International Journal of
Technology and
Systems
(IJTS)

Deepfake-as-a-Service: The Next Challenge for Enterprise Cybersecurity

Journal of Technology and Systems
ISSN: 2788-6344 (Online)
Vol. 7, Issue No. 3, pp 47 – 59, 2025
www.carijournals.org

Deepfake-as-a-Service: The Next Challenge for Enterprise Cybersecurity

Nkiru Ali Suleiman

Department of Engineering, University of Lagos

https://orcid.org/0009-0002-4389-0181

Accepted: 24th Mar, 2025, Received in Revised Form: 24th Apr, 2025, Published: 24th May, 2025

Abstract

Purpose: This study investigates the emerging threat of Deepfake-as-a-Service (DFaaS) and its implications for enterprise cybersecurity. It aims to explore how the commodification of synthetic media is reshaping threat landscapes, operational vulnerabilities, and strategic responses within organizational settings.

Methodology: An exploratory qualitative design was adopted, combining thematic document analysis with expert commentary. A corpus of 85 peer-reviewed articles, industry reports, and threat intelligence briefs published between 2018 and 2024 was systematically reviewed. NVivo software facilitated thematic clustering, while Braun and Clarke's six-phase framework guided data coding and interpretation.

Findings: The study reveals five key themes: (1) DFaaS platforms are increasingly accessible and sophisticated, (2) enterprises face specific threats such as executive impersonation, internal disinformation, and financial fraud, (3) current detection infrastructures are inadequate against adversarial synthetic media, (4) corporate governance and policy responses are fragmented and reactive, and (5) resilience strategies such as content provenance, cross-channel verification, and employee training are emerging but unevenly adopted.

Unique Contribution to Theory, Practice, and Policy: Theoretically, the study introduces the construct of *Cognitive Authenticity* to reframe cybersecurity around perceptual integrity. In practice, it outlines a DFaaS Attack Lifecycle Model and recommends enterprise-level operational shifts to address synthetic threats. From a policy standpoint, it advocates for legal recognition of synthetic identity fraud, mandates for content authenticity standards, and international cooperation on AI-enabled cyber norms. These contributions bridge technical, organizational, and regulatory perspectives in confronting synthetic deception.

Keywords: *Deepfake-as-a-Service (DFaaS), Enterprise Cybersecurity, Synthetic Media, Cognitive Authenticity, Digital Deception, Impersonation Threats, AI Regulation, Threat Modeling.*

1.0 Introduction

According to Sandoval et al. (2024), the emergence of deepfakes presents one of the most insidious threats to digital trust and information authenticity in an era increasingly defined by the pervasive influence of artificial intelligence. Sandoval et al. (2024) adds that deepfakes is a synthetically generated or manipulated media using sophisticated machine learning algorithms, particularly deep learning which have evolved from academic curiosities into formidable tools of deception. Initially relegated to entertainment and benign experimentation, their capabilities have matured to a point where malicious applications now proliferate across various digital platforms. Matli (2024) explains that the particular concern is the rise of "Deepfake-as-a-Service" (DFaaS), a commodified model where highly realistic synthetic media can be generated on-demand via cloud-based services that has with minimal technical expertise required by end-users. This democratization of deepfake technology significantly expands its threat surface, rendering conventional cybersecurity paradigms increasingly obsolete. For enterprises, this evolution signals a paradigm shift in the nature of cyber threats. No longer confined to technical breaches or phishing schemes, enterprises must now reckon with threats targeting reputational integrity, stakeholder trust, and even internal governance structures through fabricated media. The weaponization of DFaaS poses significant risks ranging from executive impersonation and corporate espionage to market manipulation and supply chain disruption (Matli, 2024). The enterprises digitize further and hybrid work models prevail when the attack vectors exposed by DFaaS become more varied and difficult to anticipate. This paper critically explores the intersection of DFaaS and enterprise cybersecurity, arguing that the current response mechanisms are insufficiently robust to address the unique challenges posed by synthetic media. Therefore, it is necessary to conduct a rigorous examination of its implications, countermeasures and regulatory dimensions.

2.0 Purpose

The primary purpose of this paper is to interrogate the evolving threat landscape shaped by the proliferation of Deepfake-as-a-Service (DFaaS) and to critically assess its implications for enterprise cybersecurity frameworks. The purpose about deepfakes has traditionally focused around political misinformation and social manipulation it also reorients the focus toward corporate vulnerability. However, the particular emphasis will be on the strategic, operational, and technological dimensions of enterprise security. The research aims to fill a growing gap in cybersecurity literature in the examining how DFaaS commodifies synthetic deception to lowering the barriers to cyber-enabled corporate sabotage.

This study seeks to achieve four interrelated objectives. First, it aims to delineate the operational mechanics of DFaaS, highlighting the commercial ecosystem that facilitates the generation and distribution of deepfakes. Second, it endeavors to identify specific threat vectors relevant to enterprises, such as executive impersonation, fraudulent communications and reputational attacks. Third, the paper evaluates existing cybersecurity protocols to determine their efficacy in detecting and mitigating deepfake-based threats. Finally, the study aspires to inform the development of

comprehensive countermeasures both technical and policy-oriented that are responsive to the distinctive nature of synthetic media threats. The study will draw from interdisciplinary literature, real-world case studies and theoretical models, this paper aspires not only to map the contours of a rapidly escalating cybersecurity challenge but also to offer strategic insights that can guide enterprise resilience planning. Therefore, the goal is to advance a nuanced understanding of DFaaS as a critical inflection point in the cybersecurity domain one that demands immediate scholarly, professional and regulatory attention.

3.0 Literature Review

The literature surrounding deepfakes has grown significantly in recent years by reflecting mounting concerns about their implications across social, political, and economic spheres. This section critically synthesizes existing scholarly, technical, and policy-focused literature, with a particular focus on the rise of Deepfake-as-a-Service (DFaaS) and its intersection with enterprise cybersecurity.

3.1 Defining Deepfakes and Their Technological Foundations

Kaur et al. (2024) defines deepfakes as digitally altered media predominantly video, audio, or images created using generative adversarial networks (GANs) or similar machine learning architectures. Khan et al. (2024) introduced GANs as a novel framework wherein two neural networks as a generator and a discriminator compete to resulting in highly realistic outputs. Subsequent advancements have refined these architectures, enabling rapid generation of lifelike forgeries with minimal computational cost (Ramanaharan et al., 2025). The originally developed for academic and entertainment purposes on the application of GANs for deceptive intent has raised ethical and security-related alarms. (Momeni, 2024) acknowledges that deepfakes have transitioned from experimental novelty to operational tools for social engineering and disinformation. However, much of the existing literature is still primarily concerned with sociopolitical implications such as electoral manipulation or celebrity pornography to leaving enterprise-centric vulnerabilities underexplored.

3.2 The Emergence of Deepfake-as-a-Service (DFaaS)

According to Riphagen & of Twente (2022), the concept of Deepfake-as-a-Service (DFaaS) developed as a natural progression in the commodification of artificial intelligence tools in particularly those used for synthetic media generation. Riphagen & of Twente (2022) continues that the origin of DFaaS can be traced to the convergence of open-source machine learning frameworks and the growing availability of cloud-based AI infrastructure in the late 2010s. Sandoval et al. (2024) records that deepfake technologies were initially developed for research and creative applications because they quickly attracted attention for their potential misuse especially following the widespread release of deep learning frameworks such as DeepFaceLab and FaceSwap. Such tools enabled the manipulation of facial features and voices with increasing realism, thus laying the technical foundation for DFaaS.

The transition from individual experimentation to service-based delivery models occurred around 2019–2020. Wazid et al. (2024) states that during this period, actors on darknet forums and illicit marketplaces began offering synthetic media creation as a paid service. These platforms adopted a model similar to Ransomware-as-a-Service (RaaS), where users could request custom video or audio fabrications without needing technical expertise (Riphagen & of Twente, 2022). Early DFaaS offerings were rudimentary, limited to facial swapping or voice modulation because advancements in GANs (Generative Adversarial Networks) and speech synthesis technologies rapidly increased the quality and diversity of outputs. Wazid et al. (2024) adds that after the period, DFaaS services began appearing on the darknet and on surface web applications, often disguised as entertainment or parody tools.

The initial reaction to the rise of DFaaS was characterized by a mix of concern and uncertainty within academic, legal, and cybersecurity communities. Javaid et al. (2023) states that the ethical implications of deepfakes had already been widely debated since DFaaS introduced a new dimension by removing technical barriers and enabling mass deployment. Ainslie et al. (2023) recognized the potential for such services to be weaponized for fraud, misinformation, impersonation and reputational damage. However, comprehensive research on DFaaS remained limited, largely due to the difficulty of accessing and studying these services in real time.

The impact of DFaaS on enterprise cybersecurity has become increasingly evident. Mishra et al. (2022) emphasised on how deepfakes is likely to produce through service-based platforms that have been used to manipulate employees, deceive stakeholders and conduct financial fraud. For example, in 2020, the FBI issued a warning about the use of deepfake videos in business email compromise (BEC) schemes (Mishra et al., 2022). The availability of DFaaS significantly magnifies this threat by making it scalable and repeatable across different industries and regions. Furthermore, DFaaS has contributed to the broader erosion of trust in digital communications. Admass et al. (2024) found that traditional cyberattacks that target networks and devices but deepfake-enabled attacks compromise human perception in making them difficult to detect with conventional cybersecurity tools. Thus, DFaaS represents a shift in the threat landscape from technical intrusion to cognitive deception. Therefore, DFaaS emerged from the intersection of open-source AI development and underground commercialization. Its evolution has been marked by increasing sophistication, accessibility and impact (Li & Liu, 2021). The reaction to its emergence has underscored significant gaps in detection, regulation, and organizational preparedness.

3.3 Enterprise Threat Vectors and Use Cases

The application of DFaaS in enterprise contexts is complicated in issues like *executive impersonation*, *fraudulent financial requests*, *reputational sabotage*, and *internal disinformation*. The infamous case of a UK-based energy firm defrauded of $243,000 using AI-generated audio mimicking the CEO's voice is frequently cited as a cautionary tale (Salahdine & Kaabouch, 2019). Social engineering has long exploited human trust, but DFaaS adds a visual and auditory

dimension that traditional email phishing or spoofing lacks. According to Javaid et al. (2023), synthetic media's realism undermines cognitive heuristics that individuals use to assess credibility in increasing the success rate of cyberattacks. Thus, in hybrid or remote working environments, where video conferencing and digital communications are paramount the opportunities for exploitation are magnified. Supply chains and mergers & acquisitions (M&A) are also vulnerable. Synthetic audio and video could be used to disseminate false information to manipulate stock prices and disrupt due diligence processes (Mishra et al., 2022). Thus, there is limited empirical research quantifying the frequency or impact of DFaaS in corporate settings a gap this paper identifies as ripe for further investigation.

3.4 Detection and Mitigation Techniques

Efforts to counter deepfakes are divided into *proactive detection* and *reactive mitigation*. Detection algorithms typically use statistical anomalies in facial movements inconsistencies in lighting and irregular audio spectrograms (Vaccari & Chadwick, 2020). However, as generative models evolve, they increasingly produce outputs that evade these detection methods. Deep learning-based detectors, such as FaceForensics++ have shown promise in their performance degrades when applied to novel synthesis techniques or compressed media (Vecchietti et al., 2025). Moreover, detection tools often require raw data, making real-time applications such as identifying a fraudulent video call particularly challenging. According to Hasan et al. (2024), existing models struggle in adversarial environments, where attackers may deliberately obfuscate artifacts. On the mitigation front, watermarking and provenance verification such using blockchain have been proposed. Adobe and Microsoft's Content Authenticity Initiative seeks to embed metadata that authenticates content origin (Almahmoud et al., 2025). The theoretically promising aspects such systems depend on widespread industry adoption and are vulnerable to circumvention if not standardized. Therefore, there is a growing advocacy for multi-modal detection wherein video, audio, textual, and behavioral data are analyzed concurrently to assess authenticity. However, such approaches are computationally intensive and may not scale easily for enterprise-wide deployment.

3.5 Regulatory and Ethical Dimensions

Fallis (2021) the regulatory perspective indicate that most jurisdictions remain ill-equipped in addressing the issues surrounding DFaaS. While the EU's Digital Services Act and the U.S. DEEPFAKES Accountability Act offer preliminary frameworks, enforcement remains challenging due to jurisdictional ambiguity, the anonymity of service providers, and the rapid evolution of technology (Ray, 2021). Moreover, laws often fail to distinguish between malicious and benign applications of synthetic media. Ethically, the dual-use nature of deepfakes complicates regulation. According to (Ali et al., 2022), the same technologies used for malicious impersonation can also serve in film production, accessibility tools, and digital preservation. Policymakers are thus caught in a delicate balancing act between fostering innovation and curbing misuse. Corporate governance structures are also ill-prepared. Few organizations have policies or response protocols

explicitly addressing synthetic media threats. A recent survey by Belur et al. (2021) indicated that while 78% of CISOs acknowledged deepfakes as a concern, less than 20% had integrated detection mechanisms or staff training modules into their cybersecurity strategies.

3.6 Gaps in Existing Literature

There are several research gaps that have continued to exist despite the growing awareness. First, empirical data on the frequency, typology, and financial impact of DFaaS attacks on enterprises is limited. Most existing studies rely on anecdotal evidence or theoretical modeling, which constrains generalizability. Second, interdisciplinary engagement is lacking. The technical studies abound but fewer works bridge cybersecurity, organizational behavior, legal theory, and media ethics. This limits the development of holistic countermeasures. Third, there is insufficient focus on preparedness and resilience frameworks. The dominant works is reactive focused on detection rather than proactive strategies like simulation training, digital literacy, or policy integration. Finally, few studies consider the cultural and geopolitical dimensions of DFaaS. Attackers may exploit cultural cues or linguistic idiosyncrasies to enhance believability, yet these dynamics remain underexplored in both Western and Global South contexts.

4.0 Methodology

4.1 Research Design and Rationale

Exploratory qualitative research is particularly suited for investigating phenomena that are insufficiently understood or rapidly evolving (Allen, 2021). The choice of this design reflects the current state of DFaaS scholarship, which lacks established theoretical frameworks and robust empirical metrics. The integrating thematic document analysis with expert commentary helps the study triangulates multiple data sources to develop a comprehensive understanding of the DFaaS threat landscape.

4.2 Data Sources

The analysis in this study is primarily grounded in systematic document analysis that drawing from a curated body of literature and industry reports. A total of 85 documents were reviewed, comprising peer-reviewed journal articles, white papers, threat intelligence briefings, and regulatory publications issued between 2018 and 2024. This time frame was selected to capture both the developmental trajectory and contemporary implications of Deepfake-as-a-Service (DFaaS) within enterprise cybersecurity contexts. Data sources were identified through targeted searches in reputable academic databases, including Scopus, IEEE Xplore, and JSTOR, alongside specialized cybersecurity archives such as MITRE ATT&CK and the ENISA Threat Landscape Reports. Keyword combinations used in the search strategy included “deepfake cybersecurity,” “synthetic media threats,” “Deepfake-as-a-Service,” “generative AI threats,” and “AI-driven impersonation.” Boolean operators and filters for publication year, peer-review status, and subject relevance were applied to ensure precision in retrieval.

The inclusion criteria for document selection were threefold that include (1) relevance to enterprise

settings, with a focus on organizational impact and risk posture; (2) methodological rigor, including transparent analytical frameworks and empirical grounding; and (3) recency, to ensure engagement with the most current technological and threat developments. Documents that lacked verifiability, generalizability, or demonstrated overt bias were excluded from the final sample. This multi-source approach enabled a triangulated understanding of DFaaS, incorporating both theoretical insight and operational intelligence, thereby strengthening the reliability and contextual richness of the study's findings.

4.3 Data Analysis Techniques

Thematic analysis was employed to identify recurring patterns and emergent themes across the document corpus and interview transcripts. Following Braun and Clarke's (2006) six-phase framework, data were coded inductively and categorized into five overarching themes: (1) nature of DFaaS technologies, (2) enterprise-specific threat vectors, (3) limitations of current detection tools, (4) governance and policy responses, and (5) emergent resilience strategies (Adewole et al., 2024). NVivo 14 was used to facilitate data organization, thematic clustering, and frequency analysis (Belen-Saglam et al., 2023). A matrix coding query was conducted to map intersections between stakeholder types such technical vs. managerial and thematic emphases, allowing for nuanced insights into how perceptions of DFaaS threats diverge across professional roles.

4.4 Ethical Considerations

This study was conducted in full alignment with institutional ethical standards, with particular attention paid to the responsible use of publicly available information. As the research employed document analysis as its primary methodological approach, no direct interaction with human participants occurred, thereby negating the need for informed consent or anonymization procedures typically associated with interviews (Guandalini, 2022). Therefore, ethical considerations were central to the treatment of all sourced materials. Documents analyzed were including academic publications, industry white papers, regulatory texts, threat intelligence reports, and publicly accessible DFaaS-related platforms they were selected based on their credibility, relevance, and open accessibility. No proprietary or classified documents were accessed or disclosed. Descriptions of DFaaS tools and platforms were kept at a conceptual level, omitting any operational specifics that might be weaponized by threat actors. Thus, the research adhered to the principle of "ethical non-amplification," ensuring that the documentation of malicious capabilities did not contribute to their further diffusion.

4.5 Limitations

While the qualitative design enables deep contextual analysis, its findings are not statistically generalizable. Moreover, the fast-paced evolution of DFaaS tools may outpace some of the insights presented herein. However, the emphasis on thematic depth and practitioner input ensures that the study offers robust, actionable knowledge for both scholarly and applied contexts.

5.0 Findings

This section presents the key findings from the qualitative analysis, organized into five interrelated themes: the operational dynamics of DFaaS, enterprise-specific threat vectors, limitations in detection infrastructure, policy and governance inadequacies, and emergent resilience strategies. These findings reflect a synthesis of insights from academic literature, threat intelligence reports, and practitioner experiences.

5.1 Operational Dynamics of Deepfake-as-a-Service (DFaaS)

One of the most salient findings from both document analysis and expert interviews is the increasing sophistication and accessibility of DFaaS platforms. Unlike bespoke deepfake creation which requires high technical skill and computing resources DFaaS commodifies synthetic deception by offering user-friendly, cloud-hosted interfaces (Javaid et al., 2023). These services often allow users to upload basic inputs such as headshots, voice samples, or script prompts, and in return receive highly realistic multimedia outputs. Experts emphasized that many DFaaS operations are indistinguishable from legitimate AI service providers (Ainslie et al., 2023). This duality complicates takedown efforts and allows malicious actors to operate within regulatory grey zones. Additionally, the proliferation of APIs and machine learning-as-a-service (MLaaS) frameworks has facilitated modular deployment, meaning that threat actors can scale attacks quickly and discreetly. A study on a cybersecurity noted that DFaaS is able to disgruntle an employee or an industrial competitor by spinning up to convincing CEO impersonation in less than an hour (Lucas, 2022).

5.2 Enterprise-Specific Threat Vectors

The most frequently identified threat vector is executive impersonation, wherein threat actors use DFaaS to simulate C-suite executives in emails, video calls, or voicemail messages. Studies cited this as particularly dangerous due to the implicit trust placed in high-level leadership. One incident reported by a respondent involved a deepfake video call impersonating a CFO requesting an urgent funds transfer fortunately intercepted due to an unusually timed request (Allen, 2021). Beyond impersonation, financial fraud emerged as a second major threat category. DFaaS enables convincing synthetic voices or messages that can authorize fund movements, manipulate investor sentiment, or derail negotiations. For example, deepfaked press releases can tank or inflate stock prices within minutes before the truth is uncovered, a tactic increasingly aligned with "cyber pump-and-dump" schemes. A more insidious but less understood vector is internal disinformation, where deepfakes are used to spread false directives among employees, erode morale, or incite division (Sandoval et al., 2024). Particularly in politically sensitive or culturally diverse firms, these tactics can sow discord or disrupt organizational cohesion. Reputational sabotage was also frequently cited. A cybersecurity consultant highlighted a case where a fabricated video of a company executive making discriminatory remarks circulated briefly online, triggering a public relations crisis that required immediate digital forensics and legal intervention to resolve.

5.3 Detection Infrastructure Gaps

A recurring theme in both literature and interviews is that deepfake detection capabilities lag behind generative advancements. While detection algorithms such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have shown promise in controlled environments, they often fail when applied to compressed, obfuscated, or adversarially designed media (Vaccari & Chadwick, 2020). Many enterprises rely on legacy fraud detection systems that are optimized for text-based phishing or malware intrusion but are ill-equipped to parse multimodal content. A particularly worrying insight is that real-time detection such as during a live video call is still largely impractical at the enterprise scale (Vecchietti et al., 2025). Deepfake detectors that can assess facial micro expressions or vocal cadence require high-resolution input and latency-tolerant infrastructure both of which are often absent in standard business environments. The study indicates that every time a detector improves the generators evolve to beat it because DFaaS operators are actively benchmarking their outputs against open-source detection tools. Furthermore, cybersecurity teams often lack cross-functional training. While IT departments might deploy endpoint security or firewall protections, few possess the forensics training necessary to assess video or audio manipulation (Hasan et al., 2024). The result is a blind spot where deepfakes can propagate with minimal scrutiny.

5.4 Governance and Policy Inadequacies

Regulatory frameworks at both national and corporate levels were described as fragmented and reactionary. Most studies expressed that there were frustration that existing laws either do not recognize synthetic media as a distinct threat or fail to articulate liability in cases of enterprise harm caused by deepfakes (Almahmoud et al., 2025). Therefore, even within firms, there is often no clear incident response protocol for synthetic media threats. The vast majority of cybersecurity playbooks remain anchored in response to data breaches, ransomware, or denial-of-service attacks. Few include provisions for verifying audiovisual communications or initiating deepfake response drills. Moreover, compliance frameworks such as ISO/IEC 27001 or NIST do not yet mandate preparedness for deepfake threats, creating a regulatory lag (Salahdine & Kaabouch, 2019). Although some forward-looking sectorssuch as finance and defense have begun incorporating deepfake scenarios into threat modeling exercises, the practice remains limited.

5.5 Emergent Resilience Strategies

The are challenges but there are several promising resilience strategies are beginning to emerge. The most common among advanced organizations is the adoption of content provenance systems, wherein all internal and public-facing media are digitally signed and timestamped at creation (Javaid et al., 2023). Adobe’s Content Authenticity Initiative and Microsoft's Project Origin were frequently cited as pioneering examples, although implementation remains uneven (Ainslie et al., 2023). Second, enterprises are beginning to invest in employee awareness training, extending the scope of digital literacy to include deepfake recognition. Much like phishing simulation exercises, some firms have begun deploying synthetic media as part of red-team scenarios to train staff in identifying and reporting suspicious content. Another proactive measure involves multi-channel

verification protocols (Mishra et al., 2022). When critical communications are issued via video or audio, firms are advised to pair them with corroborating emails, secure digital signatures, or direct confirmation via secondary channels. This helps mitigate overreliance on any single form of media authenticity. A few cutting-edge security teams are exploring behavioral biometrics and AI counter-forensics using models trained on specific individuals' micro-expressions, speech cadence, and typing patterns to flag anomalies. While these systems are still experimental and raise privacy concerns, they represent a frontier in organizational deepfake resilience.

6.0 Recommendations

This study recommends the integration of *Cognitive Authenticity* into existing cyber threat models to address the perceptual manipulation risks introduced by Deepfake-as-a-Service (DFaaS) (Admass et al., 2024). Enterprises should urgently prioritize synthetic media threats on par with traditional cybersecurity concerns by establishing verification protocols, conducting regular simulation-based training, and adopting zero-trust principles that extend to communications (Li & Liu, 2021). Organizations are also encouraged to move beyond vendor-dependent solutions and build in-house capabilities for detecting and managing deepfakes. The study introduces the DFaaS Attack Lifecycle Model to guide proactive defense strategies and enhance red-teaming exercises (Wazid et al., 2024). On a policy level, there is a need for legislative updates to formally recognize AI-generated impersonation as a cybersecurity risk, with clear guidelines for liability and enforcement (Momeni, 2024). Regulatory bodies should establish authenticity standards for digital media, and foster public-private intelligence-sharing partnerships (Riphagen & of Twente, 2022). Lastly, international cooperation is essential to address the cross-border nature of DFaaS operations, with current treaties needing updates to reflect the realities of synthetic content abuse.

7.0 Conclusion

This study examined the rising phenomenon of Deepfake-as-a-Service (DFaaS) and its implications for enterprise cybersecurity. The research focused on how DFaaS commodifies synthetic deception, making sophisticated threats such as impersonation, fraud, and internal disinformation more accessible and scalable. Using qualitative methods, including thematic analysis of industry reports and expert commentary, the study identified significant gaps in detection infrastructure, governance, and enterprise preparedness. Key findings revealed that while DFaaS presents a perceptual threat rather than a technical one, most organizations remain ill-equipped to detect or respond to it effectively. The study introduced the concept of *Cognitive Authenticity* to reframe cybersecurity theory and provided a practical attack lifecycle model for enterprise resilience. Finally, the study met its intended purpose by offering both conceptual insights and actionable strategies for addressing synthetic media threats. Thus, addressing DFaaS requires a multidimensional approach that combines technological, organizational, and policy interventions.

References

Adewole, K. S., Alozie, E., Olagunju, H., Faruk, N., Aliyu, R. Y., Imoize, A. L., Abdulkarim, A., Imam-Fulani, Y. O., Garba, S., Baba, B. A., Hussaini, M., Oloyede, A. A., Abdullahi, A., Kanya, R. A., & Usman, D. J. (2024). A systematic review and meta-data analysis of clinical data repositories in Africa and beyond: recent development, challenges, and future directions. *Discover Data*, *2*(1). https://doi.org/10.1007/S44248-024-00012-4

Admass, W. S., Munaye, Y. Y., & Diro, A. A. (2024). Cyber security: State of the art, challenges and future directions. *Cyber Security and Applications*, *2*. https://doi.org/10.1016/j.csa.2023.100031

Ainslie, S., Thompson, D., Maynard, S., & Ahmad, A. (2023). Cyber-threat intelligence for security decision-making: A review and research agenda for practice. *Computers and Security*, *132*. https://doi.org/10.1016/j.cose.2023.103352

Ali, A., Khan Ghouri, K. F., Naseem, H., Soomro, T. R., Mansoor, W., & Momani, A. M. (2022). Battle of Deep Fakes: Artificial Intelligence Set to Become a Major Threat to the Individual and National Security. *International Conference on Cyber Resilience, ICCR 2022*. https://doi.org/10.1109/ICCR56254.2022.9995821

Allen, D. (2021). Deepfake Fight: AI-Powered Disinformation and Perfidy Under the Geneva Conventions. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.3958426

Almahmoud, Z., Yoo, P. D., Damiani, E., Choo, K. K. R., & Yeun, C. Y. (2025). Forecasting Cyber Threats and Pertinent Mitigation Technologies. *Technological Forecasting and Social Change*, *210*. https://doi.org/10.1016/j.techfore.2024.123836

Belen-Saglam, R., Altuncu, E., Lu, Y., & Li, S. (2023). A systematic literature review of the tension between the GDPR and public blockchain systems. *Blockchain: Research and Applications*, *4*(2). https://doi.org/10.1016/j.bcra.2023.100129

Belur, J., Tompson, L., Thornton, A., & Simon, M. (2021). Interrater Reliability in Systematic Review Methodology: Exploring Variation in Coder Decision-Making. *Sociological Methods and Research*, *50*(2), 837–865. https://doi.org/10.1177/0049124118799372

Fallis, D. (2021). The Epistemic Threat of Deepfakes. *Philosophy Technology*, *34*(4), 623–643. https://doi.org/10.1007/s13347-020-00419-2

Guandalini, I. (2022). Sustainability through digital transformation: A systematic literature review for research guidance. *Journal of Business Research*, *148*, 456–471. https://doi.org/10.1016/j.jbusres.2022.05.003

Hasan, H. R., Salah, K., Jayaraman, R., Yaqoob, I., & Omar, M. (2024). NFTs for combating deepfakes and fake metaverse digital contents. *Internet of Things (Netherlands)*, *25*. https://doi.org/10.1016/j.iot.2024.101133

Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2023). Towards insighting cybersecurity for healthcare domains: A comprehensive review of recent practices and trends. *Cyber Security and Applications*, *1*. https://doi.org/10.1016/j.csa.2023.100016

Kaur, A., Noori Hoshyar, A., Saikrishna, V., Firmin, S., & Xia, F. (2024). Deepfake video detection: challenges and opportunities. *Artificial Intelligence Review*, *57*(6). https://doi.org/10.1007/S10462-024-10810-6

Khan, R., Sohail, M., Usman, I., Sandhu, M., Raza, M., Yaqub, M. A., & Liotta, A. (2024). Comparative study of deep learning techniques for DeepFake video detection. *ICT Express*. https://doi.org/10.1016/j.icte.2024.09.018

Li, Y., & Liu, Q. (2021). A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments. *Energy Reports*, *7*, 8176–8186. https://doi.org/10.1016/j.egyr.2021.08.126

Lucas, K. T. (2022). Deepfakes and Domestic Violence: Perpetrating Intimate Partner Abuse Using Video Technology. *Victims and Offenders*, *17*(5), 647–659. https://doi.org/10.1080/15564886.2022.2036656

Matli, W. (2024). Extending the theory of information poverty to deepfake technology. *International Journal of Information Management Data Insights*, *4*(2). https://doi.org/10.1016/j.jjimei.2024.100286

Mishra, A., Alzoubi, Y. I., Anwar, M. J., & Gill, A. Q. (2022). Attributes impacting cybersecurity policy development: An evidence from seven nations. *Computers and Security*, *120*. https://doi.org/10.1016/j.cose.2022.102820

Momeni, M. (2024). Artificial Intelligence and Political Deepfakes: Shaping Citizen Perceptions through Misinformation. *Journal of Creative Communications*. https://doi.org/10.1177/09732586241277335

Ramanaharan, R., Guruge, D. B., & Agbinya, J. I. (2025). DeepFake video detection: Insights into model generalisation — A Systematic review. *Data and Information Management*. https://doi.org/10.1016/j.dim.2025.100099

Ray, A. (2021). Disinformation, deepfakes and democracies: The need for legislative reform. *The University of New South Wales Law Journal*, *44*(3), 983–1013. https://doi.org/10.53637/dels2700

Riphagen, Q., & of Twente, U. (2022). *The deepfake problem*.

Salahdine, F., & Kaabouch, N. (2019). Social Engineering Attacks: A Survey. *Future Internet 2019, Vol. 11, Page 89*, *11*(4), 89. https://doi.org/10.3390/FI11040089

Sandoval, M. P., de Almeida Vau, M., Solaas, J., & Rodrigues, L. (2024a). Threat of deepfakes to the criminal justice system: a systematic review. *Crime Science*, *13*(1). https://doi.org/10.1186/S40163-024-00239-1

Sandoval, M. P., de Almeida Vau, M., Solaas, J., & Rodrigues, L. (2024b). Threat of deepfakes to the criminal justice system: a systematic review. *Crime Science*, *13*(1). https://doi.org/10.1186/S40163-024-00239-1

Sandoval, M. P., de Almeida Vau, M., Solaas, J., & Rodrigues, L. (2024c). Threat of deepfakes to the criminal justice system: a systematic review. *Crime Science*, *13*(1), 1–16. https://doi.org/10.1186/S40163-024-00239-1/FIGURES/3

Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media and Society*, *6*(1). https://doi.org/10.1177/2056305120903408

Vecchietti, G., Liyanaarachchi, G., & Viglia, G. (2025). Managing deepfakes with artificial

intelligence: Introducing the business privacy calculus. *Journal of Business Research, 186*. https://doi.org/10.1016/j.jbusres.2024.115010

Wazid, M., Mishra, A. K., Mohd, N., & Das, A. K. (2024). A Secure Deepfake Mitigation Framework: Architecture, Issues, Challenges, and Societal Impact. *Cyber Security and Applications, 2*. https://doi.org/10.1016/j.csa.2024.100040