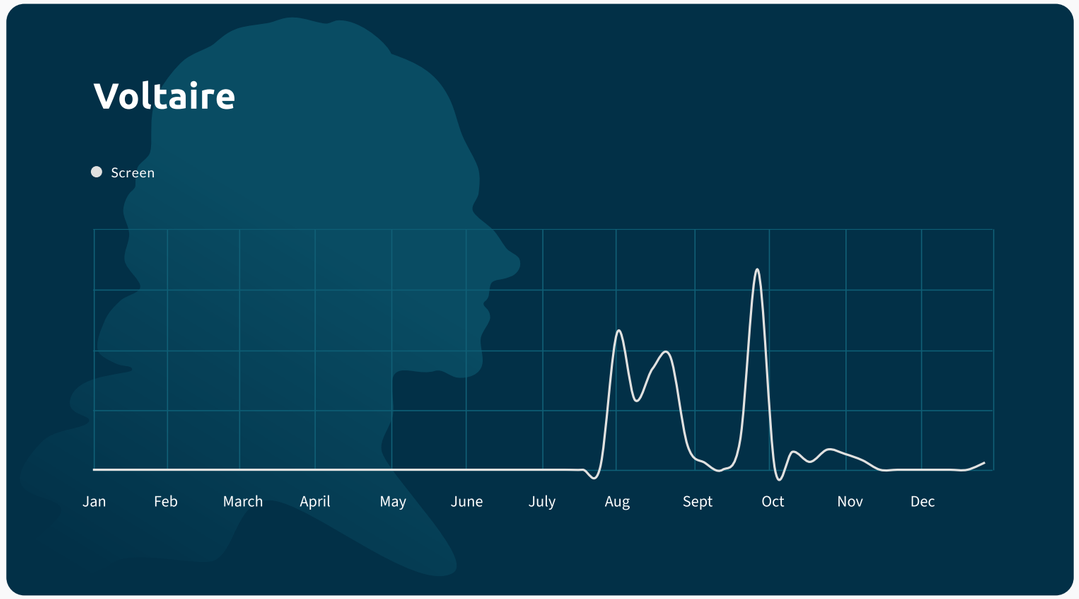
Most Wanted List Through 2023: Top Four Threat Actors

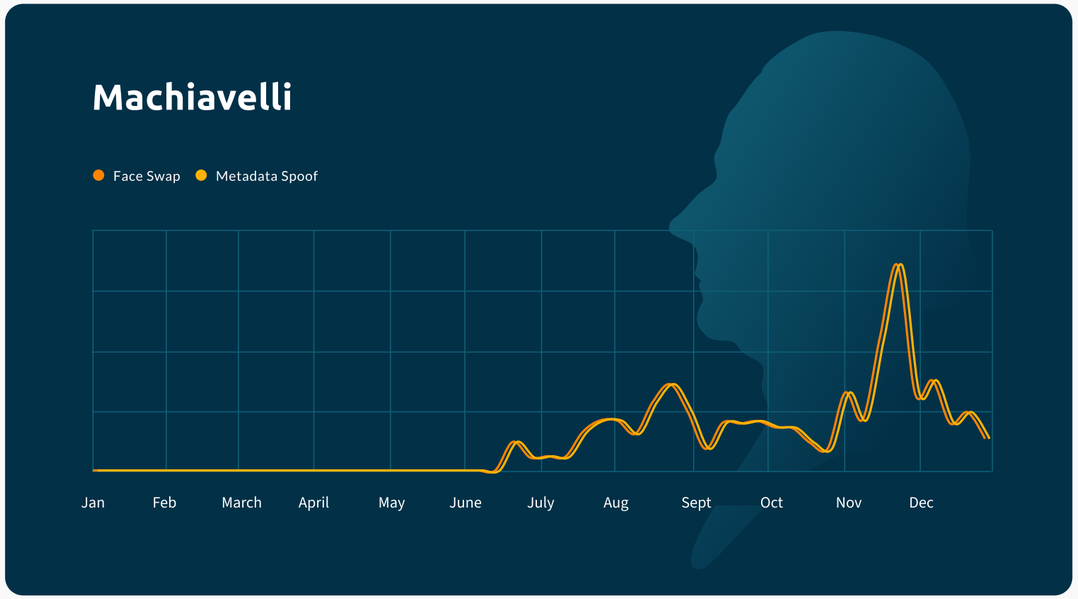
It is crucial for us to uncover the tools and techniques bad actors use to launch attacks on our biometric platform and gather pertinent information, including their methods and motives. To achieve this, our Threat Intelligence Team scrutinizes prolific threat actor personas, specifically evaluating the sophistication of their methodology, the effort, and the frequency of their attacks. Such analysis yields invaluable intelligence and enables us to continually improve our biometric platform’s security.

Our threat actor naming is internally devised and in no way aligned with or intended to represent public brands, personas, or historically renowned philosophers.

1. **Voltaire** is a highly skilled individual who employs cost-effective and reproducible methods to carry out attacks. They have been known to capture a person's facial image and overlay it digitally onto a body standing in front of a patterned backdrop. The resulting image is then displayed on a separate large-screen device, using advanced techniques to simulate life-like movements. Once this is accomplished, the image is recorded by a camera device. It is important to note that Voltaire's primary motivation is financial gain, and he frequently targets banks in Latin America.
Sophistication: Low, Effort: Low, Frequency: High

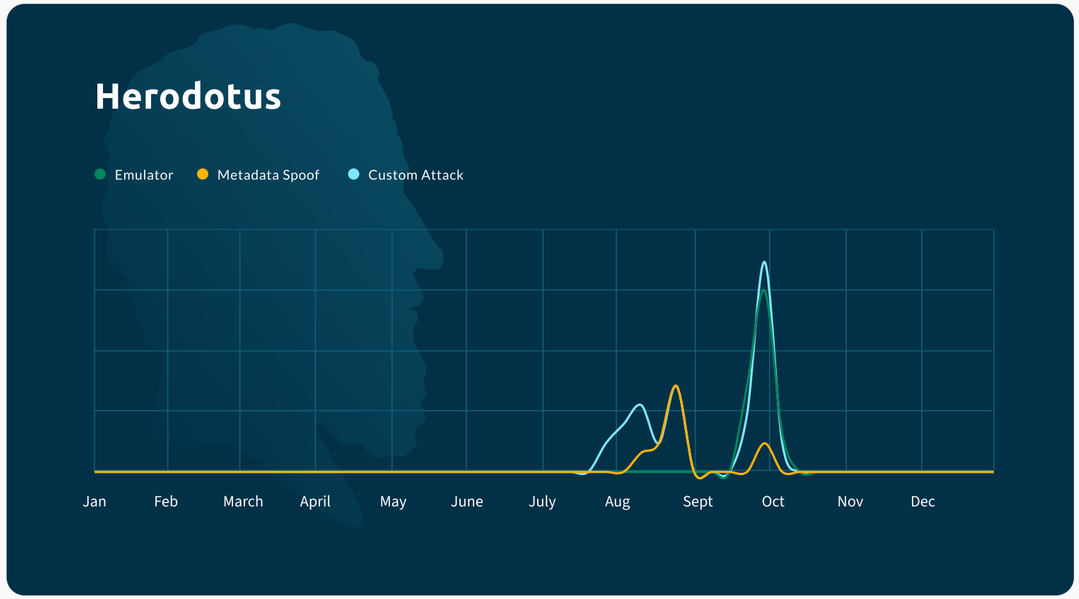
2. **Machiavelli** is an extremely perseverant actor, engaging in frequent and sizable transactions and attempting over 100 attacks daily. Their attack methods include the use of low-quality face swaps and the manipulation of data identifiers. While they tend to stick to familiar techniques, they have sometimes utilized various video filter applications. It is believed that Machiavelli is a threat actor who speaks Russian, and their primary targets are financial service websites with the goal of securing financial gains.
Sophistication: High, Effort: High, Frequency: Medium

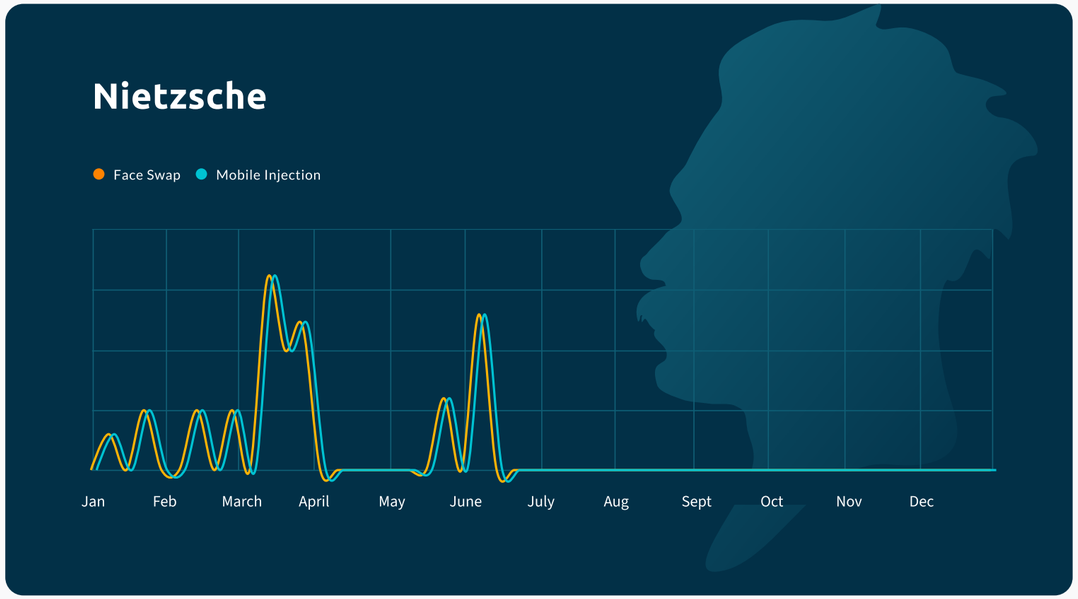




3. **Herodotus** Utilized techniques of a higher level of sophistication Despite the higher sophistication. This actor was identified via iSOC system monitoring and threat intelligence. The individual attempted to bypass our advanced system using a customized technique without success. It was discovered that their primary fraud type was the attempted exploitation of the gig economy applications. After repeated failures, the actor has not been seen to be active for 30 days. Herodotus claims to be based in South Asia, but all indications suggest that Herodotus is based in the United States.
Sophistication: High, Effort: Medium, Frequency: Low

4. There appears to be a recurring pattern of attacks by an individual or group known as **Nietzsche**, with intervals of several weeks or even months between each attack occurrence. This entity has been observed to employ sophisticated face swapping techniques via Android device injection. Additionally, it seems that this threat actor frequently employs the same fabricated identity on multiple occasions without any alterations. From the information gathered, it appears that this individual or group is based in the US and is targeting US-based cryptocurrency wallets and account providers with a motive for financial gain.
Sophistication: High, Effort: High, Frequency: Low



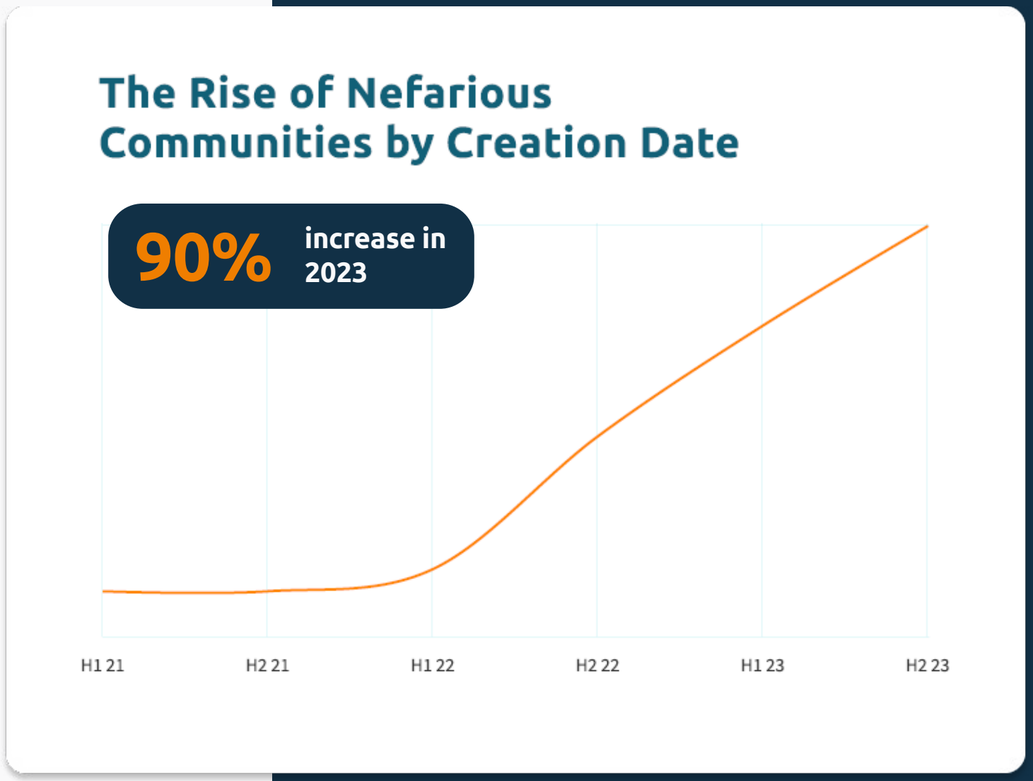


04 The Rise of Nefarious Communities

There has been a significant increase in the number of threat groups engaged in the exchange of information related to attacks against biometric and video identification systems.

Of the groups identified by our analysts, 47% had been created in 2023, implying an increase in the number of groups of over 90% through 2023.

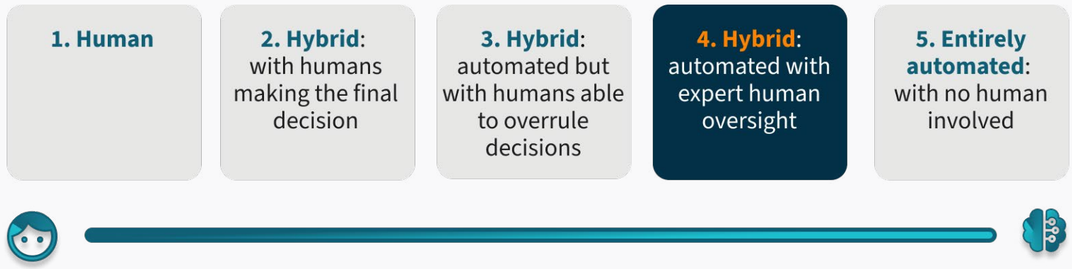
The median value of members within the groups was greater than 1000.



Impact of Cyber Attack Tools on Remote Identity Verification: Manual, Hybrid, or Automated

Organizations leveraging biometric verification technology are in a stronger position to detect and defend against these attacks than those relying solely on manual operation. As new technologies are introduced, there are numerous ways to implement them, leading to a myriad of integrations.

Human, hybrid, and automated verification processes can be divided into five categories:



It must be noted that all remote verification methods are vulnerable to synthetic media tools such as face swaps, whether entirely human-operated video calls, hybrid processes with facial biometric checks and human oversight, or fully automated.

Doubling up is not the answer. A low-cost solution with human oversight does not equal high performance. Research studies have shown that humans are inconsistent at identifying deepfakes and offer a weak line of defense against generative AI and other forms of advanced synthetic imagery attacks. Therefore, systems that rely on humans, such as methods 1, 2 and 3 above, to make the ultimate decision, in some cases overruling the technology, are not efficient or effective methods.

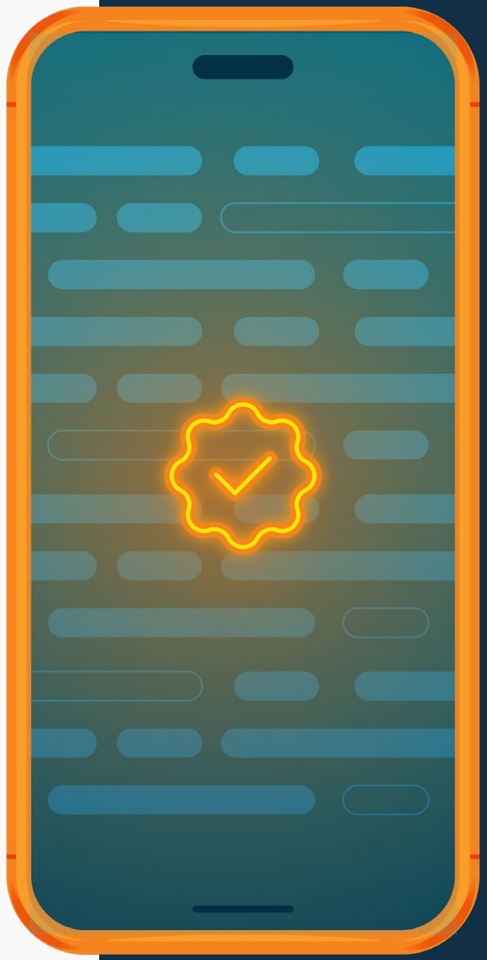
Like all cybersecurity applications, biometric technologies must constantly evolve to stay ahead of the evergreen threat of novel attacks. Therefore, it is important to understand that not all biometric face verification technologies provide equal levels of threat mitigation. Consequently, they provide varying levels of identity assurance. This can begin with remote onboarding, when a user first asserts their identity by capturing an image of a government-issued identity document and their face. Returning users can authenticate with their face biometric, which is compared to the biometric template captured at onboarding, known as facial biometrics or liveness. This can be triggered at various inflection points across the user lifecycle based on time, activity, risk threshold changes, or any other factors determined by the organization.

To address new and emerging threats, the industry must establish more rigorous certification requirements. Programs like Presentation Attack Detection (PAD) can detect attacks such as masks and paper printouts, but they do not prevent all types of biometric system attacks.

With the evidenced increase of tools, such as emulators and virtual cameras used in conjunction with imagery manipulation - even basic presentation attacks that would not have typically bypassed single-frame liveness solutions - could potentially bypass these systems. This, along with the growing threat generative AI alone poses, creates a large attack surface in which threat actors can freely operate.

It is vital organizations ensure that vendor deployments have been vigorously tested by externally accredited penetration testing agencies or a government's own Red Team.

The Center for Identification Technology, CITeR is developing a new standard that aligns with the latest attack methodologies, such as digital injection attacks and media manipulation like deepfakes. However, the industry has no standards to certify a solution's ability to detect and defend against digital injection attacks or metadata manipulation. Leaving a two-fold vacuum that threat actors are eager to fill.

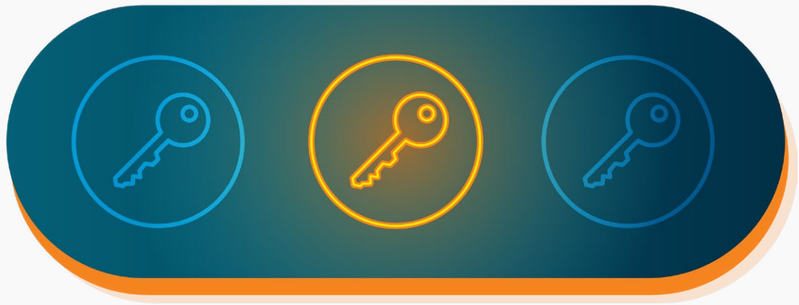


Recommendations

The time it takes for an exploit to be detected, processes to be adapted, and security to be redeployed can be long, posing great risk. This is particularly true when relying on human decisions, interventions or technologies that do not have an iSOC with active threat mitigation. The speed of retraining to prevent successful attacks as they occur in this instance is not possible. Organizations must stay alert and agile to keep their systems secure and resilient against ever-changing threats.

Key Takeaways

- Any system operating remote person-to-person verification is vulnerable to attack, even if it has not deployed any technology and relies on human judgment.
- Human operator-led systems can no longer consistently and correctly detect synthetic media such as deepfakes. Super-recognizers and forensic examiners are subject to increasing cost challenges as intensive training and more skilled examiners require higher salaries. This makes person-to-person remote video call identification obsolete.
- In order to detect synthetic media created using generative AI, verification technologies that leverage AI are essential.
- To benefit from improved efficiencies, organizations need to understand biometric verification technologies and their differentiators to obtain the appropriate level of assurance required for the use case.
- Single-frame liveness provides little protection for organizations. On-premise solutions are vulnerable to reverse engineering, are slow to deploy defenses, and are costly. In addition, on-premise is resource intensive, requiring an organization's in-house Red Team and hardware to manage the infrastructure.
- Multi-frame liveness with cloud deployment and a Security Operations Centre (SOC) is vital to detecting and preventing generative AI, deepfakes, face swaps, and metadata manipulation techniques.
- Don't rely on PAD certification alone. Perform your own testing or employ Red Team testers such as Outflank, or ask your vendor which external Red Team testing has been conducted on their platform to test their resilience against digital injection attacks.
- While high-level security may be required for certain use cases, it must not be to the detriment of the user.
- To ensure inclusion and a positive user experience, technologies should be deployed that offer high levels of assurance while actively mitigating bias.



Conclusion

Active threat intelligence plays a crucial role in regularly delivering security updates without delay when required rather than on a set schedule. Whether it be weekly or monthly, the data is evident, threat actors and groups do not adhere to these timelines. In fact, on-premise biometric solutions deployed just weeks ago risk becoming obsolete the moment a threat actor or vector is successful. This victory will be quickly shared via their communities, and within hours, a system could fall victim to multiple well-targeted attacks.

Organizations that fail to have systems in place to check for imagery manipulation or metadata spoofing are at high risk of being targeted by fraudsters, whether they use facial biometrics or not.

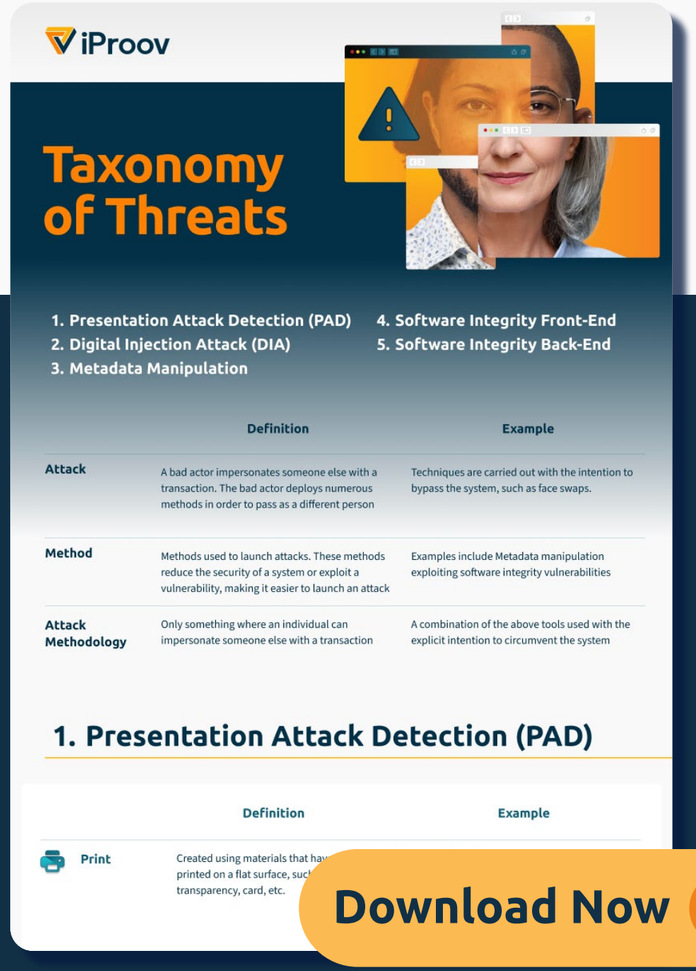
Manual remote identity verification has been proven ineffective. With the emergence of advanced technologies such as generative AI and face swaps, identity fraud has become a significant concern for organizations.

Threat actors are exploiting processes that rely on lower-cost technology as well as those that leverage human intervention. Current tools are outpacing defenses in both availability and sophistication. As a result, these new threat vectors are evading many current remote identity verification techniques faster than organizations can detect or adapt their security measures.

The need for consistent security and flexible, inclusive verification should not be a trade-off. Passive challenge-response, multi-frame technology that offers the highest level of assurance without requiring users to carry out random actions is available.

A proactive approach leveraging science is needed to identify, mitigate, and prevent potential threats before they become serious. As the threat landscape continues to evolve, organizations need to have insight into the tools available and the emerging techniques threat actors are developing. Hybrid automated identity verification processes that invest in human experts for real-time oversight deliver reliable, consistent results, as opposed to hybrid models that enable humans to override decisions. Furthermore, matching the use case with the appropriate level of assurance, risk appetite, and the threat landscape is vital. Without this balance, organizations will not reap the benefits of biometric technology.

Taking a proactive stance backed by an interdisciplinary team of biometric security experts, organizations can start to draw a picture of the threat landscape to stay a step ahead of threat actors, minimizing the risk of exploitation of both present and future remote identity verification transactions.



Find Out About the Different Attack Types With Our

Taxonomy of Threats

Download Now