

Deepfake-Enabled Infiltration: The Threat of Synthetic Identities in Corporate Environments

Joseph Lozada
College of Business, Innovation,
Leadership, and Technology
Marymount University
Arlington, VA, USA
josephilozada@gmail.com
0009-0008-4903-5710

Abstract—This paper explores the evolving threat of deepfakes in the context of insider threats, particularly how advanced persistent threats (APTs) are leveraging AI-generated audio and video to impersonate job applicants and gain access to sensitive systems. While deepfakes have legitimate applications in entertainment, education, and business, they are increasingly being weaponized for deception and cyber intrusion. The paper outlines recent incidents, assesses technical vulnerabilities, and evaluates current risk management frameworks such as NIST RMF. It concludes with policy and technology recommendations to enhance detection and prevention strategies, especially during remote hiring and onboarding processes.

Keywords—deepfake, insider threat, artificial intelligence, AI, machine learning, ML, cybersecurity

I. INTRODUCTION

January 2024, in the months leading up to the July general election, deepfake videos were posted to Facebook which falsely depicted sitting UK Prime Minister Rishi Sunak endorsing controversial policies [1]. The election held on July 4th ultimately resulted in the election of Keir Starmer [2], and it is unclear what role (if any) the deepfake videos played in the outcome, but it is clear that deepfakes have the potential to deceive, confuse, and impact the way people view and understand the information they consume. The US Government Accountability Office (GAO) defines a deepfake as a type of visual or audio media that has been generated using AI to appear real, but is in fact fake [3]. Deepfakes are typically created using deep learning techniques such as generative adversarial networks (GANs), which train two neural networks, a generator and a discriminator, in competition to produce highly realistic synthetic media [4]. These models are trained on large datasets of audio, video, or images to learn and replicate facial expressions, voice patterns, and movements with high accuracy. Facial manipulation can be accomplished in different ways including; modifying facial expressions, changing gender, altering age, or swapping identity altogether [5].

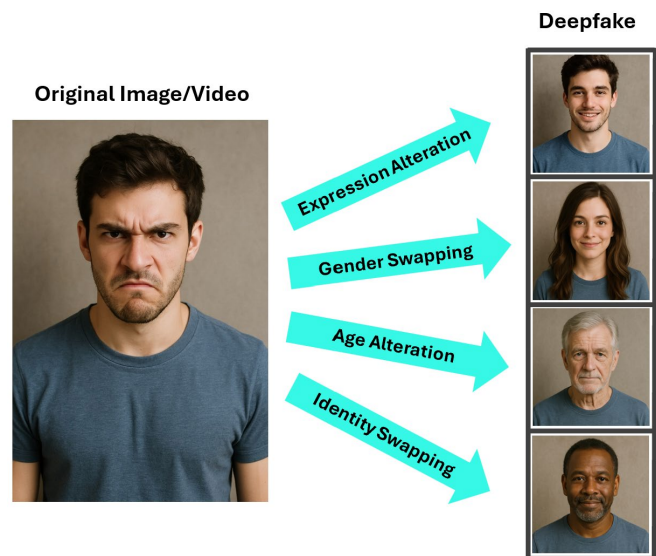


Fig. 1. Example of deepfake facial manipulations.

Although the term generally comes with negative connotations of deception and stolen identity, deepfake technology was not originally developed with malicious purposes in mind, and has many legitimate uses. For example, in the 2022 film *Top Gun: Maverick*, deepfake technology was used to replicate the voice of Val Kilmer which had been destroyed due to throat cancer [6]. In 2021 a UK startup created a demo of a movie scene with Robert De Niro's voice and lip movements dubbed in German [7]. The company behind the technology states that its mission is to elevate filmmaking, and allow for stories without borders [8]. These examples demonstrate how deepfake technology can be used to enhance storytelling, preserve performances, and localize content for global audiences. Rather than deceive, these applications aim to expand creative possibilities within the film and entertainment industry. Aside from this industry, deepfake technology has many other legitimate uses.

Deepfake technology has the potential to alter the educational landscape [9] through the creation of

personalized content, such as inserting students into training videos, or adjusting language output [10]. The technology could potentially be used to bring historical figures back to life through animations and avatars that may help students engage with a curriculum, much in the same manner as done with films. The technology has also been used experimentally in medical settings for grief counseling and post-traumatic stress disorder (PTSD) therapy by allowing patients to more vividly confront assault perpetrators [11], [12]. These early use cases demonstrate how deepfakes, when applied responsibly, can enhance learning, emotional healing, and engagement.

Deepfakes also have use cases in the business and marketing fields. Businesses can use the technology to create brand ambassadors and marketing campaigns, using the likeness of their CEO or of celebrities [13], and influencers are using deepfake technology to scale their operations by creating avatars that engage with fans from different countries and across different languages [14]. These applications allow brands and individuals to expand their global reach while maintaining a consistent and personalized presence. As a result, deepfakes are becoming a valuable tool for enhancing customer engagement and driving innovative marketing strategies.

The legitimate applications of deepfake technology are both diverse and impactful. From enhancing education and accessibility to supporting healthcare and advancing scientific research, deepfakes offer powerful tools when used ethically and transparently. However, despite these promising use cases, the same technology also raises serious concerns. In the wrong hands, deepfakes can be weaponized for deception and manipulation [15]. Nation-backed actors have reportedly begun leveraging deepfake tools to create synthetic identities, combining falsified video, voice, and documents, to impersonate job applicants during remote interviews and gain access to sensitive roles within U.S. technology firms [16]. These tactics present a new kind of insider threat, where malicious actors can embed themselves into critical organizations under entirely fabricated identities, bypassing traditional hiring safeguards and posing significant national security risks. Deepfakes also have the potential to disrupt social and political stability through the creation of disinformation using pre-trained models and tools easily accessible to users, [5] expanding the overall risk landscape. Although there are a wide variety of risks and challenges created by the growth of deepfake technology, this paper focuses on the internal threat businesses face when hiring remote workers and how bad actors can use deepfakes to create synthetic identities in an attempt to embed themselves into US companies, and the strategies organizations can use to identify these deceptions and limit the overall impact.

II. TECHNICAL CHALLENGES OF DEEPFAKES

Despite the impressive capabilities of deepfake technology, it still faces several technical challenges that limit its effectiveness in some applications. Model training requires large, high-quality datasets, and very often high computational

complexity [17]. A common issue can be the quality of the generated output. In many cases, deepfakes can suffer from low resolution and visible artifacts [4], especially when applied to footage with fast motion, complex lighting, or occlusions like glasses or facial hair [5]. Another hurdle is the “uncanny valley” effect, where deepfakes that are nearly, but not quite, realistic elicit discomfort from viewers [18]. Subtle inaccuracies in facial expressions can disrupt believability, making the content feel artificial. However, as deep learning models grow more sophisticated and datasets become more diverse and refined, these challenges are rapidly diminishing.



Fig. 2. Uncanny, poorly executed deepfake.

As research into generative AI accelerates, improvements in model architecture, training efficiency, and post-processing techniques are steadily overcoming these barriers. Advanced approaches are helping to produce deepfakes that are more emotionally expressive [19]. Additionally, the growing availability of large, diverse datasets is enabling models to better generalize across different facial features, environments, and movements. While technical imperfections once limited the practical use of deepfakes, continued innovation is making it increasingly difficult to distinguish synthetic media from real footage.

III. MALICIOUS USE OF DEEPFAKES

In September of 2024, a cyber attacker orchestrated a deepfake Zoom call impersonating Ukraine’s then-former Foreign Minister Dmytro Kuleba to communicate with U.S. Senator Ben Cardin, Chair of the Senate Foreign Relations Committee [20]. The attack appeared to have been designed to extract information, and the senator initially fell for the call, but then began suspecting the deception of the caller after they started asking pointed questions about an upcoming election [20]. Another incident with national security implications occurred in June of 2025 when a scammer used

deepfake technology to impersonate Secretary of State Marco Rubio and speak with foreign leaders and government officials, prompting the State Department to issue a memo to diplomats warning them of the activity [21]. Incidents like these underscore the growing sophistication of AI-driven impersonation tactics and the national security risks they pose. They show how deepfakes can be weaponized to manipulate high-level diplomatic communication.

Perhaps some of the fastest growth in the malicious use of deepfakes has been in the non-consensual replication of celebrity identities for scams or pornography. In January of 2024 in an infamous case of stolen identity, deepfake pornographic images of Taylor Swift were circulated online, prompting outrage from her fans and others [22]. In another case, of stolen identity, a deepfake video of TikTok influencer Molly Mae was circulated to fans prompting them to purchase a perfume from scammers [23]. Deepfake technology can be used to exploit the likeness of public figures for financial gain or to cause personal harm. As the technology becomes more accessible, the risk of such abuses targeting both celebrities and private individuals continues to grow.

Businesses have also been targets of deepfake scams and attacks. In 2024, a finance worker at a multinational firm was tricked into transferring \$25 million after attending a deepfake video call in which fraudsters impersonated the company's CFO and other staff members [24]. Although the worker was initially suspicious of the request when it came through email, they were convinced of the request's authenticity after the deepfake video call [24]. Today attacks can be highly coordinated and convincing, thereby overcoming even initial skepticism. As these tactics continue, advancing organizations must strengthen their verification protocols to protect against identity-based deception. The Deloitte Center for Financial Services projects that deepfake-related fraud losses in the U.S. could escalate to \$40 billion by 2027, up from roughly \$12.3 billion in 2023 [25], and research shows that better coordination, training, and communication is needed to effectively combat this growing threat [26]. This dramatic rise reflects not only the increasing sophistication of deepfake technology but also the vulnerability of current fraud prevention systems. Without swift intervention, businesses and institutions may face escalating financial and reputational risks from these attacks.

IV. INSIDER THREATS

In March of 2025, Rippling, an HR and payroll software provider, alleged that 2 years earlier, a rival company, Deel, hired one of Rippling's employees to covertly gather confidential business information including sales pipelines, pricing strategies, and customer data, in an operation reportedly directed by Deel's CEO Alex Bouaziz and CFO Philippe Bouaziz [27]. If true, the allegation serves as a shocking example of the type of insider threats high-stakes technology companies can face. An insider threat can be defined as the threat that an insider, someone granted

legitimate access to a system, may use that access for malicious purposes [28].

Although insider threats may be less common than other types of cybersecurity risks, they are real threats that businesses and governments both have to consider. Especially organizations with high-value business information, such as customer data, or intellectual property. For example, in March of 2024, former Google software engineer, Linwei Ding, was indicted for allegedly stealing more than 500 confidential files, including intellectual property related to Google's AI chip architecture [29]. He allegedly stole this information over a 1 ½ year period starting in May of 2022, and funneled the information to a Chinese start-up he had been hired on to [29]. This case highlights the serious damage a single employee can cause when granted access to sensitive systems and data. It also demonstrates the importance of implementing strong insider threat detection protocols, especially in organizations working with emerging technologies.

In some cases insider threats can bring a high level of sophistication to their operations by receiving funding, technology, and support from organizations backed by governments or other powerful entities. In such cases those backing organizations are referred to as Advanced Persistent Threats (APT). The National Institute of Standards and Technology (NIST) describes an Advanced Persistent Threat (APT) as an adversary who creates a higher level of threat to an organization due to its access to more sophisticated tools, technology, and techniques [30]. What if applicants backed by APTs can leverage deepfake technology to mask their true identities and gain entrance to US systems? By manipulating video, audio, and identity documents, they could convincingly impersonate qualified candidates during remote interviews, bypassing traditional background checks. Once embedded in a company, they might gain access to proprietary code, customer data, or network infrastructure. This type of synthetic infiltration could serve to undermine national security and corporate integrity from within.

In July of 2024, cybersecurity firm KnowBe4 discovered it had unknowingly hired a North Korean operative who used stolen identity documents and deepfake technology to pose as a legitimate job applicant and attempt to infiltrate the company's systems [31], and the data show this is not an isolated incident, but rather North Korea is making a coordinated effort to infiltrate US companies who have valuable information to steal [32]. In February of 2025, North Korean hackers posing as job applicants used deepfake video and forged documents during interviews to gain trust and trick targets into deploying macOS malware known as FlexibleFerret [33]. These incidents reflect a growing pattern of state-sponsored cyber actors leveraging deepfake technology not just for misinformation, but for direct access to internal systems. They accomplish these deceptions using deepfake technology during video calls and remote interviews to take on a false identity [34], and by using laptop farms

(typically coordinated through a third party) to spoof their location and make it appear as if they're in the US [16]. They use online job boards and social media such as LinkedIn to create personas that appear to legitimize their fake identities [35]. Once hired, these operatives can gain privileged access to sensitive data, intellectual property, or critical infrastructure. In some cases, they may also attempt to install malware or backdoors that facilitate longer-term espionage efforts [36]. Aside from planting malware, or stealing data, another rationale of APT-sponsored deepfake insider threats is to earn money for sanctioned regimes, who might otherwise struggle to fund their operations, by tapping into the high salaries offered by US tech companies and funneling those earnings back to their countries [16].

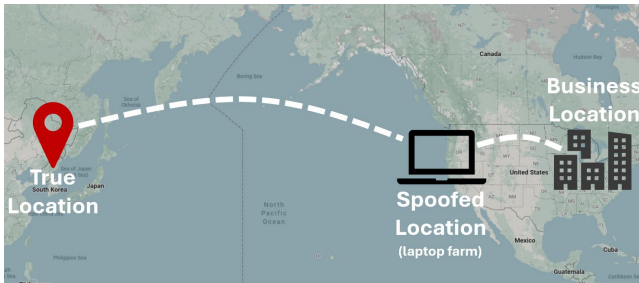


Fig. 3. Example of how a malicious actor may spoof their location.

Unfortunately, due to the growth of remote work in the US and other countries, there are more opportunities for these types of tactics to succeed. Research shows that despite fluctuations in the availability of fully-remote roles in the US, the job market has stabilized since 2023 with approximately 13% of all new job postings being for fully-remote roles [37]. In certain local markets, however, the percentage of newly created remote roles can be much higher, such as 26% in Wyoming, or 24% in Idaho [38].

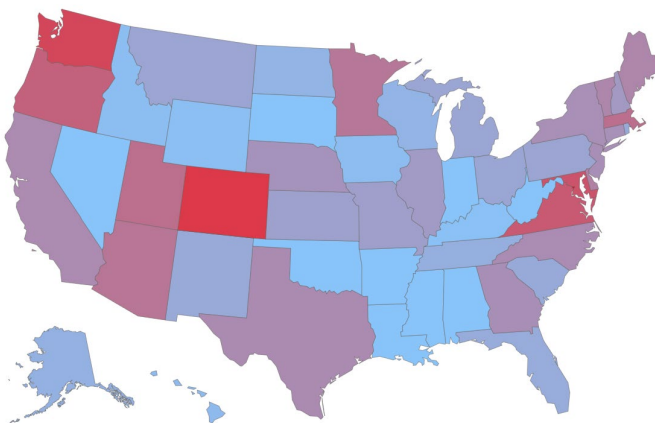


Fig. 4. Risk map of US states, based on estimated percentage of fully-remote jobs in Information, Finance, or Government [39], [40].

This sustained prevalence of remote roles means that in some cases identity verification and applicant screening must

now occur entirely online, often without in-person safeguards. As a result, it becomes significantly easier for malicious actors to fabricate identities and bypass traditional vetting procedures. Without stronger protections, the rise of remote work may continue to expand the attack surface for insider threats leveraging deepfake technology. How do the current risk management practices and guidelines address these threats? Are there any other strategies organizations can use to reduce the risk of deepfake insider threats?

V. CURRENT RISK MANAGEMENT FRAMEWORK

The NIST Risk Management Framework (RMF) provides a structure organizations can follow to position themselves to respond effectively to risks. In order, the high-level steps are: Prepare, Categorize, Select, Implement, Assess, Authorize, and Monitor, with much of the work occurring upfront in the Prepare step [41]. Important preparatory tasks (Section 3.1) include defining risk management roles and responsibilities (Task P-1), performing a risk assessment (Task P-3), and developing a continuous monitoring strategy (Task P-7) [41]. Although there are other tasks listed in the Prepare step, these three are of particular importance because they create a system of accountability and awareness that facilitates an organization identifying and responding to breaches if they occur. Roles noted by NIST that are of particular importance to this scenario include:

- **Authorizing Official:** Senior official with the authority to determine acceptable risks, and approving plans, or changes to systems [41].
- **Security Architect:** An individual responsible for developing the overall security strategy and ensuring that system security and data privacy and protection are properly configured [41].
- **Security Engineer:** A technical resource who builds the security controls designed by the security architect [41].
- **Information Owner:** An individual who has authority over a specific system or specific information, and access to that information and governs how it's generated or collected [41].

In simple terms, the authorizing official takes ownership of the big picture and overall risk posture, the architect designs the security architecture, the engineer builds the controls designed by the architect, and the information owners act as the gatekeeper to the various systems they control. While there is no specific role identified or mention of assigning responsibility for candidate screening and onboarding in the RMF, NIST SP 800-53 is referenced throughout the RMF as containing more detailed information about establishing system controls. SP 800-53 contains a few important tips for establishing effective personnel security including:

- Assign risk a designation to all organizational positions by assessing their duties and responsibilities using the

Position Designation System (PDS) and determining the degree of potential damage that role may cause [42].

- Screen individuals prior to granting access to systems. The screen procedures include guidelines for special circumstances such as citizenship requirements, or classified information, and a related control, PM-12, which details guidelines for creating an insider threat program [42].
- Processes for personnel termination and/or transfer [42].

An insider threat program, as noted by PM-12, is required under “Executive Order 13587 and the National Insider Threat Policy” when organizations handle classified information, but can also be helpful toward organizations improving security in non-classified situations as well [42]. Notable related controls included in PM-12 include:

- AC-6: Least Privilege [42]
- AT-2: Literacy Training and Awareness [42]
- AU-6,7,10,12: Guidelines for performing audits and reporting results [42]
- CA-7, SI-4: Continuous Monitoring, and System Monitoring [42]
- IR-4: Incident Handling [42]
- PS-4,5,7,8: Guidelines for personnel screening (previously discussed) [42]
- SC-7, 38: Boundary Protection and Operations Security: Setting up firewalls, routers, encryptions, or other systems to control how connections are made to external systems, and following an OPSEC approach for analyzing risk and limiting information sharing [42].

The NIST guidelines in the RMF and SP 800-53 explain the importance of establishing employee screening, monitoring, system segmentation, and response plans, but do not provide guidance for actually identifying insider threats before they are onboarded into an organization. If one wishes to avoid insider threats altogether, they must look beyond the NIST RMF and correlated publications, such as the Office of the Director of National Intelligence which runs the National Insider Threat Task Force (NITTF), whose mission is to “develop a government-wide insider threat program for deterring, detecting, and mitigating insider threats” [43]. Through work done at NITTF the President issued a memo called the National Insider Threat Policy and the Minimum Standards in 2012 [43], and a newer guide in 2014 called the Guide to Accompany the National Insider Threat Policy and Minimum Standards [44]. This guide provides recommendations for employee screening, such as background checks, and security clearance investigations, but also fails to go beyond this nor does it speak to the recent use of deepfake technology [44].

Additionally, the Cybersecurity Maturity Model Certification, created by the Department of Defense, requires

the implementation of insider threat awareness training for certification at levels 2 and 3, but also does not address the enhanced threat created by deepfakes [45]. Carnegie Mellon University’s guide on insider threats recommends the following actions at the pre-onboarding stage in addition to background checks recommended by the NIST RMF: It advises organizations to consider behavioral indicators of risk, such as prior disruptive conduct or adversarial attitudes, during interviews and through reference checks [46]. While the guide does not advocate for routine monitoring of social media during hiring, it notes that social media activity may reveal potential threats and should be evaluated in compliance with legal and ethical standards [46].

Beyond risk management guidelines, much research is being done to create systems capable of identifying deepfake media, although with varying success. For example, in 2019 Facebook (Meta) initiated the Deepfake Detection Challenge, hosted through Kaggle, wherein participants competed to build models that could detect deepfake images [47]. The best performing model by Selim Seferbekov achieved an average precision of 65.18% on unseen data [47], meaning that the percentage of the model’s positive predictions (deepfake) were correct 65% percent of the time, noticeably better than random guessing (50%), but still not great. Additionally, NIST created the Open Media Forensics Challenge (OpenMFC). The purpose of the OpenMFC is to develop and benchmark tools for distinguishing manipulated or AI-generated media from authentic content [48]. In the 2020-2021 challenge, the top performing model achieved an accuracy of approximately 69%, indicating that there was still work to be done [49].

VI. PROPOSED ENHANCED RISK MANAGEMENT FRAMEWORK

In researching the various risk management guidelines and publications on insider threats, it is clear that current publications have not caught up to the latest evolution of these threats. The following are not fully explained or are missing from the existing NIST RMF framework and related publications:

- Within the NIST RMF Prepare stage, no role is noted with the responsibility of determining onboarding procedures and acting as a liaison with an organization’s human resources (HR) team. It is important for this responsibility to be assigned to someone, since an HR team typically will not have expertise in the specific technical risks that should be considered when screening candidates.
- Even when recommendations exist to perform pre-screening activities, no specific guidelines exist for identifying insider threats prior to onboarding them into an organization.

Given the highly sophisticated nature of these deceptions, technology should also be part of the solution to identify deepfakes at the early stages of recruiting and onboarding. Enhanced strategies to address this risk should include the following:

- **Limit Fully-Remote Employment:** Although simple in concept, a move away from fully-remote employment can be an effective way of addressing the threat of deepfake-enabled infiltration. Deepfakes only work in a virtual environment where physical presence is not a requirement of employment.
- **Incorporate Insider Threat Software:** Several technology companies sell platforms for analyzing data related to insider threats such as Splunk's user behavior analytics (UBA) platform [50], and Gurucul's insider threat management software [51].
- **Check Social Media:** See if there's a legitimate presence in the location the candidate claims to be.
- **Record and Analyze Remote Interviews:** Use AI to check recorded candidate interviews. Research shows that multi-modal deep learning techniques perform better at spotting deepfakes [52]. If remote interviews must be conducted, models trained for detecting deepfakes can be used to analyze the video and audio recordings to verify authenticity. Even if the model's accuracy is limited, it will help paint an overall picture of the candidate.

While these proposed solutions offer practical steps to reduce deepfake-enabled insider threats, their implementation is not without challenges. Moving away from fully-remote roles may face resistance from both employees and recruiters, especially in industries where flexible work has become a competitive advantage. Deploying insider threat software requires careful calibration to avoid false positives, protect employee privacy, and comply with legal and regulatory standards around monitoring. Similarly, using AI to analyze recorded interviews introduces ethical concerns, including potential bias in detection models, lack of transparency in decision-making, and the risk of unjustly flagging candidates based on model errors. Organizations must therefore establish clear governance, transparency protocols, and human oversight to ensure these technologies are applied responsibly and fairly.

A phased adoption timeline is recommended to balance urgency with feasibility. In the first 3 months, organizations can conduct a risk assessment to identify high-value roles and assets most vulnerable to deepfake-enabled insider threats, followed by policy updates to reflect enhanced security. During this time, procurement teams can evaluate available insider threat detection tools and develop governance structures for their ethical use. In months 3-9 those systems and policy changes should be implemented. Organizations can begin piloting their use with limited groups if the planned changes have the potential to be disruptive. Over the following 9-12 months, these measures can be scaled across all critical business units, with staff training, further testing, and continuous monitoring programs in place. Full operationalization, including continuous evaluation and governance reviews, should be achieved within 12-18 months.

VII. CONCLUSIONS

As deepfake technology continues to evolve, so too must our strategies for detecting and mitigating the risks it introduces, especially in the context of insider threats. While legitimate uses of deepfakes span entertainment, education, and business, the same capabilities have been exploited by adversarial actors, including nation-state-backed APTs, to infiltrate organizations, steal intellectual property, and deploy malware. Existing risk management frameworks like NIST RMF and CMMC provide a foundation for addressing insider threats, but they must be enhanced to meet the challenges posed by AI-generated deception. Organizations must take a proactive stance by incorporating advanced detection tools, strengthening identity verification during hiring, and adapting policies to reflect the growing reality of remote work and digital impersonation. Failure to act now may leave critical systems and intellectual assets vulnerable to increasingly undetectable threats.

REFERENCES

- [1] B. Quinn and B. Q. P. Correspondent, "Slew of deepfake video adverts of Sunak on Facebook raises alarm over AI risk to election," *The Guardian*, Jan. 12, 2024. Accessed: July 23, 2025. [Online]. Available: <https://www.theguardian.com/technology/2024/jan/12/deepfake-video-adverts-sunak-facebook-alarm-ai-risk-election>
- [2] A. Bhatiya, "What happened in the 2024 UK general election?," Economics Observatory. Accessed: July 23, 2025. [Online]. Available: <https://www.economicsobservatory.com/what-happened-in-the-2024-uk-general-election>
- [3] "Science & Tech Spotlight: Deepfakes," Government Accountability Office. Accessed: July 16, 2025. [Online]. Available: <https://www.gao.gov/assets/gao-20-379sp.pdf>
- [4] B. Fan et al., "Generating Higher-Quality Anti-Forensics DeepFakes with Adversarial Sharpening Mask," *ACM Trans. Multimed. Comput. Commun. Appl.*, Apr. 2025, doi: 10.1145/3729233.
- [5] S. Waseem, S. A. R. S. Abu Bakar, B. A. Ahmed, Z. Omar, T. A. E. Eisa, and M. E. E. Dalam, "DeepFake on Face and Expression Swap: A Review," *IEEE Access*, vol. 11, pp. 117865–117906, 2023, doi: 10.1109/access.2023.3324403.
- [6] "Artificial Intelligence: Deepfakes in the Entertainment Industry." Accessed: July 16, 2025. [Online]. Available: <https://www.wipo.int/en/web/wipo-magazine/article-details/?assetRef=42620&title=artificial-intelligence-deepfakes-in-the-entertainment-industry>
- [7] W. Knight, "This AI Makes Robert De Niro Perform Lines in Flawless German," *Wired*. Accessed: July 22, 2025. [Online]. Available: <https://www.wired.com/story/ai-makes-de-niro-perform-lines-flawless-german/>
- [8] "Our Mission for Great Visual Storytelling | Flawless." Accessed: July 22, 2025. [Online]. Available: <https://flawlessai.com/mission>
- [9] J. Roe, M. Perkins, and L. Furze, "Deepfakes and Higher Education: A Research Agenda and Scoping Review of Synthetic Media," *J. Univ. Teach. Learn. Pract.*, vol. 21, no. 10, Nov. 2024, doi: 10.53761/2y2np178.
- [10] J. Kietzmann, L. W. Lee, I. P. McCarthy, and T. C. Kietzmann, "Deepfakes: Trick or treat?," *Bus. Horiz.*, vol. 63, no. 2, pp. 135–146, Mar. 2020, doi: 10.1016/j.bushor.2019.11.006.
- [11] A. Van Minnen, F. J. J. Ter Heide, T. Koolstra, A. De Jongh, S. Karaoglu, and T. Gevers, "Initial development of perpetrator confrontation using deepfake technology in victims with sexual violence-related PTSD and moral injury," *Front. Psychiatry*, vol. 13, Aug. 2022, doi: 10.3389/fpsy.2022.882957.

- [12] U. van Amsterdam, "Deepfakes with therapeutic effect - Informatics Institute," Informatics Institute - University of Amsterdam. Accessed: July 22, 2025. [Online]. Available: <https://ivi.uva.nl/content/news/2024/09/deepfakes-with-therapeutic-effect.html>
- [13] Y. Lu, "What brands should know before hiring a digital celebrity | Vogue Business." Accessed: July 22, 2025. [Online]. Available: <https://www.voguebusiness.com/story/technology/what-brands-should-know-before-hiring-a-digital-celebrity>
- [14] "Deepfake Technology and Its Implications for Influencer Marketing," in *Advances in Information Security, Privacy, and Ethics*, IGI Global, 2024, pp. 66–90. doi: 10.4018/979-8-3693-5298-4.ch005.
- [15] "Increasing Threats of Deepfake Identities," US Department of Homeland Security.
- [16] "Hundreds of laptops, bank accounts linked to North Korean fake IT workers scheme seized in major crackdown," POLITICO. Accessed: July 22, 2025. [Online]. Available: <https://www.politico.com/news/2025/06/30/justice-department-north-korea-it-workers-00433744>
- [17] Á. F. Gambín, A. Yazidi, A. Vasilakos, H. Haugerud, and Y. Djenouri, "Deepfakes: current and future trends," *Artif. Intell. Rev.*, vol. 57, no. 3, Feb. 2024, doi: 10.1007/s10462-023-10679-x.
- [18] E. Wilson, F. Shic, S. Jörg, and E. Jain, "Towards mitigating uncanny(eye)ness in face swaps via gaze-centric loss terms," Feb. 05, 2024, *arXiv: arXiv:2402.03188*. doi: 10.48550/arXiv.2402.03188.
- [19] D. Milmo, A. Hern, and R. Cousins, "'Inceptionism' and Balenciaga popes: a brief history of deepfakes," *The Guardian*, Apr. 08, 2024. Accessed: July 22, 2025. [Online]. Available: <https://www.theguardian.com/technology/2024/apr/08/inceptionism-and-balenciaga-popes-a-brief-history-of-deepfakes>
- [20] R. Tait, "US senator targeted by deepfake caller posing as Ukrainian diplomat," *The Guardian*, Sept. 26, 2024. Accessed: July 23, 2025. [Online]. Available: <https://www.theguardian.com/us-news/2024/sep/26/ben-cardin-dmytro-kuleba-deepfake-ukraine>
- [21] M. Lee, "Impostor uses AI to impersonate Rubio and contact foreign and US officials," AP News. Accessed: July 23, 2025. [Online]. Available: <https://apnews.com/article/rubio-artificial-intelligence-impersonation-1b3cc78464404b54e63f4eba9dd4f5a9>
- [22] "Taylor Swift deepfakes spread online, sparking outrage - CBS News." Accessed: July 23, 2025. [Online]. Available: <https://www.cbsnews.com/news/taylor-swift-deepfakes-online-outrage-artificial-intelligence/>
- [23] "Gobsmacked Molly-Mae targeted by AI scam as she issues warning to fans," *The Sun*. Accessed: July 23, 2025. [Online]. Available: <https://www.thesun.co.uk/tvandshowbiz/35954362/molly-mae-targeted-ai-scam-warning-tiktok/>
- [24] H. Chen and K. Magramo, "Finance worker pays out \$25 million after video call with deepfake 'chief financial officer' | CNN." Accessed: July 23, 2025. [Online]. Available: <https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html>
- [25] "Generative AI is expected to magnify the risk of deepfakes and other fraud in banking," Deloitte Insights. Accessed: July 23, 2025. [Online]. Available: <https://www.deloitte.com/us/en/insights/industry/financial-services/deepfake-banking-fraud-risk-on-the-rise.html>
- [26] L. S. F. Lin, "Examining the Role of Deepfake Technology in Organized Fraud: Legal, Security, and Governance Challenges," *Front. Law*, pp. 6–17, 2025.
- [27] D. Primack, "Deel and Rippling's espionage fight keeps escalating," *Axios*. Accessed: July 23, 2025. [Online]. Available: <https://www.axios.com/2025/06/06/vc-court-deel-rippling-hr-software>
- [28] "Defining Insider Threats | CISA." Accessed: July 16, 2025. [Online]. Available: <https://www.cisa.gov/topics/physical-security/insider-threat-mitigation/defining-insider-threats>
- [29] K. Freifeld and J. Stempel, "Former Google engineer indicted for stealing AI secrets to aid Chinese companies," *Reuters*, Mar. 06, 2024. Accessed: July 23, 2025. [Online]. Available: <https://www.reuters.com/technology/former-google-engineer-indicted-stealing-ai-secrets-aid-chinese-companies-2024-03-06/>
- [30] Joint Task Force Transformation Initiative, "Guide for conducting risk assessments," National Institute of Standards and Technology, Gaithersburg, MD, 2012. doi: 10.6028/nist.sp.800-30r1.
- [31] J. Tidy, "Firm hacked after accidentally hiring North Korean cyber criminal." Accessed: July 24, 2025. [Online]. Available: <https://www.bbc.com/news/articles/ce8vedz4yk7o>
- [32] B. Johnson, "Your Favorite New Coworker Is an AI-Enhanced Operative From North Korea," *Wired*. Accessed: July 24, 2025. [Online]. Available: <https://www.wired.com/story/north-korea-stole-your-tech-job-ai-interviews/>
- [33] R. Wright, "State-linked hackers deploy macOS malware in fake job interview campaign | Cybersecurity Dive." Accessed: July 24, 2025. [Online]. Available: <https://www.cybersecuritydive.com/news/north-korean-hackers-fake-interview/739165/>
- [34] E. Gordenker, "False Face: Unit 42 Demonstrates the Alarming Ease of Synthetic Identity Creation," Unit 42. Accessed: July 22, 2025. [Online]. Available: <https://unit42.paloaltonetworks.com/north-korean-synthetic-identity-creation/56:hhhh>
- [35] M. Miller and D. Nickel, "Tech companies have a big remote worker problem: North Korean operatives," POLITICO. Accessed: July 24, 2025. [Online]. Available: <https://www.politico.com/news/2025/05/12/north-korea-remote-workers-us-tech-companies-00340208>
- [36] R. Lakshmanan, "BlueNoroff Deepfake Zoom Scam Hits Crypto Employee with macOS Backdoor Malware," *The Hacker News*. Accessed: July 24, 2025. [Online]. Available: <https://thehackernews.com/2025/06/bluenoroff-deepfake-zoom-scam-hits.html>
- [37] R. Half, "Remote Work Statistics and Trends for 2025," Robert Half. Accessed: July 24, 2025. [Online]. Available: <https://www.roberthalf.com/us/en/insights/research/remote-work-statistics-and-trends>
- [38] "The State of Hybrid and Remote Work: Q1." Accessed: July 24, 2025. [Online]. Available: <https://press.roberthalf.com/The-State-of-Hybrid-and-Remote-Work-Q1-2025>
- [39] "Industry employment by state, seasonally adjusted," Bureau of Labor Statistics. Accessed: July 25, 2025. [Online]. Available: <https://www.bls.gov/charts/state-employment-and-unemployment/industry-employment-by-state.htm>
- [40] "States: Persons at work by telework status, 2023 annual averages," Bureau of Labor Statistics. Accessed: July 25, 2025. [Online]. Available: <https://www.bls.gov/lau/state-telework-table.htm>
- [41] Joint Task Force Transformation Initiative, "Risk management framework for information systems and organizations: a system life cycle approach for security and privacy," National Institute of Standards and Technology, Gaithersburg, MD, Dec. 2018. doi: 10.6028/nist.sp.800-37r2.
- [42] Joint Task Force Interagency Working Group, "Security and Privacy Controls for Information Systems and Organizations," National Institute of Standards and Technology, Gaithersburg, MD, Sept. 2020. doi: 10.6028/nist.sp.800-53r5.
- [43] "National Insider Threat Task Force (NITTF)." Accessed: July 24, 2025. [Online]. Available: <https://www.dni.gov/index.php/ncsc-how-we-work/ncsc-nitff>
- [44] National Insider Threat Task Force, "Guide to Accompany the National Insider Threat Policy and Minimum Standards." National Counterintelligence and Security Center, 2014.
- [45] "Cybersecurity Maturity Model Certification (CMMC) Model Overview." US Department of Defense - Chief Information Officer, Sept. 2024.
- [46] CERT National Insider Threat Center, "Common Sense Guide to Mitigating Insider Threats, Sixth Edition," Carnegie Mellon University - Software Engineering Institute, Dec. 2018.

- [47] "Deepfake Detection Challenge Dataset." Accessed: July 23, 2025. [Online]. Available: <https://ai.meta.com/datasets/dfdc>
- [48] "Open Media Forensics Challenge | NIST." Accessed: July 23, 2025. [Online]. Available: <https://www.nist.gov/itl/iad/mig/open-media-forensics-challenge>
- [49] H. Guan, Y. Lee, L. Diduch, and I. Ghorbanian, "OpenMFC 2020-2021 Results." Multimodal Information Group, Information Access Division, ITL, NIST, Dec. 08, 2021.
- [50] "The Essential Guide to UEBA," Splunk. Accessed: July 24, 2025. [Online]. Available: https://www.splunk.com/en_us/form/the-essential-guide-to-ueba.html
- [51] "Insider Threat Software | Gurukul Insider Threat Management," Gurukul. Accessed: July 24, 2025. [Online]. Available: <https://gurukul.com/solutions/insider-threat-management-solution/>
- [52] G. M. Dimitri, "A Short Survey on Deep Learning for Multimodal Integration: Applications, Future Perspectives and Challenges," *Computers*, vol. 11, no. 11, p. 163, Nov. 2022, doi: 10.3390/computers11110163.